

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНО-ПРАКТИЧЕСКИЙ ЖУРНАЛ



МИКРОКОНТРОЛЛЕРЫ

со склада в Санкт-Петербурге

UBICOM, CYGNAL, FUJITSU, GOAL, ATMEL, MOTOROLA, AMD

UBICOM



SX20AC, SX28AC, SX52BD, IP2022-160

ЭЛЕКТРОСНАБ

Официальный представитель
UBICOM, CYGNAL, GOAL в России



2-3/2003

2-3/2003

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель и издатель:

ФГУП «Издательство «Политехника»»

Главный редактор

М. Б. Сергеев,
доктор технических наук, профессор

Зам. главного редактора

Г. Ф. Мощенко

Редакционный совет:

Председатель А. А. Оводенко,
доктор технических наук, профессор

В. Н. Васильев,
доктор технических наук, профессор

В. Н. Козлов,
доктор технических наук, профессор

Ю. Ф. Подоплекин,
доктор технических наук, профессор

Д. В. Пузанков,
доктор технических наук, профессор

В. В. Симаков,
доктор технических наук, профессор

А. Л. Фрадков,
доктор технических наук, профессор

Л. И. Чубраева,
доктор технических наук, профессор, чл.-корр. РАН

Р. М. Юсупов,
доктор технических наук, профессор

Редакционная коллегия:

В. Г. Анисимов,
доктор технических наук, профессор

В. Ф. Мелехин,
доктор технических наук, профессор

А. В. Смирнов,
доктор технических наук, профессор

В. А. Фетисов,
доктор технических наук, профессор

В. И. Хищенко,
доктор технических наук, профессор

А. А. Шальто,
доктор технических наук, профессор

А. П. Шепета,
доктор технических наук, профессор

З. М. Юлдашев,
доктор технических наук, профессор

Редактор: О. А. Рубинова

Корректоры: Т. Н. Гринчук, Е. П. Смирнова

Дизайн: М. Л. Черненко

Компьютерная верстка: Ю. А. Окунева,

Г. А. Мирзоева, А. А. Бузов

Ответственный секретарь: О. В. Муравцова

Адрес редакции: 191023, Санкт-Петербург,

Инженерная ул., д. 6

Тел./факс: (812) 312-53-90

E-mail: asklab@aanet.ru

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций. Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.

Журнал распространяется по подписке. Подписку можно оформить в любом отделении связи по каталогу агентства «Роспечать». Индекс 15385.

Банковские реквизиты издательства:

Получатель: ИНН 7825331278, ФГУП «Издательство «Политехника»»,

Р/с 40503810819000001855 в Куйбышевском филиале

ОАО «Банк Санкт-Петербург» в г. Санкт-Петербурге,

К/с 30101810900000000790, БИК 044030790, ОКОНХ 87100,

© Коллектив авторов, 2003

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

<i>Соловьев Н. В.</i> Применение спектральных характеристик для распознавания изображений при дистанционном зондировании земной поверхности	2
<i>Яковенко М. К.</i> Метод распознавания объектов с динамическими характеристиками	8
<i>Сергеев М. Б.</i> Гибридный разрядный метод решения систем уравнений в целочисленной арифметике	16

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

<i>Соколов Б. В., Малюгин К. А.</i> Комплексное моделирование процессов управления структурной динамикой информационной системы	19
<i>Рыжиков Ю. И.</i> Оценки системы моделирования GPSS WORD	30
<i>Смирнов Ю. М., Поляков А. О., Однобоков В. В.</i> Математические методы внешнего проектирования сложных систем	39
<i>Изилов Я. Ю.</i> Многослойная перцептронная нейронная сеть в задаче моделирования речевых сигналов	44

ПРОГРАММНЫЕ И АППАРАТНЫЕ СРЕДСТВА

<i>Чьонг Динь Тяу.</i> Взаимодействие открытых систем промышленной автоматизации — состояние и проблемы	52
<i>Колбанёв М. О.</i> Формализованное описание процесса функционирования центров обработки информации и управления интеллектуальных сетей для целей оценки вероятностно-временных характеристик	58

ЗАЩИТА ИНФОРМАЦИИ

<i>Ерош И. Л.</i> Разграничение доступа к ресурсам в системах коллективного пользования	63
---	----

УПРАВЛЕНИЕ В МЕДИЦИНЕ И БИОЛОГИИ

<i>Поляков А. О., Тукабаев П. Т.</i> Информационные проблемы организации обратной связи при взаимодействии биологических и технических систем	67
<i>Коротков К. Г., Крыжановский Э. В., Муромцев Д. И., Бабицкий М. А., Борисова М. Б.</i> Автоматизированная система измерения динамических характеристик параметров изображения газоразрядного свечения	73

УПРАВЛЕНИЕ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

<i>Лукманов Ю. Х.</i> Управление инвестициями в российских регионах на основе концепции «разумного роста»	80
---	----

СВЕДЕНИЯ ОБ АВТОРАХ

89

АННОТАЦИИ

91

ЛР № 010292 от 18.08.98.
Сдано в набор 20.06.2003. Подписано в печать 27.08.2003. Формат 60×90/8.
Бумага офсетная. Гарнитура Pragmatica. Печать офсетная.
Усл. печ. л. 12,0. Уч.-изд. л. 13,3. Тираж 1000 экз. Заказ 835.
Отпечатано с готовых диапозитивов в ООО «Политехника-сервис»,
191023, Санкт-Петербург, Инженерная ул., д. 6.
Оригинал-макет изготовлен в ФГУП «Издательство «Политехника»»
191023, Санкт-Петербург, Инженерная ул., д. 6.
и отделе электронных публикаций и библиографии ГУАП.
190000, Санкт-Петербург, ул. Б. Морская, 67.

УДК: 621.865.8

ПРИМЕНЕНИЕ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ДЛЯ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ПРИ ДИСТАНЦИОННОМ ЗОНДИРОВАНИИ ЗЕМНОЙ ПОВЕРХНОСТИ

Н. В. Соловьев,

старший преподаватель

Санкт-Петербургский государственный университет
аэрокосмического приборостроения (ГУАП)

Рассмотрены некоторые проблемы распознавания изображений, полученных при дистанционном зондировании земной поверхности с помощью многополосных спектрометров. Основное внимание уделено вопросам снижения размерности пространства признаков при классификации аномалий и определению их геометрических параметров при наличии проективных искажений. Предложен метод идентификации «реперных» точек по их спектральным характеристикам и рассмотрены особенности его компьютерной реализации.

Some problems of the identification of the images obtained during remote sensing of a terrestrial surface with using of multiline spectrometers are considered. The main attention is given to the problems of reduction of dimensionality of space of indications for classifications of anomalies and calculation of their geometric parameters with availability of projective distortions. The method of identification of reference points through their spectral reflectance is offered and the features of its computer realization are considered.

Введение

Задача распознавания объектов и определения их геометрических характеристик и параметров положения в пространстве по изображениям возникает при разработке систем очувствления и управления роботом, например, системы «глаз—рука» [1]. Сходные проблемы возникают и при создании различных робототехнических комплексов, например, при автоматизации сбора и анализа визуальной информации для систем типа «летающая лаборатория». Общая концепция распознавания изображений достаточно хорошо проработана [2]. Классический подход предполагает выполнение следующих этапов:

1) предварительная фильтрация изображения с целью удаления яркостных помех и выделения наиболее информативных участков изображения, например, контура (проблема—выбор из очень большого числа разнообразных линейных и нелинейных фильтров тех, которые обеспечат на последнем этапе максимальную надежность распознавания);

2) бинаризация изображения с постоянным или адаптивным порогом (проблема—выбор способа определения порога);

3) сегментация изображения, возможно, с последующим возвратом к многоградационному изображению отдельных фрагментов (проблема—выбор

параметров сегментации, например, типа связности пикселей);

4) вычисление признаков для каждого фрагмента (проблема—выбор наиболее информативных признаков, инвариантных к неустраненным пространственным искажениям);

5) классификация фрагментов одним из известных методов, например, по расстоянию в пространстве признаков (проблема—выбор метода распознавания, дающего наилучшие результаты для имеющейся совокупности классов объектов при значительной вариативности признаков).

Следует отметить, что в результате выполнения этих этапов распознавание произойдет, скорее всего, только для плоских объектов. В свою очередь, определение их геометрических характеристик и положения в пространстве может потребовать выполнения дополнительных операций, например, компенсации пространственных искажений, связанных с изменением взаимного расположения устройства получения изображения и распознаваемого объекта. Анализ более сложных объектов и трехмерных сцен [3] предполагает выполнение еще целого ряда процедур, например, построение трехмерной модели объекта по двум изображениям (проблема—нахождение соответствия точек), анализ взаимного расположения объектов в пространстве и т. п.

Постановка задачи

Наряду с классической задачей распознавания образов и анализа сцен существуют и другие, во многом сходные с ней задачи, в частности, автоматизация обработки и анализа аэрокосмических снимков. В некоторых случаях при анализе фотоснимков поверхности Земли, полученных обычным способом, на предмет выявления каких-либо аномалий (загрязнения, заражения, косяки рыбы и т. п.) и определения их геометрических характеристик (площадь, форма и т. д.) последние или вообще не различимы даже человеком, или отличия лежат в пределах допустимых статистических отклонений, что не позволяет с достаточной степенью надежности выявить возможные аномалии [4].

В то же время, изображения этих же зон земной поверхности, сделанные с помощью щелевого многополосного спектрометра, позволяют сопоставить каждой точке изображения (в компьютерном представлении — пикселу) набор спектральных отсчетов размером N , где N — число полос спектрометра. Последующий анализ полученных спектральных изображений может повысить надежность выделения аномалий, обеспечить их идентификацию и определение геометрических характеристик. Некоторые возможные подходы к разрешению проблем, связанных с попыткой автоматизации этой задачи, рассмотрены ниже.

Снижение размерности пространства признаков и классификация объектов

Одна из проблем, возникающих при обработке спектральных изображений, — это слишком большой объем информации, поступающей от устройства ввода. В самом деле, при наличии 256-полосного щелевого спектрометра с числом отсчетов в каждой полосе 256 (1 байт на одну полосу для одного пиксела) и числом пикселов в щели 256 на передачу спектральных данных по одной щели требуется 64 Кб, что при скорости работы спектрометра в режиме телекамеры требует около 2 Мб/с. Увеличение числа пикселов до значения, стандартного для компьютерных фотоизображений, приведет к значительному росту объема передаваемой информации. Желательно понизить его без существенной потери информативности.

Если считать каждую спектральную полосу признаком объекта, относящегося к некоторому классу, то можно рассматривать эту проблему как задачу выявления наиболее информативных признаков. В качестве возможного решения может быть использован критерий информативности Фишера [3]

$$d_{ij}^n = \frac{(m_i - m_j)^2}{D_i + D_j}, \quad (1)$$

где d_{ij}^n — значение критерия по n -му признаку для пары классов i и j ; m_i , D_i , m_j , D_j — математические ожидания и дисперсии по n -му признаку для классов i и j . Критерий (1) позволяет выбрать наи-

более информативный признак для пары классов с учетом статистических характеристик разброса значений каждого признака для сравниваемых классов. Естественно, для эффективного применения этого критерия необходимо наличие репрезентативной выборки спектральных отсчетов пикселов для каждой из возможных аномалий. Применение данного метода на спектральных изображениях, адекватно моделирующих возможные реальные аномалии, показало возможность 30 % сокращения объема необходимой информации без существенного снижения качества распознавания.

Учитывая большое число признаков и их статистический характер, наиболее эффективным представляется использование в качестве параметра классификации минимальное расстояние Махаланобиса [3] в пространстве признаков

$$r_i = (X - M_i)C_i^{-1}(X - M_i)^T, \quad (2)$$

где r_i — расстояние до i -го класса; $X = (x_1, \dots, x_n)$ — вектор признаков распознаваемого образа; $M_i = (m_{i1}, \dots, m_{in})$ — вектор математических ожиданий i -го класса; C_i^{-1} — обратная матрица ковариации i -го класса.

Классификация по расстоянию (2) позволяет учесть статистические параметры всей совокупности признаков. Применение данного метода для классов, адекватно моделирующих спектральные характеристики реальных аномалий с учетом их статистических параметров, показало вполне допустимую степень надежности распознавания (около 85 %). Сокращение числа распознаваемых классов за счет априорной информации о возможных аномалиях на конкретном изображении позволяет для некоторых сочетаний аномалий повысить степень распознаваемости до 98 %.

Определение геометрических характеристик аномалии

После распознавания, т. е. отнесения подавляющего большинства пикселов изображения к тому или иному классу объектов, естественно возникает необходимость определения геометрических характеристик каждого из распознанных объектов, например площади. Возможно, для дальнейшей обработки изображения потребуется провести предварительную фильтрацию для удаления случайных малоразмерных помех и сегментацию для пространственного разделения объектов одного класса.

Вопрос определения истинных геометрических характеристик объекта фактически сводится к проблеме компенсации пространственных искажений изображения, связанных с взаимным расположением объекта и устройства получения изображения. Если оптическая ось устройства при съемке направлена вертикально вниз, то необходимо компенсировать только масштабные искажения, что не вызывает сложностей при наличии информации о расстоянии до поверхности Земли в момент съемки. Более сложным является вопрос компен-

саци пространственных искажений изображения объекта при его получении устройством бокового обзора (рис. 1). Использование подобных устройств особенно актуально при анализе возможных аномалий, так как в некоторых случаях нахождение «летающей лаборатории» непосредственно над аномалией не представляется целесообразным.

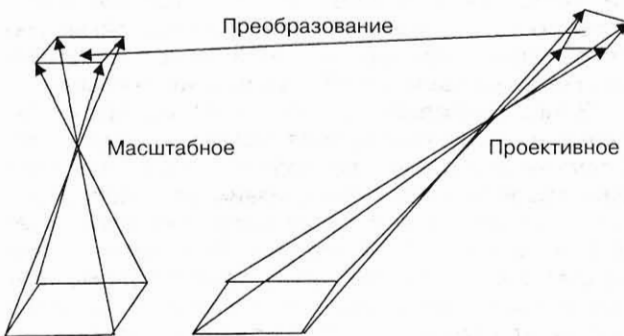
Для определения геометрических характеристик объекта и, возможно, сравнения его с «эталонным» изображением можно использовать аппарат теории групп. Как известно [5], преобразование изображения объекта при изменении взаимного расположения самого объекта и устройства получения изображения описывается при условии центрального проецирования проективной группой:

$$\begin{cases} x' = \frac{a_1x + a_2y + b_1}{a_5x + a_6y + 1} \\ y' = \frac{a_3x + a_4y + b_2}{a_5x + a_6y + 1} \end{cases} \quad (3)$$

где x, y, x', y' — исходные и преобразованные проективной группой координаты точки изображения; $a_1, \dots, a_6, b_1, b_2$ — параметры проективной группы. При параллельном проецировании преобразование изображения описывается аффинной группой — частным случаем проективной:

$$\begin{cases} x' = a_1x + a_2y + b_1 \\ y' = a_3x + a_4y + b_2 \end{cases} \quad (4)$$

В каждом конкретном случае все точки изображения $X = \{x, y\}$ подвергаются преобразованию $X' = g * X$ одним и тем же элементом группы $g \in G$, т. е. параметры $a_1, \dots, a_6, b_1, b_2$ группы постоянны при преобразовании всех точек изображения. Следовательно, для приведения искаженного изображения к исходному виду достаточно найти параметры обратного преобразования g^{-1} и применить его к искаженному изображению, компенсируя тем самым пространственные искажения. Известно несколько способов нахождения g^{-1} . Общая процедура устранения на изображениях перспективных и иных искажений описана и обоснована в работе [6].



■ Рис. 1. Пространственные искажения изображения

Пусть в пространстве U действует транзитивная группа преобразований G , каждый элемент $g \in G$ которой может быть представлен в виде произведения некоторых элементов ее подгрупп G_1, G_2, \dots, G_n , т. е. $g = g_1 * g_2 * \dots * g_n; g_i \in G_i$. Тогда каждая точка X исходного изображения (эталонного образа) преобразуется элементом $g: X' = g * X = g_1 * g_2 * \dots * g_n * X$.

Если из каких-либо соображений будет найден элемент g_i (или g_i^{-1}), тогда, подвергнув искаженное изображение преобразованию элементом g_i^{-1} , мы обеспечим «сближение» искаженного и эталонных изображений. Продолжая процесс последовательной нейтрализации подгрупп, можно либо полностью устранить искажения (распознать все объекты и определить параметры их положения), либо существенно упростить задачу и свести ее к известному решению. В качестве первых шагов обычно выбираются некоторые характеристики изображения, такие, например, как центр формы или центр массы фигуры, ее площадь и т. п.

К сожалению, на практике этот способ пока эффективно применяется только для компенсации искажений смещения и масштабирования. В самом деле, смещение предварительно центрированного изображения полностью описывается параметрами b_1, b_2 , которые определяют смещение центра объекта соответственно по оси X и Y . Масштабное искажение может быть компенсировано предварительной нормализацией площади исходных изображений всех возможных объектов, что фактически позволяет учесть смещения объекта по оси Z , перпендикулярной к плоскости изображения. При отсутствии других искажений дальнейшая классификация объекта не представляет трудностей. Например, можно использовать метод масок [1]. Однако, повороты объекта вокруг осей X, Y, Z приводят к искажениям, групповое представление которых не может быть эффективно выражено через произведение более простых групп. Именно такие искажения необходимо компенсировать при определении геометрических характеристик аномалий на изображениях, полученных методом бокового обзора.

Метод полной одновременной компенсации искажения, описываемого аффинной или проективной группой, обеспечивается использованием так называемых «реперных» точек или просто «реперов» на изображении [6].

В самом деле, как уже отмечалось, групповое преобразование всех точек изображения происходит с одними и теми же коэффициентами $a_1, \dots, a_6, b_1, b_2$. Достаточно определить эти коэффициенты для обратного преобразования, чтобы вернуть изображение к исходному виду. Из уравнений (3), (4) следует, что при наличии информации о координатах нескольких точек на исходном и преобразованном изображениях и их взаимном соответствии можно определить параметры как прямого, так и обратного преобразования, решив линейную систему уравнений. Для аффинной группы необходимое число этих точек, называемых «реперными», — три и, соответственно, необходимо решить систе-

му из шести линейных уравнений, для проективной группы число «реперов» — четыре, и необходимо решить систему уже из восьми линейных уравнений. Заметим, что для проективной группы уравнения относительно коэффициентов группы остаются линейными, хотя уравнения, описывающие преобразование, таковыми не являются.

Если искажению подвергается изображение плоского объекта, то оно полностью описывается, как максимум, проективной группой. В случае объемного изображения, например, многогранника, искажение изображения каждой из видимых граней будет описываться проективной группой со своими коэффициентами. Применительно к аэрофотоснимкам это означает, что, если высота какого-либо объекта на земной поверхности представляет собой значительную величину по сравнению с высотой до точки съемки, то следует как-то учесть эту особенность. Одним из возможных решений может быть представление преобразования в виде кусочно-проективного или кусочно-аффинного [7]. Для этого следует выделить на исходном изображении достаточное количество «реперных» точек и разбить все изображение или только интересующую нас часть на треугольные фрагменты, используя в качестве вершин выбранные «реперы». Компьютерное моделирование данного метода показало, что разбиение на треугольники, т. е. представление преобразования в виде кусочно-аффинного, более эффективно с точки зрения автоматизации, так как не приводит к существенной потере точности представления.

В качестве «реперов» естественно использовать малоразмерные и хорошо различимые объекты. На поверхности Земли это могут быть перекрестки дорог, мосты, отдельно стоящие здания, деревья или их компактные группы, характерные по форме участки береговой черты или речного русла. Очевидно, можно предложить следующую последовательность действий:

1) оператор на карте местности, которую предполагается обследовать на предмет выявления возможных аномалий, заранее выбирает «реперы» и заносит их координаты в компьютер;

2) полученная в процессе обследования информация (аэрофотоснимки и спектральные характеристики земной поверхности) обрабатывается в автоматическом режиме для классификации максимально возможного числа пикселей изображения;

3) при выявлении аномалий оператор на соответствующем аэрофотоснимке в интерактивном режиме находит «реперные» точки и определяет их соответствие выбранным ранее;

4) по полученной информации автоматически производится восстановление пространственных искажений аномальной зоны и определяются ее реальные геометрические характеристики.

Моделирование предложенной последовательности действий на компьютере показало ее достаточную эффективность. Однако наибольшие проблемы вызывает необходимость вмешатель-

ства оператора на этапе выявления и идентификации «реперов» на снимке земной поверхности. Эксперименты с реальными изображениями выявили значительный процент ошибок и резкое увеличение общего времени обработки информации, особенно у малоопытного оператора.

В свете сказанного выше представляется существенным исключение из предложенного алгоритма участие оператора на этапе выявления «реперов» на снимке с аномальной зоной. В качестве альтернативы можно предложить автоматизировать поиск и идентификацию «реперных» точек на основании спектральной информации.

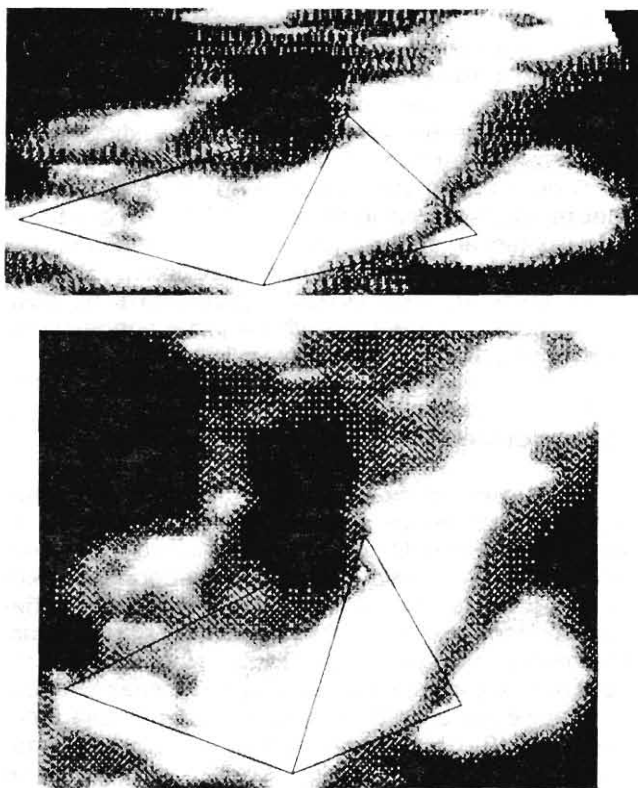
Выявление «реперных» точек

Очевидно, проблема сводится к рассмотренному выше обычному распознаванию объектов по их изображению. Следует заметить, что задача поиска соответствующих точек на стереоизображениях, возникающая при анализе сцен в случае бинокулярного зрения, только на первый взгляд близка к поставленной выше. Проблема поиска соответствия характерных точек при анализе стереоизображений достаточно хорошо освещена в литературе [2, 6]. Отличие кроется в самой постановке задачи и, как следствие, в возможном способе решения. Действительно, при работе со стереоснимками предполагается наличие двух изображений одной и той же сцены, снятых практически одновременно и отличающихся смещением точек съемки. В таком случае предполагается поиск соответствующих точек по характерным фрагментам на изображениях, возможно предварительно подвергнутым обработке тем или иным фильтром [2]. В другом варианте [6] предполагается использовать ту или иную подсветку, например лазером, для однозначного нахождения соответствия точек на стереоснимках.

В задаче выявления «реперов», во-первых, как правило, имеется только одно изображение. Если даже существует другое изображение данной области поверхности Земли, то оно чаще всего получено значительно раньше и с другого ракурса. Во-вторых, на карте местности или ранее сделанном снимке «реперы» уже отмечены и требуется идентифицировать на обрабатываемом снимке именно их.

Применение известных методов распознавания, рассмотренных выше, дает положительный результат только на сравнительно простых изображениях, например, при распознавании характерного фрагмента береговой линии, небольших островов и других объектов на водной поверхности. Поиск «реперов» на снимках, насыщенных различными объектами, в большинстве случаев не дает приемлемых результатов.

В качестве альтернативы можно предложить выявление объектов, выбранных в качестве «реперов», по их спектральным характеристикам. Если «реперы» имеют различные спектры, то их идентификация фактически заканчивается в момент



■ Рис. 2. Восстановление изображения путем преобразования аффинными группами

распознавания по этим спектрам в соответствии с расстоянием (2) в пространстве признаков. Если различные «реперы» относятся к одному классу в пространстве спектральных признаков, то их дальнейшую идентификацию можно проводить обычными методами распознавания, используя их изображения на обычном аэрофотоснимке, поскольку этап фрагментации изображения уже будет выполнен. Учитывая малые размеры объектов, используемых в качестве «реперов», практически можно не учитывать искажения формы их изображений, ограничиваясь искажениями типа смещения, изменения масштаба и поворота. Методы компенсации таких искажений и выбора признаков, инвариантных к некоторым из них, известны [6]. На рис. 2 в качестве примера приведены исходное изображение выявленной аномалии с указанием «реперов» и изображение этой же аномалии, преобразованное аффинной группой с коэффициентами, определенными по координатам «реперов».

Некоторые особенности программной реализации

Программная [8] и аппаратная [9] реализации методов обработки аэрофотоснимков имеют свои особенности. Групповые преобразования (3), (4) относятся к разряду непрерывных, т. е. не только ко-

эффициенты $a_1, \dots, a_6, b_1, b_2$ могут принимать действительные значения, но и координаты x, y, x', y' , также могут принимать действительные значения, естественно, в определенных пределах. Очевидно, последнее невозможно для компьютерного представления изображения. Следует различать абстрактное представление изображения как непрерывной действительной функции двух переменных $f(x, y)$, значения которой представляют яркость изображения в точке с координатами x, y . В таком случае естественно использовать термин «точка изображения», подчеркивая тем самым, что речь идет о не имеющем собственного размера элементе изображения. Представление изображения в компьютере носит принципиально дискретный характер. Под изображением понимается матрица, значение элемента которой соответствует яркости некоторой области поля зрения устройства получения изображения и, следовательно, носит интегральный характер. Номер элемента матрицы может быть только целым числом, и значения элемента, хотя и могут представляться в действительной форме, как правило, представляются в целочисленной форме. Представляется оправданным использовать в таком случае термин «пиксел изображения», подчеркивая наличие конечных размеров у «точки изображения».

В рассмотренном методе компенсации пространственных искажений, описываемых групповыми преобразованиями, предложено определять коэффициенты обратного преобразования и последовательно применять его к точкам искаженного изображения, приводя его тем самым к исходному виду. Непосредственная реализация этого метода на компьютере приводит к некоторым проблемам. Представление изображения как непрерывной функции фактически означает наличие взаимно однозначного соответствия между точками исходного и искаженного изображения, которое и описывается групповым преобразованием. Естественно, следует исключить вырожденные случаи — преобразование изображения в точку или линию, как не имеющие практического значения. Представление изображения в виде матрицы не дает возможности говорить о взаимно однозначном соответствии между пикселями исходного и преобразованного изображений. Как следствие, применение обратного преобразования к искаженному изображению приводит в общем случае к появлению на восстановленном изображении пропусков в виде незаполненных пикселей. Применение тех или иных методов аппроксимации для их заполнения приводит, с одной стороны, к усложнению алгоритма восстановления, а с другой — к уменьшению резкости восстановленного изображения, хотя и без изменения контрастности, что может в дальнейшем снизить надежность распознавания. Под резкостью изображения в данном случае понимается разность яркости соседних пикселей, а под контрастностью — разность максимальной и минимальной яркости на изображении. Понятно, что чем выше резкость изображения, тем точнее можно определить его контур, а, как извест-

тно [2], именно контур несет максимум информации о распознаваемом объекте.

При компьютерной реализации процесса компенсации пространственных искажений, описываемых групповым преобразованием, представляется более оправданным находить коэффициенты прямого преобразования. После их вычисления следует найти для всех пикселей восстанавливаемого изображения соответствующие им пиксели на искаженном изображении и присвоить им найденные значения яркости. В таком случае никаких пропусков в восстановленном изображении не будет и, следовательно, не потребуются их заполнение. Соответственно и резкость для соседних пикселей восстановленного изображения не

будет меньше, чем на искаженном изображении. Практическая реализация предложенного алгоритма показала вполне удовлетворительное качество восстановления изображения с точки зрения надежности распознавания имеющихся на нем образов.

Описанный метод был использован для определения характеристик аномалий земной поверхности (выявления области радиоактивного или химического заражения местности, учета природных ресурсов и т. п.) с привязкой этих новых объектов к карте местности, а также для выбора оптимальных маршрутов движения автономного мобильного робота на неизвестной пересеченной местности.

Литература

1. **Мошкин В. И., Петров А. А., Титов В. С.** и др. Техническое зрение роботов. — М., Машиностроение, 1990, 272с.
2. **Марр Д.** Зрение. Информационный подход. — М.: Радио и связь, 1987. — 400 с.
3. **Ту Дж., Гонсалес Р.** Принципы распознавания образов. — М.: Мир, 1978. — 411 с.
4. **Дистанционное зондирование: количественный подход** / Под ред. Ф. Свейна и Ш. Девис. Пер. с англ. — М.: Недра, 1983. — 396 с.
5. **Файн В. С.** Опознавание изображений. Основы непрерывно-групповой теории и ее применение. — М.: Наука, 1970.
6. **Ерош И. Л., Игнатьев М. Б., Москалев Э. С.** Адаптивные робототехнические системы (методы анализа и системы обработки изображений). — Л.: Изд-во ЛИАП, 1985.
7. **Ерош И. Л.** Восстановление рельефа неровной поверхности по плоскому изображению и реперным точкам. «Экстремальная робототехника». — Вып. VII. — СПб.: Изд-во СПбГТУ, 1998.
8. **Ерош И. Л., Гладкова И. Г., Соловьев Н. В.** Программная реализация методов обработки и анализа аэрофотоснимков // В сб.: Теория и практика создания систем технического зрения. — М., 1990.
9. **А. С. №1466542.** Устройство для ввода изображения / И. Л. Ерош, В. Л. Генкин, Н. В. Соловьев и др.

УДК 681.007:61.1

МЕТОД РАСПОЗНАВАНИЯ ОБЪЕКТОВ С ДИНАМИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ

М. К. Яковенко,

аспирант

Санкт-Петербургский государственный университет
аэрокосмического приборостроения (ГУАП)

Распознавание объектов с динамическими характеристиками — сложная задача, обусловленная высокой вариабельностью их ключевых параметров. Основная проблема — это выявление эффективного множества признаков, способных описать объект в достаточной мере для корректного распознавания с приемлемым уровнем. В данной статье в качестве примера объектов с динамическими характеристиками рассмотрены лейкоциты человека (пять основных классов).

Recognition of objects with dynamic characteristics is hard task because of very high variability of their key parameters. Main problem is extraction of effective set of features that will describe target objects enough for classification with acceptable recognition rate. In this paper is considered good example of objects with dynamic characteristics — human white blood cell (five main classes).

Введение

Распознавание объектов с динамическими характеристиками является далеко не тривиальной задачей не только для искусственного интеллекта, но и для эксперта. В качестве яркого примера подобных объектов можно отметить биологические материалы живых организмов.

Тестовыми объектами для разработки описываемого в настоящей работе метода послужили лейкоциты человека — одни из наиболее сложных для классификации биологических объектов ввиду их значительной вариабельности. Помимо разных этапов жизненного цикла самих клеток (они могут быть молодыми, зрелыми, отработанными), на вид лейкоцитов в значительной мере оказывает влияние физическое состояние и здоровье человека. Также сказываются экологические условия, вредные воздействия (например, радиоактивное) и даже время суток, в которое был взят анализ крови.

Актуальность проблемы создания автоматического классификатора лейкоцитов основывается на необходимости проведения в повседневной жизни огромного количества рутинных исследований с привлечением высококвалифицированного персонала.

Описание предметной области

В качестве анализируемых объектов в данной работе использованы основные лейкоциты периферической крови иммунной системы человека.

Это пять типов клеток: нейтрофил, эозинофил, базофил, лимфоцит и моноцит (рис. 1 — см. на 3-й стороне обложки).

Клетки распознаются врачом-гематологом по совокупности характеристик, описанных в работе [1]. Следует отметить высокую вариабельность размеров, форм и окрасок ядра и цитоплазмы каждого класса клеток. Медицинский алгоритм распознавания клеток изобилует такими терминами как зернистость, глыбчатость, сетчатость, гранулированность, равномерность, размытость. Все это описания *текстуры* ядер и цитоплазм клеток. Текстура является одним из важнейших признаков, по которым опытный гематолог с высокой вероятностью классифицирует клетки крови.

Изображения лейкоцитов для проведения исследований были получены из публичной базы данных [2], предназначенной для обучения и повышения квалификации врачей. Эта база составлена из мазков, взятых в различное время в нескольких госпиталях Швеции. Мазки принадлежат людям разного возраста и обоих полов. Лейкоциты классифицированы высококвалифицированными профессионалами-гематологами. Все клетки окрашены по методу May-Grunwald-Giemsa. Для экспериментов предоставлено по 50 клеток каждого класса.

Выбор признаков описания объектов

Из описания предметной области лейкоцитов, с учетом опыта профессиональных врачей-гематологов, был сделан вывод о высокой дискрими-

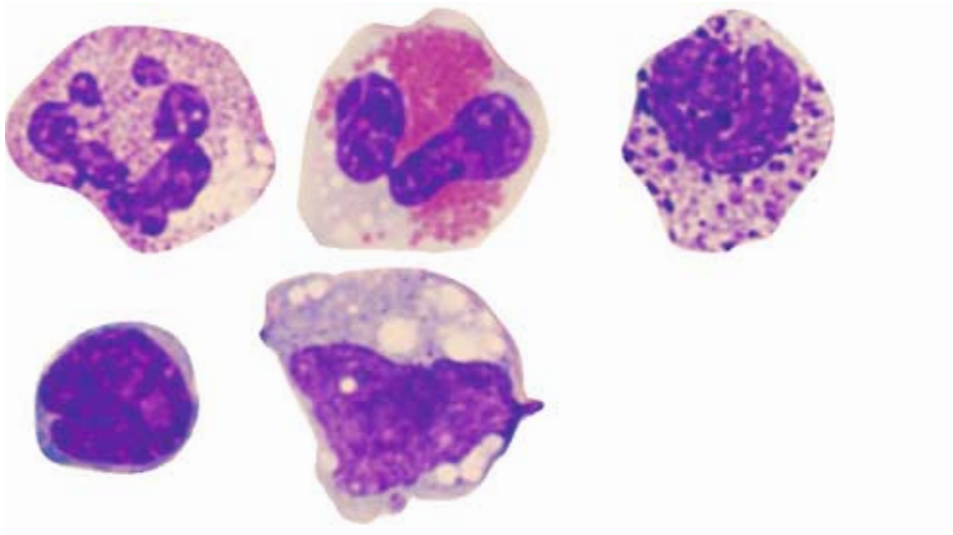


Рис.1. Объекты исследования: нейтрофил, эозинофил, базофил, лимфоцит и моноцит

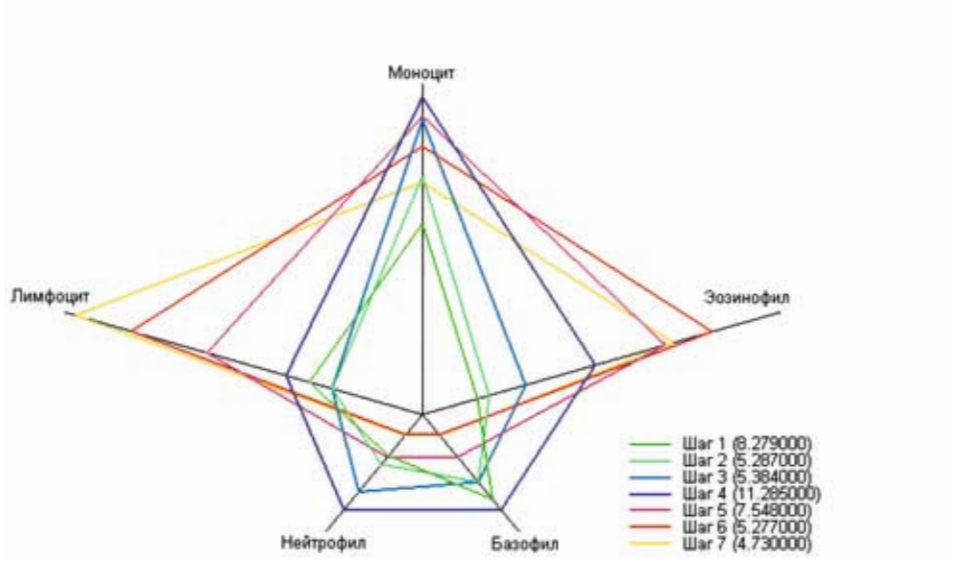


Рис.4. График для определения оптимального шага построения МСВУЯ

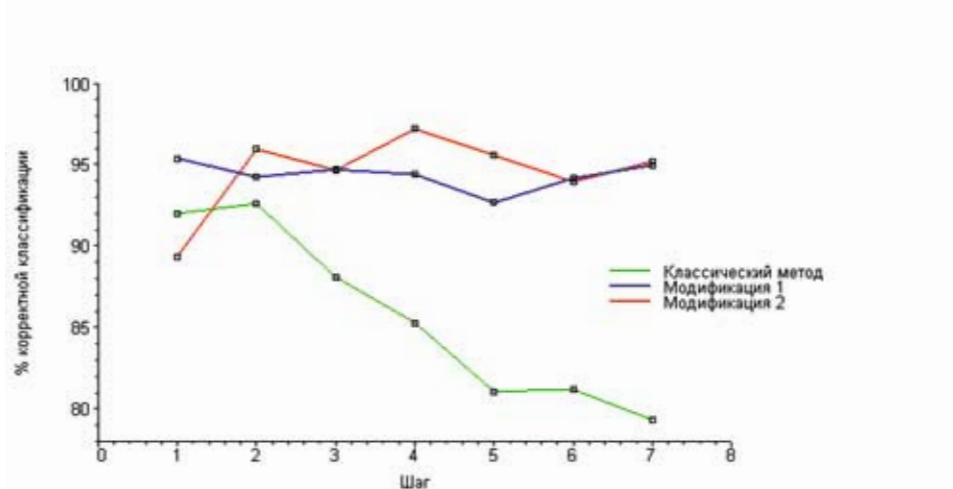


Рис.5. Сводный график результативности различных методов формирования МСВУЯ (показан средний по классам процент корректной классификации)

национной способности текстуры клеток. Следует отметить, что применение текстурных описателей при составлении пространства признаков объектов целесообразно не только в случае лейкоцитов. Исследователи целого ряда природных и, в частности, биологических объектов отмечают высокий потенциал текстуры.

В литературе рассмотрено большое количество различных методов описания текстуры. На данный момент не существует единого мнения о том, какие текстурные дескрипторы наиболее эффективны. Исследователи отдают предпочтение тем или иным методам. При анализе текстурных описателей применяются либо изображения из специальных тестовых баз данных текстур, либо искусственно созданные текстуры с теми или иными характеристиками, либо текстуры реальных объектов исследования.

В [3] производится сравнительный анализ четырех подходов к описанию текстуры: 1) анализ спектра Фурье; 2) разности уровней яркости (РУР); 3) матрицы совместной встречаемости уровней яркости (МСВУЯ); 4) и статистика длин примитивов (СДП). Авторы приходят к выводу, что статистические текстурные описатели лучше фиксируют сущность текстуры, нежели подход с использованием спектра Фурье.

В [4] проводятся исследования тех же методов описания текстуры, что и у предыдущих авторов. Тестирование осуществляется на сгенерированных текстурах на основе случайных полей второго порядка и полей Маркова. Наиболее мощным признан метод МСВУЯ.

В [5] сравниваются фрактальный подход, МСВУЯ и распределение энергии текстуры (РЭТ). По итогам исследований последние два метода дали лучшие результаты.

В [6] анализируется производительность четырех подходов к получению текстурных дескрипторов: 1) марковские случайные поля (МСП); 2) фильтр Габор; 3) фрактальный подход; 4) МСВУЯ. Для тестирования использованы фрактальные изображения, изображения кожи, рисованные картины, изображения со свойствами гауссовых МСП. Лучшими описателями признаны МСВУЯ и фракталы. Авторы отмечают, что МСВУЯ эффективнее применять в маленьких окнах, а не по всему изображению. Также показано, что использование нескольких направлений при формировании МСВУЯ увеличивает вероятность корректной классификации.

В [7] сравниваются несколько методов для классификации спутниковых изображений с использованием нейронных классификаторов. Проанализированы следующие подходы: МСВУЯ, РУР, фильтр Габор, анализ спектра Фурье. Автор делает вывод, что универсальный набор дескрипторов, показывающий наилучшие результаты на изображениях различных классов, построить невозможно. В целом предпочтение автора было отдано спектру Фурье, второе место поделили МСВУЯ и фильтр Габор.

Из приведенного анализа литературных источников можно констатировать превосходство МСВУЯ

над прочими методами построения текстурных описателей. Однако следует отметить, что у исследователей встречаются и прямо противоположные результаты. В [8] приводится сравнительный анализ следующих методов: РЭТ, фильтр Габор, вэйвлеты, спектр Фурье, МСВУЯ. Различные методы были лучшими на разных множествах текстурных изображений. При этом МСВУЯ оказались худшими во всех случаях, а предпочтение отдано вэйвлетам. Также в пользу вэйвлетов высказываются исследователи в [9]. В данном анализе помимо вэйвлетов использовались спектр Фурье и фильтр Габор. Последний метод наравне с вэйвлетами показал хорошие результаты. Интересно отметить, что исследования, проведенные авторами в [6], были повторены, а результаты — опровергнуты другим исследователем [10]. В его работе наилучшую производительность показали МСП, а худшую — МСВУЯ.

Таким образом, по всей видимости, эффективность текстурных дескрипторов во многом зависит от предметной области, объектов исследования. Также следует отметить, что не все результаты исследований можно принимать с большой достоверностью, так как среди них встречаются неглубокие проработки того или иного подхода. Например, исследователи используют в своих тестах ограниченный набор текстурных дескрипторов, которые можно получить по определенному методу. Также может оказаться недостаточной или вовсе отсутствовать оптимизация параметров метода.

Проведенные исследования и эксперименты с МСВУЯ показали достаточно хорошие результаты при применении на изображениях клеток. Далее приводится описание метода МСВУЯ.

Положим размер анализируемого изображения на $N_x \times N_y$, число уровней яркости — N_g . Пусть $L_x = \{1, 2, \dots, N_x\}$ и $L_y = \{1, 2, \dots, N_y\}$ — соответственно горизонтальная и вертикальная пространственная область, $G = \{1, 2, \dots, N_g\}$ — множество уровней яркости изображения. Тогда изображение может быть определено следующим образом: $I : L_x \times L_y \rightarrow G$.

МСВУЯ представляют собой оценку плотности распределения вероятностей второго порядка. Элемент матрицы — условная вероятность $P(i|j)$ возникновения уровня яркости i на расстоянии d от уровня яркости j . Поскольку число таких матриц может быть очень большим, необходимы некоторые упрощения. Имеет смысл высчитывать четыре матрицы для заданного шага d для $0^\circ, 45^\circ, 90^\circ$ и 135° расположения пар уровней яркости. Формально элементы матриц для указанных направлений могут быть записаны как

$$P(i|j)_0 = \#\{(k, l), (m, n) \in (L_x \times L_y) | (k - m = 0, |l - n| = d), \\ l(k, l) = i, l(m, n) = j\};$$

$$P(i|j)_{45} = \#\{(k, l), (m, n) \in (L_x \times L_y) | (k - m = d, |l - n| = d), \\ l(k, l) = i, l(m, n) = j\};$$

$$P(i|j)_{45} = \# \{ (k, l), (m, n) \in (L_x \times L_y) \mid (k-m=d, l-n=-d) \vee (k-m=-d, l-n=d), l(k, l)=i, l(m, n)=j \};$$

$$P(i|j)_{90} = \# \{ (k, l), (m, n) \in (L_x \times L_y) \mid (|k-m|=d, l-n=0), l(k, l)=i, l(m, n)=j \};$$

$$P(i|j)_{135} = \# \{ (k, l), (m, n) \in (L_x \times L_y) \mid (k-m=d, l-n=d) \vee (k-m=-d, l-n=-d), l(k, l)=i, l(m, n)=j \};$$

где знак # обозначает количество элементов в множестве.

Полученные матрицы являются квадратными с размером, равным N_g . Также они являются симметричными, т. е. $P(i|j) = P(j|i)$.

Можно провести следующую нормализацию матриц для оперирования с вероятностями:

$$N(i|j) = \frac{P(i|j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i|j)}.$$

МСВУЯ самостоятельно обычно не используются в качестве текстурных описателей. На их основе строятся разнообразные характеристики [3, 4, 11, 12]. В данной работе используется 24 признака. Для примера рассмотрим три из них:

$$T_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (P(i|j))^2; \quad (1)$$

$$T_2 = \sum_{n=1}^{N_g-1} n^2 \left(\sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} P(i|j) \right); \quad (2)$$

$$T_3 = \left(\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij)P(i|j) - \mu_x \mu_y \right) / (\sigma_x \sigma_y), \quad (3)$$

где μ_x, μ_y — и математические ожидания частных распределений, связанных с $P(i|j)$ по строкам и столбцам матриц соответственно; а σ_x, σ_y — соответствующие стандартные отклонения.

Выражение (1) описывает угловой второй момент, являющийся мерой однородности текстуры. Формула (2) определяет контраст изображения. Контраст является дифференциальным моментом матрицы совместной встречаемости и дает количественную оценку локальных изменений яркости, присутствующих на изображении. Параметр, заданный по выражению (3), — это корреляция, являющаяся мерой присутствия полутоновой линейной зависимости на изображении.

Каждый из построенных текстурных дескрипторов на основе МСВУЯ должен быть вычислен четыре раза — для каждого направления ($0^\circ, 45^\circ, 90^\circ$ и 135°). При этом теоретически каждый дескриптор из группы наиболее точно отражает лишь свое направление. Исследования [11] показали, что мож-

но использовать среднее значение в каждой группе дескрипторов в качестве текстурного описателя, инвариантного по вращению. Дополнительно к среднему рекомендуется использовать разброс в группе описателя (минимальное и максимальное значения).

Единственной сложностью остается выбор параметра шага d . Несмотря на то, что проведенные исследования и эксперименты с МСВУЯ показали достаточно хорошие результаты при применении на изображениях клеток, тем не менее, за счет корректного выбора шага d не удалось добиться высокой степени дифференцированности кластеров. Поэтому была проанализирована возможность модернизации МСВУЯ для повышения информативности получаемых из них текстурных дескрипторов.

МСВУЯ могут быть представлены в нотации дискретных цепей Маркова. Пусть система меняет свое состояние в моменты времени $t = 1, 2, \dots, n$. При этом она может находиться в одном из r состояний s_1, s_2, \dots, s_r , где $r > 2$. Пусть $A_k^{(j)}$ — событие того, что система на k -м шаге входит в состояние s_j . Тогда, если для всех $k = 2, \dots, n$ выполняется

$$P[A_k^{(jk)} | A_{k-1}^{(jk-1)}, \dots, A_1^{(j1)}] = P[A_k^{(jk)} | A_{k-1}^{(jk-1)}],$$

то данная система является цепью Маркова. Состояние такой системы в определенный момент времени зависит только от предыдущего состояния и не зависит от прошлой истории состояний системы. Таким образом, это система без памяти о прошлом. Обозначим $P(i|j)$ вероятность того, что цепь Маркова в момент времени t находится в состоянии s_i при условии, что в момент времени $t-1$ она находилась в состоянии s_j . Причем эта вероятность не зависит от конкретного значения t . Заданные таким образом вероятности являются вероятностями перехода и представляются в виде матрицы:

$$P = \begin{bmatrix} P(1|1) & \dots & P(1|r) \\ \vdots & \ddots & \vdots \\ P(r|1) & \dots & P(r|r) \end{bmatrix}.$$

Данная матрица эквивалентна введенной выше матрице совместной встречаемости уровней яркости.

В случае шага, равного 1, метод МСВУЯ функционирует очевидным образом, как описано выше. При этом мы имеем симметричную матрицу. При увеличении шага в классическом варианте рассматриваются пары пикселей на расстоянии $d > 1$. В этом случае изменяются только параметры построения матриц, но информации о текстуре в них не добавляется. Видимо, с этим связано употребление шага, равного 1, подавляющим большинством исследователей при построении МСВУЯ.

В данной работе предлагается повысить информативность МСВУЯ при использовании шага $d > 1$, несколько модифицировав при этом классическое представление о взаимодействии элементов изображения.

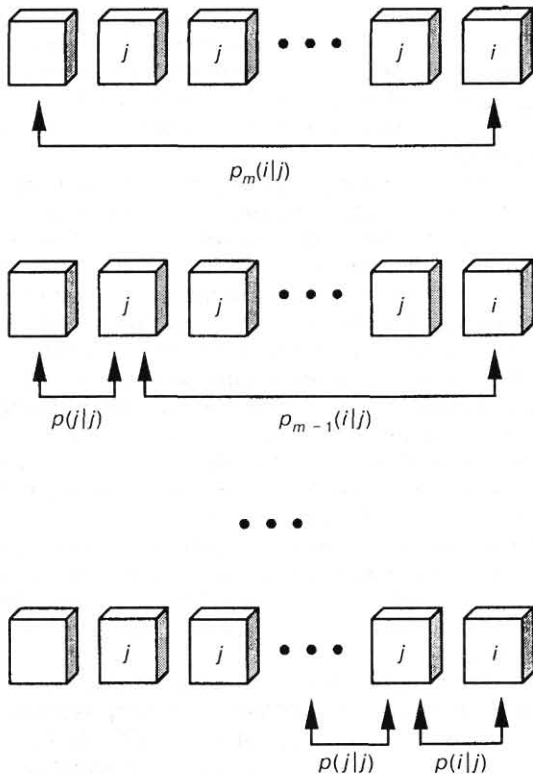


Рис. 2. Порядок вычисления вероятности перехода для формирования модифицированных МСВУЯ

В теории Маркова m -шаговая вероятность перехода определена [13] как

$$P_m = (i | j) = \sum_{k=1}^r P_{m-1}(i | k) P(k | j).$$

Это вероятность того, что цепь Маркова в момент времени t находится в состоянии s_i при условии, что в момент времени $t - m$ она находилась в состоянии s_j .

Для увеличения информативности МСВУЯ предлагается при вычислении вероятностей переходов в каждый момент времени t придерживаться следующих условий:

$$\begin{cases} p_m(i | j) \neq 0 \\ p_{m-1}(i | j)p(j | j) \neq 0 \\ \dots \\ p(i | j) \neq 0 \end{cases}$$

Таким образом, одна условная вероятность будет характеризовать переход из одного пикселя в группу пикселей изображения, т. е. под состоянием предлагается понимать не только один пиксель, но и последовательность пикселей одной яркости (рис. 2).

За счет применения такого подхода увеличивается количество информации, заложенной в переход между состояниями. При этом матрица пе-

рестает быть симметричной, что влечет за собой увеличение информативности сформированных на ее основе текстурных дескрипторов. Несимметричность матрицы также позволяет ввести дополнительные описатели текстуры, которые чаще всего не имели смысла при классическом способе формирования МСВУЯ, например, треугольную симметрию.

В рамках дальнейшего изучения возможности увеличения информативности текстурных признаков, построенных на базе МСВУЯ, рассмотрим марковские случайные поля. Как отмечалось выше, исследователи выделяют этот метод как имеющий высокий потенциал для описания текстур.

Теория случайных полей Маркова представляет математическую основу для решения проблемы получения информации о процессе в целом по имеющейся локальной информации. МСП являются двурным обобщением дискретных цепей Маркова. Обозначим $F = \{F_1, \dots, F_m\}$ множество случайных чисел, определенных на множестве позиций $S = \{S_1, \dots, S_m\}$. Набор значений $f = \{f_1, \dots, f_m\}$ называется конфигурацией F . Значения конфигурации принадлежат множеству возможных меток $L = \{L_1, \dots, L_k\}$. В случае изображений множество меток содержит возможные значения пикселей. Позиции взаимодействуют друг с другом в некотором окружении. Окружением позиции является $N = \{N_i | \forall i \in S\}$, где N_i — множество позиций, окружающих i . При этом позиция не входит в свое окружение: $i \in N_i$, и окружение имеет свойство ободности: $i \in N_i \Leftrightarrow i' \in N_i$. В случае двумерного изображения множество элементов окружения радиуса r определяется выражением

$$N_i = \{i' \in S | dist(S_i, S_{i'})^2 \leq r, i \neq i'\},$$

где $dist(A, B)$ — Евклидово расстояние между точками.

Количество анализируемых «соседей» определяется порядком модели МСП. На рис. 3 представлена структура используемых множеств «соседей» для моделей до седьмого порядка включительно [14].

		7	6	7		
	5	4	3	4	5	
7	4	2	1	2	4	7
6	3	1	<i>i</i>	1	3	6
7	4	2	1	2	4	7
	5	4	3	4	5	
		7	6	7		

Рис. 3. Структура окружения для разных порядков моделей МСП

Совместная вероятность возникновения определенной конфигурации обозначается как $P(F_1 = f_1, \dots, F_{m1} = f_m) = P(f)$. F образует МСП на S относительно системы окружения N тогда и только тогда, когда одновременно выполняются два условия:

- 1) $P(f) > 0, \forall f \in F$ — позитивность;
- 2) $P(f_i | f_{S-(i)}) = P(f_i | f_{N_i})$ — свойство Маркова, где $f_{N_i} = \{f_{i'} | i' \in N_i\}$.

Условие позитивности необходимо в силу некоторых технических причин и обычно может быть выполнено на практике. Например, когда выполняется это условие, совместная вероятность $P(f)$ любого случайного поля уникально определяется через локальные условные вероятности [15]. Свойство Маркова означает, что конфигурация позиции зависит только от ближайшего окружения. Всегда можно выбрать размер окружения N_i такой, что это свойство будет достигнуто. Окружение более высокого порядка включает в себя все позиции окружений более низких порядков.

В [15] доказывается тот факт, что совместная вероятность $P(F_1 = f_1, \dots, F_{m1} = f_m) = P(f)$ имеет функцию распределения Гиббса:

$$P(f) = \frac{e^{-\frac{1}{T}U(f)}}{Z},$$

где $Z = \sum_{f \in F} e^{-\frac{1}{T}U(f)}$ — нормализационная константа,

так называемая термодинамическая сумма; T — глобальный контролирующий параметр, так называемая температура, которая обычно выбирается равной 1; $U(f)$ — функция энергии, характеризующая текстуру.

Существует несколько подходов к построению функции $U(f)$ по имеющемуся классу текстур [14]. Взаимосвязь МСП и распределения Гиббса описывается теоремой Hammersley—Clifford [15]. Ее практическая польза состоит в возможности задания совместной вероятности для изображения. Таким образом, МСП характеризуется свойством локальности, и в то же время имеет глобальное свойство (распределение Гиббса).

Если положить нормальный закон распределения появления уровней яркости в каждой позиции (μ, σ) , то мы будем иметь гауссовское МСП. Так как распределение Гаусса принадлежит семейству распределений Гиббса, то в этом случае распределение совместной вероятности примет вид

$$P(f) = \frac{\sqrt{|B|}}{\sqrt{(2\pi\sigma^2)^m}} e^{-\frac{(f-\mu)^T B(f-\mu)}{2\sigma^2}},$$

где m — $1 \times m$ вектор из m_i ; B — $m \times m$ матрица взаимодействия, элемент которой имеет вид $b_{i,i'} = \delta_{i,i'} - \beta_{i,i'}$, где $\beta_{i,i'}$ — константа, отражающая взаимодействие пары положений i и i' .

С использованием модели гауссовского МСП в [16] были получены признаки описания текстур

практически без потерь информации. С их помощью удавалось синтезировать копии текстур, близкие к оригиналам. Это является практической иллюстрацией эффективности использования модели МСП для выявления характерных свойств изображения.

Для второй модификации способа построения МСВУЯ использованы приведенные выше доказательства высокой информативности локального окружения модели МСП для описания текстур. При этом для формирования статистики совместной встречаемости применяется окружение, описываемое в модели. Шаг построения МСВУЯ ассоциируется в данном случае с порядком модели МСП.

При таком методе формирования МСВУЯ строится только одна матрица, учитывающая все направления текстуры одновременно. Для каждого положения i определяется набор условных вероятностей в виде $\{P(f_i | f_k), \forall k \in N_i\}$.

Теоретические выкладки, описывающие представленные выше методы модификации классического подхода формирования МСВУЯ, нашли адекватное отражение в практических результатах экспериментов.

Для уменьшения вычислительной нагрузки при работе с матрицами совместной встречаемости имеет смысл понизить их размерность. Для этого необходимо квантовать изображение на меньшее количество уровней яркости. Классическое яростное изображение имеет 256 градаций серого. На практике применения МСВУЯ оказывается достаточно восьми [11]. Для корректного понижения количества уровней яркости может быть использована процедура равновероятностной квантизации.

Оптимизация параметров признаков

Матрицы совместной встречаемости в классической нотации имеют два параметра: направление и шаг. Дополнительно был разработан еще один — тип шага. В данном исследовании при расчете текстурных дескрипторов лейкоцитов используются следующие типы шагов: классический, с использованием последовательностей и порядок модели МСП. Из четырех характеристик направленных матриц можно выбрать любое направление ($0^\circ, 45^\circ, 90^\circ$ и 135°), минимальное, максимальное или среднее значение. Усреднение дескрипторов четырех направлений позволяет добиться инвариантности характеристик к вращению, что является полезным при анализе клеток. При использовании второго из предложенных методов модификации МСВУЯ (на основе модели МСП) формируется только одна матрица. Таким образом, остается необходимость оптимизировать только один параметр — размер шага. Далее предлагается оригинальный подход к достаточно эффективному решению этого вопроса.

В данном случае имеется M классов объектов и N признаков. Необходимо вычислить несколько наборов признаков для разных значений шага, например, от 1 до 7. Задача состоит в выборе одного

из этих шагов, при котором достигается максимальное дифференцирование кластеров. Воспользуемся критерием Фишера для оценки расхождения двух кластеров по признаку k :

$$f_{ij}^k = \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2}, \forall ij, i > j, f_{ij}^k = f_{ji}^k.$$

Сформируем матрицу интегральной межклассовой оценки критерия Фишера:

$$F_{ij} = \sum_{k=1}^N f_{ij}^k,$$

где N — количество признаков.

Далее для каждого класса из этой матрицы находим минимальную оценку. Для этого анализируются строка и столбец матрицы, равные номеру класса:

$$F \min_d = \left\{ \min_{i=1}^M (F_{ij}, F_{ji}), j = 1 \dots M \right\}.$$

Таким образом, для каждого шага d построения текстурных дескрипторов будем иметь набор чисел размером, равным количеству классов M . Построим специальный график для оценки эффективности выбора шага формирования МСВУЯ (рис. 4 — см. на 3-й стороне обложки). Представленный график соответствует использованию МСВУЯ, модифицированных по второму методу. Каждый класс определяет свою ось координат. По осям откладываются полученные интегральные оценки Фишера для каждого шага. Соединенные значения оценок для одного шага образуют многоугольник. Значения отношения площади многоугольника к дисперсии соответствующего вектора $F \min_d$ представлены на рис. 4.

Предлагается выбрать шаг d построения текстурных дескрипторов, соответствующий максимальной приведенной площади полученного многоугольника. Эксперименты с данным методом показали высокую корреляцию его показаний с высшим средним процентом корректной классификации. В данном случае следует выбрать шаг, равный 4.

Таким образом, предложенный подход позволяет проводить автоматическую оптимизацию параметров формирования текстурных признаков с низкими вычислительными затратами. Это особенно актуально в условиях полученной размерности пространства признаков при достаточно большой тренировочной выборке.

Классификация

Так как проверка распределений значений признаков показала их близость к нормальному распределению (т. е. априори известен вид функции распределения), то оказывается возможным применение статистического классификатора, построенного на основе формулы Байеса.

Дискриминационная функция в случае нормального распределения примет вид

$$d_i(X) = -\frac{1}{2} \left(\ln (\Sigma_i) + (X - M_i)^T \Sigma_i^{-1} (X - M_i) \right),$$

где M_i — вектор математического ожидания; Σ_i — матрица ковариации.

Каждый кластер представляется гиперэллипсоидом, параметры которого описаны в выражении $(X - M_i)^T \Sigma_i^{-1} (X - M_i)$, называемом расстоянием Махалабониса. Главные оси гиперэллипсоида заданы собственными векторами Σ_i . Собственные числа Σ_i определяют длины этих осей. Центр гиперэллипсоида задан посредством M_i .

Большим преимуществом статистического классификатора является возможность вычисления вероятности ошибки классификации, с помощью которой можно оценить производительность выбранной дискриминационной функции. Вероятность ошибки в случае нормального распределения и симметричной функции потерь примет вид

$$P_e = \frac{1}{2} \Phi \left[-\frac{1}{2} \sqrt{r_{ij}} \right] + \frac{1}{2} \left(1 - \Phi \left[\frac{1}{2} \sqrt{r_{ij}} \right] \right) = \Phi \left[-\frac{1}{2} \sqrt{r_{ij}} \right].$$

Оценим качество кластеризации для выбранного в качестве оптимального метода построения текстурных дескрипторов.

В табл. 1 представлены значения межклассовых расстояний Махалабониса, полученных при применении второй модификации метода построения МСВУЯ и шага, равного 4. Минимальное расстояние — 122,48, что соответствует уровню ожидаемой максимальной ошибки вследствие перекрытия кластеров — 0,00000157 %. Это свидетельствует о достигнутом высоком качестве дифференцирования кластеров.

Полезный тест для проверки качества кластеров осуществляется путем классифицирования той же базы данных объектов, что была задействована при формировании параметров кластеров. Результаты такого теста составили 100 % корректной классификации для всех классов. Эти значения нельзя принимать в качестве основного результата, так как имеется прямая зависимость параметров классификатора от тестового массива. Однако на их основе можно сделать однозначный вывод об абсолютном разнесении кластеров в пространстве признаков, что подтверждает их (признаков) высокое дискриминационное качество. Также можно констатировать отсутствие явных кластерных патологий, таких как, например, возникновение кластера в другом кластере.

■ Таблица 1. Расстояния между классами (вторая модификация МСВУЯ, шаг 4)

	Моноцит	Лимфоцит	Нейтрофил	Базофил	Эозинофил
Моноцит	0	2500,21	481,10	347,87	175,24
Лимфоцит	3165,91	0	2443,70	352,55	11 033,34
Нейтрофил	713,53	1255,46	0	676,58	122,48
Базофил	1984,75	402,03	205,06	0	417,52
Эозинофил	215,42	1188,18	134,15	168,23	0

■ **Таблица 2.** Результаты вычисления T-статистики (вторая модификация МСВУЯ, шаг 4)

	Моноцит	Лимфоцит	Нейтрофил	Базофил	Эозинофил
Моноцит	0	57,23	35,03	41,82	34,38
Лимфоцит	57,23	0	70,79	51,45	54,19
Нейтрофил	35,03	70,79	0	33,71	22,84
Базофил	41,82	51,45	33,71	0	31,59
Эозинофил	34,38	54,19	22,84	31,59	0

Чтобы определить, действительно ли векторы M_i , определяющие центры кластеров, различны в статистическом смысле (т. е. принадлежат независимым генеральным совокупностям), применим обобщенную версию T-теста. Указанный тест — это многомерное расширение проверки с использованием статистики Стьюдента. Будем проверять гипотезу о равенстве пары векторов математических ожиданий (M_i, M_j), т. е. предположим, что выборки построения этих векторов взяты из одной генеральной совокупности. Уровень значимости при этом положим 1 %.

Для проверки гипотезы используется следующая статистика [18]:

$$t_{ij}^2 = \frac{n_i n_j (n_i + n_j - 2)}{n_i + n_j} (M_i - M_j)^T H_{ij}^{-1} (M_i - M_j),$$

где n_i — число объектов в классе i ; $H_{ij} = n_i \Sigma_i + n_j \Sigma_j$.

Данная статистика имеет $(n_i + n_j - 4 - 1)$ степеней свободы и распределена по следующему закону [18]:

$$T(x) = \frac{2\Gamma\left(\frac{n}{2}\right)}{(n-1)^{\frac{k}{2}} \Gamma\left(\frac{n-k}{2}\right) \Gamma\left(\frac{k}{2}\right)} \cdot \frac{x^{k-1}}{\left(1 + \frac{x^2}{n-1}\right)^{\frac{n}{2}}}, \quad x > 0,$$

где n — число степеней свободы статистики; k — порядок статистики.

Для выбранного уровня значимости была вычислена квантиль распределения — 7,766. Проверяемая гипотеза должна быть отвергнута при превышении статистикой этого значения.

Результаты вычисления статистики приведены в табл. 2. На их основании гипотеза о выборке данных из одной генеральной совокупности должна быть отвергнута. Таким образом, имеется пять статистически независимых популяций.

■ **Таблица 3.** Матрица неточностей классификации (вторая модификация МСВУЯ, шаг 4)

Клетка	Лимфоцит	Нейтрофил	Эозинофил	Базофил	Моноцит	Сумма
Лимфоцит	50	0	0	1	0	51
Нейтрофил	0	48	3	0	0	51
Эозинофил	0	0	46	0	0	46
Базофил	0	1	0	49	1	51
Моноцит	0	0	1	0	49	50
Неизвестный	0	1	0	0	0	1
Сумма	50	50	50	50	50	250
Процент корректного распознавания	100,0	97,9	92,0	98,0	98,0	
Средний процент корректного распознавания						97,2

Классический классификационный тест предусматривает разделение всего множества анализируемых объектов на две группы — тренировочную и тестовую (обычно равные). После этого определяются параметры классификатора (в данном случае описатели кластеров) с использованием тренировочной базы данных. Затем осуществляется проверка работоспособности классификатора на тестовой базе. Таким образом, достигается независимость настроек классификатора от объектов реальной работы. Данный подход высокоэффективен при наличии достаточно большого исходного массива анализируемых объектов.

При имеющемся размере базы данных применение классического метода не представляется возможным, так как разделение массива на две группы в таком случае резко снижает точность оценки результатов.

В данной ситуации целесообразно воспользоваться популярным среди исследователей методом leave-one-out, описанным, например, в [17]. Смысл его заключается в полном формировании всех кластеров, кроме одного. Из массива данных оставшегося класса извлекается один объект и формируется последний кластер. Классификатору предъявляется выбранный объект, при этом сохраняется независимость параметров классификатора от тестовых образцов. Далее процедура повторяется со всеми классами и объектами в них. Таким образом, в формировании результатов классификации последовательно принимают участие все объекты базы данных. Основной недостаток данного метода — вычислительная сложность, так как на каждом шаге необходимо рассчитывать параметры одного из кластеров заново.

Результаты

Результаты экспериментов с различными способами формирования МСВУЯ (классический, предложенный в [11], и обе модификации, предложенные в данной работе) представлены на рис. 5 (см. на 3-й стороне обложки). Гра-

фики показывают усредненные межклассовые проценты корректной классификации для трех методов. Видно, что максимальная производительность достигнута при помощи второй модификации классического метода построения МСВУЯ. Максимальный средний процент корректной классификации для этого метода составляет 97,2 % на шаге 4. Это согласуется с ожидаемым оптимальным шагом, выбранным по рис. 4.

Детально оценить результаты классификации можно с помощью матрицы неточностей (confusion matrix), представленной в табл. 3. Столбцы матрицы представляют собой результаты классификации клеток одного класса (можно отследить, с какими классами были «перепутаны» клетки данного класса в процессе классификации). Строки матрицы указывают на появление данного класса при классификации других классов. На-

пример, при классификации 50 нейтрофилов 48 раз они были распознаны верно, один нейтрофил был опознан как базофил и один раз зафиксирован отказ от классификации.

По табл. 3 видно, что наименьшая вероятность корректной классификации достигнута для эозинофила — 92,0. Следует отметить, что это достаточно редко встречающаяся клетка — около 2 % от общего количества лимфоцитов в пробе. Если положить количество лимфоцитов, найденных во время анализа, — 28 %, моноцитов — 9 %, нейтрофилов — 60 %, базофилов — 1 % и эозинофилов — 2 %, то по формуле полной вероятности получим процент корректной классификации на уровне 98,4 % (указанные цифры соответствуют средним процентным соотношениям клеток [1]).

В работе [19] описано исследование, призванное оценить уровень ошибки классификации, допускаемой квалифицированным экспертом. Для этого было задействовано девять экспертов, которым было предложено осуществить классификацию базы данных из 314 лейкоцитов. В состав экспертной группы входили пять профессиональных гематологов и четыре опытных лаборанта. Ошибка эксперта оценивалась относительно мажоритарного мнения группы. По отдельным классам максимальная экспертная ошибка составила 21,06 % (для эозинофила), а максимальная средняя (по всем экспертам) ошибка — 4,09 %. Усредненная по всем классам ошибка, характеризующая работу эксперта, в итоге составила 1,97 %.

Таким образом, полученные результаты автоматической классификации хорошо согласуются с результатами работы профессионального врача-гематолога.

Литература

1. **Абрамов М. Г.** Гематологический атлас. — М.: Медицина, 1979.
2. **www.cellatlas.com**
3. **Weszka J. S., Dyer C. R., Rosenfeld A. A.** Comparative Study of Texture measures for Terrain Classification // IEEE Transactions on Systems, Man and Cybernetics. — 1976. — N 6. — P. 269–285.
4. **Connors R. W., Harlow C. A.** A Theoretical Comparison of Texture Algorithms // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1980. — N 2. — P. 204–222.
5. **Baf J. M. H., Kardan M., Spann N.** Texture feature performance for image segmentation // Pattern Recognition. — 1990. — Vol. 23. — N 3/4. — P. 291–309.
6. **Ohanian P. P., Dubes R. C.** Performance evaluation for four classes of textural features // Pattern Recognition. — 1992. — Vol. 25. — N 8. — P. 819–833.
7. **Augustejin M. F.** Performance evaluation of texture measures for ground cover identification in satellite images by means of neural-network classifier // IEEE Transactions on Geoscience and Remote Sensing. — 1995. — Vol. 33. — P. 615–625.
8. **Randen T., Husoy J. H.** Filtering for texture classification: a comparative study // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1999. — Vol. 21. — N 4. — P. 291–310.
9. **Chen C. C., Chen P. C.** Filtering methods for texture discrimination // Pattern Recognition Letters. — 1999. — Vol. 20. — P. 783–790.
10. **Smith G.** Ohanian and Dubes Comparisons // www.cssip.uq.edu.au/staff/meastex/www/comparisons/ohanianDubesCmp.html
11. **Haralick R. M., Shanmugam K., Dinstein I.** Textural features for image classification // IEEE Transactions on Systems, Man and Cybernetics. — 1973. — Vol. SMC-3. — N 6. — P. 610–621.
12. **Pressman N. J.** Markovian analysis of cervical cell images // The Journal of Histochemistry and Cytochemistry. — 1976. — Vol. 24. — N 1. — P. 138–144.
13. **Parzen E.** Modern Probability Theory and Its Applications. — New York: John Wiley, 1969. — P. 136–147.
14. **Li S. Z.** Markov Random Field Modeling in Computer Vision. — Springer-Verlag, 1999.
15. **Besag J.** Spatial interaction and the statistical analysis of lattice systems (with discussions) // Journal of the Royal Statistical Society. — 1974. — N 36. — P. 192–236.
16. **Chellappa R., Chatterjee S.** Classification of textures using Gaussian Markov random fields // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1985. — Vol. ASSP-33. — N 4. — P. 959–963.
17. **Duda O. R., Hart P. E., Stork D. G.** Pattern classification (2nd edition). — Wiley-Interscience, 2000.
18. **Cramer H.** Mathematical Methods of Statistics. — Princeton Press, 1963.
19. **Bacus J. W., Gose E. E.** Leukocyte pattern recognition // IEEE Transactions on Systems, Man and Cybernetics. — 1972. — Vol. SMC-2. — N 4. — P. 513–526.

УДК 681.3:518.5

ГИБРИДНЫЙ РАЗРЯДНЫЙ МЕТОД РЕШЕНИЯ СИСТЕМ УРАВНЕНИЙ В ЦЕЛОЧИСЛЕННОЙ АРИФМЕТИКЕ

М. Б. Сергеев,

д-р техн. наук, профессор

Санкт-Петербургский государственный университет аэрокосмического приборостроения (ГУАП)

Предложен гибридный разрядный метод решения целочисленных систем линейных алгебраических уравнений, гарантирующий абсолютную точность и реализуемый простейшими арифметическими операциями. Рассмотрен класс решаемых систем уравнений, оценена эффективность программной реализации в микропроцессорных системах.

A hybrid bit-wise method for solution of integer systems of linear algebraic equations is presented. The method guarantees absolute accuracy of the decision and uses the elementary arithmetic operations. The class of solved systems of the equations is considered and efficiency of program solution in microprocessor systems is evaluated.

Введение

Применение микропроцессоров и систем на их основе для управления процессами, протекающими в реальном масштабе времени, показало, что они обладают достаточной вычислительной мощностью с точки зрения выполнения математических операций при решении «трудоемких» задач линейной алгебры в системах управления и контроля.

Одной из важнейших проблем, возникающих при решении систем линейных алгебраических уравнений (СЛАУ), является выбор метода, обеспечивающего наилучшую точность вычислений [1]. Однако существует ряд задач, сводящихся в математической постановке к СЛАУ и требующих обеспечения абсолютной точности их решения [2], в которых системы уравнений имеют квадратную целочисленную матрицу коэффициентов \mathbf{A} ($\det \mathbf{A} \neq 0$) и целочисленные векторы свободных членов \mathbf{B} и неизвестных \mathbf{X} . Задача вычисления целочисленных решений СЛАУ с рациональными \mathbf{A} и \mathbf{B} , точно представимыми в пределах разрядной сетки процессора, при подходящем строчном масштабировании также сводится к решению поставленной задачи.

Классические прямые и итерационные методы решения СЛАУ реализуют, как правило, вычисления в арифметике с плавающей точкой и, вследствие присущих ей ошибок округления, получить целочисленное решение таких СЛАУ не представляется возможным. При этом нет гарантии, что по

завершении вычислений простым округлением результатов до ближайших целых значений будет получено правильное решение. Поэтому на практике для рассматриваемых систем уравнений применяются специальные методы: подбора (путем проб) и безошибочных вычислений [2], являющиеся достаточно трудоемкими.

Метод вычислений

Пусть

$$\mathbf{AX} = \mathbf{B} \quad (1)$$

целочисленная система уравнений порядка n с ограничением в виде

$$|a_{ij}| = \sum_{i=1, i \neq j}^n |a_{ij}|, \quad i = 1, \dots, n. \quad (2)$$

Известно, что любое целое число β можно представить конечной суммой в виде

$$\beta = \sum_{l=1}^p \beta^{(l)},$$

составленной из $\beta^{(l)} \in \{1, -1\}$. Вместе с тем одним из свойств систем линейных уравнений является то, что если

$$\mathbf{X} = \sum_{l=1}^p \mathbf{X}^{(l)}, \quad (3)$$

где $X^{(l)} = (x_1^{(l)}, x_2^{(l)}, \dots, x_n^{(l)})^t$, то каждому составляющему вектора X соответствует $B^{(l)} = AX^{(l)}$ и

$$B = \sum_{l=1}^p B^{(l)}.$$

Тогда справедлива следующая запись:

$$-B + AX^{(1)} + AX^{(2)} + AX^{(3)} + \dots + AX^{(p)} = 0. \quad (4)$$

Введем обозначение

$$\varepsilon^{(l+1)} = \varepsilon^{(l)} + AX^{(l)} \quad (5)$$

и, положив $\varepsilon^{(1)} = -B$, будем формировать $x_i^{(l)} \in \{1, -1\}$ на каждом l -м шаге таким образом, чтобы выполнялось условие $\|\varepsilon^{(l+1)}\| < \|\varepsilon^{(l)}\|$, где

$$\|\varepsilon^{(l)}\| = \sum_{i=1}^p \|\varepsilon_i^{(l)}\|,$$

а на k -м шаге — соответственно $\|\varepsilon^{(k)}\| = 0$, учитывая реализацию целочисленных вычислений.

Для этого достаточно, чтобы

$$\text{sign}(x_i^{(l)}) = -\text{sign}(a_{ij} \varepsilon_j^{(l)}). \quad (6)$$

Тогда в худшем случае для любого i при $\text{sign}(x_i^{(l)} a_{ij}) \neq \text{sign}(x_j^{(l)} a_{ij})$, где $j = 1, \dots, n$ и $j \neq i$, исходя из ограничения на соотношение целочисленных величин коэффициентов (2), гарантируется уменьшение соответствующей i -й невязки минимум на единицу, что обеспечивает сходимость метода к решению.

Однако приведенный метод гарантирует вычисление целочисленных неизвестных лишь в том случае, если все они либо четны, либо нечетны.

Теорема. Если неизвестные целочисленной системы линейных алгебраических уравнений имеют различную четность, то вычислить их предлагаемым методом нельзя.

Доказательство. Предположим обратное: за одинаковое число шагов вычислительного процесса при $x_i^{(l)} \in \{1, -1\}$ можно вычислить неизвестные системы линейных уравнений, отличающиеся четностью. Тогда, исходя из принятого способа формирования решения (3), четное неизвестное $2x_1$ можно представить как

$$\sum_{n_1} (1) + \sum_{n_2} (-1) = 2x_1, \quad (7)$$

а любое нечетное как

$$\sum_{n_1'} (1) + \sum_{n_2'} (-1) = 2x_2 + 1. \quad (8)$$

Тогда по предположению

$$n_1 - n_2 = n_1' - n_2'. \quad (9)$$

Выразив n_1 и n_1' соответственно через n_2 и n_2' из (7) и (8) и подставив в (9), получим $2x_1 = 2x_2 + 1$,

что противоречиво при любых целых значениях x_1 и x_2 .

Очевидный выход из этой ситуации, расширяющий класс решаемых систем линейных уравнений, заключается в приведении всех неизвестных перед решением к четности путем введения в (1) любого четного целого коэффициента α : $A\alpha X = \alpha B$. Истинное решение системы будет получено последующим делением вычисленных неизвестных на α .

Пусть для любого m значение неизвестного

$$|x_m| = \max_{i=1, n} \{ |x_i| \}.$$

Тогда $\text{sign}(b_m) = \text{sign}(a_{mm}x_m^{(1)})$ и $\text{sign}(\varepsilon_{mm}^{(l)}) = \text{sign}(a_{mm}x_m^{(l)})$. Следовательно x_m , а соответственно и все остальные неизвестные будут сформированы за $\alpha|x_m|$ шагов вычислительного процесса.

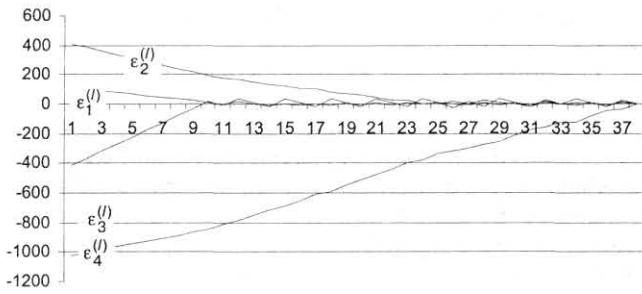
Особенности использования

Идеология разрядных методов вычислений [4, 5], к которым относится рассматриваемый метод, изначально строилась на предположении их эффективной аппаратной реализации, поэтому решение, получаемое с использованием простейших арифметических операций, может быть выполнено в любом микропроцессоре и микроконтроллере, включая простейшие [3].

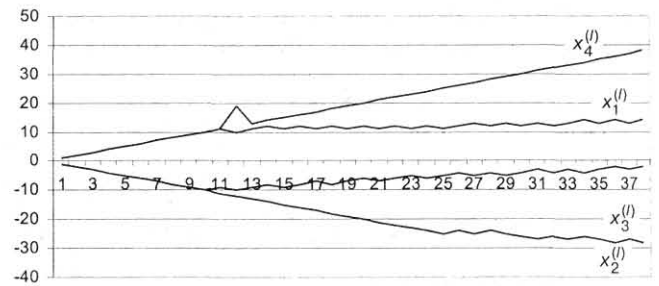
Операционная «стоимость» одного шага вычислений составляет эквивалент n операций сравнений и $n^2 + n$ операций целочисленного сложения или вычитания. Это выгодно отличает предлагаемый метод от известных, ставших уже классическими, для сравнения со «стоимостью» которых можно рекомендовать табл. 24.1 из [1] и табл. 2.3 и 2.4 из [4].

Наиболее экономичным с точки зрения минимизации времени решения и простоты реализации в целочисленной арифметике является использование $\alpha = 2$. При этом $2B$ получается арифметическим сдвигом влево на один двоичный разряд соответствующих свободных членов, а деление на двойку вычисленных значений неизвестных — сдвигом вправо на один двоичный разряд. Необходимое число шагов вычислительного процесса, определяемое в этом случае как $2|x_m|$, даже при больших значениях $|x_m|$ не должно стать существенным препятствием для программной реализации метода ввиду малой операционной «стоимости» одного вычислительного шага и высокой скорости выполнения целочисленных арифметических операций в микропроцессорах.

Операции сложения и вычитания p -разрядных целочисленных данных ($p > r$, где r — разрядность процессора, процессорного модуля) позволяют не вводить дополнительные ограничения на величины элементов систем уравнений, исключая возможность переполнения при промежуточных вычислениях значений $\varepsilon_i^{(l)}$.



■ Рис. 1. Динамика уменьшения значений невязок на шагах вычислительного процесса



■ Рис. 2. Динамика изменения приближений неизвестных на шагах вычислительного процесса

Численный пример

В качестве демонстрации характеристик предложенного метода рассмотрим решение тестовой целочисленной СЛАУ из [6], имеющей точное решение $\mathbf{X} = (7; -14; -1; 19)^t$:

$$\begin{pmatrix} -13 & 2 & -1 & 3 \\ -6 & 19 & 4 & 5 \\ -1 & -5 & -37 & 7 \\ -5 & -9 & 8 & 23 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -61 \\ -217 \\ 233 \\ 520 \end{pmatrix}$$

Изменение значений невязок на каждом шаге вычислительного процесса приведено на рис. 1. До начала вычислений невязки увеличены в два раза для выполнения требования одинаковой четности неизвестных. Момент окончания решения СЛАУ определяется в процессе основных вычислений по значению $\|\varepsilon^{(l)}\| = 0$, что является результатом 38-го вычислительного шага (максимальное значение неизвестного $x_4 = 19$).

Динамика сходимости приближений неизвестных к решению представлена на рис. 2. Для получения истинных значений неизвестных необходимо выполнить деление (сдвиг вправо на один двоичный разряд) вычисленных значений.

Заключение

Рассмотренный метод решения целочисленных систем уравнений следует рассматривать как гибридный, поскольку он не является прямым — неизвестные формируются накоплением приращений, но и итерационным — решение вычисляется за конечное число арифметических действий, кратное величине максимального неизвестного.

Литература

1. Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления. — М: Наука, 1984. — 320 с.
2. Грегори Р., Кришнамурти Е. Безошибочные вычисления. Методы и приложения. — М.: Мир, 1988. — 208 с.
3. www.microchip.com
4. Байков В. Д., Вашкевич С. Н., Сергеев М. Б. Прикладные задачи микропроцессорных систем управления и контроля. — СПб: Политехника, 1992. — 223 с.
5. Евдокимов В. Ф., Стасюк А. И. Параллельные вычислительные структуры на основе разрядных методов вычислений. — Киев: Наукова думка, 1987. — 312 с.
6. Фаддеева В. Н., Колотилина Л. Ю. Вычислительные методы линейной алгебры: Набор матриц для тестирования. — Л.: Наука, Ч. 1, 1982; Ч. 2, 3, 1983. — 387 с.

УДК: 681.324

КОМПЛЕКСНОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ УПРАВЛЕНИЯ СТРУКТУРНОЙ ДИНАМИКОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Б. В. Соколов,

д-р. техн. наук, профессор

Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН)

К. А. Малюгин,

начальник отдела интегрирования и администрирования
программного обеспечения

Главное управление Центрального банка РФ по Санкт-Петербургу

Рассмотрены методологические и методические основы организации и проведения комплексного моделирования процессов управления структурной динамикой информационной системы. Предложен оригинальный полимодельный комплекс, описывающий процессы функционирования и развития информационной системы и включающий в себя динамические аналитико-имитационные модели управления операциями, потоками, ресурсами, структурами информационной системы. Рассмотрен пример использования указанного комплекса.

Within the proposed study the methodological and methodical bases for the problem of integrated modeling of information system structure dynamics control. An original multiple-model complex for simulation of information system (IS) was proposed. The complex included analytical-simulation models of control processes for IS operation, flows, resources and structures. The utilization example of multiple-model complex was outlined.

В современных условиях успешное развитие любой крупной бизнес-системы (БС) во многом определяется эффективностью функционирования ее информационной системы (ИнС), обеспечивающей процессы обработки разнородных потоков информации и управления структурной динамикой как БС, так и как самой ИнС. В настоящее время при описании ИнС используется, как правило, их полиструктурное многомодельное описание. В качестве основных структур в этом случае рассматриваются следующие типы структур: топологическая, техническая, технологическая (функциональная), организационная структуры; структура программно-математического и информационного обеспечения (ПМО и ИО) [1, 2].

При таком варианте описания ИнС как объект исследования представляет собой интегрированную территориально-распределенную сложную организационно-техническую систему (СОТС). Одной из основных особенностей указанной системы является то, что ее параметры и структуры на различных этапах жизненного цикла изменяются под действием объективных и субъективных причин. Другими словами, на практике мы постоянно

сталкиваемся со структурной динамикой СОТС.

На рис. 1 представлены возможные варианты структурной динамики СОТС. В этих условиях для повышения (сохранения) требуемого уровня работоспособности и возможностей СОТС либо обеспечения наилучших условий их функционирования при деградации указанных систем необходимо осуществлять управление структурами СОТС (в том числе управление реконfigurацией структур СОТС). Задачи управления структурной динамикой СОТС по своему содержанию относятся к классу задач структурно-функционального синтеза облика СОТС, формирования и реализации соответствующих программ управления их развитием [1, 3–5].

Главная особенность решения задач рассматриваемого класса состоит в следующем. Определение оптимальных программ управления основными элементами и подсистемами СОТС может быть выполнено лишь после того, как будет известен перечень функций и алгоритмов обработки информации и управления, который должен быть реализован в указанных элементах и подсистемах.

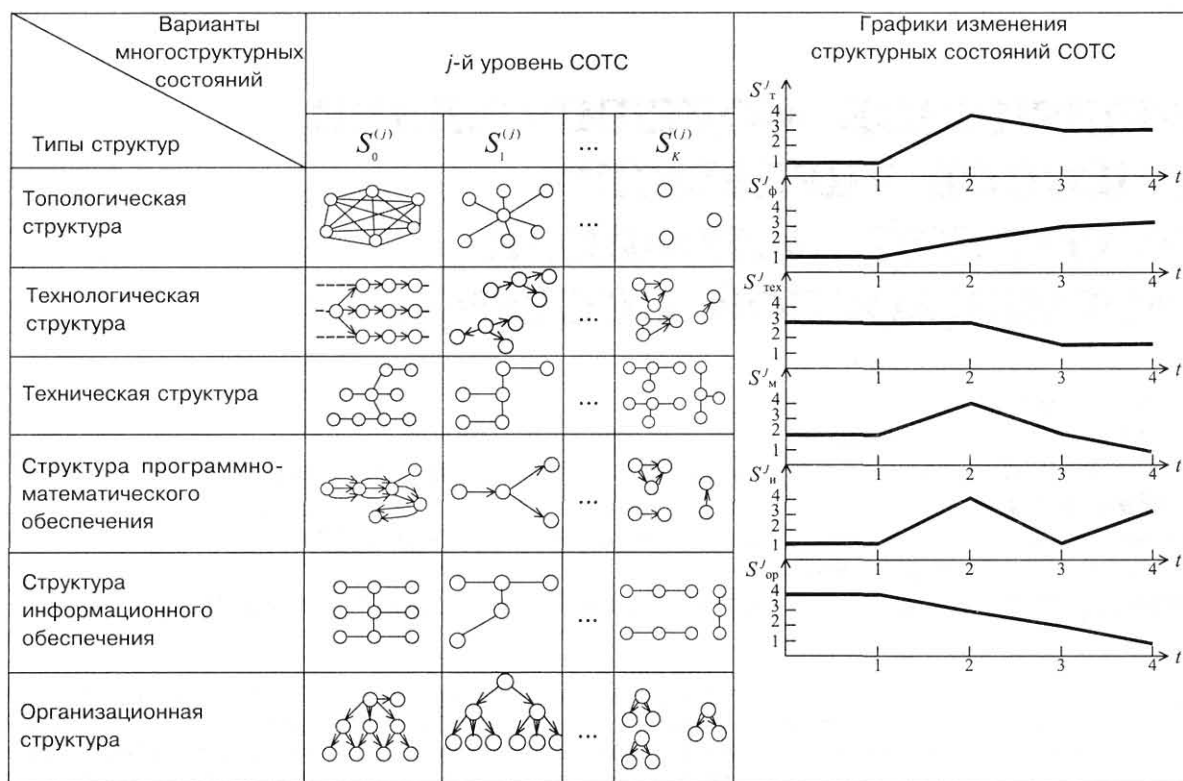


Рис. 1. Возможные варианты структурной динамики сложной организационно-технической системы:

$S_k^{(j)}$ — k-е структурное макросостояние СОТС; $S_1^j, S_{\text{ф}}^j, S_{\text{тех}}^j, S_{\text{м}}^j, S_{\text{и}}^j, S_{\text{ор}}^j$ — состояние топологической, функциональной, технической структур, состояние структур программно-математического и информационного обеспечения, состояние организационной структуры соответственно

В свою очередь, распределение функций и алгоритмов по элементам и подсистемам СОТС зависит от структуры и параметров законов управления данными элементами и подсистемами. Трудность разрешения данной противоречивой ситуации усугубляется еще и тем, что под действием различных причин (внутренних либо внешних, объективных либо субъективных) изменяется состав и структура СОТС.

К настоящему времени рассматриваемый класс задач структурно-функционального синтеза и управления развитием СОТС исследован недостаточно глубоко. Получены новые научные и практические результаты в рамках следующих направлений исследований [3–7]:

— синтез технической структуры СОТС при известных законах функционирования основных элементов и подсистем СОТС (первое направление исследований);

— синтез функциональной структуры СОТС или, по-другому, синтез законов и программ управления основными элементами и подсистемами СОТС при известной технической структуре (второе направление исследований);

— синтез программ создания и развития новых поколений СОТС без учета этапа совместного функционирования существующей СОТС и внедряемой СОТС (третье направление исследований);

— одновременный синтез как технической, так и функциональной структур СОТС (четвертое направление исследований)

Кратко остановимся на состоянии и результатах исследований, проведенных в рамках каждого из перечисленных направлений. К настоящему времени как в нашей стране, так и за рубежом выполнен достаточно большой объем научных работ, посвященных решению проблемы синтеза технических структур СОТС различных классов. При этом обобщенная задача синтеза (выбора) структуры (либо структур) СОТС сводится, как правило, к постановке и решению следующей оптимизационной задачи [3, 4]:

$$\bar{S} \{ \{ \bar{f} \in \bar{F}(\bar{\pi}) \bar{R} \{ \bar{m} \in \bar{M} \} \} \} \rightarrow \text{extr}, \quad (1)$$

$$\bar{\pi} \in \bar{P}, \quad (2)$$

$$\bar{f} \in \bar{F}(\bar{\pi}), \quad (3)$$

$$\bar{m} \in \bar{M}, \quad (4)$$

где \bar{P} — множество возможных принципов и алгоритмов управления $\bar{\pi} \in \bar{P}$, используемых при формировании облика системы или ее элементов; \bar{F} — множество взаимосвязанных функций (задач, операций), выполняемых системой. Каждому набору принципов и алгоритмов управления $\bar{\pi}$ соответ-

ствуется множество функций $\bar{F}(\bar{\pi})$, из которого при проектировании системы необходимо выбрать подмножество, достаточное для реализации выбранных принципов управления, т. е. выбрать $\bar{f} \subset \bar{F}(\bar{\pi})$; M — множество возможных взаимосвязанных элементов СОТС. Подобными элементами применительно к СОТС могут быть средства приема, регистрации, передачи и обработки информации, пункты обслуживания и управления и т. п.; R — операция отображения элементов множества \bar{F} на элементы множества M .

В этом случае данное оптимальное отображение должно обеспечивать экстремум некоторой (либо некоторых) целевой функции S при выполнении заданных ограничений.

Дальнейшее обобщение рассматриваемой задачи синтеза проводится в направлении учета факторов неопределенности, вызванных воздействием внешней среды, а также в направлении, связанном с постановкой и решением многокритериальных задач выбора оптимальных структур СОТС.

Одной из основных трудностей решения сформулированной задачи синтеза структур вида (1)–(4) является ее большая размерность, которая определяется общим числом переменных и ограничений, используемых при формализации и решении данной задачи. Поэтому на основу решения рассматриваемого класса задач положены различные варианты реализации процедур декомпозиции (композиции), агрегирования (деагрегирования), координации, с помощью которых удается «проклятие размерности». Дополнительные особенности данные процедуры приобретают еще и из-за того, что большинство переменных, используемых при решении задачи синтеза структур СОТС, должны принимать целочисленные значения. В качестве примера успешного решения задач данного класса можно привести результаты, полученные в работах [3, 4]. Авторами данных работ был предложен агрегативно-декомпозиционный подход, предусматривающий (на основе альтернативно-графовой формализации) построение многоуровневого комплекса взаимосвязанных моделей различного типа: оптимизационных аналитических и/или имитационных моделей синтеза структур СОТС.

Многочисленные исследования задач синтеза структур СОТС вида (1)–(4) показали [3, 4], что если при формировании облика СОТС для некоторых ее элементов и подсистем возникают проблемы ликвидации больших (пиковых) информационных нагрузок, то в этом случае должны оптимизироваться сами правила, алгоритмы функционирования указанных элементов и подсистем (второе направление исследований).

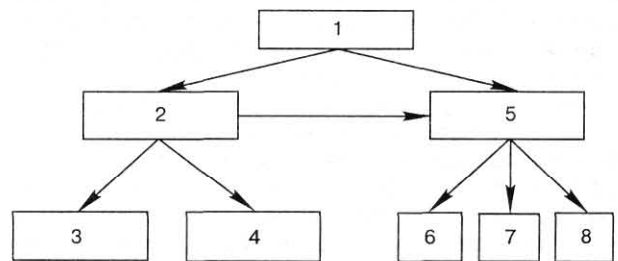
Исследование данных вопросов (задач синтеза функций, алгоритмов функционирования и правил поведения элементов заданной иерархической системы) уже более 40 лет выполняется различными научными школами в нашей стране и за рубежом в рамках интенсивно развивающейся теории управления [5–7].

В связи с этим, кратко остановимся лишь на одном из частных направлений этих исследований, непосредственно связанных с тематикой данной статьи, а именно с исследованием задач управления структурной динамикой СОТС. Эволюцией исследований данных вопросов определялась эволюцией создания и применения соответствующих СОТС с управляемой структурной динамикой.

На рис. 2, взятом из работы [5], представлена классификация СОТС, в которых к настоящему времени получила практическую реализацию концепция управляемой структурной динамики.

К настоящему моменту времени задачи управления базовыми СОТС с реконфигурацией исследованы в наибольшей степени. Полученные фундаментальные научные результаты (концепции, принципы, способы, методы управления данными СОТС) позволили получить интересные прикладные результаты в различных предметных областях [5–8]. В отличие от базовых СОТС с реконфигурацией исследования проблем создания и применения интегрированных систем активного управления к настоящему моменту находятся в начальной стадии. Это, в первую очередь, касается СОТС с управляемой структурной динамикой, в которых реализуются элементы интеллектуального управления. Данные системы функционируют в условиях существенной неопределенности, связанной, как уже указывалось ранее, с изменением содержания целей и задач, стоящих перед СОТС, воздействием возмущающих факторов со стороны внешней среды и имеющих целенаправленный и/или нецеленаправленный характер.

Для указанных систем использование традиционных подходов при формализации и решении задач управления структурной динамикой СОТС становится уже малоэффективным. Одним из наиболее перспективных путей конструктивного учета перечисленных выше факторов сложности в соот-



■ Рис. 2. Классификация СОТС с управляемой структурной динамикой:

1 — СОТС с управляемой структурной динамикой; 2 — базовые СОТС с реконфигурацией; 3 — системы координатно-параметрического (операторного, бинарного) управления (СКПУ); 4 — системы с активной управляемой технологией (САУТ); 5 — интегрированные системы активного управления (ИСАУ); 6 — системы альтернативного и многоуровневого управления; 7 — системы отказоустойчивого и самовосстанавливаемого управления; 8 — системы, в которых реализованы концепции интеллектуального управления

ветствующих задачах управления является направление исследований, получившее название «искусственного интеллекта» [6]. В рамках данного научного направления применительно к исследованию проблем управления сформировалась отдельная область исследований, связанная с созданием и эксплуатацией систем интеллектуального управления. К настоящему времени наиболее интересные научные и практические результаты, связанные с процессами формирования и реализации управляющих воздействий в СОТС, получены на основе применения производственных систем в форме нечетких и других правил, а также применения искусственных нейронных сетей [6].

Вместе с тем, до сих пор наименее исследованными остаются задачи структурно-функционального синтеза СОТС, в ходе решения которых требуется одновременно синтезировать как функциональную, так и техническую структуры указанных систем с учетом различных классов пространственно-временных, технических и технологических ограничений [1, 2, 8]. Такого рода задачи возникают, например, на этапе применения СОТС для следующих ситуаций: при изменении способов и целей функционирования СОТС, их содержания, последовательности выполнения; при перераспределении и децентрализации функций, задач, алгоритмов управления, информационных потоков между уровнями СОТС, при реконфигурации их структур в случае их деградации. Таким образом, в современных условиях назрела острая необходимость в формировании теоретических основ управления структурной динамикой СОТС. С нашей точки зрения, данная теория должна иметь междисциплинарный характер и основываться на результатах, полученных в классической теории управления, в исследовании операций, в искусственном интеллекте, в теории систем и системном анализе. Два последних научных направления позволяют в дальнейшем корректно осуществлять переход от слабоструктурированной к структурированной постановке задач управления структурной динамикой. На рис. 3 показано место разрабатываемой теории в рамках предлагаемых междисциплинарных исследований.

Учет структурной динамики СОТС (в том числе и для ИнС) особенно важен в настоящее время, так как в рамках современных транснациональных корпораций уже сейчас создаются так называемые «виртуальные производства» с постоянно изменяющейся конфигурацией. В этом случае, исходя из стратегических целей менеджмента данных корпораций, деятельность реальных производственно-экономических систем, входящих в их состав, координируется на базе соответствующих ИнС, взаимодействующих по телекоммуникационным сетям. Решение задач управления структурной динамикой как самих БС, так и ИнС представляет собой чрезвычайно сложную многоплановую проблему. Важная роль в исследовании данной проблемы отводится методам моделирования. При этом сам процесс моделирования должен иметь комплексный характер и проводиться в рамках соответствующих



■ Рис. 3. Определение места теории управления структурной динамикой СОТС в междисциплинарных исследованиях

ющих интегрированных систем поддержки принятия решений (ИСППР), в состав которых, в общем случае, должны входить: имитационные системы (ИС); интеллектуальные информационно-поисковые системы (ИИПС); экспертные системы поддержки принятия решений (ЭСППР); расчетно-логические системы (РЛС); инструментальные CASE-средства автоматизации проектирования.

Центральной проблемой при построении ИСППР на основе перечисленных средств автоматизации принятия решений была и остается проблема согласования и координации их работы. Для этого необходимо выбрать в качестве базовой структуры одну из тех математических структур, на которых строится большинство рассматриваемых систем, входящих в состав ИСППР.

Проведенный анализ показывает, что в качестве такой базовой структуры следует выбрать динамический альтернативный системный граф (АСГ) с управляемой структурой, с помощью которого можно описать многомодельную структурную динамику как БС, так и их ИнС. Более того, если провести динамическую интерпретацию данного АСГ, то в этом случае, как показали предварительные исследования, можно с единых позиций описать как задачу структурно-функционального синтеза основных элементов и подсистем ИнС, так и процессы управления их развитием [1, 8, 9]. Территориальная распределенность БС и соответствующей ИнС, а также их высокая структурная динамика потребовали от исследователей поиска новых концептуальных моделей процессов управления данными системами. Одним из наиболее перспективных в этом случае является направление, связанное с построением мультиагентных систем [6, 10].

На рис. 4 представлен один из возможных вариантов описания агента, входящего в состав соответствующей мультиагентной системы. В работах [11, 12, 13] данные агенты названы активными подвижными объектами (АПО). Указанные объек-

ты представляют собой искусственно созданные материальные объекты (аппаратно-программные комплексы), перемещающиеся в пространстве и осуществляющие взаимодействие (информационное, вещественное, энергетическое) с объектами обслуживания (ОБО), другими АПО. На рис. 4 показана обобщенная структура АПО как объекта управления. Данный объект АПО состоит из четырех подсистем, которым поставлены в соответствие четыре процесса (вида функционирования): 1) процесс движения; 2) процессы взаимодействия с ОБО, другими АПО; 3) процессы функционирования целевой и обеспечивающей аппаратуры; 4) процессы расхода и (или) пополнения ресурсов. Предлагаемая структура АПО, как показано в работах [11–13], допускает весьма многообразную интерпретацию. Так, например, в качестве АПО может выступать наземное, воздушное, надводное или подводное средство передвижения с установленным на нем аппаратно-программным комплексом, АПО можно интерпретировать также и как космическое средство (КСр), и, наконец, как мобильный агент [10].

Наряду с АПО в составе рассматриваемой концептуальной модели следует выделить ОБО, с которыми АПО осуществляют взаимодействие. Физически ОБО могут представлять собой естественные или искусственно созданные материальные объекты или естественную материальную среду, созданную природой. Так, в качестве ОБО может рассматриваться часть наземной или водной поверхности, области воздушного, подводного, подземного пространства, подвижные объекты, информационные ресурсы сети Internet. В этом случае природа и формы взаимодействия АПО с ОБО

могут носить как активный, так и пассивный характер (т. е. сопровождаться либо не сопровождаться изменением состояния ОБО).

Из вышеизложенного следует, что для каждого типа АПО должна быть сформулирована цель его функционирования, связанная с процессами взаимодействия с ОБО, с другими АПО, и определена соответствующая последовательность действий, в ходе которой будет достигнута поставленная цель. Анализ показывает, что в этом случае концептуальное описание указанной деятельности АПО удобно проводить с использованием понятия «операция», под которой в дальнейшем будем понимать действие или систему действий, объединенных общим замыслом и единой целью [1, 7, 8]. Из анализа рис. 4 следует, что цель функционирования АПО реализуется в ходе выполнения им операций, связанных с информационным, вещественным и энергетическим обменом с ОБО, другими АПО. Таким образом, операция обмена является основным системообразующим фактором, объединяющим (интегрирующим) различные виды деятельности АПО (движение, работу аппаратуры, расход ресурса и т. п.). Содержание и специфика каждой выполняемой операции АПО находит свое отражение в задании соответствующих параметров, характеризующих: результаты выполнения операций (объем, качество, время выполнения операции и т. п.); расход ресурсов при выполнении операции; информационные и материальные потоки, возникающие в ходе выполнения операции.

В работах [11–13] было предложено различать АПО первого и второго рода (АПО-I, АПО-II). При этом на АПО-I принято возлагать задачи непо-



■ Рис. 4. Обобщенная структурная схема активного подвижного объекта (АПО)

средственного взаимодействия с ОБО и внешней средой, а на АПО-II — задачи обеспечения указанного взаимодействия. Переходя непосредственно к задачам комплексного моделирования структурной динамики ИнС на основе концепции АПО, остановимся, прежде всего, на теоретико-множественном описании данных задач. Для проведения данного теоретико-множественного описания воспользуемся ранее предложенной графической интерпретацией процессов управления структурной динамикой СОТС. В рассматриваемых задачах каждая из структур может быть задана своим динамическим альтернативным мультиграфом следующего вида [9]:

$$G_{\alpha}^{t_j} = \langle T_{\alpha}^{t_j}, \Gamma_{\alpha}^{t_j}, P_{\alpha}^{t_j} \rangle, \quad (5)$$

где α — индекс, характеризующий вид структуры, $A = \{T, \Phi, \text{Тех}, \text{ПМО}, \text{ИО}, \text{ОР}\}$ — множество индексов, соответствующих топологической (Т), функциональной (Ф), технической структурам (Тех), структурам программно-математического и информационного обеспечения (ПМО, ИО), организационной структуре (ОР) ИнС; $t_j \in \{t_1, t_2, \dots, t_f\}$ — линейно упорядоченное множество моментов времени перехода ИнС от одного многоструктурного макросостояния S_{δ} к другому макросостоянию $S_{\delta'}$; $\delta, \delta' \in \{1, \dots, \Delta\}$ — множество номеров макросостояний ИнС; $X_{\alpha}^{t_j} = \{x_{\alpha}^{t_j}\}$, $\Gamma_{\alpha}^{t_j} = \{\gamma_{\alpha}^{t_j}\}$ — множество вершин и дуг графа, соответствующих фиксированному виду структуры « α » в момент времени t_j и задающих ее альтернативные варианты; $P_{\alpha}^{t_j} = \{p_{\alpha}^{t_j}\}$ — множество «весов» (приоритетов), количественно характеризующих значимость каждой из дуг в соответствующем альтернативном мультиграфе. Кроме того, задано множество допустимых (исходя из содержательной постановки каждой конкретной задачи) операций отображения указанных выше альтернативных мультиграфов с динамической структурой вида (1) друг на друга:

$$M_{\langle \alpha, \alpha' \rangle}^{t_j} : G_{\alpha}^{t_j} = G_{\alpha'}^{t_j}, \quad (6)$$

а также множество допустимых операций отображения многоструктурных макросостояний ИнС S_{δ} друг на друга:

$$\Pi_{\langle \delta, \delta' \rangle}^{t_j} : S_{\delta} \rightarrow S_{\delta'}. \quad (7)$$

При этом предполагается, что каждое многоструктурное макросостояние ИнС в момент времени t_j задается в результате операции композиции соответствующих альтернативных мультиграфов, описывающих каждый вид структуры « α ».

С учетом вышеизложенного, графическая интерпретация рассматриваемой задачи синтеза облика ИнС и управления ее структурной динамикой сводится к поиску такого макросостояния $S_{\delta}^* \in \{S_1, \dots, S_{\Delta}\}$ и такой последовательности выполнения операций отображения во времени $\Pi_{\langle \delta, \delta' \rangle}^{t_j}$, $\Pi_{\langle \delta', \delta'' \rangle}^{t_{j+1}}$, \dots , $\Pi_{\langle \delta_{f-1}, \delta_f \rangle}^{t_f}$, при которых обеспечивается выбор наилучшей (с точки зрения обобщенного показателя эффективности функционирования

ИнС) программы создания, применения и развития ИнС.

На первом шаге решения предложенной обобщенной проблемы синтеза облика ИнС (в том числе и БС) и управления ее структурной динамикой необходимо, прежде всего, рассмотреть вопрос формирования методологических основ проведения соответствующих исследований.

Масштабность и сложность рассматриваемой в этой связи проблемы требует выбора в качестве базовой методологии принципов, концепций, подходов современного обобщенного системного анализа, представляющего собой одно из главных направлений реализации системного подхода, в рамках которого на основе гармоничного сочетания формально-математических и логико-эвристических методов осуществляется конструктивное решение разнородных и разноуровневых задач анализа и синтеза ИнС на различных этапах их жизненного цикла [2, 8].

Конструктивное решение рассматриваемой проблемы поиска и выбора наилучших вариантов создания и развития ИнС предполагает, во-первых, построение соответствующего полимодельного комплекса, описывающего различные аспекты жизненного цикла существующей и создаваемой ИнС, во-вторых, разработку методов, алгоритмов и методик многокритериального синтеза структуры перспективной ИнС, и, в-третьих, разработку поэтапной интерактивной процедуры поиска решения задачи синтеза ИнС и программ их развития.

Кратко остановимся на вопросе формирования облика одной из компонент ИСППР, а именно, на составе и структуре имитационной системы (ИС), с помощью которой могут быть решены различные классы задач управления структурной динамикой ИнС. В современных ИС выбор допустимых альтернатив основывается на сужении (сжати) множества рассматриваемых вариантов экзогенных переменных путем отбраковки доминируемых по заданным отношениям предпочтения альтернатив. Указанные процедуры по своему содержанию близки к идеям, реализованным в многочисленных модификациях метода «ветвей и границ». При отбрасывании доминирующих экзогенных переменных в зависимости от этапа решения задачи выбора, обеспеченности исходными данными ЛПР каждый раз пользуется такими моделями и методами получения релаксированных решений исходной задачи, чтобы оценки затрат на реализацию полученных решений (затрат на расход используемого ресурса) становились все более и более точными по мере сужения множества допустимых альтернатив.

Исследование процессов управления структурной динамикой разнородных классов АПО, в том числе и исследование задач анализа и синтеза БС и соответствующих ИнС, показало, что данные процессы и системы имеют многоуровневый и полифункциональный характер. Данное представление процессов функционирования АПО и соответствующей системы управления (СУ АПО) повлияло на выбор струк-

туры банка моделей разрабатываемого специального программно-математического обеспечения ИС, в котором необходимо, прежде всего, выделить три основных блока (рис. 5): 1) модели функционирования СУ АПО и объектов обслуживания (ОБО) (блок I); 2) модели оценки и анализа состояния АПО, СУ АПО, оценки обстановки (блок II); 3) модели принятия решений в СУ АПО (блок III).

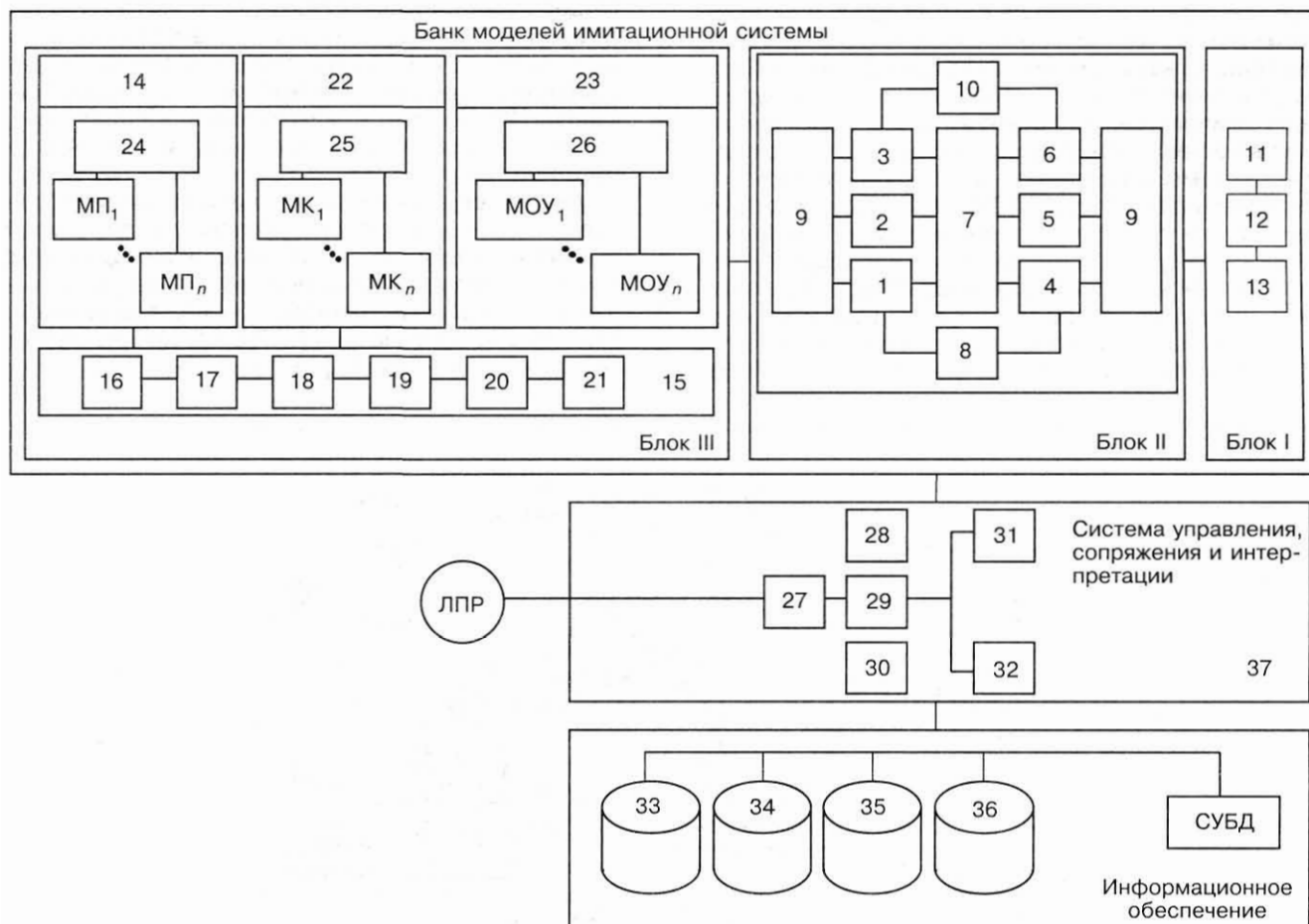
Блок моделей функционирования СУ АПО, ОБО включает в себя: модели функционирования АПО-I, системы АПО-I, группировки систем АПО-I (блоки 1, 2, 3); модели функционирования АПО-II (блок 4), системы АПО-II (блок 5), группировки систем АПО-II (блок 6); модели взаимодействия основных элементов и подсистем СУ АПО между собой и ОБО (блок 7); модели функционирования ОБО (блок 8); модели воздействия внешней среды на СУ АПО (блок 9); модели имитации результатов целевого применения СУ АПО (блок 10).

Блок моделей оценки и анализа состояния АПО, СУ АПО, оценки обстановки включает в себя: модели и алгоритмы оценки и анализа состояния

движения, аппаратуры, ресурсов и обмена АПО (блок 11); модели и алгоритмы оценки и анализа состояния ОБО (блок 12); модели и алгоритмы оценки и анализа ситуаций и обстановки (блок 13).

В блок моделей принятия решений в СУ АПО входят: модели и алгоритмы долгосрочного и оперативного планирования операций взаимодействия (ОВ) в СУ АПО (блок 14); модели и алгоритмы управления структурами СУ АПО (блок 15), а именно: топологической (блок 16), технической (блок 17), технологической (блок 18), организационной структурой (блок 19); структурой СПМО (блок 20), информационной структурой (блок 21); модели и алгоритмы коррекции долгосрочных и оперативных планов проведения ОВ в СУ АПО (блок 22); модели и алгоритмы решения задач координации в СУ АПО на этапах планирования (блок 24), коррекции (блок 25), оперативного управления (блок 26); модели и алгоритмы оперативного управления элементами и подсистемами СУ АПО (блок 23).

На рис. 5 также изображена система управления, сопряжения и интерпретации (блок 37), в которую



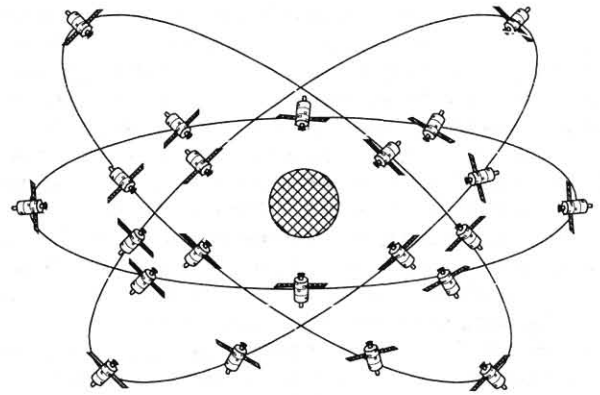
■ Рис. 5. Структура имитационной системы:

$MP_1, \dots, MP_n, MK_1, \dots, MK_n, MOY_1, \dots, MOY_n$ — соответственно модели планирования, коррекции и оперативного управления АПО, входящими в СУ АПО (1, ..., n)-го типов

входят: общая диалоговая система управления СПМО (блок 27), локальные системы управления и сопряжения (блок 28), блок обработки, анализа и интерпретации результатов планирования, управления, моделирования (блок 30), блок формализации сценариев моделирования (блок 31), блок параметрической и структурной адаптации СПМО (блок 32), блок выработки рекомендаций по организации процедур моделирования и принятия решений (блок 29).

Важную роль в решении задач анализа и синтеза СУ АПО играет информационное обеспечение, включающее в себя: базы данных о состоянии АПО (блок 33), СУ АПО (блок 35), ОБО (блок 34), в целом по обстановке (блок 35); базы данных об аналитических и имитационных моделях функционирования и принятия решений в СУ АПО (блок 36).

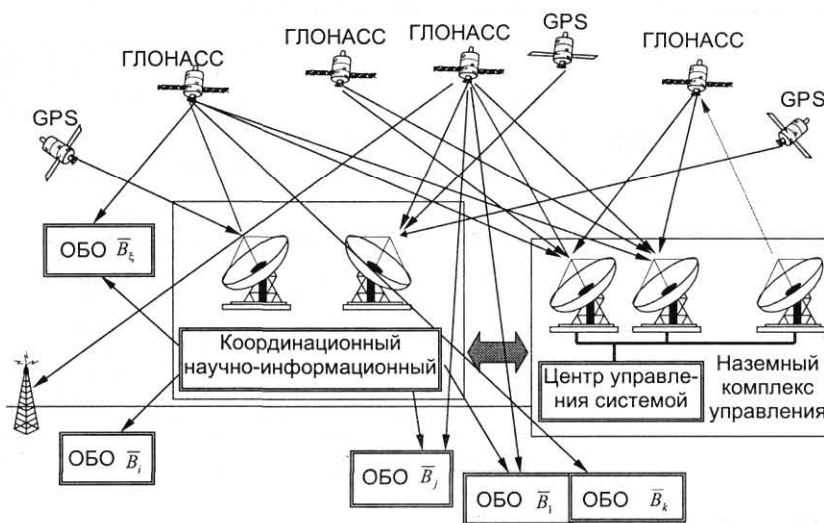
К настоящему времени, используя предложенную имитационную систему, удалось успешно решить целый ряд важных задач управления структурной динамикой СОТС в таких предметных областях, как космонавтика, информатика, менеджмент [9, 12, 13, 18]. В качестве примера, иллюстрирующего конструктивность предлагаемого подхода к организации и проведению комплексного моделирования СОТС, рассмотрим постановку и основные этапы совместного решения задачи оперативного синтеза технологии автоматизированного управления (ТАУ) космическими средствами (КСр) и задачи планирования работы бортовых и наземного комплексов управления (БКУ и НКУ), входящих в состав соответствующей автоматизированной системы управления навигационными космическими аппаратами (АСУ НКА) [19]. Базовую основу технического обеспечения указанной системы составляет территориально распределенная информационно-вычислительная сеть, топологическая структура которой постоянно изменяется за счет пространственного перемещения как НКА, так и отдельных подсистем НКУ. Вычислительные ком-



■ Рис. 6. Топологическая структура орбитальной группировки навигационных КА

плексы, входящие в состав указанной сети, располагаются как на борту каждого НКА, так и в подсистемах НКУ. На рис. 6 и 7 в качестве примеров приведены обобщенные топологические структуры существующих систем НКА и НКУ Российской Федерации (система ГЛОНАСС) и США (система GPS) [17]. В этом случае, рассматривая КСр как один из возможных вариантов реализации концептуальной модели АПО, можно, используя разработанную ИС, решить сформулированные ранее прикладные задачи управления данными космическими средствами.

Анализ содержательной и формальной постановок задач управления структурной динамикой АСУ НКА показывает, что из-за их полимодельного многокритериального описания требуется разработка специальных процедур (сценариев) интерактивного взаимодействия ЛПР с ИС для получения решения данных задач [13–15].



■ Рис. 7. Топологическая структура наземного комплекса управления навигационными КА

Рассмотрим некоторые возможные процедуры (сценарии) моделирования процессов управления структурной динамикой АСУ НКА. Для удобства дальнейшего изложения материала будем предполагать, что обобщенное полимодальное многокритериальное описание задач управления структурной динамикой АСУ НКА имеет следующий вид [12]:

$$J(x(t), u(t), o(t), t) \rightarrow \underset{u \in \Delta}{extr}, \quad (8)$$

$$\Delta = \{u | x(t) = \varphi(x(t), u(t), \xi(t), \beta, t), y(t) = \psi(x(t), u(t), \xi(t), v, t), x(T_0) \in X_0(\beta), x(T_f) \in X_f(\beta), u(t) \in Q(x(t), t), \xi(t) \in \Xi(x(t), t), \beta \in B, x(t) \in \bar{X}(t)\}, \quad (9)$$

где $J(x(t), u(t), o(t), t) = \|J^{(g)T}, J^{(o)T}, J^{(k)T}, J^{(p)T}, J^{(n)T}, J^{(e)T}, J^{(c)T}, J^{(v)T}\|$ — вектор показателей эффективности функционирования АСУ НКА, а $J^{(g)T}, J^{(o)T}, J^{(k)T}, J^{(p)T}, J^{(n)T}, J^{(e)T}, J^{(c)T}, J^{(v)T}$ — соответственно векторы показателей эффективности управления движением, операциями взаимодействия, каналами, ресурсами, потоками, параметрами операции, структурами, вспомогательными операциями в АСУ НКА, а индексы в обозначении векторов «g», «o», «k», «p», «n», «e», «c», «v» соответствуют представленным в [12, 13] моделям управления движением (M_g), управления операциями взаимодействия (M_o), управления каналами (M_k), управления ресурсами (M_p), управления потоками (M_n), управления параметрами операции (M_e), управления структурами (M_c), управления вспомогательными операциями (M_v); управления движением (M_g), $x(t), y(t)$ — соответственно обобщенные векторы состояния и выходных характеристик динамической системы, описывающей процессы управления структурной динамикой АСУ НКА; $u(t)$ — обобщенный вектор управляющих воздействий, $\xi(t)$ — обобщенный вектор возмущающих воздействий; β — вектор структурных параметров (характеристик) АСУ НКА, определяющих ее облик; $Q(x(t), t), \Xi(x(t), t), B$ — заданные области изменений значений векторов $u(t)$ и $\xi(t)$; $\beta; X_0(\beta), X_f(\beta), \bar{X}(t)$ — заданные области изменений вектора $x(t)$ соответственно в начальный, конечный и текущий моменты времени.

В выражении (9) переходная и выходная функции $\varphi(x(t), u(t), \xi(t), \beta, t)$ и $\psi(x(t), u(t), \xi(t), v, t)$ в общем случае задаются в аналитико-алгоритмическом (имитационном) виде в рамках предлагаемой ИС. Возможные варианты описания и реализации указанных функций (в более общем случае — операторов) как раз и определяют содержание тех методов и алгоритмов, которые могут быть положены в основу построения процедур получения скоординированных решений в задачах управления структурной динамикой АСУ НКА.

К настоящему времени разработаны многочисленные подходы, способы, методы, алгоритмы и методики координационного выбора на комплексах разнородных моделей, описывающих различ-

ные предметные области [1–8, 13–15]. В табл.1 и 2 приведены примеры реализации такого рода процедур.

Перечисленные в табл.1 и 2 возможные схемы координации (согласования) моделей и показателей эффективности отличаются друг от друга: способами генерации допустимых альтернативных решений в задачах анализа и синтеза технологий автоматизированного управления КА; правилами проверки алгоритмически и аналитически заданных ограничений; способами перехода от одного шага интерактивного сужения множества допустимых альтернатив к другому шагу. Предлагаемый в статье подход к организации и проведению комплексного моделирования СОТС основывается на оригинальном динамическом описании происходящих в ней процессов и отличается от существующих тем, что в рамках единого модельного описания удается совместно решать задачи синтеза как технической, так и функциональной структур СОТС [12, 13]. Применительно к АСУ НКА технология решения указанных задач включает в себя следующие этапы.

Шаг 1. Для заданной внешней обстановки и фиксированных исходных данных, при использовании блока моделей функционирования КА (блоки I, II, см. рис. 5), осуществляется (при необходимости) формирование дополнительных исходных для

■ Таблица 1

Процедуры решения задач УСД АСУ КА	Модели задач УСД АСУ КА					
	Задача 1	Задача 2	Задача 3	Задача 4	Задача 5	Задача 6
AOM → AH → K	+					
ИОМ → AH → K				+	+	+
AOM → ИОМ → AH → K		+	+			
(AOM ⊂ ИОМ) → AH → K			+			
(ИОМ ⊂ AOM) → AH → K			+		+	+
(AOM ₁ ∪ ИОМ ∪ AOM ₂) → AH → K				+	+	+

Принятые обозначения: AOM — аналитическая оптимизационная модель; AH — анализ полученных результатов (проводимый автоматически либо с привлечением ЛПР); K — коррекция полученных решений.

■ Таблица 2

Задача 1	Задача 2	Задача 3	Задача 4	Задача 5	Задача 6
$f_0^{(a)} \rightarrow \text{extr}_{\Delta^{(a)}}$	$f_0^{(a)} \rightarrow \text{extr}_{\Delta^{(u)}}$	$f_0^{(a)} \rightarrow \text{extr}_{\Delta^{(a)} \cap \Delta^{(u)}}$	$f_0^{(u)} \rightarrow \text{extr}_{\Delta^{(u)}}$	$f_0^{(u)} \rightarrow \text{extr}_{\Delta^{(u)}}$	$f_0^{(u)} \rightarrow \text{extr}_{\Delta^{(a)} \cap \Delta^{(u)}}$
Принятые обозначения: $\Delta^{(a)}, \Delta^{(u)}$ — множества (либо подмножества) допустимых альтернатив вида (9), заданные соответственно аналитически и алгоритмически; $f_0^{(a)}, f_0^{(u)}$ — обобщенные показатели эффективности функционирования АСУ НКА, полученные на основе решения соответствующих задач векторной оптимизации и заданные соответственно аналитически и алгоритмически.					

решения задачи планирования процессов функционирования АСУ НКА. С этой целью проводятся многократные машинные эксперименты с моделями движения НКА, моделями функционирования телекоммуникационных сетей, информационно-вычислительных комплексов. В результате указанного моделирования получают множество допустимых вариантов перемещения объектов, допустимых вариантов значений максимально (минимально) допустимых скоростей передачи и обработки данных в АСУ НКА. Указанная информация с помощью матричных временных функций $\varepsilon_{ij}(t)$, $\Theta_{i \in j \lambda}(t)$ (см. [11–13]) и соответствующих констант [16] вводится в детерминированные модели управления структурной динамикой (УСД) АСУ НКА. При этом, используя нелинейные преобразования, предложенные ранее в работах [12, 18], осуществляют формирование множества допустимых альтернатив вида (9) на основе ограничений вида «и—или», косвенно задающих возможные варианты ТАУ НКА.

Шаг 2. Исходя из выделяемых на планирование и синтез ТАУ временных и вычислительных ресурсов, решается задача планирования самого процесса планирования и синтеза. Для этого используются модели, входящие в состав блока III.

Шаг 3. Проводится анализ существования решений в поставленных задачах выбора. Для этого, используя заранее рассчитанное множество достижимости динамической системы, описывающей процессы функционирования АСУ НКА, можно оценить возможность удовлетворения заданных крайних условий как в задаче планирования, так и в задаче синтеза ТАУ НКА.

Шаг 4. Если решение задач существует, то автоматически (по заранее заданному алгоритму) либо в интерактивном режиме с использованием моделей, входящих в блок III, осуществляется формирование как допустимого распределения функций управления между подсистемами АСУ НКА, так и плана ее функционирования.

Шаг 5. Используя модели блока III, осуществляют многоэтапный итерационный поиск наилучшей ТАУ и соответствующего оптимального плана функционирования АСУ НКА.

Шаг 6. Полученный на шаге 5 оптимальный план (планы) функционирования АСУ НКА, с использованием моделей, входящих в блоки III, I, проверяют на устойчивость по отношению к вариантам возмущающих воздействий, которые могут возникнуть на этапе реализации плана, оценивается чувствительность составленного плана к изменению ис-

ходных данных. Для этого в интерактивном режиме в ходе проведения имитационных экспериментов с моделями АСУ НКА, входящими в состав блоков I, III, осуществляется формирование и анализ нечетко-возможных и интервальных характеристик, показателей, характеризующих устойчивость и чувствительность разработанных планов функционирования элементов и подсистем АСУ НКА. Принципиальное отличие проводимых в ИС машинных экспериментов от классических схем методов статистических испытаний состоит в том, что в ходе моделирования возмущающих воздействий с заданным уровнем рефлексии моделируются возможные варианты компенсации данных воздействий. Для этого при планировании заранее выделяются соответствующие резервы, разрабатываются алгоритмы реконфигурации структур АСУ КА. Оценка чувствительности плана функционирования АСУ НКА проводится в соответствии с алгоритмами, описанными в [19].

Шаг 7. Проводится параметрическая и структурная адаптация составленного плана и технологии автоматизированного управления КСр к возможному возмущающим воздействиям (с привлечением блоков I, II, III ИС) [19].

Шаг 8. Формирование и интерактивный анализ ТАУ КСр и составленного плана (планов) функционирования АСУ НКА

Проведенные машинные эксперименты показали [12, 18], что при использовании динамических моделей управления КСр существенно сокращается размерность задач управления, решаемых в каждый момент времени, за счет рекуррентного описания моделей, повышается оперативность их решения на основе декомпозиции и распараллеливания вычислительного процесса, широкого использования оверлейных режимов работы ЭВМ. Так, например, для 15 НКА, четырех радиотехнических станций, одного центра управления полетом типовая технология управления указанными КСр и соответствующий оптимальный оперативный план функционирования НКУ для интервала времени 12 ч были рассчитаны на ПЭВМ Pentium III в течение 15 мин.

Оригинальность предложенного варианта формализации процессов управления КСр состоит в том, что в этом случае удается в явном виде связать содержательные аспекты функционирования АСУ КСр с технологией управления КСр. Так, например, исследуя модель M_o, M_e, M_n методами теории оптимального управления, можно оценить степень взаимного влияния вариантов распределения

ресурсов КСр, схем проведения траекторных изменений и способов маршрутизации информационных потоков на качество процессов применения КА по целевому назначению. Кроме того, в этом случае можно найти оптимальные программы управления потоками, измерениями и ресурсами АСУ КСр.

Проведенный анализ показывает, что разработанный полимодельный комплекс и соответствующая имитационная система могут быть использованы для решения различных прикладных задач

в тех предметных областях, при формальном описании которых оказываются справедливыми предпосылки, положенные в основу предложенной концептуальной модели АПО [11, 12].

Результаты, представленные в данной статье, получены при выполнении исследований в рамках проекта 1992р, выполняемого совместно с European Office Of Aerospace Research and Development, и гранта # 02-07-90463 Российского Фонда фундаментальных исследований.

Литература

1. **Балашов Е. П.** Эволюционный синтез систем. — М.: Радио и связь, 1983. — 328 с.
2. **Клир Дж.** Системология. Автоматизация решения системных задач. — М.: Радио и связь, 1990. — 540 с.
3. **Добановский С. А., Озеряный Н. А.** Системы автоматического управления с реконфигурацией // Измерение, контроль, автоматизация. — 1990. — № 4(76). — С. 62–80.
4. **Васильев С. Н.** От классических задач регулирования к интеллектуальному управлению // Теория и системы управления. — 2001. — № 1. — С. 5–22; № 2. — С. 5–21.
5. **Цвиркун А. Д.** Основы синтеза структуры сложных систем. — М.: Наука, 1982. — 200 с.
6. **Цвиркун А. Д., Акинфиев В. К.** Структура многоуровневых и крупномасштабных систем. — М.: Наука, 1993. — 160 с.
7. **Моисеев Н. Н.** Элементы теории оптимальных систем. — М.: Наука, 1974. — 526 с.
8. **Николаев В. И., Брук В. М.** Системотехника: методы и приложения. — Л.: Машиностроение, Ленинградское отделение, 1985. — 199 с.
9. **Соколов Б. В., Григорьев К. Л.** Методологические основы обоснования и выбора вариантов создания и развития информационных систем // Изв. вузов. Приборостроение. — 2000. — Т. 43. — № 8. — С. 82–85.
10. **Городецкий В. И.** Информационные технологии и многоагентные системы // Проблемы информатизации. — 1998. — № 1. — С. 3–14.
11. **Калинин В. Н.** О теории управления активными подвижными объектами // Изв. вузов. Приборостроение. — 1981. — Т. 23. — № 6. — С. 26–31.
12. **Соколов Б. В.** Комплексное планирование операций и управление структурами в АСУ активными подвижными объектами. — Министерство обороны СССР, 1992. — 232 с.
13. **Калинин В. Н., Соколов Б. В.** Многомодельный подход к описанию процессов управления космическими средствами // Теория и системы управления. — 1995. — № 1. — С. 56–61.
14. **Методологические** вопросы построения имитационных систем: Обзор / С. В. Емельянов, В. В. Калашников, В. И. Лутков и др. Под научн. ред. Д. М. Гвишиани, С. В. Емельянова. — М.: МЦНТИ, 1973. — 87 с.
15. **Краснощевков П. С., Морозов В. В., Федоров В. В.** Декомпозиция в задачах проектирования // Изв. АН СССР. Техн. кибернетика. — 1979. — Т. 6. — № 2. — С. 7–18.
16. **Соколов Б. В., Курносков А. Н.** Модель и алгоритм оперативного перераспределения функций управления между техническими комплексами // Изв. вузов. Приборостроение. — 2000. — Т. 43. — № 8. — С. 66–72.
17. **Соловьев Ю. А.** Системы спутниковой навигации. — М.: ЭКО-ТРЕНДЗ, 2000. — 267 с.
18. **Калинин В. Н., Соколов Б. В.** Динамическая модель и алгоритм оптимального планирования комплекса работ с запретами на прерывание // Автоматика и телемеханика. — 1987. — № 1. — С. 106–114.
19. **Скурихин В. И., Забродский В. А., Копейченко Ю. В.** Адаптивные системы управления машиностроительным производством. — М.: Машиностроение, 1989. — 207 с.

УДК [519.872+519.876.5]

ОЦЕНКА СИСТЕМЫ МОДЕЛИРОВАНИЯ GPSS WORLD

Ю. И. Рыжиков,

д-р техн. наук, профессор

Военно-космическая академия им. А. Ф. Можайского

Дан обзор проблемы математического моделирования систем и сетей с очередями. Описаны новые возможности, предоставляемые имитационной системой общего назначения GPSS World. Приводятся результаты ее тестирования. Предлагается комбинированное (в зависимости от сложности задачи, временных ресурсов и квалификации разработчика) использование численных методов теории очередей, программирования на языках типа Фортран и упомянутой системы.

It is presented a review of the mathematical modeling of the queuing systems and networks, and the description of new possibilities of the GPSS World. The results of its testing are discussed. The combined (depending on the problem complexity, needed precision, time resources and explorer qualification) using of numerical queuing theory methods, programming in Fortran and GPSS modeling is proposed.

В разных областях техники, в организации производства, в социальной сфере и военном деле постоянно возникает необходимость решения вероятностных задач, связанных с работой систем массового обслуживания разного вида требований. Термин «массовое обслуживание» предполагает многократную повторяемость ситуаций в том или ином смысле (много прибывших в систему и обслуженных заявок, большое число находящихся в эксплуатации аналогичных систем) и статистическую устойчивость картины. Выводы и рекомендации, получаемые методами теории массового обслуживания (ТМО), применимы лишь при наличии одного или обоих перечисленных факторов повторяемости.

По справедливому замечанию виднейшего советского специалиста в этой области Б. В. Гнеденко, легче указать ситуации, где не может быть использована ТМО (она же — теория очередей), чем перечислить все сферы ее потенциального применения. Подчеркнем, что задачи указанных типов приходится решать не только при проектировании вновь создаваемых систем и сетей обслуживания (СМО, СеМО), но и в процессе эксплуатации имеющих: при увеличении нагрузки, изменении трудоемкости обработки заявок, выходе из строя, деградации или модернизации техники, снижении квалификации персонала, пересмотре требований к оперативности обработки заявок и т. п.

Многообразие приложений теории массового обслуживания определяет постоянно растущий интерес к ней, а сложность возникающих задач не позволяет получить исчерпывающие решения на базе аналитических методов [1, 4] даже при чис-

ленной реализации последних. В таких ситуациях приходится прибегать к имитационному моделированию. В [3] приведен перечень заслуживающих внимания источников, дан обзор теоретических основ построения имитационных моделей и техники их реализации на алгоритмических языках высокого уровня и специальных языках моделирования.

Ни один из языков моделирования не оказал на имитацию столь большого воздействия, как GPSS (General Purpose Simulation System — система имитационного моделирования общего назначения) [6–10]. Например, сходство языков СЛЭНГ и GPSS простирается вплоть до заимствования большинства ключевых слов.

Базовые понятия GPSS

Модель GPSS состоит из сети блоков, представляющих необходимые действия или задержки транзактов (клиентов, покупателей, ремонтируемые телевизоры, телефонные звонки, электронные сигналы и т. п.), которые последовательно проходят через блоки. Например, блок GENERATE вводит в модель новые транзакты, воспроизводя рекуррентный поток заявок с требуемым распределением интервалов между смежными заявками. Транзакты перемещаются в системных времени и пространстве, переходя от одного блока модели к другому и воздействуя на них. Транзакты возникают и уничтожаются, могут расщепляться и сливаться. Входя в блок, транзакт вызывает определяемую типом блока подпрограмму, которая обрабатывает

соответствующее событие. Далее транзакт в общем случае пытается войти в следующий блок. Продвижение продолжается до тех пор, пока не окажется, что очередной блок должен выполнить одну из следующих функций:

- удалить транзакт из модели;
- временно заблокировать его в предыдущем блоке до выполнения некоторых условий;
- задержать его на определяемое моделью время.

Тогда начинается продвижение другого транзакта и т. д. — до завершения моделирования. За один шаг работы управляющей программы производится поочередный просмотр всех имеющихся процессов и имитация тех, которые могут быть запущены. При этом может освободиться ранее занятый ресурс или оказаться порожденным процесс более высокого приоритета, и просмотр списка процессов можно начать сначала.

Описание траектории транзакта содержит порядок и имена используемых им «устройств» (приборов обслуживания) или «памятей»; временные задержки; логические условия, управляющие продвижением транзактов; точки маршрута, в которых производится сбор данных об ожидании или контролируется время прохождения и т. п. Некоторые из названных объектов должны быть предварительно описаны специальными командами. Параметры транзакта могут менять свои значения на траектории, оказывать влияние на маршрут и собираемую статистику. Маршруты следования транзактов, временные характеристики, условие прекращения моделирования, требуемые показатели эффективности задаются пользователем.

При описании этих процессов и сборе итоговой статистики используются таймер модельного времени, стандартные числовые атрибуты (СЧА) и параметры транзактов, а также определяемые пользователем переменные, выражения и таблично задаваемые функции. Наиболее сложные аспекты модели — планирование предстоящих событий и очередности их обработки — автоматически реализует планировщик событий (интерпретатор) GPSS.

GPSS World

Система GPSS прошла длинный путь развития. До недавних пор ее главными недостатками были низкий уровень входного языка, невозможность реализации нестандартных вычислений, скрытость от программиста внутренней логики поведения модели. Они резко ограничивали реальную полезность системы. Радикальные изменения произошли одновременно с переходом к ее Windows-версиям. «Студенческая» версия GPSS World 4.2.1 (2000 г.), далее сокращенно именуемая GPSS/W, бесплатно распространяется через Интернет (электронный адрес www.minutemansoftware.com); она отличается от коммерческой лишь количественными ограничениями (до 150 блоков в составе модели) и несовместимостью получаемых программ с составленными ранее на GPSS/PC.

Инсталляция

Для работы с GPSS/W необходимо иметь операционную систему Windows 95, 98 или 2000; ПЭВМ класса Pentium; оперативную память не менее 32 Мб; 10 Мб места на винчестере. От пользователя требуются опыт работы с персональной ЭВМ, навыки программирования на языке типа Паскаль или Фортран, знакомство с основными идеями имитационного моделирования.

Заархивированная система укладывается в пять дискет. Tutorial должен загружаться отдельно. Инсталляция запускается по Setup. Транслированные программы (Имитации) и отчеты помещаются в ту же директорию, где находятся исходные модели.

В состав электронной документации к системе входят Справочник (GPSS World Reference Manual), Руководство программиста (Tutorial) и подробно прокомментированные примеры тестовых программ (все — с расширением GPS), собранные в папке Sample Models. Объектами моделирования являются:

- производственные задачи (ASSEMBLY — процесс сборки насосов; BICYCLE — велосипедная фабрика; MANUFACT — система производства электроники; FOUNDRY — литейный цех; TVREPAIR — мастерская по ремонту телевизоров; TEXTILE — текстильная фабрика; QCONTROL — система контроля качества; TAPEPREP — подготовка управляющих программ для металлорежущего оборудования с ЧПУ; FMSMODEL — гибкое автоматизированное производство; ROBOTFMS — роботизированное производство);

- транспортные задачи (TRAFFIC — транспортные потоки на Т-образном перекрестке; TURNSTIL — турникет на футбольном стадионе; SHIPPORT — порт; LOCKSIMN — канал и шлюз);

- задачи управления запасами (OILDEPOT — нефтебаза; ORDERPNT — запас с критическим уровнем возобновления; PERIODIC — запас с периодическим контролем уровня; STOCKCTL — иерархическая система складов);

- проблемы торговли (POWDER — торговля стиральным порошком; SUPERMRK — универсам);

- системы и сети связи (TELEPHON — простая телефонная система; EXCHANGE — телефонный коммутатор с внутренними и внешними линиями связи; ETHERNET — сеть Ethernet);

- абстрактные постановки общеметодического характера (MULTIWAY — сеть Ethernet; GRAECOLATIN — планирование эксперимента с применением греко-латинских квадратов; QTHEORY — задача теории очередей с известным решением; PREDATOR — динамика популяций).

В той же папке имеются многочисленные примеры моделей из нестарейшей книги Т. Шрайбера [9], перепрограммированные с учетом новых возможностей GPSS/W. Соответственно в упомянутой книге можно найти развернутые пояснения к этим моделям (за ними сохранены данные Шрайбером номера). Разумеется, надо иметь в виду возможности аналогий (Barber Shop может служить моделью не только парикмахерской, но и космети-

ческого салона, мастерской по изготовлению ключей, ремонтных служб, служб спасения, пунктов первой медицинской помощи и т. п.). Кроме того, подходящим может оказаться только название проблемной области, но не специфика ситуации. Поэтому предпочтительны абстрактная формулировка задачи (в терминах теории очередей) и инвариантные к приложениям приемы моделирования.

Попутно отметим, что специалистам в области техники связи и компьютерных систем будут полезны примеры из учебного пособия Б. Я. Советова и С. А. Яковлева [7]. В частности, там разобраны задачи моделирования абстрактной ациклической сети обслуживания, участка транспортного цеха автомобильного завода, процесса передачи данных в информационно-вычислительной сети с коммутацией пакетов, работы концентратора сети интегрального обслуживания, испытательного участка сборочного цеха предприятия, локальной сети связи ЭВМ со случайным доступом к моноканалу, вычислительного центра коллективного пользования, а также приведены 45 разнообразных заданий на моделирование.

Принципиальные особенности *GPSS/W* можно сгруппировать в несколько разделов.

1. Каждая разработка в духе современных технологий программирования рассматривается как *проект* (эта терминология в электронной документации к системе не используется). Проект состоит из объектов четырех типов: 1) исходный модуль; 2) Имитация; 3) отчеты; 4) текстовые файлы. Модель может иметь иерархическое строение и тогда включает в себя командой *INCLUDE* ранее разработанные фрагменты. В рамках проекта может существовать множество автоматически нумерованных версий. Размеры задачи физически доступным объемом оперативной памяти не лимитируются.

Для каждого из шагов проекта есть свои *инструменты*: полноэкранный Текстовый редактор, пункт *Create Simulation* в меню *Command*, множество окон контроля и команд управления моделированием, автоматический генератор отчетов, средства настройки отчетов, статистического анализа. Команда *CONDUCT*, начиная *Эксперимент*, берет на себя управление Имитацией.

В системе реализованы элементы объектного программирования в классическом его понимании: *наследование свойств* (установок) Имитации от исходной модели и *полиморфизм* — автоматическое преобразование операндов выражений (в частности, строки символов в число и обратно) к требуемому типу.

2. Интерфейс ориентирован на пользователя. Однострочный текстовый редактор для набора модели заменен на полноэкранный с реализацией стандартных функций *Windows* — табуляция, выделение текста, работа с буфером *Clipboard*, управление типом и размером шрифтов *True Type*, операции с файлами. Есть возможность замены традиционно используемого вместо знака умножения *#* на привычную ***. Компилятор формирует

список синтаксических ошибок, обеспечивает автоматический подвод курсора к месту очередной ошибки и вывод диагностики в статусную строку.

Динамика модели может отслеживаться как в числовой, так и в графической форме с помощью окон около 20 разных типов (в зависимости от класса отображаемых сущностей), причем в двух режимах: детальном и обзорном. Размер и положение окон изменяются стандартными средствами. Степень загруженности ресурсов показывается цветом, можно наблюдать за перемещением активного транзакта. В окна можно выводить СЧА и накопленную статистику. Отображение внизу главного окна значения *таймера* облегчает диагностику зависания. В особо сложных случаях можно получить *Snapshots* — снимки мгновенного состояния *цепей событий*, используемые для отладки логики модели. Имеются средства интерактивного ввода команд, контрольных остановов и трассировки. Результаты моделирования могут быть автоматически представлены в виде гистограмм соответствующих распределений.

3. Резко расширены вычислительные возможности. Отменены обязательная целочисленность операндов и принудительное округление результатов. Максимальное число измерений матрицы увеличено до шести. Используемый при инициализации матриц атрибут *UNSPECIFIED* предупреждает работу с «неприсвоенными значениями». Язык *GPSS* дополнен языком программирования расчетов *PLUS*, близким к традиционным алгоритмическим. На нем можно записывать как выражения для операндов блоков, так и процедуры пользователя, в том числе содержащие циклы и разветвления. Это практические сняло ограничения на формируемые результаты и способы их обработки. Перечень встроенных математических функций дополнен функциями генерации случайных чисел для 24 законов распределения, что избавляет пользователя от необходимого ранее мучительно-го построения таблиц обратной функции распределения для каждой комбинации параметров.

Пользователь может написать собственную *PLUS*-процедуру формирования псевдослучайных чисел с требуемым законом распределения. При этом аргумент *RNj* следует умножить на 0,001. Обычно используемая в подобных программах замена $1-U$ на одинаково распределенное с ним U здесь недопустима, поскольку в среднем одно из тысячи $\{U_i\}$ окажется нулевым. Поэтому при вычислении $\ln U$ для обычной генерации показательного распределенных величин методом обратной функции гарантируется соответствующее число аварийных остановов модели.

Операции над строками позволяют редактировать стандартный отчет и применять процедуры с текстовыми аргументами. Появились операции с потоками данных (открытие и закрытие файлов, запись и считывание информации из них).

4. Система берет на себя стандартную статистическую обработку результатов моделирования. Появившаяся еще в *GPSS/PC* команда *ANOVA* однофакторного дисперсионного анализа заменена

многофакторным эквивалентом. Имеются средства автоматической разработки и проведения статистических экспериментов (включая дробные факторные) по регрессионному анализу и оптимизации моделей.

По результатам прогона программы автоматически формируется стандартный отчет. Содержание отчета зависит от состава модели и при необходимости корректируется через Установки модели. Нестандартные вычисления могут быть запрограммированы с помощью процедур встроенного языка программирования и выведены в файл.

Работа с системой

Типичный проект начинается с создания или модификации модели. Следует иметь в виду, что модель — не просто текстовый файл: она содержит также настройки Settings, закладки Bookmarks и результаты, которые не могут читаться внешним редактором.

Все настройки задаются при создании модели через меню Edit/Settings. Последующие фазы проекта эти установки только наследуют. На вкладке Simulation задаются число попыток входа в блок до прерывания и ограничение глубины ссылок (для предупреждения зацикливания). Посредством вкладки Reports можно вывести результаты в окно вместо файла, указать желаемые дополнительные подотчеты, выводить результаты с шестью знаками после десятичной точки вместо стандартных трех. Вкладка Random Numbers распределяет по типам блоков потоки случайных чисел. Вкладка Function Keys связывает операторы с функциональными клавишами. Вкладка Expressions позволяет задать список выражений, часто используемых в процессе диалога с моделью, и при необходимости выбирать их из списка через меню Select вместо повторного набора.

Этапы моделирования

Работа начинается с составления GPSS-программы. Посредством команды INCLUDE «<имя_файла>» в нее могут вставляться ранее отлаженные фрагменты (например, PLUS-процедуры). Вставки из файлов автоматически нумеруются, и диагностика ошибок в них сопровождается указанием номера вставки.

Компиляция организуется по команде Command/Create Simulation. Синтаксический контроль достаточно развит (в частности, контролируется парность скобок). При обнаружении ошибок создается их циклический список. Каждый выбор элемента списка вызывает сообщение о типе ошибки в статусную строку и помещает точку вставки точно перед ошибочным элементом. Характер «претензий» системы обычно достаточно ясен: Line 10, Col 20, Invalid Delimiter. Expecting a GPSS Verb.

Наибольшие трудности вызывает использование некорректных имен: слишком коротких или совпадающих со стандартными числовыми атрибутами. Оперативное запоминание внесенных в модель исправлений не требуется: компиляции подвергается правленный текст.

Команда START определяет окончание моделирования: ее счетчик должен уменьшиться до нуля. Любой оператор модели TERMINATE с непустым полем операнда A уменьшает счетчик при входе в него транзактов. Такой оператор в программе должен быть единственным.

Динамические окна могут быть открыты только после компиляции модели командой Window/Simulation Window/ <тип_окна>. Одновременно могут быть открыты несколько окон. Открытие online-окон замедляет моделирование, поскольку система вынуждена ждать их обновления. Эти задержки обычно недооцениваются. Для иллюстрации степени замедления приведем данные по длительности прогонов одной и той же модели:

- все окна закрыты — 4 с;
- выводится таймер — 97 с;
- открыты окно блоков и таймер — 5271 с (почти полтора часа!).

Для сравнения укажем, что та же модель на Фортране была реализована за 0,05 с. Очевидно, что открывать динамические окна следует лишь для режима пошаговой отладки.

Начатое моделирование можно прервать командой HALT, кнопка которой имеется во всех динамических окнах. После ее выполнения можно выбирать блоки и задавать контрольные остановы в них (кнопка с красным светом). При отработке очередного останова первоначальное условие останова, если оно задано по времени, удаляется, а все остальные условия сохраняются. Продолжение моделирования с использованием команды STEP позволяет автоматически остановить его точно после заданного числа входов активного транзакта в блоки модели.

Контролировать промежуточные результаты (и просматривать окончательные) можно через окна Графики (Plot) и Выражений (Expression). Окно Таблиц позволяет наблюдать сходимость исследуемых распределений в смысле гистограммы, среднего значения и среднеквадратического отклонения. В меню Windows поддерживается список всех открытых окон, включая отчеты по прогонам предыдущих версий модели.

Сбор статистики

Для типовых имитаций автоматически собираемая статистика представляется вполне достаточной. При необходимости в более подробных сведениях можно определить в командах TABLE или QTABLE таблицы (гистограммы) распределений. Фактический сбор данных в первом случае обеспечивается специальными блоками TABULATE, а во втором — при входе в блок DEPART. Упомянутые таблицы (если отметить их в установке отчета),

а также средние значения и среднеквадратические отклонения будут выводиться автоматически и могут просматриваться в окне таблиц.

Еще более детальные сведения можно собирать и обрабатывать с помощью скалярных и матричных *сохраняемых значений* и *PLUS*-процедур. Затем «историческая» база данных может быть обработана с помощью ANOVA для получения доверительных интервалов. Результаты нестандартной обработки статистики можно вывести в файл потока данных.

Эксперименты над моделями

Задача проектирования, например, сети обслуживания должна решаться с позиций системного подхода, учитывающего ее многоаспектность и иерархичность. Естественно, что те же свойства приобретает и процесс проектирования, проходящий уже на фазе моделирования несколько этапов и требующий применения разнообразного математического аппарата. В процессе имитационного *эксперимента* над моделью с учетом случайных исходов необходимо:

- классифицировать факторы на существенные и несущественные;
- разделить и оценить количественно влияния факторов и их комбинаций на целевую функцию;
- найти наивыгоднейшую комбинацию факторов.

Структура эксперимента (количество прогонов, выбор исходных данных для них и характер обработки результатов) определяется его целью. Первая проблема решается средствами дисперсионного анализа, вторая — регрессионного и факторного анализа (скрининг), третья — методами статистической оптимизации. *GPSS/W* имеет встроенные средства для решения всех трех.

Язык *PLUS* имеет особый класс процедур (*EXPERIMENT*) с широкими возможностями, включая исследование поверхности отклика. Эксперимент организуется записанной на языке *PLUS* процедурой управления прогонами, которая для каждого прогона сама готовит исходные данные и вызывает модель. Программа Эксперимента может быть составлена как вручную, так и автоматически — с помощью диалога, реализуемого через меню *Edit/Insert Experiment*. Управление прогонами реализуется с помощью встроенной процедуры *DoCommand(<строка>)*. Аргументы определяют, какая часть поверхности отклика подлежит исследованию в процессе Эксперимента. *EXPERIMENT* заполняет глобальную Матрицу Результатов, которая затем передается процедуре ANOVA.

Дисперсионный анализ

К большинству сложных систем применим принцип Парето, согласно которому 20 % факторов определяют свойства системы на 80 % (аналогичное соотношение известно для авторов научной про-

дукции коллективов [13], наименований предметов снабжения в системах управления запасами [4] и др.). Поэтому первоочередной задачей исследователя имитационной модели является отсеивание несущественных факторов, позволяющее уменьшить размерность задачи оптимизации модели.

Анализ дисперсии (*Analysis of Variation — ANOVA*) оценивает отклонение наблюдений от общего среднего. Затем вариация разбивается на части, каждая из которых имеет свою причину. Остаточная часть вариации (*residual*), которую не удается связать с условиями эксперимента, считается его случайной ошибкой. Для подтверждения значимости используется специальный тест — *F*-статистика.

Дисперсионный анализ определяет, есть ли эффект. Регрессионный анализ позволяет прогнозировать отклик (значение целевой функции) в некоторой точке пространства параметров. Ожидаемое значение отклика вычисляется по формуле регрессии с известными коэффициентами для значений факторов, лежащих внутри выбранных диапазонов. Непосредственной задачей регрессионного анализа является оценка коэффициентов регрессии.

Число уровней каждого фактора обычно ограничивается двумя (тогда объем полного эксперимента с *k* факторами составит 2^k прогонов). Кроме того, желательно уменьшить число прогонов, не жертвуя наиболее важной информацией. Последнее соображение приводит к *дробным* факторным экспериментам. Дробность плана определяется долей числа требуемых прогонов от объема полного факторного эксперимента и может быть целой отрицательной степенью двойки.

Для каждой компоненты плана с учетом расстановки плюс- и минус- единиц в его матрице записывается линейное уравнение регрессии относительно расчетных эффектов и свободного члена. Правой частью уравнения становится результат соответствующего прогона. Компоненты решения делятся на размах (произведения размахов) отображаемых факторов, что дает коэффициенты их влияния относительно середин диапазонов. На этой основе можно планировать дальнейшие действия: игнорировать незначимые факторы, более тщательно исследовать существенные, оптимизировать последние.

Оптимизирующий эксперимент для нахождения оптимума в пространстве непрерывных параметров базируется на методе скорейшего спуска (подъема). Это предполагает построение эмпирической аппроксимации поверхности отклика; вычисление градиента (его проекции являются коэффициентами чувствительности целевой функции к изменению параметров, т. е. коэффициентами регрессии); определение направления спуска и величины шага по нему — с учетом ограничений на значения параметров.

Факторные и оптимизационные эксперименты по автоматически создаваемым Имитациям реализуются очень медленно (на ПЭВМ Pentium I/166 — часами).

Тестирование GPSS/W

Общим правилом работы с новым программным обеспечением математического характера является его тестирование на задачах с известным решением (предпочтительно аналитическим). Для системы массового обслуживания *M/M/1* теоретическая средняя длина очереди $q = \rho^2 / (1 - \rho) = 0,9^2 / (1 - 0,9) = 8,1$. Среднее время ожидания w в соответствии с формулой Литтла (закон сохранения стационарной очереди) должно быть равно q/λ и в данном случае ($\lambda = 1$) численно совпадать со средней длиной очереди. Теоретическая вероятность незанятости системы равна $1 - \rho = 0,1$. Следовательно, в среднем десятая часть заявок должна получать обслуживание без ожидания. Сопоставим с этими ожиданиями результаты моделирования (табл. 1).

В составлении приведенной программы было трудно ошибиться, к тому же она отличается от обсуждаемой у Т. Шрайбера [9] (с. 57 и далее) только типом исходных распределений. Формула Литтла и доля заявок, принимаемых к обслуживанию без ожидания, подтверждаются с достаточно высокой точностью. Это дает основания доверять интерпретатору *GPSS/W*. Показатели, определяемые только средними значениями моделируемых первичных величин (среднее время обслуживания и коэффициент загрузки), также вполне приемлемы.

Время ожидания зависит уже от двух моментов распределения времени обслуживания:

$$W = \frac{\lambda b_2}{2(1-\rho)}$$

Результаты, связанные с ожиданием, оставляют желать лучшего. Причина может быть только одна: недостаточно качественный генератор псевдослучайных чисел. Увеличение числа проведенных испытаний в общем приближает результаты к ожидаемым, но даже при 500 тыс. испытаний погрешность составляет около 17%. Судя по динамике версий, *GPSS/W* проходит процесс интенсивной отладки. Можно надеяться, что встроенные в систему датчики равномерно распределенных чисел будут усовершенствованы. В этом направлении можно приложить и собственные усилия.

В дополнение к приведенным результатам отметим, что для системы *M/M/1* распределение времени пребывания заявки в системе также подчинено показательному закону со средним $1/(1/0,9 - 1/1) = 9$. При ограничителе таймера 200 000 была

■ Таблица 1. Тестирование GPSS/W на модели M/M/1

Показатель	Теория	Число испытаний		
		50 000	200 000	500 000
Коэффициент загрузки	0,900	0,896	0,897	0,899
Среднее время обслуживания	0,900	0,899	0,902	0,901
Число входов	—	49 843	199 042	499 032
из них с нулевым ожиданием	—	5 340	21 001	50 691
Средняя длина очереди	8,100	5,132	5,757	6,690
Среднее время ожидания	8,100	5,148	5,785	6,703

заказана гистограмма распределения времени пребывания; получены среднее значение 9,264 и среднеквадратическое отклонение 9,147. Требуемое для показательного распределения их равенство приближенно выполняется, однако погрешность самих значений составляет около 2%.

Выше приведены контрольные результаты для системы *M/M/1*. Сопоставим с этими ожиданиями результаты моделирования при различных способах задания показательных распределений (встроенная Exponential, табличная ДФР и теоретически точное обращение $-\ln(1 - U)$) и числе испытаний (табл. 2). Результаты, связанные с ожиданием, оставляют желать лучшего даже в случае с теоретически точным переходом от равномерного к экспоненциальному распределению. Причина та же: недостаточно качественный генератор равномерных псевдослучайных чисел. Эти ошибки усугубляются погрешностями перехода к показательному распределению в Exponential и отчасти нейтрализуются (случайный эффект) при использовании табличного преобразования. Увеличение числа проведенных испытаний в общем приближает результаты к ожидаемым.

Обсудим, наконец, временные показатели сравниваемых версий ДСЧ. Для достижения отметки таймера 500 000 при табличной реализации логарифма и при работе с Exponential потребовалось практически одинаковое время — около 135 с (процессор Pentium 1/166). По-видимому, Exponential реализована таблично. При точном обращении показательного распределения было затрачено 187 с.

Прогон аналогичной программы на *Фортране* 77 при тех же исходных данных дал среднюю длину очереди 8,683. Это заметно ближе к теоретической оценке, чем лучший из приведенных в таблице результатов. Прогон занял всего 6,26 с. Столь

■ Таблица 2. Влияние способов генерации показательных распределенных чисел

Показатель	Exponential		Табличная ДФР		$-\ln(1 - U)$	
	50 000	500 000	50 000	500 000	50 000	500 000
Коэффициент загрузки	0,896	0,899	0,900	0,900	0,909	0,902
Среднее время обслуживания	0,899	0,901	0,899	0,901	0,903	0,896
Число входов	49 843	499 032	50 048	499 390	50 320	503 176
из них с нулевым ожиданием	5 340	50 691	5 022	49 823	4 764	50 498
Средняя длина очереди	5,132	6,690	7,770	7,911	7,163	6,831
Среднее время ожидания	5,148	6,703	7,763	7,920	7,118	6,788

большая разница является естественной платой за универсальность интерпретатора *GPSS*. Окупается ли она удобством программирования — зависит от опыта и квалификации пользователя, его желания и возможностей освоить *GPSS*, наличных вычислительных мощностей, сложности моделей и частоты потребностей в моделировании.

Перед моделированием реальной ситуации желательно убедиться, что выбранная процедура действительно воспроизводит случайную величину с требуемым распределением. Гистограмма распределения, порожденная функцией *Exponential* и построенная по 200 тыс. транзактов, оказалась близкой к ожидаемой. Среднее значение вместо заказанной единицы составило 1,005, а среднеквадратическое отклонение — 1,006.

Общая оценка *GPSS World*

GPSS/W является весьма ценным инструментом имитационного моделирования, свободным от ограничений аналитических и численных методов, достаточно «прозрачным», допускающим нестандартную обработку данных и снимающим с программиста множество нетривиальных проблем программирования и отладки моделей. Тем не менее, приходится отметить наличие у нее ряда серьезных недостатков.

1. Громоздкость системы и явная перегруженность встроенными возможностями (многообразие примитивов); непомерное разнообразие графических обозначений блоков, не поддерживаемое даже новейшими версиями «чертежной» системы *Visio*.

2. Медленная работа интерпретатора.

3. Отсутствие концептуального единства. Достаточно указать, например, различие в обращении к элементам матрицы при простой ссылке и изменении значения, круглые индексные скобки в основном тексте и квадратные — в *PLUS*-выражениях; обязательность приставки *MX\$* при ссылке на глобальную матрицу в тексте модели и столь же обязательное ее отсутствие внутри процедур; контекстно зависимый вид ссылок на параметры активного транзакта; потерявшее смысл в *GPSS/W* различие переменных *VARIABLE* и *FVARIABLE*; выражение коэффициента использования устройства в тысячных долях.

4. Путаная семантика операндов: в разных операторах одно и то же значение закреплено за разными операндами. Было бы гораздо легче работать с ключевыми операндами, имеющими значимые имена.

5. Неудачные обозначения операторов отношения *L*, *G*, *E* (было бы лучше согласовать с фортрановскими *LT*, *GE*, *EQ*); арифметическое *SQR* используется для квадратного корня (в *Паскале* так обозначается квадрат); в связи с «числовым» представлением логических значений и объединением понятий числового равенства и логической эквивалентности нарушено общепринятое старшинство операций; состояние логических ключей описывается как *SET* и *RESET* (буквальный перевод «уста-

новлен» и «установлен заново») вместо *ON*, *OFF*; операнд *RE* (традиционный смысл этой приставки — повторение действия) означает удаление.

6. Однократность прерываний устройств и недопустимость прерываний для памяти, которые могут быть использованы для моделирования многоканальных устройств обслуживания и, следовательно, окажутся подвержены прерываниям.

7. Дополнительные установки действий функциональных клавиш для каждой модели приходится задавать заново; было бы удобнее иметь прямой доступ к заданию и корректировке установок по умолчанию.

8. Отсутствуют средства подбора параметров теоретических вероятностных распределений по заданным моментам, а также формулы для моментов логарифмического распределения Лапласа, логарифмически логистического, распределений крайних значений и обратного гауссова. Нет непосредственных возможностей генерации коррелированных случайных величин.

9. Использование кириллических символов даже в комментариях исключает правильную работу Имитации.

10. Невозможно непосредственно определить вектор. Нижняя граница индексов матрицы по любому измерению равна 1, что может нарушить естественность индексации (например, при расчете вероятностей состояний системы). Крайне неудобен просмотр матриц через динамическое окно: одновременно можно видеть только два столбца. Ненулевая инициализация матрицы требует отдельного оператора *INITIAL* на каждый элемент.

11. Отсутствует возможность регулировать темп автоматического моделирования; он оказывается слишком быстрым, чтобы разглядеть нужные эффекты, а пошаговое продвижение транзактов — чересчур медленным.

12. Иконки свернутых динамических окон неразличимы, и при их восстановлении трудно выбрать желаемое.

13. Нельзя менять тип шкал графиков (только линейные) и их разметку, цвет и структуру линий, что может сделать последние неотличимыми друг от друга и/или от фона при черно-белом выводе.

14. Аргументом графиков может быть только время, так что распределение вероятностей состояний системы автоматически построить нельзя. К стати, для такого графика потребовалась бы логарифмическая шкала.

15. При работе с *ANOVA* нельзя менять расчетный уровень значимости.

16. Оставляет желать лучшего электронная документация к системе. При вызове *Help'a* по ключевому слову пользователю представляется «слепое» множество альтернатив, так что приходится выполнять их перебор. Часто встречаются пережитки предыдущих версий (*GPSS/PC*). Отсутствует упоминание о наличии в системе операции отрицания 'NOT'. Заявление о числовом значении «his value has a magnitude limited to 306 decimal digits» можно понять как длину мантиссы, хотя на самом деле речь идет о модуле десятичного порядка.

Ничего (кроме факта существования) не сказано о библиотеке. В разделе определения табличных функций понятие обратной функции заменено *кумулянтной*. В заголовке определения команды MATRIX неверно указано число операндов, не выделен строкой и буквой неиспользуемый первый операнд. Этот перечень можно было бы продолжить.

17. «Универсальность» интерпретатора GPSS оказывается ограниченной. Так, обслуживание с динамическим приоритетом (линейный рост приоритетов по времени ожидания с угловыми коэффициентами, зависящими от типа заявки), в GPSS даже с помощью мощного аппарата цепей пользователя обеспечить не удастся. Дело в том, что цепь может быть переупорядочена произвольно только в момент входа в нее новой заявки, тогда как нам для реализации динамического приоритета необходимо переупорядочение в момент выборки. На такой момент максимум возможного — это инверсия, т. е. реализация дисциплины LIFO (Last In — First Out). Предложенная в книге Шрайбера ([9], с. 391) схема «динамического приоритета» на самом деле реализует другую дисциплину выбора из очереди: по возрастанию времени обработки.

Трудности возникают также при моделировании «кругового» обслуживания. Для организации обслуживания всех заявок, накопившихся к началу обслуживания данной очереди, необходима переменная, запоминающая начало обслуживания текущей очереди. Ее значение наряду со значением указателя Pointer учитывается по схеме логического умножения при допуске к серверу очередной заявки. Не пропущенные заявки приходится записывать в цепь пользователя. При завершении обслуживания текущей очереди для каждого источника анализируется суммарное число заявок в очереди и в цепи. Однако после выбора нового положения стрелки блок UNLINK требует указания имени блока, в который адресуются освобождаемые заявки. Это имя в нашей модели можно получить с помощью функции Catenate присоединением к стандартному префиксу Inc номера выбранного источника. Но по синтаксису GPSS имя блока не может быть строкой, и моделирование завершается аварийно. Безуспешными оказались и попытки использовать числовые обозначения блоков с помощью команды EQU (было получено сообщение «You can assign your own value to entity label except for block labels»).

Трудности возникают и при моделировании замкнутой системы, в которой средний интервал между смежными заявками зависит от числа заявок в ней. Отсутствие в GPSS системного атрибута, представляющего это число, обходится введением соответствующей сохраняемой величины, на которую можно сослаться в блоке GENERATE. Проблема состоит в том, что к моменту изменения числа заявок в системе в цепь будущих событий уже занесен сформированный по прежней интенсивности момент прибытия очередной заявки. Доступа же к содержимому этой цепи программист не

имеет. Значит, точное моделирование замкнутых СМО в GPSS с помощью одного GENERATE невозможно. Остается громоздкий вариант с отдельными сегментами генерации заявок по максимальному числу активных источников и циркуляцией в каждом сегменте единственной заявки.

Кстати, заметим, что Дж. Шрайбер скептически оценивает полезность цепей пользователя и в особенности их индикаторов связи ([9], с. 385). В его книге объемом в 70 авторских листов приведены лишь два не очень убедительных примера работы с цепями пользователя, а в [7] — ни одного.

18. Оптимизацию систем обслуживания иногда приходится проводить по дискретным аргументам. Выбор решения часто осуществляется из дискретных рядов устройств с фиксированными характеристиками (быстродействием, емкостью памяти, набором алгоритмов функционирования и т. п.) и состоит в назначении целого числа устройств или расширяющих блоков. Это обстоятельство определяет необходимость использования при оптимизации методов дискретного программирования вплоть до элементарного перебора сначала на редкой решетке, затем — на более частой в перспективной окрестности оптимума.

С другой стороны, при определении тенденций оптимизации часто помогает рассмотрение компонент целевой функции. Например, в задачах управления запасами [4] минимизируется сумма затрат на хранение избыточного (среднего) запаса, организацию поставок и «штрафы» за неполное или несвоевременное обеспечение спроса. Большой удельный вес одной из упомянутых составляющих указывает на необходимость соответственно уменьшения уровня заказа, предельного уровня запаса или объема заказываемой партии; увеличения максимального запаса или объема заказываемой партии, увеличения порога заказа. Оптимизирующие эксперименты GPSS/W такой возможности не предусматривают. Ясно, что никакая «универсальная» система подобную специфику учесть не может.

В общем, работа на GPSS и других подобных языках безусловно оправдана при массовом создании моделей средней сложности. Эпизодическую разработку моделей умеренной сложности и создание проекта уникальной сложности программист с опытом работы на любом языке высокого уровня может вести на этом языке, причем знакомство с основными понятиями GPSS подскажет ему ряд полезных технических приемов. Наконец, всегда остается желательным расчет упрощенных вариантов исследуемых моделей с помощью численных методов теории очередей и реализующих их пакетов прикладных программ: он необходим хотя бы для получения эталонных значений, требуемых при отладке имитационных моделей.

Таким образом, практически необходимо использовать все три упомянутых подхода. Автор подготовил монографию по численным методам и учебник — по имитационным и надеется опубликовать их в ближайшем будущем.

Литература

1. **Рыжиков Ю. И.** Машинные методы расчета систем массового обслуживания: Учеб. пособие. — Л.: ВИКИ им. А. Ф. Можайского, 1979. — 177 с.
2. **Рыжиков Ю. И.** Эффективность и эксплуатация программного обеспечения ЭЦВМ: Учеб. пособие. — МО СССР, 1985. — 263 с.
3. **Рыжиков Ю. И.** Имитационное моделирование систем массового обслуживания: Учеб. пособие. — Л.: ВИКИ им. А. Ф. Можайского, 1991. — 111 с.
4. **Рыжиков Ю. И.** Теория очередей и управление запасами: Учеб. пособие. — СПб.: Питер, 2001. — 376 с.
5. **Киндлер Е.** Языки моделирования/Пер. с чешского. — М.: Энергоатомиздат, 1985. — 288 с.
6. **Советов Б. Я., Яковлев С. А.** Моделирование систем. — М.: Высшая школа, 1998. — 320 с.
7. **Советов Б. Я., Яковлев С. А.** Моделирование систем: практикум. — М.: Высшая школа, 1999. — 224 с.
8. **Харин Ю. С.** и др. Имитационное и статистическое моделирование: Учеб. пособие. — Минск: Белорусский гос. ун-т, 1992. — 176 с.
9. **Шрайбер Т. Дж.** Моделирование на GPSS/Пер. с англ. — М.: Машиностроение, 1980. — 592 с.
10. **Gordon G.** The Application of GPSS V to Discrete System Simulation. — Prentice-Hall, Englewood Cliffs, N. J. — 1975. — 389 p.
11. **Bobiller P. A., Kahan B. S., Probst A. R.** Simulation with GPSS and GPSS V. — Prentice-Hall, Englewood Cliffs, N. J., 1976. — 495 p.
12. **Янбых Г. Ф., Столяров Б. А.** Оптимизация информационно-вычислительных сетей. — М.: Радио и связь, 1987. — 232 с.
13. **Добров Г. М.** Наука о науке. — Киев: Наукова думка, 1966. — 270 с.

МИКРОКОНТРОЛЛЕРЫ

со склада в Санкт-Петербурге

Широкий спектр микроконтроллеров для различных приложений, отладочные средства, софт со склада в Санкт-Петербурге. Поставки микроконтроллеров на основе ядра 8051, RISC, ARM. DSP процессоры, микроконтроллеры со встроенными АЦП, ЦАП. Импортные электронные компоненты двойного назначения.

UBICOM, CYGNAL, FUJITSU, GOAL, ATMEL, MOTOROLA, AMD

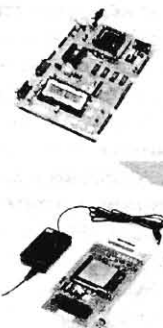


SX20AC, SX28AC, SX52BD, IP2022-160

ЭЛЕКТРОСНАБ

Официальный представитель
UBICOM, CYGNAL, GOAL в России

С-Петербург, наб. реки Фонтанки, д.38, тел./ф.: (812) 380 16 60
info@electrosnab.ru www.electrosnab.ru



УДК 577.4

МАТЕМАТИЧЕСКИЕ МЕТОДЫ ВНЕШНЕГО ПРОЕКТИРОВАНИЯ СЛОЖНЫХ СИСТЕМ

Ю. М. Смирнов,

д-р техн. наук, профессор

Институт интеллектуальных систем и технологий СПбГПУ

А. О. Поляков,

д-р техн. наук, профессор

Санкт-Петербургский институт информатики и автоматизации РАН

В. В. Однобоков,

старший преподаватель

Псковский политехнический институт СПбГПУ

Предлагается рассмотрение особенностей постановки и решения задач оптимального распределения ресурсов. Предложены и обоснованы алгоритмы многошагового улучшения составляющих интегрального показателя. Определены требования к характеру зависимостей интегрального показателя от частных и частных показателей от затрат на их улучшения, обеспечивающие сходимость и оптимальность магистрального решения. Дается схема оптимального распределения ресурсов.

The particularities of formulation and solving the problems of optimum resource distribution are marked out. The algorithms of multi-steps improvement of integral index's components are proposed and grounded. There are determined requirements to the nature of dependencies of integral index from particular factors and particular ones from costs for their improvement, ensuring convergence and optimum of a main solution. The sheme of optimum distribution is given.

Введение

В работах [1, 2] дан анализ состояния и направлений развития методологии системного проектирования. Одним из важных ее разделов является разработка моделей и методов оптимального распределения ресурсов при статистической оценке показателей функционирования проектируемых комплексов; параметрическом синтезе, т. е. выборе характеристик устройств проектируемого комплекса; формировании и реализации проектов создания новой техники или развития предприятий.

Постановка и методы решения задач этого класса рассмотрены в публикациях трудов СПбГПУ [2, 5, 6 и др]. Общими их чертами являются алгоритмическое задание и наличие приближенных аналитических выражений для критериальной функции F и функции ограничения Φ , дискретный характер изменения переменных, необратимость начальных или текущих затрат. Эти особенности затрудняют непосредственное использование хорошо известных методов поисковой оптимизации, но позволяют предложить и обосновать достаточно эффективные процедуры итеративного решения задач.

Задача статистической оценки показателя функционирования сводится к минимизации дисперсии относительной погрешности такой оценки при ограничении на проведение экспериментов. Если вероятность P сложного события оценивается по частотам элементарных событий, то

$$F = D\left(\frac{dP}{P}\right) \approx \sum_s \frac{y_s^2}{n_s} \rightarrow \min; \quad \Phi = \sum_s x_s^2 n_s \leq \Phi_{\text{зад}}, \quad (1)$$

где при $P = \prod_s p_s$ $y_s^2 = \frac{1-p_s}{p_s}$.

В формулах (1) n_s — число испытаний для оценки вероятности p_s элементарного события, $x_s = x_s^2$ — затраты на проведение одного эксперимента для оценки p_s . В работе [3] предложены два подхода к решению задачи в несколько этапов: 1) использование формального решения (зависящего от искомым величин) для квазипотенциального распределения ресурсов на каждом этапе по оценкам искомым величин, полученным на предыдущем; 2) замена исходной задачи последовательностью эквивалентных задач линейного программирования, имеющих на каждом этапе правильное решение.

Формальное решение (1) дается формулами $\alpha_s n_s = \alpha_s \Phi_{зад}$ и $D_{min} = \frac{(xy)^2}{\Phi_{зад}}$, где $a_s = \frac{x_s y_s}{(xy)}$. Поэтому при первом подходе на начальном этапе (когда нет оценок p_s) принимают $\alpha_s n_{s1} = \gamma_s \Phi_1$, где $\Phi_1 \leq \frac{\Phi_{зад}}{k}$, $k = \frac{y_{max}}{y_{min}}$, $\gamma_s = \frac{x_s}{\sum_i x_i}$.

Для последующих этапов квазиоптимальное распределение определяются с учетом оценок $y_s^{(l)}$, полученных на предыдущем этапе l :

$$\alpha_s n_{s/l+1} = a_s^{(l)} \Phi_{l+1}, \quad (2)$$

где $a_s^{(l)} = \frac{x_s y_s^{(l)}}{(x y^{(l)})}$, $\Phi_{l+1} = \Phi_l + \Phi^{(l+1)}$ — затраты для первых $(l+1)$ этапов.

Недостатками первого подхода являются неустойчивость оценок $y_s^{(l)}$ к результатам экспериментов и трудность учета требования о необратимости текущих затрат $a_s^{(l)} \Phi_{l+1} \geq a_s^{(l-1)} \Phi_l$. При некоторых допущениях можно показать, что для его выполнения должно быть

$$\Phi^{(l+1)} = g \Phi^{(l)} \text{ при } \frac{1-g^N}{1-g} = k, \quad (3)$$

где N — общее число этапов. Например, для $k = 3,0$ и $N = 4$ должно быть $g = 0,8105$ и $\Phi_1 = 0,333$, $\Phi_2 = 0,604$, $\Phi_3 = 0,823$, $\Phi_4 = 1,000$ (в долях от $\Phi_{зад}$). Формулы (2) и (3) не обеспечивают целочисленных значений n_{sl} , что является недостатком этого подхода.

При замене исходной задачи (1) на последовательность эквивалентных задач необходимо на каждом этапе решать задачу

$$D_l - D_{l+1} \approx \sum_s \lambda_s^{(l)} z_s^{(l)} \rightarrow \max; \quad (4)$$

$$\Phi_{l+1} - \Phi_l \approx \sum_s z_s^{(l)} \leq \Phi^{(l+1)},$$

где $z_s^{(l)} = \alpha_s n_s^{(l)}$ — возможные затраты на эксперименты по уточнению $p_s^{(l)}$; $\lambda_s^{(l)} = \frac{x_s^2 (y_s^{(l)})^2}{(z_s^{(l)})^2}$ — именной множитель Лагранжа для переменной n_s ; $(y_s^{(l)})^2 = \frac{n_{ls} - m_{ls}}{m_{ls} + 1}$ — малосмещенная оценка y_s^2 по числу m_{ls} появления элементарного события в n_{ls} опытах. Решением задачи (4) на конкретном $(l+1)$ -м этапе является $z_j^{(l)} \leq \Phi^{(l+1)}$, $z_j^{(l)} = 0$ ($j \neq l$), где $\lambda_j^{(l)} = \max \lambda_s^{(l)}$, т. е. использование ресурсов, выделенных для этапа, только на проведение экспериментов по уточнению вероятности того элементарного события, которому соответствует именной множитель Лагранжа (ИМЛ) с максимальным значением. Так как с ростом затрат ИМЛ будет уменьшаться, при достаточно большом $\Phi_{зад}$ значения всех ИМЛ $\lambda_s^{(l)} \rightarrow \lambda$ и $\alpha_s n_{sl} \rightarrow \alpha_s \Phi_l$.

Задача параметрического синтеза с использованием гипотезы об экспоненциальной зависимости стоимости устройств от дисперсии ошибок d_s и вероятности отказов q_s формируется как

$$F = \sum_s \gamma_s \left(1 + \frac{\alpha_s}{\gamma_s} \ln \frac{d_s^H}{d_s} \right) \left(1 + \frac{\beta_s}{\gamma_s} \ln \frac{q_s^H}{q_s} \right) \rightarrow \min; \quad (5)$$

$$\Phi(d, q) \leq \Phi_{зад},$$

где α, β, γ — параметры, определяемые методом наименьших квадратов по ряду устройств с одинаковыми функциями; $\Phi(d, q)$ — значение показателя функционирования (например, вероятность невыполнения системой тактической задачи) при фиксированных характеристиках устройств. Обычно $\Phi(d, q)$ задана алгоритмически, т. е. определяется на имитационной модели, но может приближенно описываться аналитической зависимостью, например, $\bar{\Phi} = 1 - (1 - e^{-y}) e^{-u} \leq 1 - P$, где P — вероятность выполнения тактической задачи, $u = \sum_s q_s$, $v = \sum_s d_s$, $y = \frac{p}{v}$.

В работе [3] предложены две стадии решения (5): 1) поиск начального приближения $x^0 = \{d_s^0, q_s^0\}$; 2) последовательное уточнение x^0 , т. е. итерационное построение последовательности $x^l \rightarrow x^{opt}$. На первой стадии, используя свойство робастности оптимального решения, находят начальное приближение x^0 как решение упрощенной задачи:

$$\bar{F} = T - \sum_s (\alpha_s \ln d_s + \beta_s \ln q_s) \rightarrow \min; \quad (5')$$

$$\bar{\Phi} = 1 - (1 - e^{-y}) e^{-u} \leq 1 - P.$$

Из уравнений Лагранжа получаем

$$d_s^0 = \frac{\alpha_s}{A} v, \quad q_s = \frac{\beta_s}{B} u, \quad (6)$$

где $A = \sum_s \alpha_s$, $B = \sum_s \beta_s$.

Из условия $\bar{\Phi} = 1 - P$ получаем

$$\ln \frac{z-1}{Pz} - \frac{B}{A} \times \frac{\ln z}{z-1} = 0,$$

где $z = e^y > \frac{1}{1-P}$.

Получив методом Ньютона решение z^* этого уравнения, можно найти для (6) значения $y = \ln z^*$ ($y > 1$ при $P > 0,63$), $v = \frac{P}{y}$, $u = \frac{B}{A} \times \frac{y}{z^* - 1}$. Например, значения $\frac{v}{u}$ для различных P и $\frac{B}{A}$ приведены в табл. 1.

■ Таблица 1.

$\frac{B}{A}$	P	0,6	0,7	0,8
0,5		0,42/0,23	0,39/0,18	0,34/0,12
1,0		0,32/0,34	0,30/0,25	0,275/0,17
2,0		0,235/0,43	0,231/0,31	0,220/0,20

На второй стадии используют гипотезу о предсказуемости F и Φ в окрестности начального приближения x^0 отрезками ряда Тейлора:

$$F - F_0 \approx (f^0 \epsilon) + \frac{1}{2} \epsilon Z_0 \epsilon;$$

$$\Phi - \Phi_0 \approx (\Phi^0 \epsilon),$$

где f, φ — градиенты функций, Z — матрица чувствительности, т. е. вторых производных F от компонент вектора x . При этом уточнение начального приближения осуществляется в несколько этапов, на каждом из которых $x^{l+1} = x^l + e^l$, ($l = 0, 1, 2, \dots$), где $f^l + Z_l e^l + \lambda_l \varphi^l = 0$, $(\varphi^l \epsilon^l) = \Phi_{\text{зад}} - \Phi_l$. При выбранном в начале $(l+1)$ -го этапа значении x^l можно вычислить компоненты векторов f, φ и матрицы Z , пользуясь приближенными аналогичными выражениями для \bar{F} и $\bar{\Phi}$, а значение Φ_l определять алгоритмически на модели.

Тогда

$$\epsilon^l = b^l - \lambda_l a^l, \text{ где } a = Z^{-1} \varphi, \text{ } b = Z^{-1}(-f). \quad (7)$$

Подстановка в ограничение дает

$$\lambda_l = \frac{B_l - (\Phi_{\text{зад}} - \Phi_l)}{A_l}, \text{ где } A = (\varphi a), \text{ } B = (\varphi b). \quad (8)$$

Формулы (7) и (8) являются основой итерационной процедуры для уточнения начального приближения x^0 с аналитическим вычислением поправок. При замене исходной задачи на последовательность эквивалентных необходимо на каждом этапе решать задачу

$$F_l - F_{l+1} \approx \sum_s \lambda_s^l z_s^l \rightarrow \max;$$

$$\Phi_{l+1} - \Phi_l \approx \sum_s z_s^l = g_l,$$

где $\lambda^s = \frac{-f_s}{\varphi_s} > 0$, $z_s = \varphi_s \epsilon_s$, $g_l = \frac{1}{N} (\Phi_{\text{зад}} - \Phi_l)$.

Решением такой задачи является

$$z_j^l = g_l, \text{ } z_i^l = 0 \text{ (} i \neq j \text{)}, \quad (9)$$

где при $g_l > 0$ $\lambda_j^l = \max_s \lambda_s^l$, а при $g_l < 0$ $\lambda_j^l = \min_s \lambda_s^l$.

Процедура обесценивает сближение ИМЛ, если

$$\frac{\partial \lambda^s}{\partial x_s} < 0 \text{ или } h_{ss} \equiv \frac{F_{ss}}{-F_s} + \frac{\Phi_{ss}}{\Phi_s} > 0.$$

Например, для задачи (5') при произвольном s имеем:

$$\lambda^d = \frac{\alpha v e^u e^y}{d y}, \text{ } \lambda^q = \frac{\beta e^u e^y}{q e^y - 1}$$

$$\text{и } h_{dd} = \frac{1}{d} + \frac{y-2}{v} = \left(\frac{1}{d} - \frac{1}{v} \right) + \frac{y-1}{v} > 0 \text{ при } y \geq 1,$$

так как $d < v$; $h_{dd} = \frac{1}{q} - 1 > 0$.

После реализации нескольких этапов стадии сближения ИМЛ можно определять дальнейшие поправки ϵ^l к ранее найденному приближению x^l из условий

$$\lambda_j^s(x^l + \epsilon^l) \equiv \lambda_{j+1}^s, \text{ (} s = 1, 2, \dots, n \text{)},$$

а именно

$$\lambda_j^s(x^l + \epsilon^l) \equiv \lambda_j^s + \sum_i \left(\frac{\partial \lambda^s}{\partial x_i} \right) \epsilon_i^l = \lambda_j^s \left(1 - \sum_i h_{si}^l \epsilon_i^l \right) = \lambda_{j+1}^s,$$

где

$$h_{si}^l = -\frac{1}{\lambda^s} \frac{\partial \lambda^s}{\partial x_i} = \frac{F_{si}}{-F_s} + \frac{\Phi_{si}}{\Phi_s}.$$

В векторно-матричной форме полученные соотношения можно представить как

$$H_l \epsilon^l = e - \lambda_{l+1} p^l, \text{ где } e_s = 1, \text{ } p_s = \frac{1}{\lambda^s}.$$

Отсюда

$$\epsilon^l = b^l - \lambda_{l+1} a^l, \text{ где } a = H^{-1} p, \text{ } b = H^{-1} e \quad (10)$$

и с учетом ограничения

$$\lambda_{l+1} = \frac{B_l - g_l}{A_l}, \text{ } A = (\varphi a), \text{ } B = (\varphi b). \quad (11)$$

Замечания.

1. Для упрощенной задачи (5') оба метода [линеаризации уравнений Куна—Таккера с определением поправок по формулам (7), (8) и линеаризации ИМЛ с определением поправок по формулам (10), (11)] дают одинаковые аналитические выражения ϵ_s^l .

2. При отсутствии аналитических выражений для \bar{F} и $\bar{\Phi}$ необходимо использовать в качестве оценки ИМЛ отношение возможных изменений F и Φ (со знаком «минус») при выборе значений «своей» переменной, соответствующих ближайшим техни-

ческим решениям $\lambda^s \equiv -\left(\frac{F^+ - F^-}{\Phi^+ - \Phi^-} \right)_s$, где $\epsilon_s = x_s^+ - x_s^-$.

Задачи распределения ресурсов при формировании и реализации проектов, их роль в методологии управления научно-производственной деятельностью и подходы к формализации отдельных процедур рассмотрены в [4]. С использованием гипотезы об экспоненциальном росте затрат на улучшение частных показателей, задачу распределения ресурсов можно представить так:

$$F = \frac{1}{2} \sum_s (1 - x_s y_s)^2 \rightarrow \min;$$

$$\Phi = \sum_s \left(\alpha_s \ln \frac{1 - x_s^0}{1 - x_s} + \beta_s \ln \frac{1 - y_s^0}{1 - y_s} \right) \leq \Phi_{\text{зад}}, \quad (12)$$

где x_s — относительная функциональная характеристика подразделения (например, отношение текущей производительности труда в подразделении к предельно достижимой), y_s — относительная

структурная характеристика подразделения (например, отношение текущей численности персонала или оборудования в подразделении к предельно достижимой). Тогда $x_s y_s$ — отношение мощности подразделения к предельно достижимой (или требуемой для реализации программы НПД), а задача заключается в минимизации суммарного отклонения мощностей подразделений от предельно достижимых или требуемых при ограничении затрат на их увеличение.

Для произвольного подразделения s можно определить пару ИМЛ:

$$\mu = \lambda^x = \frac{y(1-x)}{\alpha}(1-xy),$$

$$\nu = \lambda^y = \frac{x(1-y)}{\beta}(1-xy).$$

В [5] доказаны два положения о текущем расходовании необратимых ресурсов: 1) если $\mu > \nu$, то увеличение y^0 при равном эффекте приводит к увеличению затрат (аналогичный результат от увеличения x^0 при $\mu < \nu$); 2) если $\mu = \nu$, т. е.

$$y = \frac{\frac{x}{\beta}}{\frac{x}{\beta} - \frac{1-x}{\alpha}},$$

то отступление от этой кривой («магистральной») при равном эффекте ведет к увеличению затрат. Отсюда следует принцип оптимальности магистрального решения (МР) — расходование ограниченных ресурсов должно иметь две стадии:

1) при неравенстве ИМЛ необходимо увеличивать только ту переменную, для которой ИМЛ максимален;

2) при достижении равенства ИМЛ необходимо увеличивать одновременно все соответствующие переменные, не сходя с магистрали, т. е. сохраняя это равенство.

Итак, при построении модели распределения, т. е. при математической формулировке задачи, используется экспериментально подтвержденная гипотеза об экспоненциальной связи затрат со значениями частных показателей. При выборе и обосновании итерационного метода решения задачи распределения используется свойство робастности оптимального решения для обоснования линеаризации условий оптимальности или замены исходной задачи последовательностью эквивалентных задач линейного программирования. В последнем подходе итерационные процедуры поиска и уточнения начального приближения к решению основаны на понятиях ИМЛ и магистрального их улучшения за счет увеличения соответствующих переменных. Ниже рассматриваются свойства критериальной функции и функции ограничения, при которых магистральный способ улучшения ИМЛ обеспечивает оптимальное распределение ресурсов, и дается подход к оценке эффекта от реализации магистрального решения применительно к трем типам задач распределения указанного выше класса.

Недопустимость увеличения переменных с меньшими значениями ИМЛ означает, что при одинаковом эффекте, т. е. при одной значении F в этом случае затраты Φ должны быть больше. Обозначим $x = (x_1, x_2, \dots, x_j)$, $y = (y_1, y_2, \dots, y_n)$. Пусть $\lambda'_j = \lambda_j(x', y') = \lambda$ при $(i = 1, 2, \dots, l)$, $\lambda'_j = \lambda_j(x', y') < \lambda$ при $(j = 1, 2, \dots, n)$, где $x' = x + \varepsilon$, $y' = y + \delta$ при $\delta_j > 0$.

В приращениях (с точностью до величин первого порядка малости) в окрестностях точки (x', y')

$$F - F' \cong -\sum_i f_i \varepsilon_i - \sum_j f_j \delta_j = \lambda \sum_i z_i + \sum_j \lambda'_j z_j$$

и

$$\Phi' - \Phi \cong \sum_i \varphi_i \varepsilon_i + \sum_j \varphi_j \delta_j = \sum_i z_i + \sum_j z_j,$$

где $z_i = \varphi_i \varepsilon_i$, $z_j = \varphi_j \delta_j$ при одинаковом эффекте

$$F - F' = 0 \text{ и } \sum_i z_i = -\frac{1}{\lambda} \sum_j \lambda'_j z_j, \text{ а } \Phi' - \Phi = \sum_j z_j \left(1 - \frac{\lambda'_j}{\lambda}\right) > 0$$

при $z_j = \varphi'_j \delta_j > 0$ и $0 < \lambda'_j < \lambda$. Эти условия выполняются, если первые производные у критериальной функции отрицательные, а у функции ограничения — положительные:

$$f_s = \frac{\partial F}{\partial x_s} < 0, \quad \varphi_s = \frac{\partial \Phi}{\partial x_s} > 0. \quad (14)$$

Недопустимость отступления от магистрали означает, что нарушение равновесия $\lambda_j = \lambda$ при одинаковом эффекте ведет к увеличению затрат. Заметим, что соблюдение этого равенства, как следует из предыдущего, — необходимое условие оптимальности для задачи распределения (уравнения Лагранжа). Следует найти, какими должны быть свойства F и Φ , чтобы для $x' = x + \varepsilon$ при $\lambda_s(x) = \lambda$ и $dF = F(x + \varepsilon) - F(x) = 0$ было $d\Phi = \Phi(x + \varepsilon) - \Phi(x) > 0$.

С учетом величин второго порядка малости

$$dF \cong (f\varepsilon) + \frac{1}{2}(\varepsilon, Z\varepsilon) = 0 \text{ или } \lambda(\varphi\varepsilon) \cong \frac{1}{2}(\varepsilon, Z\varepsilon) \quad d\Phi \cong (\varphi\varepsilon) + \frac{1}{2}(\varepsilon, Z\varepsilon) = \frac{1}{2}(\varepsilon, Z\varepsilon) + \frac{1}{2}(\varepsilon, Z\varepsilon),$$

где Z и W — матрицы вторых производных F и Φ , соответственно.

МР будет оптимальным, если

$$(\varepsilon, H'\varepsilon) > 0, \text{ где } H' = \frac{1}{\lambda}Z + W. \quad (15)$$

Достаточным условием выполнения (15) при произвольном ε является требование, чтобы матрица H' была положительно определенной, т. е. ее основные (диагональные) определители должны быть больше нуля. Необходимые условия должны учитывать ограничения на выбор ε с учетом $dF = 0$.

Для произвольных i и j при $\lambda^i = \lambda^j$ имеем

$$h'_{ij} = \frac{z_{ij}}{\lambda} + w_{ij} = \varphi_i h_{ij} = \varphi_j h_{ij}, \quad (16)$$

$$\lambda^i_j = \frac{\partial}{\partial x_j} \left(-\frac{f_j}{\varphi_i} \right) = -\lambda^i h_{ij},$$

где $h_{ij} = \frac{z_{ij}}{-f_j} + \frac{w_{ij}}{\varphi_i}$.

Таблица 2.

t	R	x ₁	x ₂
0,780776	0,000	0,3904	0,3904
0,7808	0,009	0,395	0,386
0,785	0,122	0,453	0,331
0,79	0,18	0,485	0,305
0,80	0,26	0,530	0,270
0,83	0,43	0,630	0,200
0,86	0,55	0,705	0,155
0,90	0,69	0,795	0,105
0,95	0,85	0,900	0,050
1,00	1,00	1,000	0,000

Например:
 в задаче оценки показателя (1) $h_{ss} = \frac{2}{n_s}$, $h_{ij} = 0$
 при $i \neq j$, т. е. матрица H и H' — диагональные;
 в задаче распределения требований (5) для
 произвольной пары d_s, q_s , характеризующих одно
 устройство, $h_{dd}^{(s)} = \frac{1}{d_s} + \frac{y-2}{v}$, $h_{qq}^{(s)} = \frac{1}{q_s} - 1$ и незави-
 симо от s $h_{dq} = -1$, $h_{qd} = -\frac{y}{v(e^y - 1)}$;

в задаче распределения ресурсов (12) для про-
 извольной пары x_s, y_s , характеризующей одно под-
 разделение предприятия или направление проекта:

$$\begin{aligned} h_{xx}^{(s)} &= \frac{1+y-2xy}{(1-x)(1-xy)}, & h_{yy}^{(s)} &= \frac{1+x-2xy}{(1-y)(1-xy)}, \\ h_{xy}^{(s)} &= \frac{2xy-1}{y(1-xy)}, & h_{yx}^{(s)} &= \frac{2xy-1}{x(1-xy)}. \end{aligned} \quad (16)$$

Для переменных из разных пар $h_{ij} = 0$ ($i \neq j$), т. е.
 H' — квазидиагональная матрица. В общем случае

$$(H'\epsilon)_i = \varphi_i U_i,$$

где $U_i = \sum_j \frac{h_{ij}}{\varphi_j} z_j$ и $(\epsilon, H'\epsilon) = \sum_i z_i U_i$.

Если H' квазидиагональная, то квадратичная
 форма $d\Phi \approx \frac{1}{2}(\epsilon, H'\epsilon)$ является полусуммой блоков
 для каждой пары x_s, y_s :

$$(\epsilon, H'\epsilon)_s = h'_{xx} \epsilon_x^2 + 2h'_{xy} \epsilon_x \epsilon_y + h'_{yy} \epsilon_y^2.$$

При произвольном выборе ϵ_x, ϵ_y достаточными
 условиями $d\Phi > 0$ являются

$$h'_{xx}, h'_{yy} > 0 \text{ и } h'_{xx}h'_{yy} - (h'_{xy})^2 > 0,$$

что в силу (16) эквивалентно требованиям

$$h_{xx}, h_{yy} > 0 \text{ и } h_{xx}h_{yy} - h_{yx}h_{xy} > 0 \quad (17)$$

для произвольных переменных;

при специальном выборе для всех пар x_s, y_s ,
 $\epsilon_y = -\frac{f_x}{f_y} \epsilon_x$ или $z_y = -z_x$ в силу $\lambda^x = \lambda^y$ (что обеспе-
 чивает с точностью до величин первого порядка

малости выполнение равенства $dF = 0$), и исходя
 из (16)

$$(\epsilon, H'\epsilon)_s = \frac{z_x^2}{\varphi_x \varphi_y} \{ \varphi_y (h_{xx} - h_{yx}) + \varphi_x (h_{yy} - h_{xy}) \}$$

и необходимыми условиями $d\Phi > 0$ являются

$$h_{xx}, h_{yy} > 0, \quad h_{xx} - h_{yx} > 0, \quad h_{yy} - h_{xy} > 0, \quad (18)$$

для произвольных переменных.

Для задачи распределения ресурсов (12) не-
 обходимые условия (18) имеют вид $h_{xx} > \frac{1}{1-x}$,
 $h_{yy} > \frac{1}{1-y}$, $h_{xx} - h_{yx} = \frac{1}{x(1-x)}$, $h_{yy} - h_{xy} = \frac{1}{y(1-y)}$ и
 выполняются во всей области допустимых значе-
 ний $0 < x, y < 1$.

Достаточные условия (17) эквивалентны требо-
 ванию

$$\frac{(1+y-2xy)(1+x-2xy)}{(1-x)(1-y)} - \frac{(2xy-1)^2}{xy} > 0,$$

которое заведомо выполняется при $t = x + y > 0$.
 Точнее, должно быть $(t-1) + (3-2t)xy \geq 0$ и при не-
 которых $0 < t < 1$ в этом диапазоне

$$xy = x(t-x) \geq \frac{1-t}{3-2t} \quad (19)$$

для $x_2 < x < x_1$ и $y = t - x$, где $x_{1,2} = \frac{t \pm R}{2}$,

$$R^2 = t^2 - 4 \frac{1-t}{3-2t} \geq 0.$$

Так как $R^2 = \frac{(2-t)(2t^2+t-2)}{3-2t}$, то в указанном ди-
 апазоне $R = 0$ при $t' = \frac{\sqrt{17}-1}{4} \approx 0,781$ и $R > 0$ при

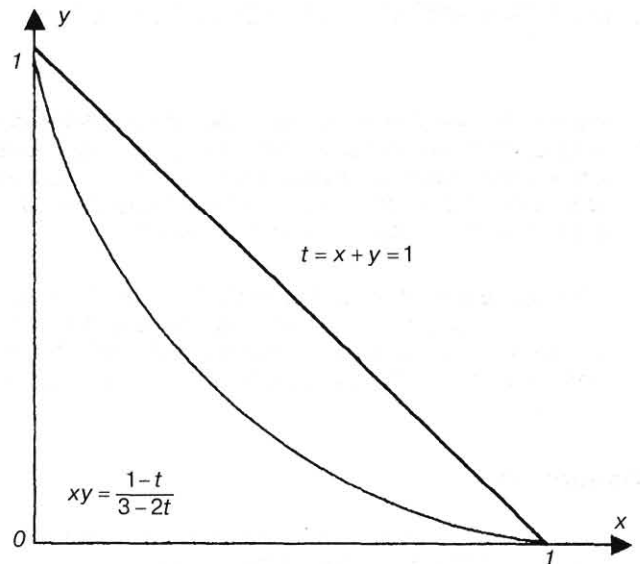


Рис. 1. Подобласть допустимых значений с выполне-
 нием условия оптимальности МР

$t > t'$. В табл. 2 приведены координаты точек кривой (19) на рисунке, которая выделяет подобласть допустимых значений x , y , где выполняется доста-

точное условие оптимальности МР (в силу симметрии выражений относительно x и y : $y = x_2$ при $x = x_1$ и $y = x_1$ при $x = x_2$).

Литература

1. **Смирнов Ю. М.** Состояние и перспективы развития методов системного моделирования // Методы кибернетики и информационные технологии: Сб. научных трудов. — Вып. 1., — Саратов: Изд-во Саратовского университета, 1994. — С. 34–40.
2. **Смирнов Ю. М.** Направления развития методологии системного проектирования // Вычислительная техника, автоматика и радиоэлектроника. Труды СПбГТУ. — СПб.: Изд-во СПбГТУ, — 1998. № 472. — С. 109–123.
3. **Смирнов Ю. М.** Системное проектирование комплексов управления летательными аппаратами: Уч. пособие. — СПб.: Изд-во СПбГТУ, 1996. — 120 с.
4. **Поляков А. О., Смирнов Ю. М., Турчак А. А.** Информационно-динамические основы организации управления крупными предприятиями и холдингами. — СПб.: Изд-во СПбГПУ, 2002. — 192 с.
5. **Смирнов Ю. М., Швырков В. Г.** Математические методы внешнего проектирования систем // Вычислительная техника, автоматика и радиоэлектроника. Труды СПбГТУ. — СПб.: Изд-во СПбГТУ, 2002. — № 488 — С. 32–58.
6. **Смирнов Ю. М.** Планирование испытаний в условиях неопределенности // Управление в условиях неопределенности: Сб. Под ред. А. Е. Городецкого. — СПб.: Изд-во СПбГТУ, 2002. — С. 339–360.

УДК 621-52:004.52; 629.78; 681.3

МНОГОСЛОЙНАЯ ПЕРСЕПТРОННАЯ НЕЙРОННАЯ СЕТЬ В ЗАДАЧЕ МОДЕЛИРОВАНИЯ РЕЧЕВЫХ СИГНАЛОВ

Ю. Я. Изилов,

канд. техн. наук

Санкт-Петербургский государственный политехнический университет

Рассматриваются возможности применения искусственной нейронной сети перцептронного типа в задаче моделирования речевых сигналов. Приводятся структуры и алгоритм, позволяющий осуществить функционирование многослойной перцептронной нейросети на персональном компьютере. Обсуждаются проблемы, возникающие при практической реализации, а также достоинства и недостатки использования данного подхода.

This article considers the opportunities of application artificial neural network based on perceptron type for speech signals modeling task. Structures and the algorithm, which allows to carry out functioning of multilayer perceptron neural network on personal computer, are resulted. The problems arising at practical realization, and also limits and advantages of the given approach usage are discussed.

Введение

Речь представляет собой важнейшее и самое удобное средство взаимодействия между людьми. В этой связи, в условиях компьютеризированного общества, понятно стремление специалистов ос-

настить автоматизированные системы различного назначения средствами речевого ввода—вывода информации.

Речевой сигнал (РС) представляет собой многоуровневую структуру и характеризуется определенной иерархией: фонемы, слова, фразы и т. д.

Основу такой иерархии составляют акустические сигналы, моделирование которых вызывает наибольшую трудность.

С точки зрения акустики, РС состоит из быстрых и очень неустойчивых колебаний воздушного давления [3], что, в свою очередь, порождает значительную неустойчивость параметрического описания.

Неустойчивость РС учитывают путем слежения за его вероятностными характеристиками (такими, как функция распределения вероятностей и/или моментные функции), для оценки которых необходимо задать модель сигнала. На практике апробированные сведения о вероятностном характере РС не являются полными для того, чтобы можно было явно задать его вероятностную модель. Поэтому поведение РС приходится изучать по опытным данным. С учетом сказанного, в качестве альтернативной модели может быть выбрана искусственная нейронная сеть.

Нейронные сети (НС) являются важным дополнением традиционных методов обработки данных [1, 2, 4, 6, 13]. Наибольшее распространение среди них получили так называемые многослойные персептронные НС, или многослойные персептроны. Они могут использоваться при решении различного рода задач, таких, как прогнозирование, сжатие данных, аппроксимация функций, распознавание и восстановление образов.

Достоинства многослойных персептронов (МП) — возможность построения сложных разделяющих поверхностей, возможность осуществления любого отображения входных векторов в выходные, способность к обобщению входных данных. При этом не требуется знания закона распределения входных векторов. Вместе с тем, МП сводит задачу моделирования РС к параметрической идентификации входных векторов. Рассмотрим модель МП более подробно.

Многослойный персептрон представляет собой модель НС с прямыми связями [6, 14, 17, 18]. Она включает в себя входной слой, выходной слой и

находящиеся между ними промежуточные, так называемые скрытые слои нейронов (рис. 1).

Входной слой нейронов не производит никаких вычислений. Он осуществляет распределение данных и работает в качестве разветвителя. Скрытые слои, минимальное число которых должно быть не менее одного, предназначены непосредственно для обработки входной информации и внутреннего представления данных. Выходной слой предназначен для обработки информации, поступающей от предыдущего слоя, и выдачи соответствующих результатов.

Топология МП предусматривает соединение выхода каждого нейрона предыдущего слоя с входом каждого нейрона последующего слоя.

Особенностью МП является возможность модификации формы областей решений посредством изменения количества слоев нейронов. При этом двухслойная НС (с одним слоем нейронных связей) формирует области решений в виде гиперплоскостей; трехслойная НС (с двумя слоями нейронных связей) позволяет формировать любые выпуклые области в пространстве решений; четырехслойная НС (с тремя слоями нейронных связей) позволяет получать область решений произвольной формы и сложности [2, 13].

МП предполагает выполнение процесса обучения. Наиболее распространенный метод обучения — это алгоритм обратного распространения ошибки [15, 20]. Его можно применить к МП с любым числом слоев нейронных связей. Однако для того, чтобы понять идею алгоритма, достаточно рассмотреть его работу на примере МП с двумя слоями нейронных связей (рис. 2).

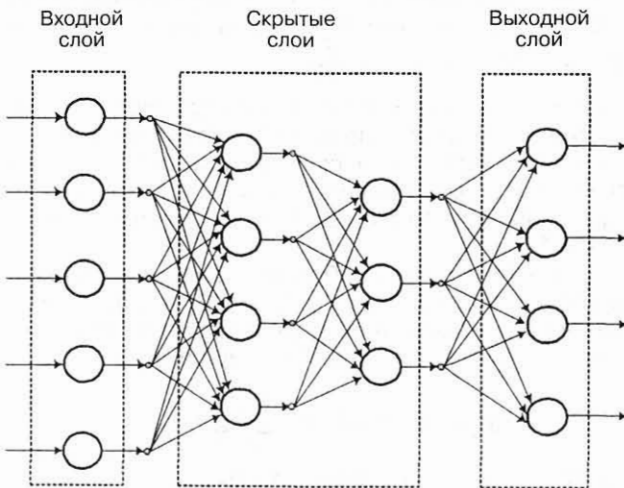
Алгоритм обратного распространения ошибки производит настройку весовых коэффициентов в многослойной НС на основе использования метода градиентного спуска. Он предполагает, что обучение происходит с учителем, и требует предъявления обучающей выборки.

Обучающая выборка представляет собой множество пар входных и выходных векторов $\{(X_m, G_m) | m = 1, 2, \dots, M\}$, между которыми нужно установить соответствие.

Множество нейронных связей в данном случае (см. рис. 2) образуется двумя матрицами весовых коэффициентов: $W^{(1)}, W^{(2)}$ с элементами $w_{ik}^{(1)}$ и $w_{kj}^{(2)}$ соответственно. Порог возбуждения k -го нейрона в скрытом слое формируется связью с весовым коэффициентом $w_{0k}^{(1)}$, по которой поступает сигнал $u_0 = 1$. Порог возбуждения j -го нейрона в выходном слое формируется связью с весовым коэффициентом $w_{0j}^{(2)}$, по которой поступает сигнал $h_0 = 1$.

Эффективность функционирования НС для некоторого конечного набора обучающих пар можно определить с помощью общей ошибки сравнения фактических и требуемых выходных векторов для каждого случая. Для этого вначале вычислим ошибку сравнения каждой пары векторов типа «вход—выход»:

$$E_m = \frac{1}{2} \sum_{j=1}^J (g_{j,m} - y_{j,m})^2, \quad (1)$$



■ Рис. 1. Многослойный персептрон

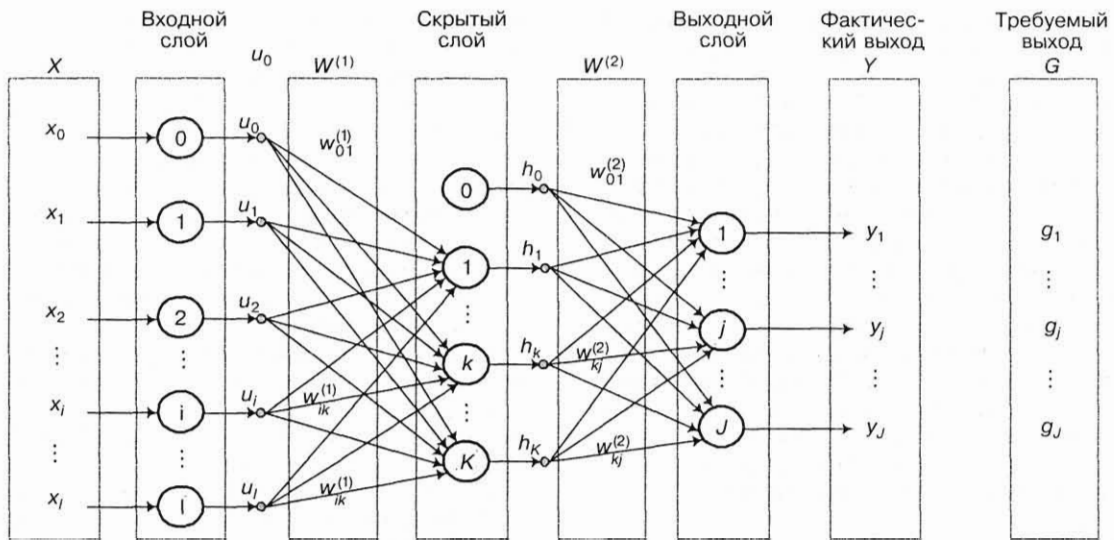


Рис. 2. Многослойный перцептрон с двумя слоями нейронных связей

где $g_{j,m}$ — требуемое значение выходного сигнала j -го выходного нейрона для m -й обучающей пары; $y_{j,m}$ — фактическое значение выходного сигнала j -го выходного нейрона для m -й обучающей пары.

Теперь общую ошибку функционирования НС, с учетом (1), можно вычислить по формуле

$$E = \sum_{m=1}^M E_m = \frac{1}{2} \sum_{m=1}^M \sum_{j=1}^J (g_{j,m} - y_{j,m})^2. \quad (2)$$

Целью обучения НС является такая настройка ее весовых коэффициентов, при которой для каждого входного вектора найдется выходной вектор, сгенерированный НС и представляющий собой требуемый выходной вектор или достаточно близкий к нему.

Другими словами, процесс обучения НС сводится к нахождению таких матриц весовых коэффициентов $W^{(1)}, W^{(2)}$, которые минимизируют общую ошибку функционирования нейросети E . Для решения данной задачи алгоритм обратного распространения ошибки использует метод градиентного спуска. При этом поиск минимального значения производится по всему пространству весовых коэффициентов НС.

Метод градиентного или наискорейшего спуска предполагает, что коррекция весовых коэффициентов происходит итерационным способом: малым перемещением и в том направлении, при котором ошибка функционирования НС минимизируется наискорейшим образом. Обычно используется следующая форма:

$$W(t+1) = W(t) + \Delta W(t), \quad (3)$$

где t — шаг итерации; $\Delta W(t)$ — коррекция весовых коэффициентов; $W(t)$ — значения весовых коэффициентов до коррекции; $W(t+1)$ — значения весовых коэффициентов после коррекции.

Согласно этому методу, частная производная меры ошибки относительно каждого веса пропорциональна изменению соответствующего веса с отрицательным коэффициентом пропорциональности:

$$\Delta W = -\eta \frac{\partial E}{\partial W}, \quad (4)$$

где ΔW — коррекция весового коэффициента; η — коэффициент скорости обучения.

Коэффициент η в формуле (4) обозначает небольшое положительно определенное число, которое обычно находится в пределах от 0,01 до 1. Он влияет на шаг итерации и, соответственно, позволяет управлять скоростью сходимости алгоритма обучения.

Результаты вычислений согласно выражениям (3), (4) будут соответствовать линии, проходящей по поверхности ошибок в пространстве весовых коэффициентов, которая ведет к наискорейшему спуску. Следовательно, задача коррекции весовых коэффициентов эквивалентна задаче вычисления частных производных $\frac{\partial E}{\partial W}$.

Задача алгоритма обучения сводится к минимизации функции ошибки НС относительно всего множества нейронных связей. При этом функцию ошибки НС, как правило, задают формулой (2), а уровень допустимых ошибок ограничивают величиной ξ .

В соответствии с обозначениями сигналов и весовых коэффициентов (см. рис. 2) выходное значение j -го нейрона в выходном слое для m -й обучающей пары определяется по формуле

$$y_{j,m} = F_j(S_{j,m}) = F_j \left(\sum_{k=0}^K w_{kj}^{(2)} h_{k,m} \right), \quad (5)$$

где $S_{j,m}$ — значение уровня возбуждения j -го нейрона выходного слоя для m -й обучающей пары;

$F_j(\cdot)$ — функция активации j -го нейрона выходного слоя; $w_{kj}^{(2)}$ — весовой коэффициент для связи, идущей от k -го нейрона скрытого слоя к j -му нейрону выходного слоя; $h_{k,m}$ — выходное значение сигнала k -го нейрона скрытого слоя для m -й обучающей пары, $h_{0,m} = 1$ соответствует сигналу, формирующему пороговое значение.

Выходное значение k -го нейрона скрытого слоя для m -й обучающей пары формируется следующим образом:

$$h_{k,m} = F_k(S_{k,m}) = F_k\left(\sum_{i=0}^l w_{ik}^{(1)} u_{i,m}\right) = F_k\left(\sum_{i=0}^l w_{ik}^{(1)} x_{i,m}\right), \quad (6)$$

где $S_{k,m}$ — значение уровня возбуждения k -го нейрона скрытого слоя для m -й обучающей пары; $F_k(\cdot)$ — функция активации k -го нейрона скрытого слоя; $w_{ik}^{(1)}$ — весовой коэффициент для связи, идущей от i -го нейрона входного слоя к k -му нейрону скрытого слоя; $u_{i,m}$ — значение сигнала, поступающего от i -го нейрона входного слоя для m -й обучающей пары, $u_{i,m} = x_{i,m}$; $x_0 = 1$ соответствует сигналу, формирующему пороговое значение.

В случаях, когда используется нелинейная функция активации (ФА) и для минимизации общей ошибки функционирования НС применяется метод градиентного спуска, для реализации процесса обучения существует важное требование. Оно предъявляется к самой функции активации. Необходимо, чтобы она была неубывающей и имела ограниченную производную. На практике это требование часто удовлетворяют посредством применения нелинейной ФА сигмоидного вида:

$$F(S) = \frac{1}{1 + e^{-S}}. \quad (7)$$

Производная такой функции имеет следующий вид:

$$F'(S) = F(S)(1 - F(S)). \quad (8)$$

Несложно определить, что производная максимальна при $F(S) = 0,5$ и равна нулю при $F(S) = 0$ или $F(S) = 1$. Дополнительное преимущество функции (7) состоит в автоматическом контроле степени усиления. Это значит, что для слабых сигналов (близких к нулю) кривая «вход—выход» имеет сильный наклон, который обеспечивает большое усиление. С увеличением сигнала степень усиления падает. Следовательно, большие сигналы воспринимаются сетью без насыщения, а слабые сигналы проходят по сети без особого ослабления.

Процесс обучения включает в себя два этапа вычислений: прямого и обратного распространения сигнала по нейросети.

На первом этапе происходит определение ошибки НС при распространении входного сигнала по сети в прямом направлении, т. е. от входа к выходу. Суть этого этапа заключается в следующем. Для заданного входного вектора вычисляют фактический выход НС с помощью выражений (5) и (6). Затем производят расчет ошибки между фактическим и требуемым выходом НС по формуле (2).

На втором этапе выполняют минимизацию полученной ошибки, которая сводится к последовательной настройке всех весовых коэффициентов НС. С этой целью вычисленный сигнал ошибки распространяется по сети в обратном направлении, т. е. от выхода к входу, и используется для подстройки весовых коэффициентов. Для определения величины, на которую необходимо скорректировать весовые коэффициенты, применяется метод градиентного спуска.

Вычисления повторяют до тех пор, пока общая ошибка функционирования НС не достигнет заданной величины ξ , устанавливающей точность процесса обучения.

Рассмотрим второй этап вычислений более подробно. Настройка весовых коэффициентов $w_{kj}^{(2)}$ между k -м нейроном скрытого слоя и j -м нейроном выходного слоя производится на основе использования выражений (3) и (4). При этом

$$\Delta w_{kj}^{(2)} = -\eta \frac{\partial E_m}{\partial w_{kj}^{(2)}}, \quad (9)$$

где $w_{kj}^{(2)}$ — коррекция весового коэффициента для связи, идущей от k -го нейрона скрытого слоя к j -му нейрону выходного слоя для m -й обучающей пары; η — коэффициент скорости обучения.

Используя правила частных производных, представим производную ошибки НС относительно выходных нейронов, принимая во внимание выражение (1):

$$\begin{aligned} \frac{\partial E_m}{\partial w_{kj}^{(2)}} &= \frac{1}{2} \sum_j \frac{\partial (g_{j,m} - y_{j,m})^2}{\partial w_{kj}^{(2)}} = \\ &= \frac{1}{2} \frac{\partial (g_{j,m} - y_{j,m})^2}{\partial w_{kj}^{(2)}} = -(g_{j,m} - y_{j,m}) \frac{\partial y_{j,m}}{\partial w_{kj}^{(2)}}. \end{aligned} \quad (10)$$

Принимая во внимание выражение (5), определим частные производные $\frac{\partial y_{j,m}}{\partial w_{kj}^{(2)}}$:

$$\begin{aligned} \frac{\partial y_{j,m}}{\partial w_{kj}^{(2)}} &= \frac{\partial F_j(S_{j,m})}{\partial S_{j,m}} \frac{\partial S_{j,m}}{\partial w_{kj}^{(2)}} = \\ &= F'_j(S_{j,m}) \frac{\partial}{\partial w_{kj}^{(2)}} \sum_k w_{kj}^{(2)} h_{k,m} = F'_j(S_{j,m}) h_{k,m}, \end{aligned} \quad (11)$$

где F'_j — производная ФА j -го нейрона выходного слоя.

С учетом (10) и (11), производную ошибки относительно выходных нейронов можно представить как

$$\frac{\partial E_m}{\partial w_{kj}^{(2)}} = -(g_{j,m} - y_{j,m}) F'_j(S_{j,m}) h_{k,m}. \quad (12)$$

Введем обозначение $\delta_{j,m}^{(2)}$:

$$\delta_{j,m}^{(2)} = (g_{j,m} - y_{j,m}) F'_j(S_{j,m}). \quad (13)$$

Переменная $\delta_{j,m}^{(2)}$ в выражении (13) представляет собой ошибку j -го нейрона выходного слоя для m -й обучающей пары.

Теперь формулу (9) можно записать в сжатом виде:

$$\Delta w_{kj,m}^{(2)} = \eta \delta_{j,m}^{(2)} h_{k,m}. \quad (14)$$

Настройка весовых коэффициентов $w_{ik}^{(1)}$ между i -м нейроном входного слоя и k -м нейроном скрытого слоя производится также на основе использования выражений (3) и (4):

$$\Delta w_{ik,m}^{(1)} = -\eta \frac{\partial E_m}{\partial w_{ik}^{(1)}}, \quad (15)$$

где $\Delta w_{ik,m}^{(1)}$ — коррекция весового коэффициента для связи, идущей от i -го нейрона входного слоя к k -му нейрону скрытого слоя для m -ой обучающей пары; η — коэффициент скорости обучения.

Используя правила частных производных, представим производную ошибки НС относительно скрытых нейронов, учитывая выражение (1):

$$\frac{\partial E_m}{\partial w_{ik}^{(1)}} = \frac{1}{2} \sum_j \frac{\partial (g_{j,m} - y_{j,m})^2}{\partial w_{ik}^{(1)}} = \sum_j (g_{j,m} - y_{j,m}) \frac{\partial y_{j,m}}{\partial w_{ik}^{(1)}}. \quad (16)$$

Определим в формуле (16) частные производные $\frac{\partial y_{j,m}}{\partial w_{ik}^{(1)}}$:

$$\begin{aligned} \frac{\partial y_{j,m}}{\partial w_{ik}^{(1)}} &= \frac{\partial F_j(S_{j,m})}{\partial S_{j,m}} \frac{\partial S_{j,m}}{\partial w_{ik}^{(1)}} = \\ &= F'_j(S_{j,m}) \frac{\partial}{\partial w_{ik}^{(1)}} \sum_k w_{kj}^{(1)} u_{k,m} = F'_j(S_{j,m}) w_{kj}^{(2)} \frac{\partial h_{k,m}}{\partial w_{ik}^{(1)}}. \end{aligned} \quad (17)$$

Вычислим в формуле (17) оставшиеся частные производные $\frac{\partial h_{k,m}}{\partial w_{ik}^{(1)}}$:

$$\begin{aligned} \frac{\partial h_{k,m}}{\partial w_{ik}^{(1)}} &= \frac{\partial F_k(S_{k,m})}{\partial S_{k,m}} \frac{\partial S_{k,m}}{\partial w_{ik}^{(1)}} = \\ &= F'_k(S_{k,m}) \frac{\partial}{\partial w_{ik}^{(1)}} \sum_i w_{ik}^{(1)} x_{i,m} = F'_k(S_{k,m}) x_{i,m}, \end{aligned} \quad (18)$$

где $F'_k(\)$ — производная ФА k -го нейрона скрытого слоя.

Подставим результаты преобразований, полученные в выражениях (16), (17), (18), в формулу (15):

$$\Delta w_{ik,m}^{(1)} = \eta \sum_j (g_{j,m} - y_{j,m}) = F'_j(S_{j,m}) w_{kj}^{(2)} F'_k(S_{k,m}) x_{i,m}. \quad (19)$$

Упростим формулу (19) с учетом зависимости (13):

$$\Delta w_{ik,m}^{(1)} = \eta \{F'_k(S_{k,m}) \sum_j w_{kj}^{(2)} \delta_{j,m}^{(2)}\} x_{i,m}. \quad (20)$$

Введем новое обозначение $\delta_{k,m}^{(1)}$:

$$\delta_{k,m}^{(1)} = F'_k(S_{k,m}) \sum_j w_{kj}^{(2)} \delta_{j,m}^{(2)}. \quad (21)$$

Переменная $\delta_{k,m}^{(1)}$ в соотношении (21) представляет собой ошибку k -го нейрона скрытого слоя для m -й обучающей пары.

Теперь формулу (20) можно представить в компактном виде:

$$\Delta w_{ik,m}^{(1)} = \eta \delta_{k,m}^{(1)} x_{i,m}. \quad (22)$$

Таким образом, формула (14) представляет собой дельта-правило, которое позволяет произвести настройку весовых коэффициентов между нейронами выходного и скрытого слоев, а формула (22) — между нейронами скрытого и входного слоев, соответственно.

Более подробное описание алгоритма обратного распространения ошибки для рассмотренного примера приводится ниже.

Шаг 1. Инициализация. Вводим обозначения и начальные данные:

m — номер обучающего примера, $1 \leq m \leq M$, где M — количество примеров обучения;

j — номер выходного нейрона, $1 \leq j \leq J$, где J — количество выходных нейронов;

k — номер скрытого нейрона, $0 \leq k \leq K$, где K — количество скрытых нейронов;

i — номер входного нейрона, $0 \leq i \leq I$, где I — количество входных нейронов;

$W^{(1)}$, $W^{(2)}$ — матрицы весовых коэффициентов скрытого и выходного слоев, соответственно;

$X = (x_0, x_1, \dots, x_i, \dots, x_j, \dots, x_J)^T$ — вектор входного сигнала, где $x_0 = 1$ соответствует сигналу, формирующему пороговое значение;

$H = (h_0, h_1, \dots, h_k, \dots, h_K)^T$ — вектор выходного сигнала скрытого слоя нейронов, где $h_0 = 1$ соответствует сигналу, формирующему пороговое значение;

$Y = (y_1, y_2, \dots, y_j, \dots, y_J)^T$ — вектор выходного сигнала;

T — символ транспонирования;

Всем весовым коэффициентам нейросети — $w_{ik}^{(1)}$ и $w_{kj}^{(2)}$ присваиваем небольшие начальные значения, которые выбираем случайным образом. Устанавливаем коэффициент скорости обучения η в пределах от 0,01 до 1 и задаем минимальное значение ошибки функционирования нейросети ξ .

Шаг 2. Прямое распространение сигнала. Вводим m -й обучающий пример и вычисляем для него фактическое значение выходного сигнала j -го нейрона:

$$y_{j,m} = F_j \left(\sum_{k=0}^K w_{kj}^{(2)} h_{k,m} \right), \quad h_{k,m} = F_k \left(\sum_{i=0}^I w_{ik}^{(1)} x_{i,m} \right).$$

Шаг 3. Расчет ошибки нейросети. Производим расчет ошибки между требуемым $g_{j,m}$ и фактическим $y_{j,m}$ значениями выходного сигнала j -го нейрона для m -го обучающего примера:

$$E_m = \frac{1}{2} \sum_{j=1}^J (g_{j,m} - y_{j,m})^2.$$

Затем вычисляем общую ошибку функционирования НС:

$$E = \sum_{m=1}^M E_m.$$

Шаг 4. Проверка окончания обучения. Если $E \leq \xi$, то вычисления завершаются. Переходим к шагу 6, в противном случае — переходим к шагу 5.

Шаг 5. Обратное распространение сигнала. Вычисляем ошибки нейронов выходного слоя и корректируем связи между выходным и скрытым слоями нейронов для m -го обучающего примера:

$$\begin{aligned} \delta_{j,m}^{(2)} &= (g_{j,m} - y_{j,m}) F'_j(S_{j,m}); \\ \Delta w_{kj,m}^{(2)}(t) &= \eta \delta_{j,m}^{(2)} h_{k,m}; \\ w_{kj,m}^{(2)}(t+1) &= w_{kj,m}^{(2)}(t) + \Delta w_{kj,m}^{(2)}(t). \end{aligned} \quad (23)$$

Вычисляем ошибки нейронов скрытого слоя и корректируем связи между скрытым и входным слоями нейронов для m -го обучающего примера:

$$\begin{aligned} \delta_{k,m}^{(1)} &= F'_k(S_{mk}) \sum_j w_{kj}^{(2)} \delta_{j,m}^{(2)}; \\ \Delta w_{ik,m}^{(1)}(t) &= \eta \delta_{k,m}^{(1)} x_{i,m}; \\ w_{ik,m}^{(1)}(t+1) &= w_{ik,m}^{(1)}(t) + \Delta w_{ik,m}^{(1)}(t), \end{aligned} \quad (24)$$

где t — шаг итерации; $\Delta w_{kj,m}^{(2)}(t)$, $\Delta w_{ik,m}^{(1)}(t)$ — коррекция весовых коэффициентов; $w_{kj,m}^{(2)}(t)$, $w_{ik,m}^{(1)}(t)$ — значения весовых коэффициентов до коррекции; $w_{kj,m}^{(2)}(t+1)$, $w_{ik,m}^{(1)}(t+1)$ — значения весовых коэффициентов после коррекции.

После вычислений согласно (23) и (24) переходим к шагу 2.

Шаг 6. Завершение обучения. Условие $E \leq \xi$ выполнено. Процесс обучения завершен.

Представленный алгоритм, основу которого составляет метод градиентного спуска, имеет ряд проблем при его практическом применении.

1. Незнание выбора числа нейронных слоев и количества нейронов для каждого слоя.

2. Влияние случайной инициализации весовых коэффициентов НС на поиск минимума функции ошибки. Данная проблема отражает то, что при различной инициализации весовых коэффициентов НС могут получаться различные решения. Это характеризует неустойчивость алгоритма обучения, когда НС в одних случаях может обучаться до требуемой ошибки, а в других нет.

3. Сложность выбора подходящего коэффициента скорости обучения η . Так, слишком малое значение коэффициента η медленно изменяет весовые коэффициенты, что увеличивает время процесса обучения и приводит к скатыванию НС в локальный минимум. Большое значение η быстрее изменяет весовые коэффициенты. С одной стороны, это позволяет ускорить процесс обучения, с другой стороны — вызывает резкие скачки при перемещении по поверхности ошибок НС. Это может привести к пропуску глобального минимума, а в случае сложной формы поверхности ошибок — к выбору неправильного направления минимизации ошибки.

4. Невозможность определения точек локального и глобального минимумов, так как метод градиентного спуска их не различает.

Рассмотрим эти проблемы более подробно. Первая проблема, которая появляется при практическом применении МП, связана с его конфигурированием. Совершенно ясно, что размерность входного и выходного слоев нейронов определяется из условия задачи. Неясность возникает с выбором количества скрытых нейронных слоев и количества нейронов в каждом из них.

Ранее отмечалось, что НС с двумя слоями нейронных связей позволяет аппроксимировать любые выпуклые области в пространстве решений. При этом точность аппроксимации зависит от числа нейронов в скрытом слое. Чем больше число нейронов в скрытом слое, тем больше точность. Однако при слишком большой размерности скрытого слоя может наступить явление, называемое «переобучением» сети, что приводит к ухудшению обобщающей способности НС. При слишком малой размерности скрытого слоя нейросеть может либо попасть в нежелательный локальный минимум, либо существенно увеличить время процесса обучения.

Обеспечить требуемую точность и повысить обобщающую способность можно также с помощью увеличения количества скрытых слоев нейронов. Однако НС с несколькими скрытыми слоями обучаются еще медленнее.

Вместе с этим, существуют методы подбора конфигурации нейронной сети [6, 7, 9, 11, 12]. Применение таких методов приводит к минимизации количества весовых коэффициентов и нейронов в сети. Этим уменьшается сложность НС и повышается ее способность к обобщению. Однако решение по выбору окончательной конфигурации сети может быть принято только после полноценного обучения (с уменьшением ошибки функционирования до уровня, признаваемого удовлетворительным) различных ее вариантов.

Вторая проблема возникает при определении начальных значений (инициализации) весовых коэффициентов. К сожалению, не существует универсального метода выбора начальных значений весовых коэффициентов, который гарантировал бы нахождение наилучшей начальной точки для любой решаемой задачи. Поэтому без какой-либо априорной информации они часто выбираются случайным образом, согласно критерию минимальной энтропии [15, 16, 18]. При этом важен размер случайно инициализируемых весовых коэффициентов. Например, для сигмоидной ФА нейронов, если весовые коэффициенты будут иметь большие значения (положительные или отрицательные), выходная активность нейронов, учитывая формулы (8), (5), (6), будет близка к единице или нулю. В данном случае весовые коэффициенты будут изменяться незначительно. Это приведет к тому, что процесс обучения остановится в ближайшем от начальных значений локальном минимуме. В работах [15, 16, 18, 20] начальные значения весовых коэффициентов предлагается случайно выбирать в диапазонах $[-0,05, 0,05]$, $[-0,1, 0,1]$, $[-0,3, 0,3]$. В работе [8] рекомендуется случайно выбирать значения весовых коэффициентов следующим образом:

$$w_{ij} \approx \frac{1}{\sqrt{n(i)}}$$

где $n(i)$ — число нейронов в слое i .

Не менее важная проблема, связанная с применением МП, возникает вследствие тенденции к стабилизации метода градиентного спуска в локальных минимумах. Метод градиентного спуска требует выполнения бесконечно малых шагов при перемещении по поверхности ошибок. Это неосуществимо на практике, так как выполнение данного условия ведет к бесконечно длительному времени процесса обучения МП. По этой причине в целях ускорения процесса обучения значение коэффициента скорости обучения η в определенной степени увеличивают.

Однако увеличение η вызывает резкие скачки при перемещении по поверхности ошибок НС. В свою очередь, перемещение с увеличенным коэффициентом скорости обучения может привести к пропуску глобального минимума, а в случае сложной формы поверхности ошибок — к выбору неправильного направления минимизации ошибки функционирования НС.

Для того чтобы сгладить резкие скачки в процессе коррекции весовых коэффициентов при перемещении по поверхности функции ошибок НС, правило обновления весовых коэффициентов дополняют значением коррекции весового коэффициента, выполненной на предыдущем шаге итерации:

$$\Delta w_{ij,m}^{(s)}(t) = \eta \delta_{j,m}^{(s)} u_{i,m}^{(s)} + \mu \Delta w_{ij,m}^{(s)}(t-1), \quad (25)$$

где s — номер слоя нейронных связей, $1 \leq s \leq S$; t — шаг итерации; $\Delta w_{ij,m}^{(s)}(t)$ — текущая коррекция весовых коэффициентов в s -м слое нейронных связей для m -й обучающей пары; $\Delta w_{ij,m}^{(s)}(t-1)$ — коррекция весовых коэффициентов, выполненная на предыдущем шаге итерации; $\delta_{j,m}^{(s)}$ — ошибка j -го нейрона в s -м слое нейронных связей для m -й обучающей пары; $u_{i,m}^{(s)}$ — значение выходного сигнала нейрона в s -м слое; η — коэффициент скорости обучения; μ — коэффициент момента.

В формуле (25) коэффициент μ называют *моментным параметром*, а последнее слагаемое — *моментным термом* (momentum term) [8, 15]. Его значение выбирают в пределах от 0 до 1. Введение его позволяет выходить из небольших локальных минимумов.

Для обучения многослойной НС, содержащей более одного скрытого слоя нейронов, алгоритм обратного распространения ошибки требует определения производной ошибки НС относительно весовых коэффициентов всех слоев сети. Вычисления, связанные с этим процессом, становятся более сложными и менее эффективными в практическом применении. Поэтому на практике очень часто ограничиваются НС с двумя слоями нейронных связей.

Существуют адаптивные методы подбора коэффициента скорости обучения η [5, 10, 19], которые более прогрессивны по сравнению с постоянным значением. Однако они тоже не могут считаться наилучшим решением, особенно при моделирова-

нии динамических процессов. Учитывая эти особенности, выбор коэффициента η на практике чаще всего осуществляют опытным путем.

Таким образом, решить большинство проблем применения МП представляется возможным в процессе экспериментов — путем поиска компромисса между длительностью обучения и ошибкой функционирования НС. Тем не менее, возможности МП не позволяют в полной мере осуществить моделирование изменяющейся во времени последовательности векторов данных. В этой связи можно заключить, что МП имеет смысл использовать для кодирования псевдостационарных сегментов РС.

Литература

1. **Минский М. Л., Пейперт С.** Перцептроны. — М.: Мир, 1971.
2. **Скурихин А. Н.** Нейронные сети: определения, концепции, применение. — М.: ЦНИИ управления экономики и информатики, 1991.
3. **Фланеган Дж.** Анализ, синтез и восприятие речи. — М.: Связь, 1968.
4. **Bishop C.** Neural Networks for Pattern Recognition. — Oxford University Press, 1995.
5. **Gill P., Murray W., Wright M.** Practical Optimization. — N. Y.: Academic Press, 1981.
6. **Haykin S.** Neural Networks: A comprehensive Foundation. — N. Y.: Macmillan College Publishing Company, 1994.
7. **Hassibi B., Stork D.** Second order derivatives for network pruning: Optimal brain surgeon // Advances in NIPS2 / Ed. D. Touretzky. — San Mateo: Morgan Kaufmann, 1993. — P. 164–171.
8. **Horz J., Krogh A., Palmer R.** Introduction to the theory of neural computation. — Addison Wesley Publishing Company, 1991.
9. **Hush D., Horne B.** Progress in supervised neural networks // IEEE Signal Processing Magazine, 1993, January. — P. 8–39.
10. **Klirmauskas G.** Neural Ware — User manual. Natick, USA: Neural Ware Inc., 1992.
11. **Le Cun Y., Denker J., Solla S.** Optimal brain damage // Advances in NIPS2 / Ed. D. Touretzky. — San Mateo: Morgan Kaufmann, 1990. — P. 598–605.
12. **Li Q., Tufts D.** Synthesizing neural networks by sequential addition of hidden nodes // IEEE Proc. ICNN, Orlando, 1994. — P. 708–713.
13. **Lippmann R. P.** An Introduction to Computing with Neural Nets // IEEE Transactions ASSP Magazine, April 1987. — N 2. — L4. — P. 4–22.
14. **Rosenblatt F.** Principle of neurodynamics. — N. Y.: Spartan, 1992.
15. **Rumelhart D. E., Hinton G. E., Williams R. J.** Learning representation by Back-Propagating Errors // Nature, October 1986. — N 323 (9). — P. 533–536.
16. **Rumelhart D. E., McClelland J. L.** Parallel Distributed Processing. — Cambridge, MA: MIT Press, 1986.
17. **Sankar K. P., Sushmita M.** Multilayer Perceptron, Fuzzy Sets and Classification // IEEE Transactions on Neural Networks, 3(5). — September 1992. — P. 683–697.
18. **Thimm G., Fiesler E.** High Order and Multilayer Perceptron Initialization // IEEE Transactions on Neural Networks, 8(2). — 1997. — P. 1045–9227.
19. **Wasserman P. D.** Experiments in translating Chinese characters using backpropagation // Proceedings of the Thirty-Third IEEE Computer Society International Conference. — Washington, D. C.: Computer Society Press of the IEEE, 1988.
20. **Werbos P. J.** Backpropagation through time: what it does and how to do it // Proceedings of the IEEE, 1990, 78(10). — P. 1550–1560.

ИЗДАТЕЛЬСТВО «ПОЛИТЕХНИКА» ПРЕДСТАВЛЯЕТ

Куприянов М. С., Матюшкин Б. Д.

Цифровая обработка сигналов: процессоры, алгоритмы, средства проектирования. — 2-е изд., перераб. и доп. — СПб.: Политехника, 2002. — 592 с.: ил.

Книга содержит три части. Первая часть «Процессоры цифровой обработки сигналов» посвящена архитектуре и особенностям организации DSP. Во второй части «Алгоритмы цифровой обработки сигналов» рассматриваются основы теории дискретных систем, методы анализа эффектов квантования сигналов при реализации алгоритмов обработки на DSP, базовые алгоритмы ЦОС и их реализация на DSP. Третья часть «Инструментальные средства проектирования систем ЦОС» содержит описание программных и аппаратных средств, используемых для решения задач проектирования и входящих в стартовый комплекс разработчика систем ЦОС. В приложении приведена система команд семейств DSP5600x и DSP5630x.

Книга рассчитана на инженерно-технических работников, занимающихся проектированием систем ЦОС, а также студентов соответствующих специальностей технических университетов.



Информационно-управляющие системы для подвижных объектов. Семинары ASK Lab 2001 / Под общ. ред. М. Б. Сергеева. — СПб.: Политехника, 2002. — 234 с.: ил.

В книге представлены статьи, посвященные актуальным проблемам в области разработки информационно-управляющих систем для подвижных объектов, вопросам их надежности, алгоритмического и аппаратного обеспечения, защиты информационных каналов.

Книга ориентирована на научных и инженерно-технических работников, специалистов в области встраиваемых систем управления не только авиационных комплексов, но и наземных подвижных дистанционно управляемых объектов различного назначения.

Ляликов А. П.

Трактат об искусстве изобретать. — СПб.: Политехника, 2002. — 416 с.: ил.

В книге изложены основные аспекты — философский, исторический, психологический, системный и эвристический — важнейшей отрасли общечеловеческой культуры, которая является источником и основой бытия, личного и социального, — технического творчества.

Книга предназначена для широкого круга читателей: от учащихся и студентов до умудренных жизнью и размышлениями о ее сущности специалистов, собирающихся изобретать, уже изобретающих и даже совсем никогда и ничего не изобретавших.



УДК 681.325.5

ВЗАИМОДЕЙСТВИЕ ОТКРЫТЫХ СИСТЕМ ПРОМЫШЛЕННОЙ АВТОМАТИЗАЦИИ: СОСТОЯНИЕ И ПРОБЛЕМЫ

Чыонг Динь Тяу,

аспирант

Санкт-Петербургский государственный
политехнический университет

Рассматривается технология OPC (OLE for Process Control) — основной стандарт взаимодействия средств современных систем сбора данных и диспетчерского управления – Supervisory Control and Data Acquisition (SCADA), затрагиваются механизм обмена данными между OPC-приложениями и вопрос производительности OPC-серверов, OPC сравнивается с другими стандартами, перечисляются преимущества и недостатки стандарта. В качестве задачи, подлежащей решению, отмечается создание универсального OPC-сервера.

This article examines the OLE for Process Control (OPC) technology — the basic standard of systems interaction in modern Supervisory Control and Data Acquisition (SCADA) system. The mechanism of data communication between OPC-applications and performance of OPC servers are mentioned. The article compares OPC by other standards. The advantages and disadvantages of the standard are listed. As a task subject to the decision, the creation universal OPC server is marked.

Введение

В настоящее время информационная автоматизированная система управления промышленным предприятием имеет открытую архитектуру (рис. 1), включающую следующие неотъемлемые уровни [1].

1. Нижний уровень, который включает полевые шины и отдельные контроллеры (Field Management), представляющие собой «интеллектуальные» технологические устройства: датчики, контроллеры, механизмы и т. д. Поток информации от нижнего уровня должен быть предоставлен пользователю и всем приложениям верхнего уровня, использующим ее посредством цифровых коммуникационных протоколов связи. При этом в системе не должно возникать проблем несовместимости.

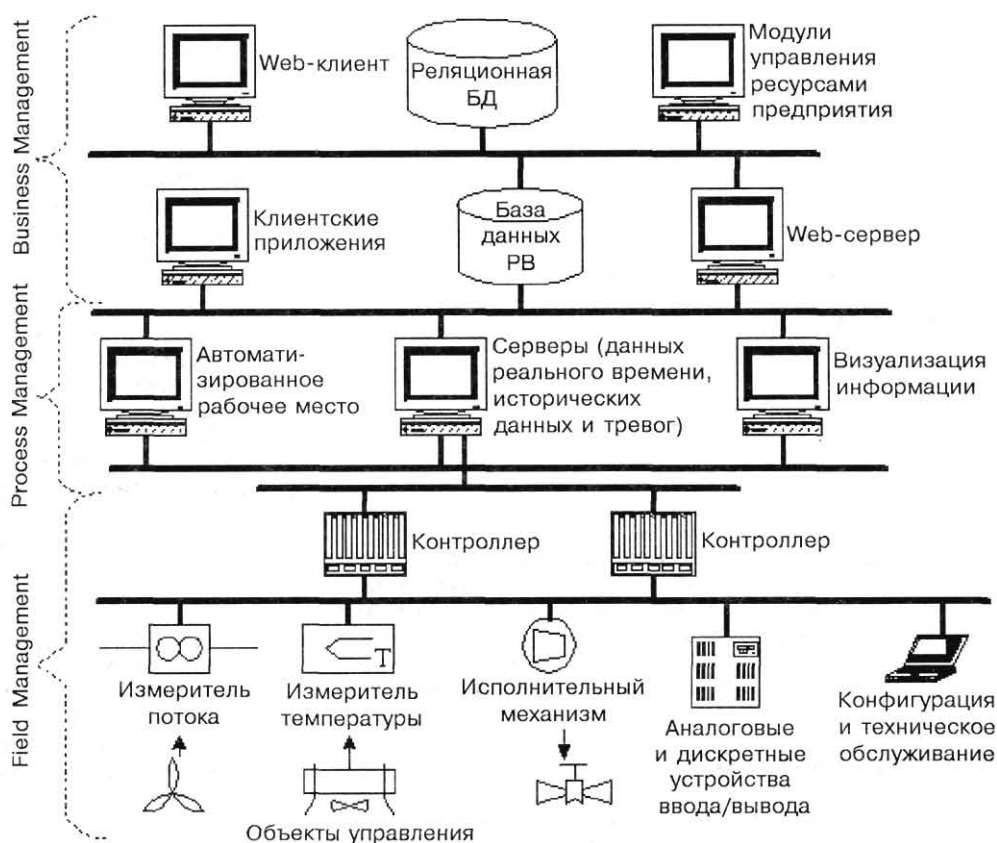
2. Уровень АСУТП (Process Management) — уровень работы систем типа SCADA (систем для сбора данных и диспетчерского управления технологическими процессами на производстве). Этот уровень обеспечивает вторичную обработку данных, которые получены с нижнего уровня, сохранение данных и их доступность приложениям и пользователям верхнего уровня.

3. Уровень АСУП — уровень приложений управления ресурсами предприятия (Business Management). Информация с уровня АСУТП должна быть доступной для уровня АСУП, т. е. доступ к данной информации с помощью прикладных программ не должен вызывать проблемы несовместимости.

Для обеспечения совместимости между уровнями и создания эффективной интегрированной системы управления предприятием системный интегратор или разработчик АСУТП должен извлекать данные технологического процесса в реальном времени с самого нижнего уровня и выстраивать «прозрачный» путь получаемым данным к самым верхним уровням. Чтобы получить систему, отвечающую всем требованиям заказчика, системному интегратору или разработчику необходимо использовать инструментальные средства управления различных уровней — SCADA-пакеты, базы данных, электронные таблицы. Ключ к этому — открытая и эффективная коммуникационная архитектура взаимодействия между приложениями, которую предлагает стандарт OPC.

Что такое OPC?

OPC — это технология связывания и внедрения объектов для систем промышленной автоматизации, предназначенная для обеспечения универсального механизма обмена данными между датчиками, исполнительными механизмами, контроллерами, устройствами связи с объектом, системами представления технологической информации, оперативного диспетчерского управления, а также системами управления базами данных [1, 2, 3].



■ Рис. 1. Архитектура информационной автоматизированной системы управления промышленным предприятием

На сегодняшний день технология OPC в определенной степени реализована и продолжает развиваться. Консорциум «OPC Foundation» пытается охватить все аспекты, связанные с взаимодействием между компонентами программного обеспечения, между программным обеспечением и между системами типа SCADA и технологическим оборудованием. «OPC Foundation» привлекает к разработке спецификаций ведущих производителей оборудования и систем автоматизации. Существует много спецификаций OPC — Data Access (доступ к данным реального времени), Alarms & Events (обработка тревог и событий), Historical Data Access (доступ к историческим данным) и т. д. Поэтому Д. Теркель [4] определяет OPC как стандарт взаимодействия между программными компонентами сбора данных и управления, основанный на объектной модели COM/DCOM фирмы Microsoft. Через интерфейсы OPC одни приложения могут читать или записывать данные в другие приложения, обмениваться событиями, оповещать друг друга о нештатных ситуациях, осуществлять доступ к данным, зарегистрированным в архивах. Эти приложения могут как располагаться на одном компьютере, так и быть распределенными в сети. При этом, независимо от фирмы-поставщика, стандарт OPC, признанный и поддерживаемый всеми ведущими фирмами-производителями SCADA-систем

и оборудования, обеспечит их совместное функционирование.

Популярный класс OPC-приложений представляют собой OPC-серверы конкретных аппаратных устройств, обеспечивающие предоставление информации о состоянии параметров многочисленных устройств технологического процесса, полученной OPC-приложением, OPC-клиентам на локальном компьютере или в компьютерной сети. OPC-сервер создает свою абстракцию устройств (это загружаемый модуль, предоставляющий интерфейс нижнего уровня с оборудованием, на котором исполняется OPC-сервер). OPC-сервер скрывает аппаратно-зависимые детали, такие как интерфейс ввода/вывода, контроллеры прерываний и механизм коммуникации в многопроцессорных системах — любые функции, зависящие от конкретной архитектуры и машин, — позволяя любому OPC-клиенту записывать и считывать данные с устройства. Устройство, для которого есть OPC-сервер, может использоваться вместе с любой современной SCADA-системой.

Современные SCADA-системы не ограничивают выбор аппаратных средств нижнего уровня, так как SCADA-пакеты предоставляют большой набор драйверов или серверов ввода/вывода (в том числе OPC-серверы) и имеют хорошо развитые средства создания собственных программных модулей или драйверов для новых и нестандартных уст-

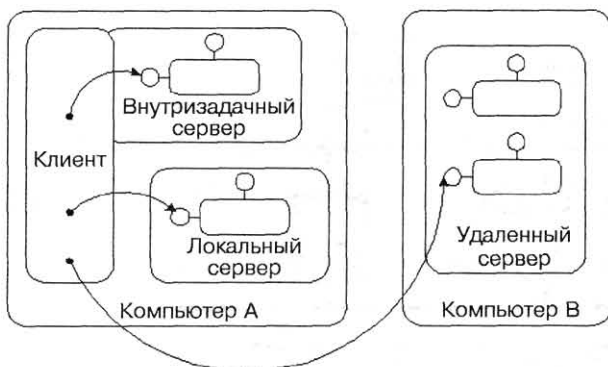


Рис. 2. Различные серверы COM-объектов

роиств нижнего уровня. Производители аппаратных средств нижнего уровня используют спецификации OPC, имеют возможность разрабатывать свой OPC-сервер для обеспечения доступа к данным реального времени и передачи данных приложениям-клиентам различных производителей программного обеспечения для промышленной автоматизации (в том числе SCADA-систем).

Механизм обмена данными OPC

Технология OPC основана на модели распределенных компонентных объектов DCOM (Distributed COM). Модель представляет собой сетевое расширение COM или распределенную COM (Component Object Model). Технология COM — модель компонентных объектов — создана Microsoft для совместного использования различных офисных приложений в Windows, например, электронных таблиц Excel, редактора Word и т. п.

COM поддерживает модель «клиент—сервер». Объекты, называемые серверами, предоставляют некие функции в распоряжение объектов, называемых клиентами. Серверы всегда являются COM-объектами, т. е. объектами, которые подчиняются спецификации COM. С другой стороны, клиенты могут быть COM-объектами или не быть таковыми. Это значит, что некоторые объекты могут быть простыми объектами C++, приложениями Visual Basic и т. п. [5]. Каждый COM-объект существует внутри конкретного сервера. Этот сервер содержит программный код реализации операций, а также данные активного COM-объекта. Один сервер может обеспечивать несколько объектов и даже несколько COM-классов. Как показано на рис. 2, используются три типа серверов [6]:

1) внутризадачный сервер (in-process) — объекты находятся в динамически подключаемой библиотеке и, следовательно, выполняются в том же процессе, что и клиент;

2) локальный сервер (local) — объекты находятся в отдельном процессе, выполняются в том же компьютере, что и клиент;

3) удаленный сервер (remote) — объекты находятся в DLL или в отдельном процессе, который расположен на удаленном от клиента компьютере.

Клиенты и локальные серверы находятся в различных адресных пространствах. OPC берет на себя передачу данных между ними. В случае внутризадачных серверов родительский процесс загружает DDL-файл, содержащий COM-сервер. Это означает, что внутризадачный сервер находится в том же адресном пространстве, что и вызвавший его процесс. В случае DCOM вызов любой функции объекта перехватывается специальным агентом-посредником, называемым Proxy Manager или Stub Manager. Proxy Manager получает все запросы на транспортировку и определяет, стандартные ли данные запрашивает или предлагает клиент. Если это так, Proxy Manager передает запрос подходящему Proxy-объекту. Подходящие Proxy и Stub — это объекты, предназначенные для транспортировки данного интерфейса и/или типа данных. Аналогичный процесс происходит на стороне Stub Manager (рис. 3). Встроенные Proxy и Stub называются транспортировщиками интерфейса.

Популярным вариантом OPC-клиента является SCADA-система. С помощью OPC-серверов SCADA-система может обращаться к любому аппаратному устройству (рис. 4) путем вызова определенных функций OPC-сервера, подписываться на получение определенных данных какого-либо канала или устройства с какой-либо частотой. В свою очередь, OPC-сервер опрашивает аппаратное устройство, вызывает известные функции клиента, подтверждает получение данных и посылает клиенту требуемые данные.

Обмен данными между OPC-клиентом и OPC-сервером может быть реализован в трех режимах: 1) синхронного чтения/записи; 2) асинхронного чтения/записи; 3) в режиме подписки (только чтение).

При синхронном чтении/записи клиент посылает серверу запрос со списком интересующих его переменных и ждет, когда сервер его выполнит.

При асинхронном чтении/записи клиент посылает серверу запрос, а сам продолжает работать. Когда сервер выполнил запрос, клиент получает уведомление. Этот режим обладает большей гибкостью.

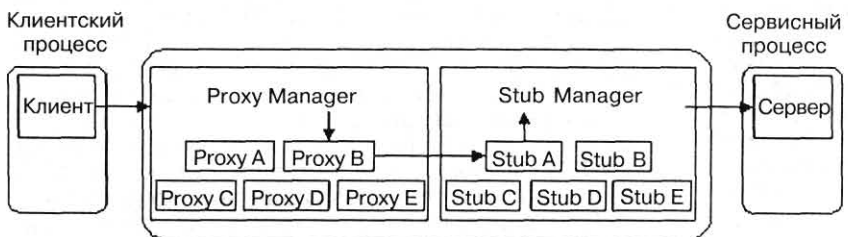


Рис. 3. Вызов функции объекта в случае DCOM

тью и рекомендуется в большинстве случаев, так как он минимизирует затраты процессорного времени и сетевых ресурсов, особенно при передаче большого количества элементов данных.

В случае подписки клиент передает серверу список интересующих его переменных, а сервер присылает клиенту информацию об изменившихся переменных из этого списка. При этом сервер будет передавать клиенту информацию только тогда, когда данные изменились, причем эти данные передаются единым блоком. Эти меры позволяют существенно ускорить обмен данными, особенно если речь идет о взаимодействии через сеть.

OPC и другие стандарты

Для подключения устройства ввода/вывода к SCADA-системе в настоящее время используются следующие механизмы:

Dynamic Data Exchange (DDE) — стандартный динамический обмен данными;

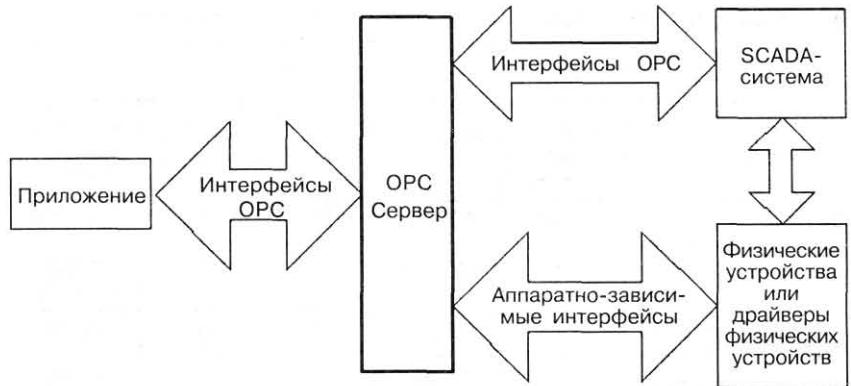
собственные протоколы фирм-разработчиков SCADA-систем; OPC-протокол.

В SCADA-системах основным механизмом, используемым для связи с внешними уровнями (верхними и нижними), изначально являлся механизм DDE, использовавшийся в качестве механизма разделения данных между прикладными системами и аппаратными устройствами. Но DDE присущ ряд ограничений по скорости и надежности. Стандарт DDE плохо подходит для обмена данными в реальном масштабе времени. Для преодоления недостатков DDE разработчики SCADA-систем предложили свои собственные протоколы, такие как FastDDE (Wonderware), AdvancedDDE (Rockwell), I-NET (Adastra). Подобные протоколы обеспечивают самый быстрый обмен данными с контроллерами, но имеют ряд недостатков [3]:

1) для каждой SCADA-системы пишется свой драйвер для каждого поставляемого устройства;

2) в общем случае две SCADA-системы не могут иметь доступ к одному устройству в одно и то же время, поскольку каждая из них поддерживает обмен через свой драйвер;

3) данные протоколы не поддерживают удаленного доступа к устройству.



■ Рис. 4. Обмен между SCADA-системой, приложениями и физическими устройствами через OPC-интерфейс

Взамен DDE компания Microsoft предложила более эффективное и надежное средство передачи данных между процессами COM/DCOM. На базе COM/DCOM появляется новый стандарт OPC, ориентированный на рынок промышленной автоматизации. Стандарт OPC поддерживается во многих современных SCADA-системах. В табл. 1 перечислены некоторые из популярных на западном и российском рынках SCADA-систем поддержки OPC.

Важной характеристикой является производительность OPC-сервера. Она зависит от ряда факторов — вида OPC-сервера (внутризадачный, локальный или удаленный), возможности группировки данных для отправки клиентам (передача 100 элементов раз в 1 с займет значительно меньше ресурсов, чем одного элемента через каждые 10 мс [4]), операционной системы (по результатам исследований центра Doculabs [7], любое приложение выполняется в среде Windows 2000 быстрее, чем на Windows NT, на 16–122 %) и т. д. Сказанное иллюстрирует табл. 2, в которой представлены результаты тестирования производительности OPC-серверов, разработанных консорциумом «OPC Foundation» [8]. OPC-клиентом в данном случае является простая специализированная программа, разработанная консорциумом для тестирования производительности OPC-серверов.

Преимущества и недостатки OPC

Как уже говорилось, основная цель OPC-стандарта заключается в определении механизма доступа к данным с любого устройства из различных приложений. OPC позволяет производителям оборудования стандартизовать свои продукты, обеспечивает подключение оборудования различных производителей, в том числе подключение к SCADA-системам.

■ Таблица 1. Популярные SCADA-системы поддержки OPC

SCADA-системы	Фирма-изготовитель	Страна
InTouch	Wonderware	США
GenieDAQ	Advantech	США
Genesis32	Iconics	США
Trace Mode	Adastra	Россия
Citect	Ci Technologies	Австралия
FactoryLink	United States DATA Co.	США
RSView	Rockwell Software Inc.	США
Lookout	National Instruments	США
Fix Dynamics	Intellution	США
Контур	Объединение ЮГ	Украина
Eclipse SCADA	Eclipse Software	США

■ **Таблица 2.** Производительность OPC-серверов «OPC Foundation» (на компьютере с процессором Pentium 233, ОС Windows NT 4.0 Workstation)

Тип сервера		Число элементов транзакций	Кратность передач	Время, с	Затраты процессорного ресурса, (%)	Скорость опроса элементов, с
Внутри-задачный	Клиент	100	100000	9	100	1111000
	Сервер	—	—	—	—	—
	Клиент Сервер	1 —	5000000 —	22 —	100 —	227000 —
Локальный	Клиент	100	10000	16	40	62500
	Сервер	—	—	—	60	—
	Клиент Сервер	1 —	50000 —	15 —	50 50	3333 —
Удаленный (Ethernet 10Base-T)	Клиент	100	1000	15	10	6666
	Сервер	—	—	—	12	—
	Клиент Сервер	1 —	3000 —	9 —	10 10	333 —

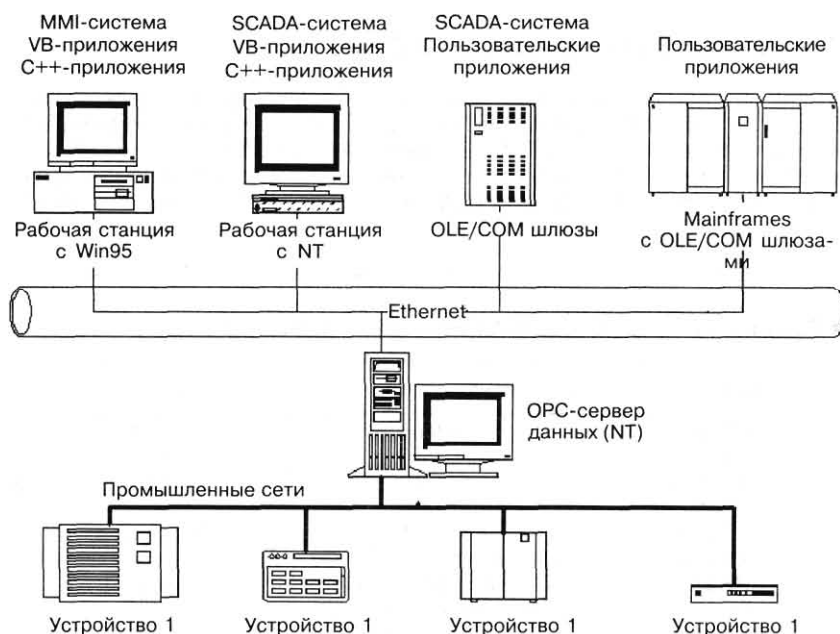
Идеальной выглядела бы следующая картина. Все в мире признают OPC своим стандартом. При этом все поставщики оборудования снабжают свои продукты OPC-серверами. Все поставщики программ для систем управления делают свои продукты OPC-клиентами — а все производители операционных систем поддерживают в своих ОС технологии COM/DCOM, а также предоставляют сервисный инструментарий. Все это они делают на высоком профессиональном уровне и очень грамотно рассказывают сборщикам систем, как это все собирать и конфигурировать. Тогда пользователи будут иметь широкие возможности, чтобы выбирать оптимальные для своей системы компоненты, а не создавать программу «с нуля» и не выбрасывать имеющийся контроллер. Таким образом, по образному выражению Теркеля [4], OPC — это свобода выбора.

Преимущество применения стандарта OPC с точки зрения интеграции достаточно прозрачно. С применением OPC-стандарта появляется возможность интегрировать в единую систему неоднородные узлы (рис. 5). Каждый из представленных на рисунке узлов может обслуживаться OPC-сервером или OPC-клиентом, которые обеспечивают простой приток данных в программу визуализации, в базы данных и т. п. с нижнего уровня на более высокий уровень.

OPC-стандарт можно рассматривать как «универсальный переходник». Если мы заменяем какой-нибудь компонент клиентского приложения, то нет необходимости корректировать серверное приложение, поскольку COM обеспечивает эффективное управление изменением программы. Если мы хотим добавить в систему новое оборудование, то достаточно включить его в OPC-сервер. При этом клиентские программы не требуют изменений, в том числе для SCADA-систем.

В настоящее время консорциум «OPC Foundation» набирает силу в разработке открытых промышленных стандартов на основе OPC-стандарта на базе PC и операционных систем Microsoft. Сейчас в состав «OPC Foundation» входят более 300 членов, среди которых практически все мировые ведущие производители технологического оборудования, систем автоматизированного управления и программного обеспечения. Членами организации являются, например, фирмы Iconics Inc. (США), Wonderware (США), Adastrа (Россия), Siemens (Германия), Rockwell Software (США), Intellution (США), Ci Technologies (Австралия), Indusoft Russia (Россия), Fastwel Inc. (Россия), ABB Automation (США), Fieldbus Foundation (США), Toshiba Corp. (Япония), Hitachi (Япония), National Instruments (США). Организация пытается охватить все аспекты, связанные с взаимодействием с технологическим оборудованием. Ведущие производители стараются максимально учесть свой опыт и предоставить все необходимое пользователям OPC. Этот факт показывает большой авторитет OPC-технологии. Это перспективная технология для использования в развитии автоматизированных систем управления.

Однако OPC-технология несет в себе и ряд недостатков.



■ **Рис. 5.** Стандарт OPC — путь к интеграции разнородных систем

1. OPC не обеспечивает работы в жестком реальном времени, поскольку в DCOM отсутствуют понятия качества обслуживания, крайних сроков и т. п. В то же время контроль за «устареванием» данных имеется: каждое передаваемое значение сопровождается меткой времени происхождения.

2. OPC может использоваться только там, где установлен Microsoft DCOM. В настоящее время технология COM/DCOM поддерживается в ОС Microsoft, в Unix-подобных ОС, в ОС реального времени VxWorks; обеспечивается фирмой WindRiver, имеется поддержка OPC, встроенная в систему разработки Tornado [9]. В других операционных системах поддержки COM/DCOM нет. Это не очень отрадный факт, поскольку разработчиков систем автоматизации в первую очередь интересуют ОС реального времени.

3. OPC-интерфейс — это интерфейс для системы верхнего уровня. Это означает, что нижележащие слои — контроллеры, полевые шины и т. д. — недоступны для самого OPC-сервера и для пользователей и представлены в виде «черного ящика».

4. Программирование COM непростое, даже с применением ATL. Сами OPC-объекты и их OPC-интерфейсы достаточно сложны и громоздки. Необходимо владеть вопросами системного уровня, такими как асинхронный обмен, многозадачность, синхронизация. Кроме того, в COM допускается выделение памяти в сервере, а удаление ее возлагается на клиента. Малейшая неточность вызывает трудно устранимые утечки памяти. А учитывая, что OPC-сервер обычно должен работать стационарно, рано или поздно крах системы неизбежен.

5. Технология DCOM не работает в глобальных сетях. Поэтому идея применения OPC-технологии в Internet нереальна. Вместе с тем, организовать обмен в этом случае также возможно следующим образом. Расширение Web-сервера является OPC-клиентом, собирающим данные от OPC-серверов. На стороне клиентов запускается динамическая html- или xml-страница, которую даже можно сделать OPC-сервером для других приложений. Указанная страница получает данные от Web-сервера [9].

Заключение

OPC-технология содержит стандарты, обеспечивающие взаимодействие программных средств в промышленной автоматизации. В технологии заложены богатые возможности, которые дают организаторам предприятия возможность интеграции разнородных систем и обеспечивают разработчикам свободу выбора с применением OPC-драйверов, позволяют не задумываться по поводу поддержки аппаратуры завтрашнего дня. Вместе с тем можно отметить, что процесс становления технологии еще далеко не завершен и есть много проблем, которые предстоит решить. Одной из таких проблем является отсутствие серьезных универсальных OPC-серверов, обеспечивающих подключение любого оборудования (основные корпорации не заинтересованы в поддержке конкурентов). Решение этой задачи представляется весьма актуальным.

Литература

1. **OLE for Process Control Data Access Standard Version 1.0A.** — September 11, 1997 OPC Foundation, Austin, Texas.
2. **Fastwel Universal OPC-сервер.** Версия 1.0. Руководство пользователя. — Fastwel Inc., 1999–2001.
3. **Куцевич Н. А., Жданов А. А.** Программное обеспечение систем контроля и управления и Windows-технологии // Мир компьютерной автоматизации. — 1999. — № 3.
4. **Теркель Д.** OLE for Process Control—свобода выбора // Современные технологии автоматизации. — 1999. — № 3. — С. 28–32.
5. **Пьюполо Д.** OLE: создание элементов управления: / Пер. с англ. — Киев: Издательская группа BHV, 1997. — 432 с.
6. **Технология** разработки программного обеспечения: Учебник / С. Орлов. — СПб.: Питер, 2002. — 464 с.
7. **М. Pendleton, G. Desai,** @Bench Test Report: Performance and Scalability of Windows 2000. Doculabs, 2000. <http://www.doculabs.com>
8. **DCOM,** OPC and Performance Issues, Intellution Inc, 1998, http://www.opcfoundation.org/07_download/performance.doc
9. **Куцевич И. В., Григорьев А. Б.** Стандарт OPC — путь к интеграции разнородных систем // Мир компьютерной автоматизации. — 2001. — № 1.

УДК 621.391.28

ФОРМАЛИЗОВАННОЕ ОПИСАНИЕ ПРОЦЕССА ФУНКЦИОНИРОВАНИЯ ЦЕНТРОВ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СЕТЕЙ ДЛЯ ОЦЕНКИ ВЕРОЯТНОСТНО – ВРЕМЕННЫХ ХАРАКТЕРИСТИК

М.О. Колбанёв,

канд. техн. наук, доцент

Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М. А. Бонч-Бруевича

Приводятся требования к методам описания процессов функционирования центров обработки информации и управления, ориентированным на построение математических моделей оценки вероятностно-временных характеристик. Излагаются основы метода составления описания, отвечающего этим требованиям.

The methods of the description of the processes functioning of centers switching and information processing focused on construction of mathematical models to an estimation of time and probabilistic characteristics are considered. The bases of a method of drawing up of the description are stated.

Количественный рост инфраструктуры интеллектуальных сетей связи и реализуемых ими приложений ведет к качественным изменениям в направлении все большей интеллектуализации центров обработки информации и управления (ЦОУ).

Модели, ориентированные на исследование процессов функционирования ЦОУ в условиях реального времени, имеют иерархическую структуру (рис. 1). В настоящей статье излагается метод составления формализованного описания процесса функционирования (ПФ) при взаимодействии ЦОУ с внешним окружением (ВО), ориентированный на использование при исследовании вероятностно-временных характеристик (ВВХ) [1–5]. Описание процесса взаимодействия ЦОУ с внешним окружением основывается на следующих предположениях о функционировании ЦОУ и структуре его ВО.

ЦОУ предоставляет информационные услуги нескольких видов. Под информационной услугой понимается либо процесс установления соединения (например, внутреннего, исходящего, входящего, транзитного), либо процесс обслуживания вызова (например, изменение скорости передачи информации или состава участников соединения), либо заказ какой-нибудь услуги, не относящейся к установлению соединений или обслуживанию вызова. В процессе предоставления информационной услуги каждого вида задействована определенная группа элементов управления и реализуется определенный процесс (сценарий) обмена управляющими сообщениями сигнализации. С каждой услугой связана определенная управляемая элементами управления линия поведения ЦОУ:

внешнее окружение ЦОУ, реализующего некоторую услугу, состоит из группы элементов управления разных типов. Тип такого элемента определяется типом используемой терминальной аппаратуры и программного обеспечения (например, пользователи с аналоговыми или цифровыми терминалами, линии связи с той или иной системой сигнализации и т.п.), а также той функцией, которую выполняет элемент в процессе предоставления информационных услуг (например, вызывающий или вызываемый пользователь сети). Тип элемента управления находит свое конкретное выражение в множествах входных и выходных сообщений, которые могут поступать от него на ЦОУ и соответственно быть приняты им от ЦОУ;

в процессе взаимодействия ЦОУ и его ВО находятся в определенных внутренних состояниях, которые из-

меняются под воздействием входных и выходных сообщений соответственно. Состояние, в которое переводятся ЦОУ и ВО при поступлении очередного сообщения, определяется как содержанием этого сообщения, так и последовательностью входных/выходных сообщений, выработанных ранее в процессе предоставления информационной услуги;

внутреннее состояние ЦОУ характеризуется составом и состоянием занятых ресурсов исполнительной системы, внутреннее состояние ВО — составом и состоянием элементов управления, взаимодействующих с центром в процессе предоставления информационной услуги. Эти состояния в обобщенном виде отражают всю предысторию ПФ ЦОУ и отображаются в базе данных управления управляющей системы.

Процесс функционирования ЦОУ удобно сначала описать изолированно для каждого вида информационной услуги и представить в виде совокупности графов процесса обмена сообщениями (ПОС) $G_{\text{ПОС}}^d$, где d — номер, идентифицирующий вид информационной услуги, предоставляемой ЦОУ. Граф $G_{\text{ПОС}}^d$ является двудольным, задает процесс обмена сигнальными сообщениями между центрами и ВО при предоставлении услуги d -го вида и строится в следующем порядке.

Шаг 1. Определить состав элементов управления, участвующих в процессе реализации информационной услуги d -го вида, и множества входных и выходных сообщений, которые могут вырабатываться и приниматься каждым из этих элементов.

Шаг 2. Определить множества входных (X_d) и выходных (Y_d) сообщений, которыми ЦОУ обменивается с ВО. Множество X_d определяется путем объединения множеств входных сообщений, которые могут поступить от каждого элемента управления, участвующего в процессе реализации услуги; множество Y_d — путем прямого произведения множеств выходных сообщений, передаваемых этими элементами. В результате каждое выходное сообщение представляет собой упорядоченную по типам элементов управления совокупность выходных сообщений, которые одновременно (в одном цикле выдачи) могут быть переданы во ВО всем участникам процесса реализации информационной услуги.



■ Рис. 1. Иерархическая структура моделей

Шаг 3. Поставить в соответствие каждому входному и каждому выходному сообщениям ЦОУ вершину графа $G_{\text{ПОС}}^d$. Вершины, соответствующие входным сообщениям, составляют первую, а выходным — вторую долю вершин графа.

Шаг 4. Вершину первой (второй) доли соединить дугой с вершиной второй (первой) доли, если в соответствии с логикой предоставления услуги в ответ на входное (выходное) сообщение ЦОУ может выдать с ненулевой вероятностью выходное (входное) сообщение.

В тех случаях, когда процесс взаимодействия не предусматривает процедуры «рукопожатия», т. е. когда одна из двух взаимодействующих сторон передает несколько сообщений, не дожидаясь ответной реакции другой стороны, для сохранения двудольной структуры графа $G_{\text{ПОС}}^d$ могут быть введены фиктивные ответные сообщения.

Пример графа обмена входными и выходными сообщениями при установлении соединения в Ш-ЦСИО в интерфейсе UNI (User-Network Interface) между пользователями и сетью приведен на рис. 2. Дуги, пронумерованные цифрами от 1 до 12, задают порядок следования сообщений при «нормальном» установлении и разрушении соединения. Состав и содержание сообщений сигнализации, которые циркулируют в интерфейсе UNI между пользователями и ЦОУ, соответствует версии 3.1 ATM Forum.

После построения графов процесса обмена входными/выходными сообщениями появляется возможность представить процесс взаимодействия в виде последовательности действий, выполняемых ЦОУ и его ВО в интервалах времени между моментами появления входных и выходных сообщений. Множество интервалов времени определяет множество этапов ПФ ЦОУ при предоставлении информационных услуг и разделяется на две группы. Этапы переходного состояния начинаются в моменты поступления входных сообщений и заканчиваются в моменты выработки ЦОУ ответных, выходных; этапы устойчивого состояния начинаются в моменты выдачи во ВО выходных сообщений и заканчиваются в моменты поступления входных сообщений из ВО.

Описание ПФ на уровне этапов взаимодействия ЦОУ с ВО представляется в виде древовидного графа $G_{\text{ПФ}}(A, B, X, Y)$, где A и B — множества вершин графа, которые соответствуют этапам устойчивого и переходного состояний; X и Y — множество дуг, которые соответствуют входным сообщениям, инициирующим этапы переходного состояния, и выходным, инициирующим этапы устойчивого состояния.

Множества X и Y были определены в процессе построения графов $G_{\text{ПОС}}^d$ ($d = 1, 2, \dots, D$). Чтобы определить множества A и B , необходимо

выявить всевозможные последовательности входных/выходных сообщений, которые могут быть выработаны в процессе взаимодействия ЦОУ с ВО. Таким образом, для получения графа $G_{ПФ}$ из двудольных графов $G_{ПФ}^d$ достаточно воспользоваться известными алгоритмами построения деревьев. Пример графа ПФ, соответствующий установлению внутреннего соединения в ЦОУ широкополосной цифровой сети интегрального обслуживания, представлен на рис. 3. Он получен путем преобразования в дерево двудольного графа процесса обмена сигналами, который приведен на рис. 2.

В полученном древовидном графе каждый полный путь, который начинается в корневой и заканчивается в висячей вершине, будет соответствовать одной определенной и допустимой последовательности сообщений взаимодействия. Граф процесса функционирования ЦОУ $G_{ПФ}$ обладает следующими свойствами.

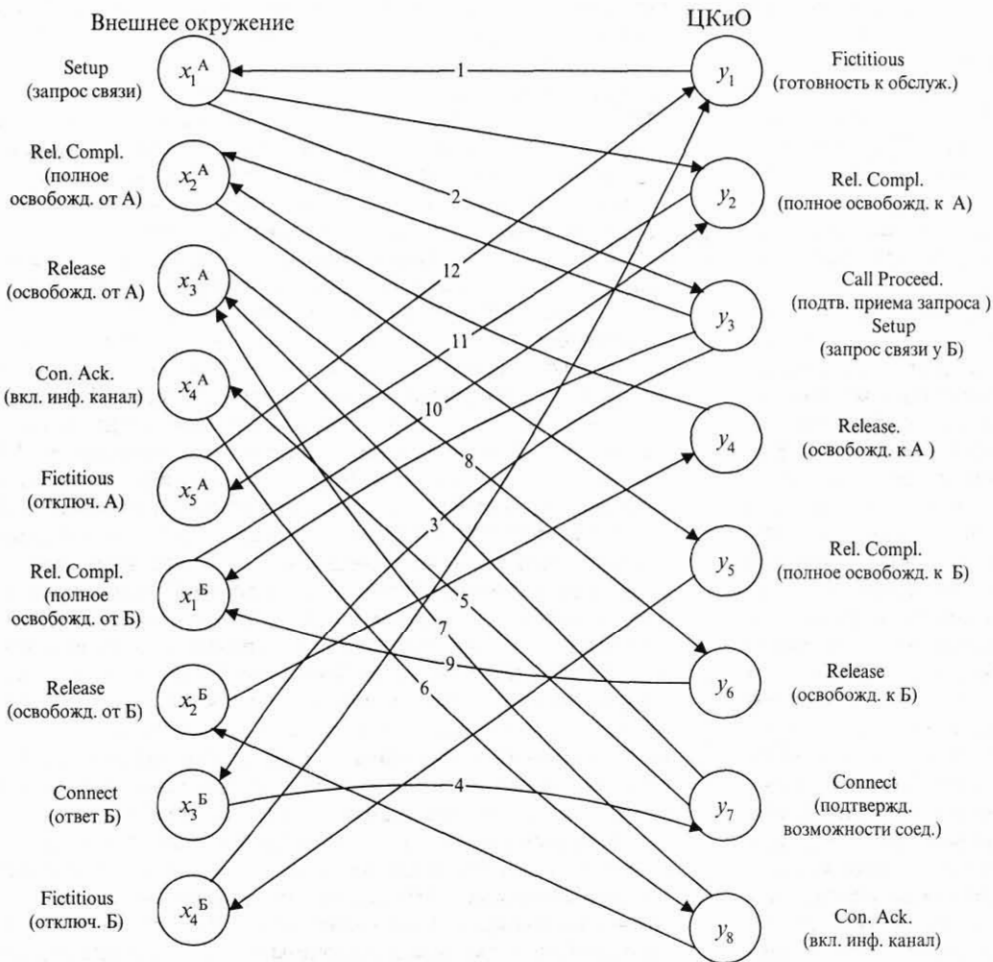
1. Он является деревом и содержит две доли вершин. Первую долю составляют вершины $a \in A$, сопоставленные этапам устойчивого состояния, вторую — вершины $b \in B$, сопоставленные этапам переходного состояния ПФ.

2. Дуги графа соответствуют входным и выходным сообщениям, которыми обмениваются центры и их ВО в соответствии с протоколами сигнализации. Дугам, исходящим из вершин типа a и входящим в вершины типа b , сопоставляются входные сообщения $x \in X$, а дугам, исхо-

дящим из вершин типа b и входящим в вершины типа a , — выходные сообщения $y \in Y$.

3. На первом уровне графа расположена корневая вершина a_0 , соответствующая начальному этапу устойчивого состояния. Количество дуг, исходящих из вершины a_0 , равно числу видов информационных услуг, предоставляемых ЦОУ, причем переход по d -й дуге ($d = 1, 2, \dots, D_0^a$) соответствует началу процесса предоставления d -й услуги.

4. На втором уровне расположены вершины типа b . Они соответствуют начальным этапам переходного состояния, на которых реализуются транзакции, связанные с обработкой начальных входных сообщений, инициирующих разные виды информационных услуг.



■ Рис. 2. Граф обмена в интерфейсе UNI

5. На третьем, пятом и т. д. уровнях графа расположены вершины типа a . Количество дуг, исходящих из вершины a_i определяется числом различных входных сообщений, которые могут поступить от элемента управления на i -м этапе устойчивого состояния в соответствующем состоянии ВО. В общем случае в число этих входных сообщений включается также сигнал срабатывания выдержки времени, которая устанавливается на i -м этапе и ограничивает время реакции элемента управления на выработанное выходное сообщение.

6. На четвертом, шестом и т. д. уровнях графа расположены вершины типа b . Количество дуг, исходящих из вершины b_j зависит от числа возможных исходов процесса выполнения транзакции на i -м этапе переходного состояния. Переход по одной из дуг соответствует «нормальному» протеканию процесса предоставления информационной услуги, переход по другим дугам — блокированию (прерыванию) процесса информационного обслуживания по различным причинам.

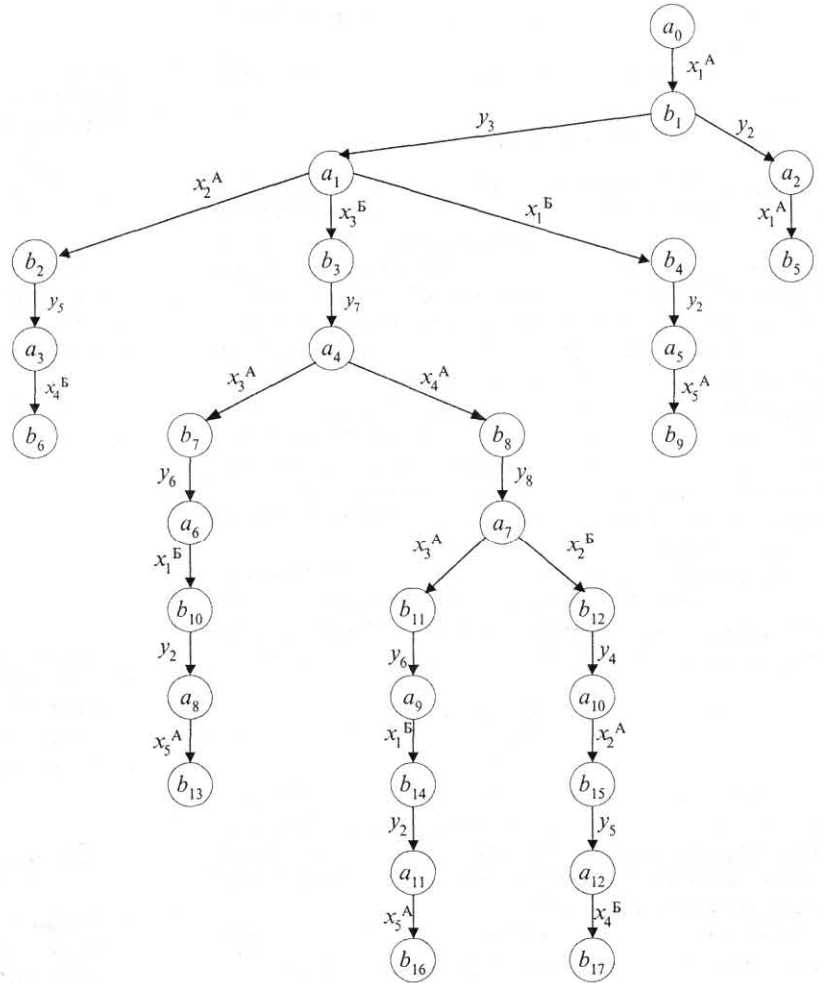
7. Конечными (висячими) вершинами графа являются вершины типа b . Транзакции, выполняемые на этих этапах, завершают процесс предоставления информационной услуги. Граф $G_{ПФ}$ построен таким образом, что каждая висячая вершина однозначно соответствует одному из возможных исходов процесса информационного обслуживания. Количество разных исходов этого процесса равно количеству висячих вершин графа.

Обозначим множество висячих вершин графа $G_{ПФ}$ через W . Это множество можно разделить на подмножества вершин по их принадлежности к тому или иному виду информационной услуги. Обозначим W^d подмножество висячих вершин, соответствующих d -й информационной услуге. Для оценки ВВХ подмножество W^d , в свою очередь, может быть разбито на подмножества:

$$W^d = W_1^d \cup W_2^d \cup W_{6л}^d;$$

$$W_1^d \cap W_2^d \cap W_{6л}^d = \emptyset,$$

где W_1^d — подмножество вершин, соответствующих успешному завершению процесса предоставления информационной услуги d -го вида; W_2^d — подмножество вершин, достижение которых соответствует от-



■ Рис. 3. Граф установления внутреннего соединения в ЦОУ широкополосной цифровой сети

казу элемента управления от обслуживания «по собственной инициативе»; $W_{6л}^d$ — подмножество вершин, соответствующих потере (блокированию) процесса обслуживания по разным причинам.

Блокирование процесса обслуживания может произойти как во внешнем окружении ЦОУ, так и в его исполнительной и управляющей системах из-за перегрузки ресурсов, сигнала срабатывания таймера, переполнения буферов различных устройств. В таблице сгруппированы обозначения для подмножеств висячих вершин, соответствующих блокировке

■ Таблица

Причина блокировки	Система, в которой произошла блокировка			
	Внешнее окружение	Исполнительная система	Управляющая система	
			Обработка данных	Доставка данных
Перезагрузка ресурсов	W_3^d	W_4^d	-	-
Срабатывание таймера	W_5^d	-	W_6^d	-
Переполнение буфера	-	W_7^d	W_8^d	W_9^d

П р и м е ч а н и е. W_1^d и W_2^d — множества вершин, которые соответствуют нормальному завершению обслуживания и отказу пользователя.

обслуживания в подсистемах ЦОУ по всем перечисленным причинам.

С учетом этих обозначений для подмножества $W_{6л}^d$ справедливо

$$W_{6л}^d = W_3^d \cup W_4^d \cup W_5^d \cup W_6^d \cup W_7^d \cup W_8^d \cup W_9^d,$$

где $W_3^d - W_9^d$ — подмножества вершин, которые завершают процесс обслуживания при d -й информационной услуге блокировкой соответственно из-за занятости вызываемого пользователя; из-за перегрузки ресурсов исполнительской системы, обслуживающих заявки по системе с потерями; из-за сигнала срабатывания таймера на этапе устойчивого состояния; из-за срабатывания таймера при выполнении программного модуля в управляющем устройстве; из-за переполнения входных буферов ресурсов исполнительской системы, которые обслуживают заявки по системе с ожиданием и, наконец, из-за переполнения входных буферов управляющих устройств и системы передачи данных управления.

Обозначим $p_{a,j}^b$ вероятности достижения требованиями в процессе обслуживания вершин b_j графа $G_d^{ПФ} \subseteq G_{ПФ}$. Тогда с учетом свойств графа для информационной услуги d -го вида получаем соотношения

$$P_{n,d}^{ЦкиО} = \frac{1}{P_{d,1}^b} \sum_{j|b_j \in W_n^d} p_{a,j}^b, n=1,2,\dots,9,$$

где индекс n задает номер исхода процесса обслуживания в соответствии с данными таблицы, и суммирование осуществляется по тем конечным вершинам, достижение которых соответствует n -му исходу информационной услуги d -го вида.

Объединение подмножеств $W_n^d (n=1,2,\dots,9)$ позволяет определить подмножества вершин $W_{ВО}^d, W_{ОбрС}^d, W_{УпрС}^d, W_{ИспС}^d$, достижение которых означает потерю (блокирование) требований во внешнем окружении ЦОУ, в обрабатывающей, управляющей и исполнительской системах соответственно:

$$W_{ВО}^d = W_2^d \cup W_3^d \cup W_5^d;$$

$$W_{ОбрС}^d = W_6^d \cup W_8^d;$$

$$W_{УпрС}^d = W_{ОбрС}^d \cup W_9^d;$$

$$W_{ИспС}^d = W_4^d \cup W_7^d.$$

В соответствии с этим для вероятностей потери заявок из-за блокировок требований во внешнем окружении ЦОУ, в его управляющей и исполнительской системах, а также в системе обработки имеем

$$P_{д}^{ВО} = P_{2,d}^{ЦкиО} + P_{3,d}^{ЦкиО} + P_{5,d}^{ЦкиО};$$

$$P_{д}^{ОбрС} = P_{6,d}^{ЦкиО} + P_{8,d}^{ЦкиО};$$

$$P_{д}^{УпрС} = P_{д}^{ОбрС} + P_{9,d}^{ЦкиО};$$

$$P_{д}^{ИспС} = P_{4,d}^{ЦкиО} + P_{7,d}^{ЦкиО},$$

а вероятность блокировки произвольного требования в процессе предоставления информационной услуги d -го вида равна

$$P_{д}^{6л} = P_{2,d}^{ЦкиО} + P_{3,d}^{ЦкиО} + P_{4,d}^{ЦкиО} + P_{5,d}^{ЦкиО} + P_{6,d}^{ЦкиО} + P_{7,d}^{ЦкиО} + P_{8,d}^{ЦкиО} + P_{9,d}^{ЦкиО}.$$

Путем объединения подмножеств W_n^d разных видов информационных услуг можно получить подмножество всяких вершин, соответствующих всевозможным исходам процесса обслуживания заявок для ЦОУ в целом. Поэтому вероятности различных исходов процесса предоставления информационной услуги для ЦОУ в целом равны:

$$P_n^{ЦкиО} = \sum_{d=1}^{D_0^a} P_{d,1}^b P_{n,d}^{ЦкиО};$$

$$\sum_{n=1}^9 P_n^{ЦкиО} = 1;$$

$$P_{\#\#}^{ЦкиО} = \sum_{d=1}^{D_0^a} P_{d,1}^b P_{\#\#}^d.$$

Древовидный нагруженный граф $G_{ПФ}$ полно и однозначно задает описание ПФ ЦОУ при предоставлении всевозможных информационных услуг, располагается на четвертой страте иерархии моделей ЦОУ (см. рис. 1) и позволяет определить ВВХ ПФ.

Литература

1. Колбанев М. О., Яковлев С. А. Модели и методы оценки характеристик обработки информации в интеллектуальных сетях связи. — СПб.: Изд-во СПб ГУ, 2002. — 230 с.
2. Колбанев М. О. Модели и методы оценки вероятностно-временных характеристик центров коммутации и обработки информации сетей связи. — СПб.: Изд-во ГУТ им. проф. М. А. Бонч-Бруевича. — Ч. 1. — 2001. — 81 с.
3. Колбанев М. О. Графическая модель процесса взаимодействия центров коммутации и обработки информации с внешним окружением // Тр. учеб. завед. связи. — 2001. — № 167. — С. 7-17.
4. Колбанев М. О., Кузьма Э. А., Беляева Е. В. Модель транзакции центров коммутации и обработки информации // Междунар. науч.-технич. конф. «Информатизация процессов формирования открытых систем на основе СУБД, САПР, АСНИ и систем искусственного интеллекта». — Вологда, 2001. — С.51-56.
5. Колбанев М. О. Анализ задачи и методов исследования вероятностно-временных характеристик центров коммутации и обработки информации интеллектуальных сетей связи // Информационно-управляющие системы. — 2003. — №1. — С. 11-25.

УДК 681.391.1

РАЗГРАНИЧЕНИЕ ДОСТУПА К РЕСУРСАМ В СИСТЕМАХ КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ

И. Л. Ерош,

д-р техн. наук, профессор

Санкт-Петербургский государственный университет
аэрокосмического приборостроения (ГУАП)

В статье рассмотрена задача на разграничение доступа в системах коллективного пользования, основанная на булевых преобразованиях двоичных последовательностей. Для обеспечения доступа используется метод сравнения запроса с ответом. Приведены примеры построения булевых функций, применение которых позволяет гибко менять доступ для каждой группы пользователей в зависимости от выбираемых запросов.

In the article task on differentiation of access in shared systems based on Boolean transformations of binary sequences is considered. For providing of access the method of comparison of inquiry with the answer is used. The examples of construction Boolean of functions are given, in which access for each group of the users flexibly varies depending on chosen inquiries.

Введение

При работе нескольких пользователей в компьютерной сети часто требуется ограничивать доступ некоторых пользователей к ресурсам системы. Рассмотрим задачу на разграничение доступа, которую условно назовем «Солдаты и офицеры». Пусть требуется снабдить каждого солдата, офицера и генерала одним электронным ключом, причем таким, чтобы солдат своим ключом мог открыть все помещения, предназначенные для солдат, офицер мог попасть во все солдатские и в офицерские помещения, а генерал мог попасть во все помещения, предназначенные для солдат и офицеров, а также в свой генеральский кабинет. При этом солдат своим ключом не может открыть помещения, предназначенные для офицеров и генералов, а офицер не может открыть генеральский кабинет.

Для обеспечения доступа будем использовать пароли, основанные на булевых преобразованиях, описанных в работах [1, 2]. Напомним общую идею этого метода. Пусть имеются две двоичные последовательности $A(a_1, a_2, \dots, a_n)$ и $B(b_1, b_2, \dots, b_n)$. При этом $a_i, b_i \in \{0, 1\}, i = 1, 2, 3, \dots, n$. Элементы последовательности B будем получать преобразованием последовательности A булевой функцией F , при этом b_i будет результатом булевого преобразования, зависящего от a_i и некоторых элементов из его окружения. Например, если $A = 10110001101$ и выбрана функция $F = 10^{*-1} \vee 2$, то это означает, что каждый элемент b_i последовательности B равен $b_i = |a_i \& a_{i-1} \vee a_{i+2}$. В данном примере из последовательности A будет получена последовательность $B = 11001110110$.

В настоящей статье по набору запросов и ответов определим соответствующие функции F , которые позволят выполнить требуемые преобразования.

Примеры решения задачи доступа

Пример 1. Для доступа в помещения требуется сравнение с известным паролем обработанного (ответного) кода на запрос. Иными словами, для случайного кода запроса, булева функция ключа выработывает пароль на открывание помещения. Если пароль совпадает с требуемым для данного помещения, оно открывается, если нет — не открывается. Таким образом, нужно найти три разные булевы функции (для солдат, офицеров и генералов), такие, чтобы на запросы при доступе к солдатским помещениям все функции выработывали один и тот же код, соответствующий коду открывания солдатских помещений. При требовании доступа к офицерским помещениям солдатские ключи должны выработывать ложный код, а офицерские и генеральские — требуемый для открывания помещений. При доступе к генеральскому кабинету солдатские и офицерские ключи должны выработывать ложные коды и только генеральский — требуемый для открывания.

Для примера выберем коды для открывания в следующем виде: 4С — для солдатских (двоичный код 01001100); D8 — для офицерских (11011000) и E5 — для генеральских (11100101).

Перед доступом в помещение вырабатывается случайный запрос. Для упрощения примера выберем по два байтовых кода для случайного запроса,

например, в таком виде: А6, С9 — перед входом в солдатские помещения; 7Е, В5 — перед входом в офицерские и 6D, 4А — перед входом в генеральский кабинет.

Используя методику, описанную в работе [1], получим один из вариантов функций, обеспечивающих ответы на запросы: $F_1 = 3^*0^* \neg 2 \vee 1^*0^* \neg 3 \vee 2^*1^*0^* \neg 1 \vee 1^*0^* \neg 3$ — функция солдата; $F_2 = 3^*1^*0^* \neg 1^* \neg 2 \vee 2^*1^*0^* \neg 1 \vee 2^*0^* \neg 1^* \neg 2 \vee 3^*1^*0^* \neg 1 \vee 1^*0^* \neg 2^* \neg 3 \vee 1^*0^* \neg 1^* \neg 3 \vee 2^*0^* \neg 1 \vee 1^*0^* \neg 1^* \neg 3$ — функция офицера; $F_3 = 1^*0^* \neg 2^* \neg 3 \vee 2^*1^*0^* \neg 2 \vee 3^*1^*0^* \neg 1^* \neg 2 \vee 2^*1^*0^* \neg 1 \vee 2^*1^* \neg 1^* \neg 2^* \neg 3 \vee 3^*1^*0^* \neg 1 \vee 2^*0^* \neg 1^* \neg 2 \vee 3^*1^*0^* \neg 1 \vee 2^*0^* \neg 1 \vee 1^*0^* \neg 2^* \neg 3$ — функция генерала.

Число различных случайных запросов может быть выбрано сколь угодно большим. По два запроса на каждом пункте проверки выбрано лишь для простоты иллюстрации излагаемого материала. Результаты представлены в табл. 1.

В данном примере для прохода требовалось совпадение с паролем ответного кода на запрос. Это означает, что система проверки просто сравнивает ответ с паролем. Система может быть усложнена, если на каждом уровне проверки ввести функцию, которая будет проверять соответствие запроса ответу.

Особенность рассмотренной системы заключается в том, что при доступе в солдатские помещения солдат, офицер и генерал неразличимы, при проходе в офицерские — неразличимы офицер и генерал. Иными словами, системе присуща неотслеживаемость доступа через системы проверки.

Пример 2. Задача доступа решается совместно с проверкой «информированности» каждой группы пользователей. Например, при проходе в солдатские помещения задаются k_1 вопросов и сверяются с правильными ответами, при проходе в офицерские и генеральское помещения, соответственно, k_2 и k_3 вопросов. В общем случае число вопросов может быть произвольным, и они могут периодически меняться. Для рассматриваемого примера возьмем $k_1 = 2, k_2 = 3, k_3 = 4$. Вопросы и правильные ответы выберем следующим образом.

Для солдатского контроля: 6С → 18; D2 → 3Е.

Для офицерского контроля: D5 → С6; 47 → В3; 9А → 67.

Для генеральского контроля: 58 → DA; EC → A3; B4 → 7C; 91 → 66.

Как и в предыдущей задаче, генерал должен правильно ответить на все запросы, офицер оши-

Таблица 1. Результаты ответов на запросы, сформированные для солдата, офицера и генерала

Запрос	Солдатская функция F_1	Офицерская функция F_2	Генеральская функция F_3	Комментарий
А6 С6	4С 4С	4С 4С	4С 4С	Все ответили верно
7Е В5	ЕС СО	D8 D8	D8 D8	
CD 4А	Е8 Е4	FE Е4	E5 E5	Верно ответил только генерал

Примечание. Здесь и далее в таблицах жирным шрифтом отмечены неверные ответы.

бается в ответах только на генеральские запросы, солдат ошибается в ответах как на офицерские, так и на генеральские запросы. Заметим, что запросы и соответствующие им ответы выбираются случайным образом и периодически могут меняться.

По методике, изложенной в работах [1, 2], получим функции: для солдат $F_1 = 1^*0^* \neg 1 \vee 2^*0^* \neg 1 \vee 1^* \neg 2$; для офицеров $F_2 = 1^* \neg 1^* \neg 2 \vee 3^*1^* \neg 1 \vee 2^*1^*0^* \neg 1^* \neg 3 \vee 1^*1^*0^* \neg 1^* \neg 3 \vee 2^*1^*0^* \neg 1 \vee 2^*1^*0^* \vee 0^* \neg 1^* \neg 2 \vee 5^*1^*0^* \vee 4^* \neg 1 \vee 2^*1^* \neg 1^* \neg 3$; для генералов $F_3 = 2^*1^*0^* \neg 1 \vee 4^*0^* \neg 1 \vee 2^*0^* \neg 1^* \neg 3 \vee 3^*2^*1^*0^* \vee \neg 1^* \neg 2^* \neg 6 \vee \neg 3^* \neg 7 \vee 5^*2^*1^* \neg 1^* \neg 2 \vee 2^*1^* \neg 1^* \neg 4 \vee 1^*0^* \neg 3^* \neg 6 \vee 1^*0^* \neg 1^* \neg 4 \vee 3^*1^*0^* \neg 1 \vee 5^*1^*0^* \vee 4^*1^* \neg 1 \vee 2^* \neg 2^* \neg 3 \vee 4^*1^*0^* \neg 1^* \neg 3 \vee 1^*0^* \neg 2^* \neg 4$.

Результаты ответов на запросы показаны в табл. 2. Из таблицы видно, что система функционирует правильно. Предположим, что злоумышленник овладел информацией как о запросах, так и ответах на них на всех пунктах пропуска. При этом будем считать, что задавались только первые запросы из каждой группы, а именно: 6С → 18 — при солдатском контроле; D5 → С6 — при офицерском контроле; 58 → DA — при генеральском контроле.

Если злоумышленник владеет способом построения булевых функций, обеспечивающих такое преобразование, а это естественно предположить, то он может найти, например, такую функцию:

$$F_{\text{злоум}} = 2^*1^*0^* \neg 1 \vee 2^*0^* \neg 1^* \neg 4 \vee 1^*0^* \neg 1^* \neg 6 \vee 4^*1^*0^* \vee 3^*2^*0^* \neg 1 \vee 2^*0^* \neg 1^* \neg 3 \vee 3^*2^*1^* \vee 2^*0^* \neg 2$$

С ее помощью злоумышленник на первые запросы из каждой группы ответит верно, поскольку из этих данных и была сформирована функция $F_{\text{злоум}}$. На вторые запросы из каждой группы злоумышленник ответит неверно: D2 → E4; 47 → 97; EC → D8 и не пройдет проверку ни на одном пункте контроля.

При нахождении булевой функции в таблицу истинности можно ввести специальную строку, отмечающую время нахождения функции, и по тестовому запросу этой строки определять, когда был получен доступ к данным ресурсам.

Пример 3. Рассмотрим задачу распределения доступа, в котором при ответах на некоторые запросы преимущество имеет генерал, при ответах на другие запросы нет преимуществ ни у кого, а при ответе на третью группу запросов преимущество имеет солдат.

Таблица 2. Результаты ответов на запросы, сформированные для солдата, офицера и генерала, с проверкой «информированности»

Запрос	Солдатская функция F_1	Офицерская функция F_2	Генеральская функция F_3	Комментарий
6С	18	18	18	Все ответили верно
D2	3Е	3Е	3Е	
D5	35	C6	C6	Солдат ответил неверно
47	40	B3	B3	
9А	Е6	67	67	Верно ответил только генерал
58	70	F0	DA	
EC	18	F8	A3	
B4	ЕС	ЕЕ	7C	
91	F1	76	66	

Выберем восемь пар «запрос—ответ» и найдем функции F_1 , F_2 и F_3 такие, чтобы обеспечить сформулированные выше требования.

D2 → 3E — все должны ответить правильно.

47 → B3 — правильно должны ответить офицер и генерал.

EC → A3 — правильно должен ответить только генерал.

В этой первой группе запросов преимущество имеет генерал.

5E → 6A — правильно отвечает только солдат.

F2 → 3D — правильно отвечает только офицер.

B4 → 9C — правильно отвечает только генерал.

В этой второй группе запросов преимуществ нет ни у кого, и каждый отвечает правильно только при доступе в свое помещение.

C5 → 39 — правильно отвечает солдат и офицер.

82 → CD — правильно отвечает только солдат.

В этой третьей группе запросов преимущество имеет солдат, и он может пройти как в офицерские помещения, так и в генеральский кабинет. Если к третьей группе запросов добавить пару D2 → 3E, то офицер сможет пройти в свое помещение и в генеральский кабинет, а генерал — только в свой кабинет.

Булевы функции для солдата (F_1), офицера (F_2) и генерала (F_3) имеют следующий вид:

$$F_1 = [2^*]1^*0^*] - 1^*] - 6 \vee [1^*]0^* - 1^*] - 4 \vee \\ \vee 2^*]0 \vee [3^*]1^*]0^*] - 1 \vee 1^* - 3 \vee [1^*]0^* \\ * - 2 \vee 3^*]1^*0 \vee 1^*]0^* - 2 \vee 0^* - 3;$$

$$F_2 = -2 \vee [2^*]0^*] - 1 \vee [1^*]0^* - 5 \vee 3^* - 1 \vee [2^*]1^*0^* \\ * - 1 \vee 2^*]0^* - 1 \vee 0^* - 3^*] - 4;$$

$$F_3 = 2^*0^* - 3 \vee [2^*]1^*]0 \vee 2^*0^*] - 1^*] - 4 \vee [1^* - 1^* \\ * - 2 \vee [1^*] - 1^* - 3 \vee 3^*]0^* - 1 \vee 2^*]1^*]0 \vee [3^*]1^*0^* - 1.$$

Результаты проверки приведены в табл. 3.

Последний пример показывает, что, не меняя функций доступа у солдат, офицеров и генералов, можно только сменой наборов «запрос—ответ» гиб-

■ **Таблица 3.** Результаты ответов на запросы, сформированные для солдата, офицера и генерала при наличии преимуществ

Запрос	Солдатская функция F_1	Офицерская функция F_2	Генеральская функция F_3	Комментарий
D2	3E	3E	3E	Все ответили верно
47	F9	B3	B3	Солдат ответил неверно
EC	3C	7B	A3	Верно ответил только генерал
5E	6A	BF	FF	Верно ответил только солдат
F2	1E	3D	FE	Верно ответил только офицер
B4	D4	7D	9C	Верно ответил только генерал
C5	39	39	3A	Верно ответил солдат и офицер
82	CD	34	1C	Верно ответил только солдат

ко менять права доступа в различные помещения. Если есть требование, чтобы в помещение имели доступ одновременно несколько человек (например, трое: солдат, офицер и генерал), то можно выбрать группу запросов, аналогичную второй группе рассмотренного примера, и одновременно предъявить все три функции. В этом случае на каждый запрос правильно ответит хотя бы один и помещение станет доступно.

Пример 4. В данном примере рассмотрим пороговую задачу доступа из N доверенных лиц при пороге h . Такая задача может решаться при любом числе N и любом пороге h [3]. Пусть, например, N офицеров получили фрагменты ключа от генеральского кабинета. При этом требуется обеспечить возможность доступа в кабинет лишь в случае, когда соберутся вместе любые h офицеров из N . Если соберется меньше, то открыть его они не могут. Из [3] следует, что если число офицеров, имеющих фрагменты ключа, равно N , и требуется обеспечить порог доступа h , то можно построить таблицу равновесных кодов длины N веса $d = N - h + 1$ и распределять пары «запрос—ответ» в соответствии с таблицей кодов.

Число таких кодов определяется как $P(h - 1, N - h + 1) = \frac{N!}{(h - 1)!(N - h + 1)!}$.

Для $N = 5$ и $h = 3$ равновесные коды длины 5 веса $d = 3$ для десяти пар «запрос—ответ» приведены в табл. 4.

Значение «1» в таблице для каждого офицера означает, что он правильно отвечает на соответствующие запросы. Из табл. 4 следует, что, если соберутся вместе любые три офицера, они правильно ответят на все десять запросов, если их будет меньше трех? они не смогут правильно ответить на все запросы.

Выберем произвольные пары «запрос—ответ» и найдем булевы функции для каждого из пяти офицеров.

$$F_1 = [2^*]1^*]0^*] - 1 \vee [2^*0^* - 1^*] - 4 \vee [1^*]0^*] - 1^*] - 6 \vee \\ \vee 3^*]1^*] - 1 \vee 2^* - 1^* - 4 \vee [2^*]1^*]0^*] - 1^* - 3 \vee \\ \vee 3^*]1^*]0 \vee 5^*]1^*0 \vee 4^* - 1 \vee 2^*]1^*] - 1^*] - 3 \vee \\ \vee [3^*]2^*]1 \vee 2^*]0^* - 2;$$

$$F_2 = 2^*]0^* - 1^*] - 2 \vee 2^*0^* - 1^* - 3 \vee [3^*]2^*]1^*]0 \vee - 1^* - 2^* \\ * - 6 \vee - 3^* - 7 \vee 5^*]1^*0 \vee 4^*]1^* - 1 \vee 2^* - 2^*] - 3 \vee [4^*]1^*0^* \\ *] - 1^* - 3 \vee 1^*0^* - 2^*] - 4;$$

$$F_3 = [3^*]2^*]1^*] - 2 \vee 2^*]1^*]0 \vee [3^*0^* - 1^*] - 3 \vee 2^*]1^*0^* \\ *] - 4 \vee - 1^* - 2^* - 6 \vee - 3^* - 5^* - 7 \vee 3^*]2^*]1^*] - 1 \vee 1^* - 2^* \\ *] - 3 \vee [3^*]2^*0^* - 1 \vee 0^* - 2^*] - 4 \vee [0^* - 1^*] - 2^*] - 3;$$

■ **Таблица 4.** Равновесные коды для $N = 5$ и $h = 3$

Номера запросов	Номера офицеров				
	1	2	3	4	5
1	1	1	1	0	0
2	1	1	0	1	0
3	1	1	0	0	1
4	1	0	1	1	0
5	1	0	1	0	1
6	1	0	0	1	1
7	0	1	1	1	0
8	0	1	1	0	1
9	0	1	0	1	1
10	0	0	1	1	1

■ **Таблица 5.** Ответы пяти офицеров на запросы системы доступа

Запрос	Правильный ответ	1-й офицер	2-й офицер	3-й офицер	4-й офицер	5-й офицер
6C	18	18	18	18	03	FD
D2	3E	3E	3E	7D	3E	F6
D5	C6	C6	C6	72	3F	C6
47	B3	B3	B1	B3	B3	F7
9A	EE	EE	67	EE	46	EE
58	DA	DA	30	30	DA	DA
EC	A3	D8	A3	A3	A3	AD
B4	7C	CD	7C	7C	BD	7C
91	66	BE	66	EE	66	66
CC	AE	F8	E3	AE	AE	AE

$$F_4 = \{3^* \} 2^* 1^* \} - 2\sqrt{2^* \} 1^* \} 0\sqrt{\} 3^* \} 2^* 0^* - 1^* \} - 3\sqrt{\} 1^* \} 0^* - 1^* \} - 4\sqrt{\} 2^* \} 1^* \} 0\sqrt{2^* 0^* \} - 1^* \} - 4\sqrt{\} 0^* \} - 5\sqrt{0^* - 2^* \} - 3\sqrt{0^* - 1^* - 3^* \} - 4\sqrt{\} 4^* \} 1^* 0^* \} - 1^* - 3;$$

$$F_5 = \{3^* 0^* - 1\sqrt{\} 1^* \} 0^* - 1^* \} - 4\sqrt{4^* \} 0\sqrt{2^* 0^* - 2^* \} - 3\sqrt{\} 5^* 2^* 0^* \} - 2\sqrt{-3^* - 5\sqrt{4^* \} - 1^* \} - 3\sqrt{1^* \} - 1^* \} - 3\sqrt{2^* \} - 2^* \} - 3\sqrt{\} 3^* \} 2^* 1^* .$$

Результаты ответов на запросы системы доступа представлены в табл. 5.

Любые три столбца в приведенной таблице содержат в любой строке хотя бы один правильный ответ. Для любых пар столбцов найдется хотя бы один запрос, на который оба офицера ответят неверно.

Влияние длины последовательностей на сложность реализации

В приводимых примерах обрабатывались запросы длиной в 8 бит. Не представляет труда построить функцию и для более длинных запросов. Пусть, например, запросы и ответы имеют длину в 128 бит и имеют вид в шестнадцатеричном коде:

A63CBE9F91BCDA6385BDF2C917ACEBD2 — запрос;
A8BD72910CFDE893DFCBAE7619DEFBA6 — ответ.
Функция *F*, преобразующая *A* в *B*, имеет вид:

$$F = \{3^* 2^* 1^* \} 0^* - 2\sqrt{5^* 2^* \} 1^* \} 0^* \} - 1^* \} - 3\sqrt{\} 2^* \} 1^* 0^* - 1^* \} - 4\sqrt{3^* \} 2^* 1^* \} 0^* - 2\sqrt{3^* 1^* \} 0^* \} - 1^* \} - 3\sqrt{2^* \} 1^* 0^* - 1^* - 4\sqrt{4^* 0^* \} - 1^* - 2^* \} - 4\sqrt{\} 3^* \} 2^* 1^* 0^* - 1\sqrt{4^* - 2^* \} 1^* \} 0^* \} - 2\sqrt{3^* \} 1^* \} - 1^* \} - 3\sqrt{2^* 0^* \} - 1^* - 3^* \} - 5\sqrt{3^* 1^* 0^* \} - 2^* \} - 3\sqrt{\} 6^* \} 1^* \} 0^* - 1^* - 2\sqrt{\} 2^* \} 1^* \} 0^* - 1^* \} - 3\sqrt{2^* 1^* \} 0^* - 1^* \} - 4\sqrt{3^* \} 2^* 0^* - 2^* \} - 3\sqrt{1^* 0^* \} - 1^* \} - 2^* - 4\sqrt{\} 4^* 2^* 0^* \} - 2\sqrt{\} 1^* 0^* - 1^* \} - 2^* - 3\sqrt{2^* \} 1^* \} - 1^* \} - 3\sqrt{\} 2^* \} 0^* \} - 1^* - 2^* \} - 4\sqrt{\} 2^* \} 0^* \} - 1^* \} - 3\sqrt{4^* \} 3^* \} 1^* 0^* - 2\sqrt{\} 3^* 1^* \} 0^* - 1^* \} - 2^* \} - 3\sqrt{4^* 2^* 0^* \} - 1^* - 2\sqrt{0^* - 1^* \} - 2^* - 5\sqrt{1^* \} 0^* - 2^* \} - 5\sqrt{\} 4^* \} 3^* \} 1^* \} - 1.$$

Число разных аргументов функции *F* в данном примере равно 12. Эксперименты показывают, что число используемых аргументов функции возрастает значительно медленнее длины последовательности *n*. При *n* = 256 их количество составляет примерно 10 %, т. е. математическое ожидание числа аргументов равно 25.

Заключение

Приведенные примеры иллюстрируют метод распределенного доступа в системах коллективного пользования, который может быть применен для управления доступом к вычислительным ресурсам компьютерной сети, для обеспечения доступа с ограничениями в помещения и др.

Литература

1. **Ерош И. Л.** Дискретная математика. Булевы функции, комбинационные схемы, преобразование двоичных последовательностей. Учебное пособие. — СПб: СПбГУАП, 2001. — 38 с.
2. **Ерош И. Л.** Защита информационных потоков в системах распределенного контроля и управления // Информационно-управляющие системы. — 2002. — № 1. — С. 40-46.
3. **Ерош И. Л.** Система передачи данных с закрытыми ключами // В кн.: Информационные системы в промышленности и экономике / Под ред. И. Л. Ероша СПб, 1999. — С. 114-116.

ИНФОРМАЦИОННЫЕ ПРОБЛЕМЫ ОРГАНИЗАЦИИ ОБРАТНОЙ СВЯЗИ ПРИ ВЗАИМОДЕЙСТВИИ БИОЛОГИЧЕСКИХ И ТЕХНИЧЕСКИХ СИСТЕМ

А. О. Поляков,

д-р техн. наук, профессор

Санкт-Петербургский институт информатики и автоматизации РАН

П. Т. Тукабаев,

канд. мед. наук

Новороссийский филиал Современного гуманитарного института

Рассмотрена возможность новых подходов к организации обратной связи при взаимодействии биологических («биосоциальных») и технических компонент высокого уровня организационной сложности. В основу исследования положена необходимость контекстно-зависимого и структурного представления систем для повышения организации уровня управления. Указывается возможность и целесообразность введения понятия биотехнической обратной связи как совокупности сигнального и структурного обмена информацией между компонентами сложной системы.

It is proposed to consider a possibility of developing new approaches to arranging for a feedback at the interaction between biological («bio-social») and technical components of high organizational complexity. The research is based upon the necessity to implement context-dependent and structural representations of the systems to receive a new higher level in control organization. Possibility and advantage of introducing new notions for biotechnical feedback as an aggregate of signal and structural information exchange between the complex system components are indicated.

Введение

Исследования сложных систем, начиная еще со времен Л. фон Берталанфи [1], а также работы современных выдающихся ученых ([2] и др.) прямо указывают на необходимость расширения понимания «живого» на небиологические системы высших порядков сложности. Такое утверждение справедливо распространить и на «разнокомпонентные» системы. Это означает, что для сложных систем уровня «живого» или уровня «симбиоза» живого и технического требуется управление, точно соответствующее тому, что существует в открытых живых системах — взаимодействие на основе контекстно-зависимого языка, принципиально недоступное для представления в фон Неймановской машине без серьезных семантических потерь. Например, по [2], это общество и город или, скажем, население и обслуживающая его медицинская клиника, или даже отдельный человек и обслуживающая его информационно-техническая система.

Задача повышения уровня представления взаимодействующих компонент систем такого рода является одной из важнейших в развитии совре-

менного расширенного понимания биотехнических систем. Исходя из сказанного, предлагается обсуждение проблем организации обратной связи, связанной с решением задач повышения уровня управления в сопряженных «неоднородных по материалу» системах при взаимодействии биологических («биосоциальных») и технических компонент высокого уровня организационной сложности.

Биологическая обратная связь

Прямой перенос методов управления, разработанных для замкнутого представления технических систем, на системы открытые, активно взаимодействующие с внешним миром, порождает множество проблем, решать которые до сих пор пытаются все большим усложнением постановок задач оптимизационного управления. Но того, что необходимо для системы технической и программной, — системы формальной, удовлетворяющей условиям ее реализации с использованием автоматов ($U_{\text{внеш}} = 0$), никогда не будет достаточно для системы природной (неформальной по сущности самой Приро-

ды [$U_{\text{внеш}} = U(p)$, где p — априорно неизвестная вероятность произвольного внешнего воздействия на систему]. Основная причина всех проблем — замкнутость представлений, обеспечиваемых математическим аппаратом, контекстной независимостью его языка.

Термин «управление» обычно используется одновременно в двух смыслах: 1) управление как достаточно свободная организационная деятельность, направленная на достижение определенных

целей: $X \xrightarrow{U_{\text{орг}}} X_{\text{целевое}}$; 2) управление в смысле использования конкретного управляющего воздействия, т. е. некоторой физической величины, изменение которой производится по алгоритму дости-

жения априорно поставленной цели: $X \xrightarrow{U_{\text{алг}}} X_{\text{опт}}$. Обратим внимание на то, что для организационного целевого управления требуются решения в открытой постановке, а аппарат выработки управляющего воздействия декларируется в подходе замкнутом, обеспечиваемом алгоритмически. Кроме учета этой коллизии для нашего исследования требуется еще и некоторое новое специфическое понимание управления, учитывающее особенности симбиоза систем с принципиально разными свойствами, разными собственными языками представления:

$$(X_{\text{биол}} \Leftrightarrow X_{\text{техн}})_{t1} \xrightarrow{U} (X_{\text{биол}} \Leftrightarrow X_{\text{техн}})_{t2}. \quad (1)$$

Управление в технических системах всегда понималось как выражение воли стороннего для нее управленца, воли, на которую накладывались лишь технические ограничения управляемой системы. Однако при любых попытках построения технических систем с биологической компонентой на управление накладываются еще и не технические ограничения. Речь идет о предварительной самостоятельной оценке биологической компонентой целесообразности управляющего воздействия. Тогда управление U в (1) представляется в виде

$$U = F(U_{\text{внеш}}, (X_{\text{биол}} \Leftrightarrow X_{\text{техн}})_{t1}). \quad (2)$$

Такого рода оценка не может быть выражена математически (без учета контекстной зависимости восприятия управления) и сегодня наиболее реальным методом ее учета (хотя и не вполне адекватным для случая взаимодействия сложных систем) является введение известного понятия биологической обратной связи (БОС).

При этом под активной БОС чаще всего понимается обучение человека управлению собственными физиологическими процессами посредством их косвенной визуализации с использованием технических приборов или иных средств. В работах же с тематикой биотехнического характера делается все возможное для сведения проблемы к выделению управляющего физического сигнала с последующим взаимодействием на уровне сигнальной обратной связи. Все это совершенно справедливо, но, по сути рассматриваемой проблемы,

недостаточно. Не всегда действие БОС можно свести к текущему обучению, часто это выработка отложенного действия. Не всегда сигнальной обратной связи достаточно для полного представления изменения состояний компонентов в системах, требующих их восприятия на уровне «живого». Известные определения БОС не отвечают реалиям функционирования систем рассматриваемого нами уровня. В них не хватает учета выраженного влияния контекстно-зависимого аспекта взаимодействия биологического («биосоциального») и технического компонентов системы.

Необходимость новой постановки исследований для биотехнических систем (как, впрочем, и осознания их на новом уровне сложности) в некоторой степени следует и из работ Е. П. Попечителя [4], впервые указавшего на необходимость глубинного осмысления взаимосвязи множества различных представлений систем, возникающего при использовании различных математических аппаратов, различных приборов и наборов признаков для их описания. Тем более это важно во взаимодействии сложных разнокомпонентных по представлению и материалу «живых» системах.

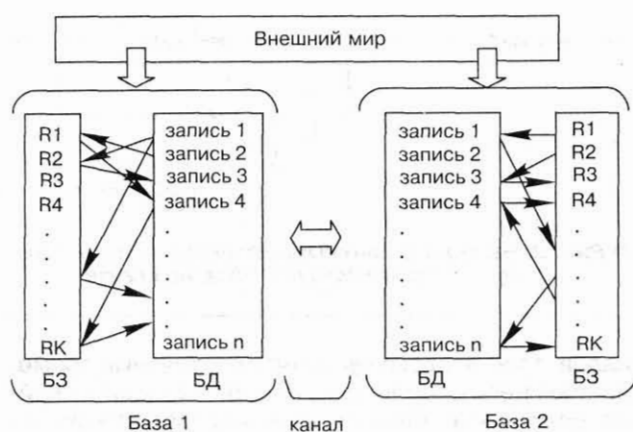
Управление в многокомпонентных «живых» системах

В настоящее время при создании биотехнических систем, составной частью которых является человек (биотехническая компонента), все реалии управления сводятся к необходимости принятия решений за счет использования систем с базами знания. Однако априорно проектируемая структура всех современных баз не позволяет учитывать в управлении требования (2). Поэтому нас здесь интересует только направление, связанное с организацией управления в сложных системах, ориентированное на специальные конфигурации баз [6–8], для которых структура хранимых записей является функцией входного информационного потока $S_{\text{хр}} = f(R, A, B)$, где A, B — данные, факты, сведения, ситуации, а R — текущие отношения между ними (текущее знание о взаимодействии). Отметим, что такие базы могут характеризовать как биологическую, так и техническую составляющую сложных систем. Такого рода построения обеспечивают минимизацию семантических потерь и выполнение условия (2) по критерию

$$S_{\text{хр}} \xrightarrow{t} S_{\text{ест}} \text{ при } f(R, A, B) \rightarrow \psi, \quad (3)$$

где ψ — предельно достижимые возможности представления естественного языка в машине фон Неймана (например, с использованием формально-логического или сетевого его представления).

В соответствии с (3) структурное управление (управление на уровне переопределения во времени знания R) должно предопределять возможность равнопонятого (воспринимаемого в одном контексте) двум и более системам использования



■ Рис. 1. Двухбазовый вариант актуализации хранимых структур

входной информации, на языке, способном нести и передавать данные и знание о предметной области. Достаточно ясно, что предложенный подход на уровне возможностей машины фон Неймана реализует известные гипотезы об управлении на основе использования контекстно-зависимых языков [2].

В [8] показано, что предложенное определение знания как структуры связи данных, которую можно наблюдать, изучать и изменять, приводит к инженерно понятному подходу к работе со знанием, к системам «с управлением на структурах». В [9 и др.] на основе (3) предлагается двухкомпонентная конструкция (рис. 1). Такая двухбазовая архитектура функционирует как система управления структурами представления взаимодействующих подсистем по условиям (1), (2) и (3) как

$$S_{\text{База 1}} \xleftrightarrow{U_t} S_{\text{База 1}}, \text{ где } U_t \equiv U(2) \quad (4)$$

и имеет возможность хранения в реляционном виде всех необходимых сведений — данных (включая числовые и текстовые) и связей между ними с выполнением (2) по критерию (3).

Структурная парадигма управления и структурная обратная связь

Введенное понятие сдвоенной базы в составе взаимодействующих баз данных и знания является конструктивным, ибо в такой формулировке обеспечивается введение и формирование новой парадигмы управления [9]. Классическая парадигма управления («сигнальная» парадигма) происходит от понятия управляющего сигнала, выражаемого в аналоговой или цифровой форме, контроля достигнутого результата управления и его коррекции по отклонению за счет системы управления с обратной связью. Представим ее в виде

$$y = F\left(\sum x_{\text{вх}}, x_{\text{ос}}\right), \quad (5)$$

что допустимо для систем, описательная сложность которых допускает их функциональное представление. При этом мы считаем, что в некоторой области допустимого управления передаточные функции системы обладают, по крайней мере, непрерывностью, т. е. имеет смысл понятие «величины сигнала управления».

Сохранение этой парадигмы для системно-сложных объектов автоматически сводит нас к уровню их кибернетической модели (модели, не создающей в себе индивидуальный информационный образ системы с изменяемой структурой связей). Системный взгляд указывает, что для разрешения этих проблем надо перейти к другой парадигме управления, которую здесь мы условно запишем сравнительно с (5) как

$$y = F\left(YS_{\text{База 1}}^{\text{База 1}} \xleftrightarrow{U_t} S_{\text{База 1}}^{\text{База 1}} \middle| K\right), \quad (6)$$

где K — выбранный критерий взаимного преобразования (контекстного «доверия») взаимодействующих баз. Такую парадигму предложено называть парадигмой согласования структур данных или парадигмой структурного управления [9] (структурной обратной связи). Итак, обратная связь, как сигнальный способ коррекции передаточной функции, в системах, требующих для своего представления естественного языка, не существует в связи с отсутствием выраженной передаточной функции. Ее место занимает структура данных и аппарат согласования структур данных.

Отметим, что здесь управление, как стабилизация состояния системы, сводится к удержанию ее на так называемом гомеокинетическом плато [7 и др.]. Некоторые ученые уже давно отмечали ограниченность систем управления с обратной связью как формы управления [10]. В нашем случае ее место занимает аппарат согласования структур данных.

Фундаментальное отличие указанных парадигм заключается отнюдь не только в способах выработки обратной связи — измерение отклонения выходного сигнала или оценку текущей структуры системы. Ясно, что второй вариант эквивалентен учету «собственной точки зрения» активной системы, а это означает выбор для постановки исследовательской работы системного подхода, концептуального понимания замкнутости сигнальной и системной открытости структурной парадигмы.

Пример сигнальной обратной связи в биологических системах

Достаточно понятно, что «внутри» любой реальной системы вполне можно считать существующей обыкновенную сигнальную обратную связь, обеспечивающую выполнение «механистических» функций отдельных подсистем биологической или биотехнической системы. Роль БОС здесь сводится, в конечном счете, к выработке механизма достижения тех или иных конечных результатов для одной или группы отдельных подсистем.

Но и даже просто с сигнальной обратной связью в биологических системах все обстоит несколько сложнее, чем в технических. Понятие сигнальной обратной связи, конечно, может быть продуктивно использовано при исследовании конкретных биологических систем в самых простых случаях, но часто эта связь существует в «усложненном» виде. Например, на основе введения в рассмотрение «сдвоенного кольца обратной связи» [11], становится ясен и без дофаминовой теории механизм болезни Паркинсона, зачастую преследующей спортсменов высшей квалификации после ухода из спорта.

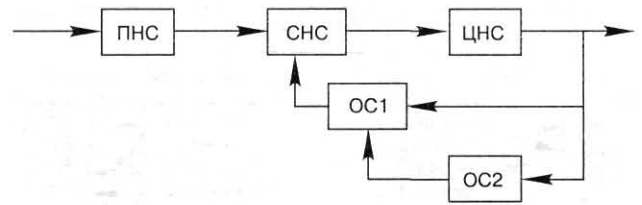
Можно утверждать, что нервная система человека в своей макроструктуре содержит минимум три функционально и морфологически отделенные компоненты — периферическую (ПНС), стволую (СНС) и центральную (ЦНС) нервную системы, т. е. три подсистемы и, по крайней мере, два кольца обратной связи (рис. 2).

Человек воспринимается как механическое устройство, в котором надо выработать относительно других спортсменов лучшее движение по параметрам скорости и траектории достижения цели. Фактически это прямой системотехнический подход к человеку в чисто техническом понимании этого подхода. Единственное отличие — механизмы выработки конечных результатов реализованы на уровне живого, что и позволяет нам говорить здесь о биотехнической проблеме биологической системы.

Тренировка (проводимая на уровне совершенствования технических возможностей биологической системы) состоит в том, чтобы подготавливая образом построить набор операторов обратной связи (ОС1 и ОС2) так, чтобы сложные связки движений, выполняемые в «нормальном режиме» с помощью осознанного конструирования из траекторий различных динамических систем, объединились в одно движение.

Это возможно только благодаря очень точной настройке ОС2, вовремя (в нужный момент движения, управляемого изначально ОС1) переключающей операторы ОС1. Фактически ОС1 вводится в «ненормальный» режим обратной связи, на которую наложено влияние ОС2 (управляющей другой составляющей того же сложного движения). При резком снижении тренировочных нагрузок нарушается точность настройки ОС2, однако ОС1 уже настроена на работу в режиме «переключения от ОС2» и если «алгоритм переключения» нарушен, то динамическая система становится неустойчивой даже в пределах простой траектории движения.

Излечить такое «техническое» заболевание в биологической системе практически невозможно, для этого требуется вернуть в прежнее состояние обратные связи, реализованные после огромных затрат организма на их создание в неестественном для него виде, как «настройки биологического механизма». Иначе надо резко замедлить «внутрисистемное время», подавив активность нейронов и снизив тем самым скорость распространения сиг-



■ Рис. 2. Выработка сложных рефлексов на уровне возможностей сигнальной обратной связи

налов, то есть резко увеличить постоянные времени всех регуляторов; фактически, «задать» нейрореплетиками активность, ставшую избыточной. Конечно, подобная подмена функционального назначения ОС может быть вызвана и патогенными причинами, и тогда остается только насильственное подавление активности элементов и подсистем биохимическими средствами.

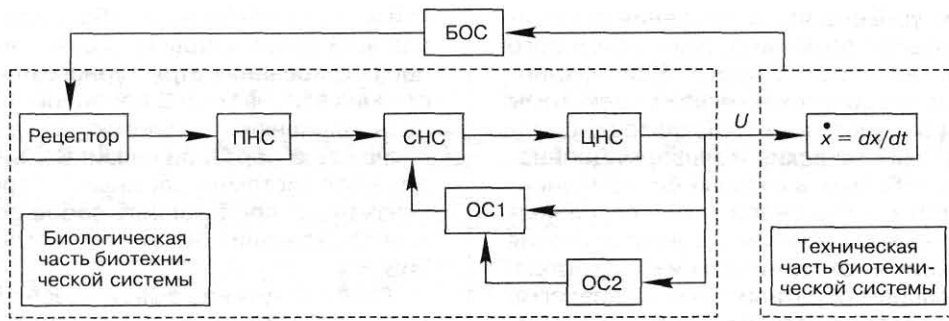
В принципе, мы пришли здесь к проблеме реального использования тренировок на основе использования активной БОС, к реализации тех самых рекомендаций, которые иногда предлагаются в литературе как «положительные перспективы». Далеко не всегда последствия могут быть столь разрушительными, но один из вариантов подстерегающей нас опасности прямого использования БОС для «внутренних тренировок» достаточно ясен.

Биотехническая обратная связь

Вернемся к реалиям существования явления, названного БОС, и оценим правомочность и целесообразность его использования для представления биотехнических систем в нашем расширенном, многокомпонентном их понимании. Рассмотрим некоторые возможности повышения уровня информационного обмена биологических и технических компонент таких систем за счет использования рассмотренного выше подхода. Начнем с того, что, исходя из смысла и содержания традиционного определения БОС, можно дать достаточно подробную схематическую иллюстрацию биотехнической системы, разделив биологическую и техническую части (рис. 3). Отметим, что выделенное в биологической части множество сигнальных обратных связей, взятое из рис. 2, может быть другим и является лишь иллюстрацией их принципиального наличия.

Однако целостное понимание всей цепи БОС здесь остается на уровне ее замыкания от технической части к центральной нервной системе на сигнальном уровне через рецепторы, ибо, в традиционном понимании БОС, контролировать можно только изменение измеряемых величин. Соответственно уровень связи двух составных частей биотехнической системы остается в такой цепи в информационном смысле даже ниже фактографического.

При таком подходе биологической части реально доступна для восприятия техническая часть си-

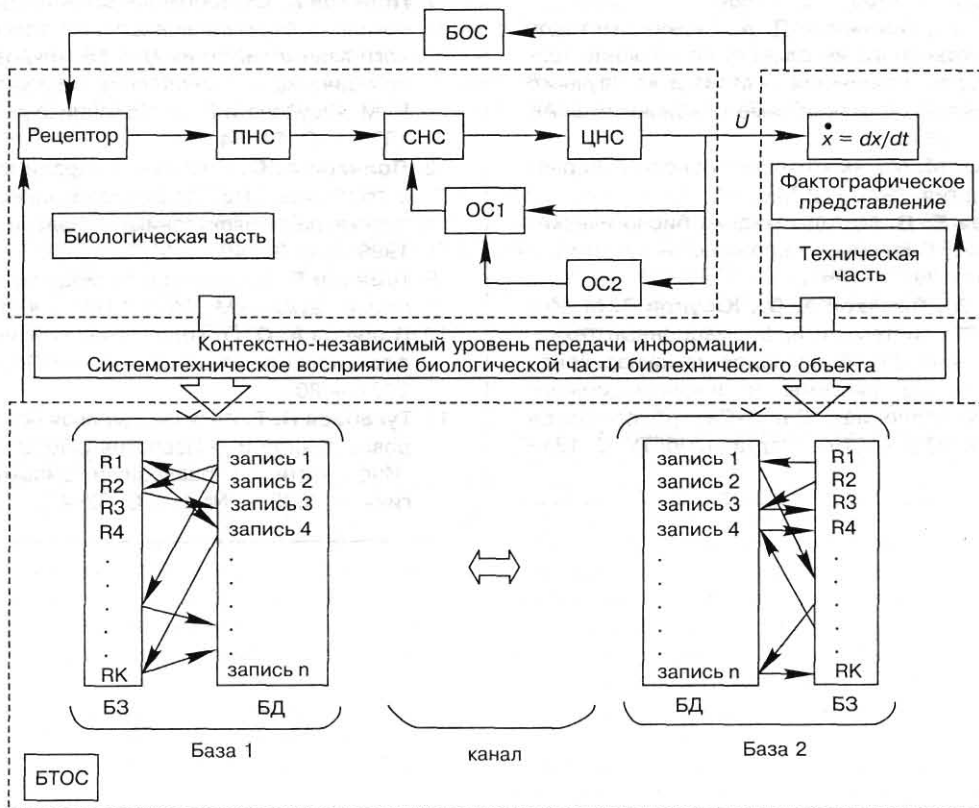


■ Рис. 3. Организация обратных связей в биотехнической системе на уровне традиционного понимания БОС

стемы лишь в объеме ее функциональной (модельной) представимости сигналами, доступными рецепторам. В этом смысле правомерно утверждать, что технической части доступны лишь измеримые на ее входе сигналы u , выработанные в активной биологической части системы, под действием которых техническая часть изменяет свое состояние. Характер происходящих при этом процессов определяется скоростью изменения переменной состояния объекта $\dot{x} = dx/dt$, которая представляет собой многомерную величину $x = (x_1, \dots, x_N)$, где x_1, \dots, x_N — скорости изменения компонент многомерной переменной x . Если бы мы могли на том

же языке адекватно описать биологическую часть биотехнической системы, то многие проблемы кибернетического представления такого комплекса были бы решены.

Однако искать совместимость собственных языков двух частей биотехнической системы вполне можно не на кибернетическом, а, например, на фактографическом уровне представления знания во исполнение положений (4) и (6), т. е. на уровне языка, гораздо более близкого к естественному, что вполне соответствует все той же гипотезе [2] об адекватной работе со сложными объектами (рис. 4). Здесь БОС в традиционном понимании не исчезает, но дополняется ее языковым представ-



■ Рис. 4. Организация биотехнической обратной связи на основе двухбазовой архитектуры хранения информации

лением, причем уровень языка общения определяется предельными возможностями текстового представления технической части. Все механистические составляющие технической подсистемы могут и должны образовывать БОС по традиционной схеме, а прогностические и информационно-аналитические проблемы, в которых биокомпонента не может обойтись без своей технической части, реализуются через сравнение структур описания «от биологической» и «от технической» подсистем, т. е. с использованием обратной связи в парадигме структурного управления. Исходя из общей терминологии, такую конструкцию можно назвать биотехнической обратной связью (БТОС), обеспечивающей взаимодействие, стремящееся к минимизации семантических потерь при реализации контекстно-зависимого управления на контекстно-независимом автомате.

Обратим внимание на конфигурацию БТОС (см. рис. 4). Реально она представляет собой две обратные связи, выполненные по рассмотренному выше структурному принципу. Каждая связь обеспечивает работу своей подсистемы — биологичес-

кой и технической, но их общность на уровне двухбазовой архитектуры позволяет вести непрерывное согласование структурного представления состояния этих частей. В принципе, здесь можно увидеть внешнюю аналогию с двумя кольцами сигнальной обратной связи. Весьма вероятно, что для всех системно-сложных построений обратная связь будет представлять собой достаточно сложные образования, но это тема отдельной публикации.

Главный для нас вывод здесь достаточно понятен — в биотехнической системе надо искать способ сведения языка биологической части к языку, доступному технике, но напротив, делать все возможное, чтобы системы, существующие в неразрывном симбиозе (простое общение человека, например, с компьютером здесь интересно в основном при организации автоматических аналитических систем [11]), общались на не программистском языке возможно высокого уровня, не отрицая и их сигнальное взаимодействие там, где такой язык оказывается достаточно экономным и представительным.

Литература

1. **Берталанфи Л. фон.** Общая теория систем: Критический обзор // Исследования по общей теории систем. М.: Прогресс, 1969. — С. 23–82.
2. **Поспелов Г. С., Поспелов Д. А.** Влияние методов теории искусственного интеллекта на решение традиционных задач управления. — М.: Изд-во Научного совета по комплексной проблеме «Кибернетика» АН СССР, 1977. — 36 с.
3. **Коршунов Ю. М.** Математические основы кибернетики. М.: Энергия, 1972. — 176 с.
4. **Попечителев Е. П.** Методы медико-биологических исследований. Системные аспекты. — Житомир: Изд-во ЖИТИ, 1997. — 186 с.
5. **Ерофеев А. А., Поляков А. О., Юсупов Р. М.** Интеллектуальные системы и информационные технологии управления // В сб. докладов междунар. конф. «Интеллектуальные системы и информационные технологии управления», Санкт-Петербург—Псков 19–23 июня 2000 г. — СПб.: Изд-во СПбГТУ, С. 131–134.
6. **Ерофеев А. А., Поляков А. О.** Интеллектуальные системы управления. — СПб.: Изд-во СПбГТУ, 1999. — 256 с.
7. **Поляков А. О.** Проблемы организации межбазового обмена информацией в условиях телеконференциального взаимодействия // В сб. трудов семинара «Телемедицина — становление и развитие» / Под ред. Р. М. Юсупова и Р. И. Полонникова. — СПб.: Омега, 1999. — С. 61–64.
8. **Поляков А. О.** Системная парадигма управления // В сб.: Труды СПбГТУ: Вычислительная техника, автоматика, радиоэлектроника. — Издание СПбГТУ № 472, 1998. — С. 32–49.
9. **Шеннон Р.** Имитационное моделирование — искусство и наука. — М.: Мир, 1978. — 418 с.
10. **Поляков А. О.** От количественной информации к информодинамической машине. — СПб.: Изд-во СПбГТУ, 2001. — 80 с.
11. **Тукабаев П. Т.** Информационная безопасность и здоровье личности // Известия СПбГЭТУ (ЛЭТИ), серия «Информатика, управление и компьютерные технологии». — 2002. — № 2. — С. 32–40.

УДК 612(075)

АВТОМАТИЗИРОВАННАЯ СИСТЕМА ИЗМЕРЕНИЯ ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК ПАРАМЕТРОВ ИЗОБРАЖЕНИЯ ГАЗОРАЗРЯДНОГО СВЕЧЕНИЯ

К. Г. Коротков,

д-р техн. наук, профессор

Э. В. Крыжановский,

аспирант

Д. И. Муромцев,

аспирант

М. А. Бабицкий,

аспирант

Санкт-Петербургский институт точной механики и оптики

М. Б. Борисова,

аспирант

Санкт-Петербургский государственный университет

Статья посвящена описанию автоматизированной системы для изучения изменений во времени характеристик изображений газового разряда, получаемых в электромагнитном поле высокой напряженности вокруг капель жидкости. Для описания газоразрядных изображений (ГРВ-грамм) предложено использовать площадь свечения, среднюю интенсивность свечения, средний радиус, нормализованное среднеквадратическое отклонение радиуса и коэффициент формы. Выявлены статически значимые различия между образцами жидкостей, что не обнаруживалось при изучении их одиночных (статических) ГРВ-грамм. Построена математическая модель для описания поведения трендов параметров ГРВ-грамм.

This article devoted to the study of changing of the characteristics of gas discharge glow images (GDV images) around the drop of liquid in the strong electromagnetic field. Several parameters to describe the GDV images were proposed: Glow Area, Average Intensity, etc. This approach was named "Dynamic GDV-graphy". The statistically significant differences were revealed between different samples of liquids in the case when conventional "static" GDV-graphy was unable to reveal ones. The mathematical model of behavior of the time series trends of the GDV-parameters was suggested.

Введение

Свечение газового разряда, возникающего вблизи поверхности исследуемого объекта при помещении последнего в электрическое поле высокой напряженности, известно уже более двух столетий. В современной научной практике исследование этого явления получило название газоразрядная визуализация (ГРВ) [1]. В настоящее время для регистрации и анализа газоразрядных изображений применяется аппаратура с микропроцессорным управлением и цифровой оптикой, а также вычислительная техника. Разработаны приборы и инструментальные средства для исследования различных физических и биологических объектов.

Сущность процесса визуализации можно описать следующей теоретической схемой. Первичным являет-

ся процесс взаимодействия магнитного поля с объектом исследования, в результате которого при определенной напряженности поля возникает эмиссия с поверхности объекта заряженных частиц, участвующих в ионизации начальных фаз газового разряда. Газовый разряд, в свою очередь, может влиять на состояние объекта, вызывая вторичные эмиссионные, деструктивные и тепловые процессы. Неоднородность поверхности и объема, процессы эмиссии заряженных частиц или выделения газов оказывают влияние на параметры электромагнитного поля, за счет чего изменяются параметры газового разряда, а именно ток разряда и оптическое излучение. При этом основная информация извлекается из характеристик получаемого видеоизображения газоразрядного свечения, которое представляет собой пространственно распределенную группу участков различной яркости.

На сегодняшний день практически значимый и воспроизводимый результат получен в исследовании жидкофазных объектов.

Проблема выявления индивидуальных различий химически близких жидкостей, а также водных растворов слабой концентрации актуальна для медицины, биологии и косметологии [2]. В [3, 4] была разработана методика исследования жидкостей путем изучения характеристик газового разряда вокруг капель. В этих работах было показано, что растворы сильных электролитов, таких как NaCl, KCl, NaNO₃, KNO₃, имеют различия по параметрам газоразрядных изображений (ГРВ-грамм), как между соседними концентрациями одного раствора, так и между одинаковыми концентрациями различных растворов.

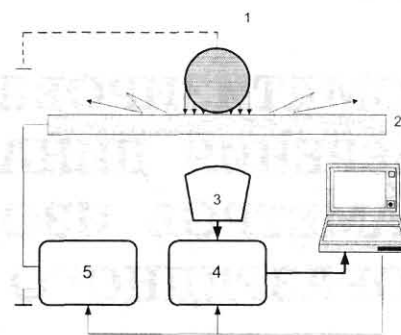
В данной статье показано, что при экспонировании капель проводящих (сильные электролиты) и слабо проводящих (эфирные масла) жидкостей в электромагнитном поле (ЭМП) в течение 2–10 с наблюдается закономерная динамика изменений характеристик газоразрядного свечения. Исследование характеристик динамических рядов ГРВ-изображений различных жидкостей показало их значимость для выявления индивидуальных различий химически близких жидкостей, как проводящих, так и слабо проводящих. Это позволяет ввести понятие динамической ГРВ-график как метода исследования жидкостей путем изучения изменения характеристик газоразрядного свечения вокруг капель жидкости, т. е. изменения параметров ГРВ-грамм во времени.

Методика эксперимента

Методика газоразрядного исследования объектов основана на включении объекта, расположенного на прозрачном электроде, покрытом диэлектриком, в цепь прибора, и получении, таким образом, газоразрядных изображений. На рис. 1 схематично изображены прибор и объект исследования. На исследуемый объект 1 и диэлектрическую пластину 2 подаются импульсы напряжения от генератора электромагнитного поля 5, для чего на обратную сторону пластины нанесено прозрачное токопроводящее покрытие. При высокой напряженности поля в газовой среде пространства контакта объекта и пластины развивается лавинный разряд и/или скользящий газовый разряд, параметры которого определяются свойствами объекта исследования. Свечение разряда с помощью оптической системы и ПЗС-камеры преобразуется в видеосигналы, которые через цифровой видеопреобразователь поступают на вход компьютера для дальнейшей обработки.

Цифровые видеосигналы записываются на жесткий диск компьютера в виде видеофайлов или отдельных растровых изображений. Типичный вид газоразрядных изображений для капель жидкостей и металлического объекта цилиндрической формы представлен на рис. 2.

Спектр изображений газоразрядного свечения объектов стандартной формы имеет достаточно стабильный вид. На рис. 3 представлен пример спектра ГРВ-изображения металлического объекта цилиндрической формы. Для таких спектров характерно нали-

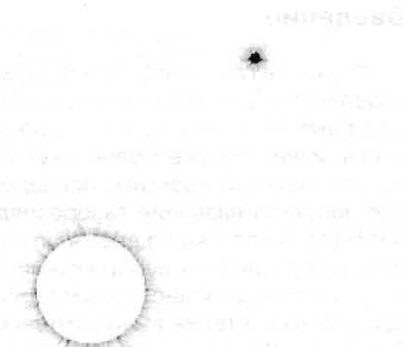


■ Рис. 1. Схематическое изображение прибора для исследования характеристик ГРВ-изображений: 1 — объект исследования; 2 — прозрачный электрод, покрытый диэлектриком; 3 — оптическая система; 4 — цифровой видеопреобразователь; 5 — электронные блоки

чие пика в темной области и резкое снижение количества точек более высокой интенсивности. Такая форма спектра у всех ГРВ-изображений позволила разработать алгоритм автоматического выявления значимого сигнала из исходного, зашумленного изображения, получаемого с прибора.

Стандартная последовательность получения, обработки и анализа динамических ГРВ-изображений включает ряд этапов, показанных на рис. 4. Каждый кадр динамического ГРВ-изображения перед началом обработки приводится к формату 8 бит/пиксель с палитрой в градациях серого. Таким образом, кадр представляет собой матрицу $F[w \times h]$, где w — размер кадра по горизонтали (ширина), h — размер кадра по вертикали (высота). Каждый элемент матрицы $F[x, y]$ — это интенсивность свечения от 0 (отсутствие свечения) до 255 (максимальная яркость свечения) в точке с координатами (x, y) , где $x \in [0, w]$, $y \in [0, h]$.

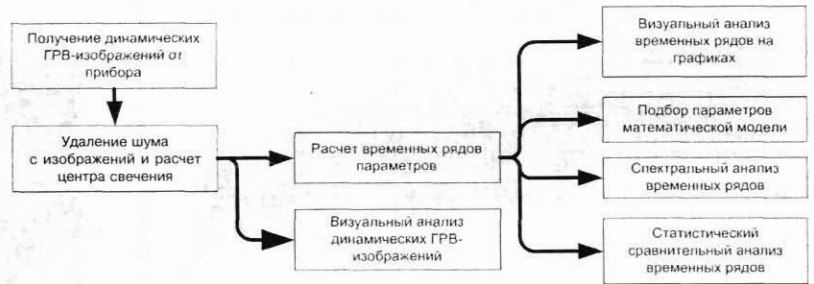
Перед расчетом параметров отдельных кадров производится фильтрация шума в следующей последовательности:



■ Рис. 2. Типичные ГРВ-изображения для капель дистиллированной воды, раствора NaCl и металлического объекта



■ Рис. 3. Спектр ГРВ-изображения



■ Рис. 4. Последовательность обработки и анализа динамических ГРВ-изображений

- 1) удаление всех точек изображения, интенсивность которых меньше или равна заданному уровню шума;
- 2) удаление фрагментов свечения, площадь которых строго меньше заданной минимальной площади фрагмента;
- 3) удаление всех точек изображения, расстояние от которых до центра свечения строго больше заданного рабочего радиуса.

После проведения фильтрации выполняется расчет следующих числовых характеристик свечения:

- 1) площади свечения S количества точек с ненулевой интенсивностью;
- 2) средней интенсивности свечения \bar{I} — средней интенсивности свечения всех точек с ненулевой интенсивностью;
- 3) среднего радиуса \bar{R} ;
- 4) нормализованного среднеквадратического отклонения радиуса $\sigma_{\bar{R}}$;
- 5) коэффициента формы K_f .

Для расчета трех последних характеристик строится функция $R(\alpha)$ следующим образом. Из центра свечения проводятся лучи с углом наклона $\alpha \in [0^\circ; 360^\circ)$, с таким шагом, чтобы покрывалась вся площадь свечения. По каждому лучу определяется положение двух точек (фактически расстояние от центра) с интенсивностью, на большей уровня шума — самой близкой к центру свечения (R_{\min}) и наиболее удаленной от него (R_{\max}). Затем вычисляется функция $R(\alpha) = R_{\max} - R_{\min}$. Таким образом, получаемая функция отражает характер изменения внешнего контура свечения в зависимости от угла. Средний радиус \bar{R} является математическим ожиданием значений функции $R(\alpha)$. Нормализованное отклонение вычисляется как $\sigma_{\bar{R}} = \frac{\sigma_R}{\bar{R}}$, где σ_R — среднеквадратическое отклонение функции $R(\alpha)$.

Коэффициент формы $K_f = \frac{L}{2\pi\bar{R}}$, где L — длина кривой функции $R(\alpha)$. K_f отражает степень «изрезанности» контура свечения: $K_f = 1$ — соответствует идеальной круглой форме свечения; чем больше значение K_f , тем больше степень «изрезанности» контура. Каждый из перечисленных параметров может быть рассчитан не только для всего изображения, но и для отдельных секторов (с центром в центре свечения). Из рассчитанных параметров отдельных кадров формируются временные ряды, которые описывают динамику изменения каждого параметра во времени.

мируются временные ряды, которые описывают динамику изменения каждого параметра во времени.

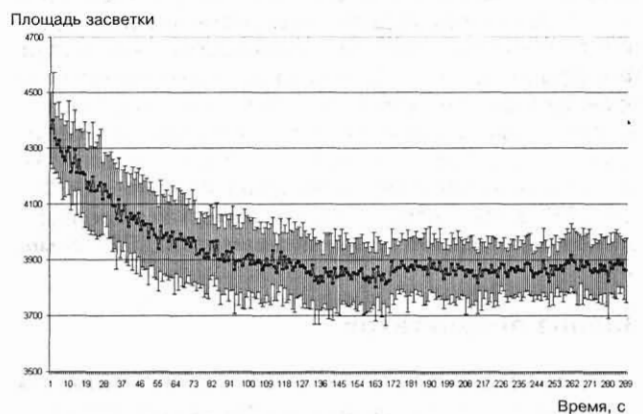
Описанная выше последовательность обработки динамических изображений реализована авторами в комплексе специализированных компьютерных программ.

Методика и цели исследования

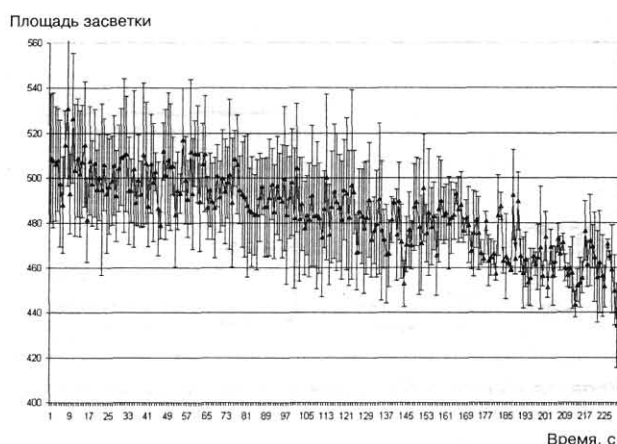
Целью настоящего анализа явилось следующее.

1. Исследование с помощью динамической ГРВ-графии особенностей характеристик процесса ГРВ проводящих жидкостей (сильных электролитов) во времени. Это дополняет результаты статистического анализа параметров одиночных ГРВ-изображений [3].

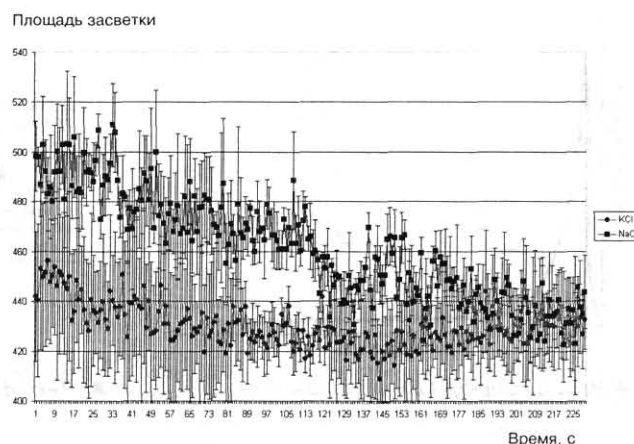
2. Исследование особенностей динамики ГРВ-грамм слабо проводящих жидкостей, что позволяет выявлять различия между образцами, имеющими близкий химический состав. В исследовании слабо проводящие жидкости были представлены маслами натурального и синтетического происхождения и органического происхождения, а также маслами, имеющими различия в оптической активности, месте получения и способе обработки. Всего рассматривалось 42 пары масел.



■ Рис. 5. Зависимость площади засветки тест-объекта от времени



■ Рис. 6. Зависимость площади засветки ГРВ раствора NaCl от времени



■ Рис. 7. Зависимость площади засветки от времени для растворов NaCl и KCl

Воспроизводимость данной методики была продемонстрирована на примерах исследования тест-объекта — сплава из титана, имеющего форму цилиндра, и раствора KCl. Для каждого объекта было сделано по 30 измерений. На рис. 5 и 6 видно, что погрешность измерения в каждой точке временных серий составляет не более 5 %.

Изучаемые в настоящей работе временные ряды параметров ГРВ изображений представляют собой изменения параметра площади засветки и средней интенсивности изображения во времени. Данные параметры были выбраны в силу наименьших значений погрешностей, при нахождении среднего значения временного ряда для выборки измерений одного исследуемого объекта, а также наибольшей чувствительности к свойствам исследуемого объекта при разрешении 160×120 пикселей. Получение изображений указанного разрешения диктовалось возможностями технического обеспечения.

В силу малой частоты дискретизации (30 Гц) по сравнению с частотой подаваемых на объект импульсов (1000 Гц) для выявления значимых различий временных рядов характеристик ГРВ-грамм жидкостей были исследованы детерминированные составляющие тренды (или только низкочастотные составляющие) методом сравнения усредненных временных серий по соответствующим выборкам исследуемых жидкостей. Фильтрация шума и сглаживание временного ряда проводились методом скользящего среднего. При одновременном анализе нескольких временных серий случайного процесса для средних значений в различные моменты времени использовался критерий Фишера [5]. Была построена математическая модель для описания поведения трендов параметров ГРВ-грамм.

Анализ результатов

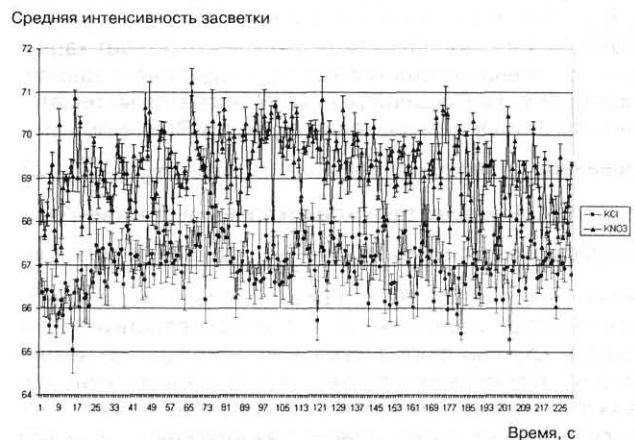
Анализ данных по исследованию сильных электролитов. Как показано в работе [3], параметры ГРВ-грамм сильных электролитов имеют различия не при всех концентрациях растворов. При одноно-

мальных разведениях растворов NaCl и KCl соответствующие средние по выборкам одиночных ГРВ-грамм не имеют статистически значимых различий.

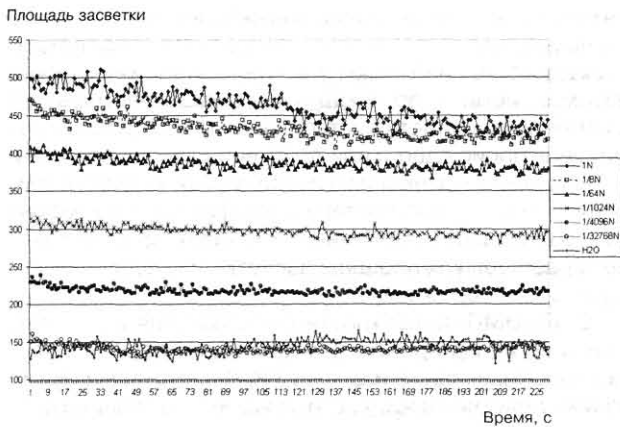
Анализ соответствующих детерминированных составляющих временных серий площадей засветок ГРВ-грамм показал, что данные растворы имеют статистически значимые различия, которые начинают проявляться примерно после третьей секунды начала воздействия ЭМП (рис. 7).

В работе [3] при сравнении однонормальных растворов KCl и KNO_3 не было выявлено статистически значимых различий по выборке одиночных ГРВ-грамм. Не видно это различие и при сравнении соответствующих временных серий площади засветки. Однако различия наблюдаются при рассмотрении динамики другого параметра — средней интенсивности засветки (рис. 8).

Вид трендов для высоких концентраций растворов электролитов принципиально отличается от вида тренда для дистиллированной воды. В первом случае тренд имеет вид монотонно убывающей функции,



■ Рис. 8. Зависимость средней интенсивности засветки ГРВ от времени для растворов KCl и KNO_3



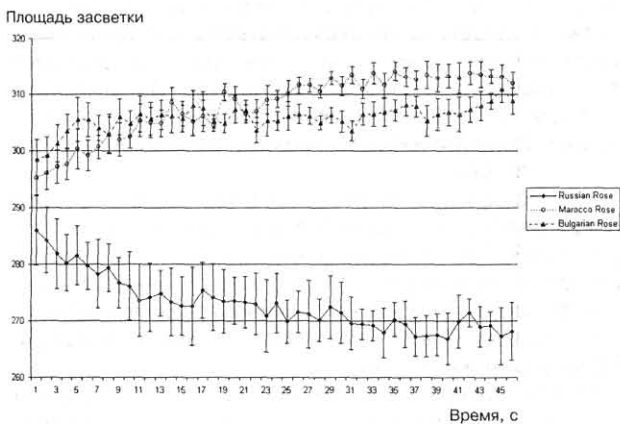
■ Рис. 9. Зависимость площади засветки ГРВ от времени для различных разбавлений раствора NaCl и дистиллированной воды

а во втором — монотонно возрастающей функции. Оказалось, что тренд площади засветки для растворов электролита меняет свое направление при разбавлении 2^{-14} , при котором соответствующие усредненные временные серии не имеют значимых различий с усредненными временными сериями дистиллированной воды (рис. 9).

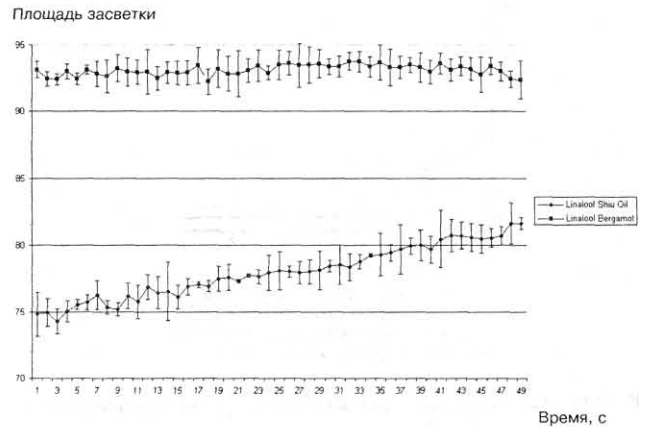
Надо отметить, что одно из основных отличий сильно проводящих жидкостей от слабо проводящих проявляется в виде тренда функции площади засветки от времени. Эта же закономерность наблюдается и в случае с такими слабо проводящими жидкостями, как эфирные масла.

Анализ данных по исследованию масел. Из более чем 60 видов рассмотренных масел только у трех тренды имели вид убывающей функции. Однако, в отличие от сильно проводящих жидкостей, у них отмечена квазилинейная временная зависимость площади засветки и небольшая скорость изменения тренда (малый угол наклона).

Метод динамической ГРВ-графии имеет большое преимущество в исследовании слабо проводящих жид-



■ Рис. 10. Зависимость площади засветки ГРВ для масел розы различного происхождения



■ Рис. 11. Зависимость площади засветки ГРВ для масел Linalool Shiu Oil и Linalool Bergamot

костей перед стандартной съемкой одиночных ГРВ-грамм.

На рис. 10 видно, что площади засветки одинаковы в начальный момент для масел Marocco, Russian и Bulgarian Rose. Однако менее чем через секунду тренды начинают разбегаться во времени и становятся статистически значимо различимы. В то же время для некоторых масел усредненные временные серии по площади засветки не имеют значимых различий в течение всего времени регистрации ГРВ процесса, но имеют различия по значениям средней интенсивности. Эта группа масел представлена Linalool Shiu Oil и Linalool Bergamot (рис. 11).

Математическая модель

Для описания поведения системы жидкость—плазма газового разряда—ЭМП необходимо смоделировать наблюдаемые изменения в данной системе. В рамках математического описания задача сводится к динамическому представлению процессов ГРВ жидкофазных объектов. При построении модели нужно учитывать три основных вида направлений для трендов: монотонно убывающие, монотонно возрастающие и периодические. Первым двум удовлетворяет общее решение уравнения Бесселя

$$z^2 u'' + zu' - (z^2 + v^2)u = 0 \quad (1)$$

в виде

$$u_v(z) = d_1 I_v(z) + d_2 K_v(z), \quad (2)$$

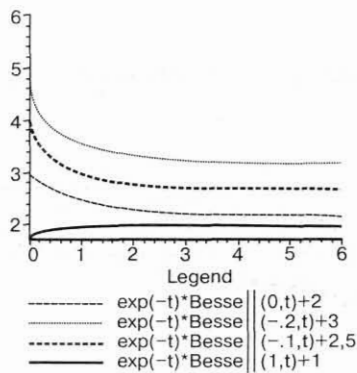
где $I_v(z)$ и $K_v(z)$ — модифицированные функции Бесселя первого и второго рода.

Заменой переменной $z \rightarrow i\tilde{z}$ уравнение (1) переходит в

$$z^2 u'' + zu' + (z^2 - v^2)u = 0, \quad (3)$$

общим решением которого является

$$u_v(z) = c_1 J_v(z) + c_2 Y_v(z), \quad (4)$$



■ Рис. 12. Функции $f(z)$ для различных коэффициентов a и B

где $J_\nu(z)$ и $Y_\nu(z)$ — функции Бесселя первого и второго рода.

Решение (4) имеет вид периодической функции при различных параметрах $\nu \in R$. Следовательно, условиям направления трендов удовлетворяют решения уравнений Бесселя (1) и (3), переходящие друг в друга заменой переменной $z \rightarrow iz$. Однако наиболее точное поведение описывает следующая функция, включающая в себя функции Бесселя:

$$f(z) = y(z) + B, \tag{5}$$

$$y(z) = Ae^{ax}u_\nu(\eta z), \tag{6}$$

где $a, A, B \in R$ и параметр $\eta = 1, i \in C$ соответствует решениям первого (1) и второго (3) уравнений Бесселя.

Обобщая доводы в пользу использования функций Бесселя и вида функции (5), приведем более общее уравнение, решением которого является функция $y(z)$:

$$z^2 y'' - (2ax - 1)xy' + [(a^2 + \eta^2)x^2 - \nu^2 - ax]y = 0. \tag{7}$$

Функция (5) показала высокую степень корреляции с полученными временными рядами параметров ГРВ изображений.

Например, на рис. 12 изображена функция (5) со значениями $A = d_1 = \eta = 1, a = -1, d_2 = 0$ и $\nu = 0; -0,2; -0,1; 1, B = 2; 3; 2,5; 1$. Из графика видно, что зависимость функции $y(z)$ при различных параметрах a и B соответствует зависимости площади засветки от времени при различных разведениях электролитов.

Обсуждение результатов

Для адекватного описания полученных результатов необходимо разобраться во всем комплексе физико-химических процессов, проходящих в системе жидкость—плазма газового разряда (ГР) — ЭМП.

После пробоя возникает канал плазмы в парах исследуемой жидкости, в котором генерация химически активных частиц происходит в газовой фазе под действием электронов, ускоряемых полями, характерными для импульсных разрядов. Прохождение тока разряда через жидкость приводит к возникновению на границах раздела плазма — раствор скачков потенциала, следствием чего является перенос через границу раздела фаз из раствора в плазму нейтральных и за-

ряженных частиц (эмиссия заряженных частиц). Действие разряда на растворы электролитов сопровождается диссоциацией молекул растворителя. Химические взаимодействия образующихся активных частиц ведут к изменению свойств раствора в целом. Подобные изменения происходят и при действии разряда на масла.

Поскольку изменения в исследуемой жидкости приводят к изменению состава и конфигурации плазмы, то это вызывает изменение таких характеристик газового разряда, как площадь засветки и средняя интенсивность изображения.

Силы ЭМП, поверхностного натяжения и гравитации, а также испарение, давление и температура влияют на образование и форму ГР. Физико-химические взаимодействия в жидкости, плазме и на границе раздела фаз с указанными силами и испарением определяют динамику ГР, а следовательно, и структуру временных серий площади засветки ГР.

В результате бомбардировки раствора инжестируемые из плазмы энергичными ионами происходит ионизация и диссоциация молекул воды с образованием радикалов H^\cdot, OH^\cdot и сольватированных электронов. Это приводит к генерации и накоплению в растворе пероксида водорода H_2O_2 . Следует также учитывать возможность образования ионов NO_2^\cdot . Однако известно, что спектры поглощения всех исходных исследованных растворов, а также пероксида водорода и нитрит-ионов лежат в области длин волн $\lambda < 254$ нм. Оптический спектр экспериментальной установки составляет 400–800 нм. Поэтому возникновение указанных соединений не может определять формирование ГРВ-грамм. Однако, при исследовании масел возможна генерация и накопление других соединений со спектром поглощения, совпадающим со спектром регистрируемым оптической системой. Такие соединения могут явиться причиной убывания тренда площади засветки ГРВ-грамм исследуемого масла.

Известно, что в открытой неустойчивой, нелинейной и неравновесной системе могут возникать периодически повторяющиеся структуры. Предполагая, что под влиянием импульсного ЭМП и последующих физико-химических взаимодействий в жидкости возникают периодические структуры с суммарным диполем, изменяющим значения скачков потенциала на границах раздела плазма — жидкость, логично предположить, что плазма ГР будет синхронно реагировать на соответствующие периодические изменения. Такие изменения должны отражаться на спектральных функциях временных серий площади засветки ГРВ-грамм. Однако для получения высоковоспроизводимых спектральных функций необходимо иметь большую частоту дискретизации.

Можно предположить, что описанные в работе тренды являются низкочастотными компонентами ГРВ процессов. Таким образом, воспроизводимые различия временных серий, как различных концентраций электролитов, так и масел различного происхождения проявляются за счет специфической самоорганизации исследуемых жидкостей.

Математическая модель, выбранная на основании видеоизмененной функции Бесселя, отражает волновую природу исследуемых ГРВ процессов. Такое описание позволяет приписать параметру ν значение

внутренней частоты системы, что соответствует определенному энергетическому уровню ее состояния, а также интерпретировать параметр a в выражении (6) как коэффициент возбуждения ($a > 0$) или затухания ($a < 0$).

Выводы

1. Динамическая ГРВ-графия позволила выявить различия во временных характеристиках проводящих жидкостей, не отличающихся по параметрам одиночных ГРВ-грамм. Было выявлено разведение 2^{-14} , при котором пропадает различие между растворами электролитов и дистиллированной водой.

2. При исследовании слабо проводящих жидкостей, имеющих близкий химический состав, из 42 пар масел в 33 были выявлены статистически значимые отличия по отдельным методам анализа. В десяти случаях были выявлены статистически значимые различия как по площади засветки, так и по средней интенсивности свечения. В 12 случаях различия были найдены только по площади засветки, в 11 случаях — по средней интенсивности засветки.

3. Очевидно, что только совместное использование различных методов ГРВ анализа позволит достоверно и статистически значимо выявлять различия параметров химически близких жидкостей.

4. Исследование как проводящих, так и слабо проводящих жидкостей показало, что рассмотренный метод позволяет выявлять статистически значимые различия при сравнении широкого спектра жидкостей.

Различия проявляются в изменении вида детерминированных составляющих временных рядов площади засветки и средней интенсивности ГРВ-грамм. Получаемые данные позволяют выявлять свойства исследуемых объектов, которые не проявляются при исследовании статических изображений. Продемонстрирована высокая воспроизводимость метода динамической ГРВ-графии.

5. Математическая модель динамики ГРВ процессов позволяет описывать получаемые данные с помощью новых параметров, соответствующих представлению различных видов исследуемых систем, с учетом лежащих в основе процесса волновых свойств.

6. Основываясь на принципах стандартной ГРВ-графии, динамическая ГРВ-графия позволяет выявлять дополнительные свойства объектов за счет измерения и анализа временных рядов параметров ГРВ-свечения во времени.

Литература

1. **Коротков К. Г.** Основы ГРВ биоэлектрографии. — СПб.: Изд-во СПбГИТМО, 2001.
2. **Ultra High Dilutions — Physiology and Physics.** (Endler, Ed. Kluuiver Acad. Pub., 1994.
3. **Skarja M., Berden M., Jerman J.,** J.Appl.Phys., 84, 2436, 1998.
4. **Korotkov K., Korotkin D.,** J.Appl.Phys., 89, 4732, 2001.
5. **Fisher R. A.** Statistical Methods for Research Workers (6th ed.). Edinburgh: Oliver and Boyd, 1936.

УДК 338.24

УПРАВЛЕНИЕ ИНВЕСТИЦИЯМИ В РОССИЙСКИХ РЕГИОНАХ НА ОСНОВЕ КОНЦЕПЦИИ «РАЗУМНОГО РОСТА»

Ю. Х. Лукманов,

канд. экономич. наук,

глава территориального управления

Адмиралтейского административного района г. Санкт-Петербурга

Статья посвящена проблеме инновационного управления экономическими процессами в регионах. На основе анализа статистического и методического материала проведен анализ факторов, влияющих на экономический рост промышленности. Автором предлагаются к рассмотрению основные принципы создания системы привлечения капитала в регион на основе концепции «разумного роста».

The issue of the article is innovation of control over the economic processes at the local level with using of statistics and methodology material of the analysis of factors affecting economic growth of the industry. The author suggests basic principles of investments into the local economy, basing on the concept of «reasonable growth».

Введение

Инвестиции, являясь одним из инструментов управления развитием социально-экономических систем (СЭС), приобретают особую роль как эффективный инструмент поддержания оптимальных параметров их динамической устойчивости. При этом следует исходить из понимания того, что инвестиции и инновации — две неразрывно связанные сферы экономической деятельности. Как справедливо отмечено в работе [1], капиталовложения без инноваций не имеют смысла, поскольку нет смысла воспроизводить устаревшее оборудование, не пользующееся спросом и, к тому же, ресурсорасточительное. Однако и инновации без капиталовложений нереальны. *Поэтому эффективный инвестиционный менеджмент возможен только на основе комплексного инновационно-инвестиционного подхода, обновления основного капитала на принципиально новой, конкурентоспособной основе.*

Сегодня проблему обновления производственного аппарата не решить с помощью старых подходов, когда научно-техническая и инвестиционная политика осуществлялась централизованно и, главным образом, за счет бюджетных источников и фондов министерств. Предприятия же выступали в роли просителей и пользователей государственных средств, занимающихся множеством мелких улучшающих инноваций. На фоне реального разгосударствления формируются инвестиционный и инновационный рынки. Федеральные и местные

бюджеты выделяют минимум средств на техническое перевооружение производства. Центр тяжести в принятии решений переместился к собственникам капитала и товаропроизводителям, которые озабочены выживанием и мало внимания уделяют производственным инвестициям и нововведениям, цикл которых выходит за границы краткосрочных интересов.

Процесс первоначального накопления капитала происходит преимущественно в сфере обращения и в малой степени затрагивает производство, которое и должно обеспечить самовозрастание капитала в нормальной рыночной экономике. Поэтому нужны новые подходы к инновационно-инвестиционной политике и механизму ее реализации, создание и последовательная реализация инновационно-инвестиционного механизма, отвечающего новым экономическим и социально-политическим реалиям, обеспечивающего крупномасштабный приток капитала для модернизации производства. Исходный элемент данного механизма — выработка селективной инновационно-инвестиционной стратегии, которая найдет отражение в федеральных и региональных инновационных и инвестиционных программах и крупных проектах, рассчитанных на длительный период и охватывающих все стадии технологического цикла. Главная трудность — выбрать те узкие приоритетные производства, где концентрация ресурсов сможет в сравнительно короткий срок обеспечить успех и цепную реакцию перелома в более широких сферах на-

родного хозяйства («стратегия лазерного луча»). Важно при этом не попасть в «черные дыры» экономики — предприятия, поглощающие в избытке ресурсы, но не способные производить конкурентоспособную продукцию.

В связи с этим представляет интерес выявление и анализ основных факторов экономического роста.

Факторы экономического роста

Вопрос об основных факторах экономического роста проанализирован экспертами Международного валютного фонда (МВФ) на основе агрегированной производственной функции Р. Солоу и приведен в работе [2]. В рассматриваемой модели экономического роста фигурируют три основных фактора: технологический прогресс, рабочая сила, инвестиции.

Комментируя приведенные в табл.1 данные, В. Клавдиенко [2] отмечает, что вклад технологического прогресса в экономический рост, измеряемый темпами прироста внутреннего валового продукта (ВВП), в последние два десятилетия XX в. возрос во всех развитых странах (35–50 %). Технологический прогресс при этом может рассматриваться и как материализация накопленной информации и знаний в традиционных производственных ресурсах, и как улучшение организации и управления экономикой.

На протяжении последних двух десятилетий в экономическом развитии в Японии ведущую роль неизменно играл инвестиционный фактор, существенным он являлся и для экономики европейских стран. В США доля инвестиций в приросте ВВП в последние два десятилетия оставалась почти неизменной.

Роль рабочей силы в увеличении производства ВВП снизилась в последние два десятилетия во всех развитых странах и в настоящее время имеет наименьшее влияние. Следует согласиться с В.Клавдиенко [2], что эксперты МВФ не учитывали роль естественно-природных факторов в развитии экономики. Благополучие промышленно

развитых стран неразрывно связано, например, с максимальным объемом потребления сырья, в том числе ввозимого из других стран. Так, постиндустриальная экономика США является крупнейшим в мире потребителем углеводородного топлива, и, прежде всего, нефти (20 баррелей на душу населения в год), 65% потребности в которой покрывается за счет импорта. Аналогична ситуация в Японии (13 баррелей) и Германии (10 баррелей) [1].

Поэтому любой анализ, не учитывающий влияние естественно-природных условий производства (земля, вода, полезные ископаемые, ресурсы растительного и животного происхождения и т. д.), а также политических, социальных и других факторов (многие из которых в настоящее время не могут быть идентифицированы, а их качественные и количественные характеристики — точно оценены), дает приближенную оценку степени влияния рассматриваемых факторов на экономический рост. В связи с этим, представленные в табл. 1 данные можно рассматривать как ориентировочные, позволяющие сделать вывод о том, что обусловленное переходом к так называемой информационной экономике повышение роли технологического прогресса (в его материализованном и нематериализованном видах) в последние два десятилетия происходило в основном за счет снижения роли живого труда, но существенно не уменьшало вклада инвестиционного фактора в обеспечение экономического роста. В группе наиболее развитых стран мира, в том числе в постиндустриальных хозяйственных системах, роль инвестиционной составляющей характеризовалась устойчивостью и была весьма значительной — от 1/3 до 1/2 прироста ВВП.

Неизменна значимость инвестиционной составляющей экономического роста, являющейся одной из основ современной мировой хозяйственной тенденции — опережающего роста международной инвестиционной активности по отношению к росту мирового ВВП и международной торговли. Так, за период 1990–2000 гг. при росте мирового ВВП в 1,3 раза и увеличении мирового товарного экспорта в 1,5 раза размер прямых иностранных инвестиций (ПИИ) в мире вырос не менее чем в 6 раз.

■ Таблица 1. Роль основных факторов экономического роста в приросте ВВП*

	США		Страны Европы		Япония	
	80-е гг.	90-е гг.	80-е гг.	90-е гг.	80-е гг.	90-е гг.
Среднегодовые темпы прироста ВВП, %	2,9	2,6	2,2	2,0	3,6	2,6
Роль рабочей силы, % (в процентах от суммы прироста ВВП)	1,1 (38,0)	0,9 (34,6)	0,3 (13,6)	0,2 (10,0)	0,7 (19,4)	0,3 (11,5)
Роль основного капитала, % (в процентах от суммы прироста ВВП)	0,9 (31,0)	0,8 (30,8)	0,9 (40,9)	0,8 (40,0)	1,8 (50,0)	1,2 (46,2)
Роль технологического прогресса, % (в процентах от суммы прироста ВВП)	0,9 (31,0)	0,9 (34,6)	1,0(45,5)	1,0 (50,0)	1,1 (30,6)	1,1 (42,3)

* Составлено и рассчитано по справочнику «World Economic Outlook». October 1999. IMF. — Washington, 1999. P. 74.

Взаимосвязь роста ВВП и динамики занятости населения в России

В работе [2] исследована зависимость темпа прироста (уменьшения) ВВП от динамики занятости, инвестиций в основной капитал и ПИИ в экономику России. Анализ выполнен с использованием эконометрической модели следующего вида:

$$y_t = \sum_{i=1}^3 a_i x_{it} + a_0 + \xi,$$

где y_t — среднегодовой темп прироста ВВП (в сопоставимых ценах); a_1, a_2, a_3 — параметры функции; a_0 — параметр, отражающий смешанное влияние технического прогресса и не идентифицированных факторов; ξ — случайное отклонение.

В результате обработки статистических данных методом наименьших квадратов [4] была получена зависимость темпа прироста (уменьшения) ВВП в 1993–2000 гг. от указанных выше факторов:

$$Y = -0,240 + 0,833x_1 + 0,023x_2 - 0,011x_3,$$

где x_1 — темп прироста численности занятых в экономике; x_2 — темп прироста инвестиций в основной капитал (в сопоставимых ценах); x_3 — темп прироста ПИИ (коэффициент корреляции равен 0,78, коэффициент Дарбина-Уотсона — соответственно 1,84). По результатам расчета с соответствующими значениями переменных для данного уравнения могут быть сделаны, в частности, следующие выводы:

— основным фактором сокращения ВВП в 1993–2000 гг. являлось уменьшение численности занятых в экономике. Уменьшение занятости в этот период в среднем на 1,37 % в год, вызвавшее сокращение ВВП на 1,14 %, на 65,1 % объясняло темпы спада в экономике;

— коэффициенты при x_2 и x_3 статистически малозначимы. С определенной вероятностью можно лишь отметить, что отрицательные темпы прироста инвестиций в основной капитал (в среднем — 8,20 % в год) содействовали сокращению ВВП почти на 17 % в год. Влияние притока ПИИ на динамику ВВП было незначительным, при этом отрицательный коэффициент при x_3 указывает на сдерживающее влияние ПИИ на спад производства, противодействуя ему; определенное противодействие спаду производства оказывает научно-технический прогресс и не идентифицированные факторы (отрицательное значение a_0).

Таким образом, исследование экономических факторов кризисного функционирования российской экономики выявило в качестве основного параметра, коррелируемого с темпами падения ВВП, занятость населения. При этом влияние научно-технического прогресса на экономический рост [2] требует дополнительных исследований, которые проведены, например, в [3]. Используемая модель кооперации хозяйствующих субъектов отражает прямую зависимость эффективности занятости от уровня технологического прогресса, который, в свою

очередь, можно оценить на основе соотношения в затратах (себестоимости) производимой продукции живого и овеществленного (прошлого) труда.

Для расчета относительной эффективности занятости работников по отраслям промышленности воспользуемся данными статистической отчетности в России. Эффективность занятости $E_{занj}$ определим как отношение величины прибыли Π_j к затратам живого труда $Z_{жj}$:

$$E_{занj} = \Pi_j / Z_{жj} \quad (1)$$

где i — натуральное число, отражающее отраслевую структуру экономики России по системе действующего статистического учета. Преобразуем выражение (1), представив Π_j в виде $\Pi_j = Z_j \cdot Np_j = (Z_{жj} + Z_{овj}) Np_j$. Здесь $Z_{овj}$ — затраты овеществленного труда, Z_j — занятость в отраслевой структуре.

Тогда $E_{занj} = (1 + Z_{овj}/Z_{жj}) Np_j$. При этом для анализа тенденций важно определение относительной оценки эффективности занятости во времени $E_{занj}^t / E_{занj}^{t-m}$. Здесь $E_{занj}^t$ и $E_{занj}^{t-m}$ — эффективность занятости в i -й отрасли соответственно в t -м и в $(t-m)$ -м годах. Принимая в качестве допущения, что за время m норма прибыли в отрасли Np_j не меняется или меняется незначительно, определим относительную оценку эффективности занятости во времени по i -й отрасли как:

$$E_j = E_{занj}^t / E_{занj}^{t-m} = (1 + Z_{овj}^t / Z_{жj}^t) / (1 + Z_{овj}^{t-m} / Z_{жj}^{t-m})$$

В качестве примера рассчитанные значения E_j для промышленности России за 1995 и 1998 годы приведены в табл. 2.

Как видно из таблицы, помимо абсолютного сокращения занятых в промышленности на 3,1 млн чел. наблюдается и снижение эффективности занятых в промышленности за счет неблагоприятных структурных сдвигов. Это равносильно дополнительному выводу численности в 2,6 млн чел., что говорит об отсутствии положительной структурной перестройки промышленности. Напротив, потери от неблагоприятных структурных сдвигов сопоставимы с абсолютным сокращением численности работающих в промышленности. Учитывая изложенное выше, полученную оценку E_j можно использовать как цель и индикатор реформирования экономики. Например, по отношению к промышленности она может быть сформулирована как проведение прогрессивной структурной перестройки промышленности страны при опережающем росте эффективности занятости, обеспечивающем компенсацию временных неблагоприятных последствий, связанных с остановкой и закрытием отдель-

■ Таблица 2. Эффективность занятости населения в промышленности России

Годы	Численность занятых, млн чел.	$Z_{ов}/Z_{ж}$	E_j
1995	17,2	4,8	1,0
1998	14,1	3,9	0,84

ных производств и сокращением численности их работников. Данные расчетов по отраслям промышленности приведены в табл.3.

Как следует из табл. 3, наиболее эффективной отраслью промышленности в 1995 г. была нефтеперерабатывающая, а в 1998 г. — электроэнергетика. К сожалению, именно лидер — нефтепереработка понесла за анализируемый период наибольшие потери по эффективности. Видимо, это обеспечило неблагоприятную тенденцию — проникновение на отечественный рынок импортной продукции нефтепереработки. Угольная промышленность — единственная отрасль, которая стабилизировалась, но на очень низком значении соотношения овеществленного и живого труда. Наименьшие потери понесли энергетика, нефтедобывающая, лесная, деревообрабатывающая и целлюлозно-бумажная промышленность, а также промышленность строительных материалов. Незначительные изменения произошли в легкой, пищевой и газовой промышленности, где сохраняется потенциал роста. Резкое падение наблюдается в цветной металлургии, где относительная эффективность занятости сократилась более чем в два раза.

Общая картина весьма неблагоприятная. Динамика отраслевых изменений имеет неблагоприятную тенденцию спада с достижением «дна» при значениях соотношения овеществленного и живого труда в районе 2,0 (пример — угольная промышленность). Полученные результаты еще раз подтверждают необходимость корректировки инвестиционной политики в промышленности. Показатель относительной эффективности занятости E_i , включая одновременно технологическую составляющую и составляющую численности занятых, с учетом вышеизложенного можно выбрать как важный критерий оценки отраслевой инновационно-инвестиционной привлекательности. Наложение его на шкалу приоритетов регионального (территориального) развития дает простой инструмент регионально-отраслевого управления инновационно-инвестиционной деятельностью.

Экономическая теория также подтверждает огромное влияние занятости и ее структуры на экономику страны (территории, региона). Как отмечается в работе [3], в складывающейся кооперации и разделении труда наукоемкие технологии, маркетинг и финансы сосредоточены в богатых странах (регионах). Структура занятости в них по форме напоминает перевернутую пирамиду и обеспечивает им получение сверхприбылей. В основании пирамиды — начальные технологические переделы (добыча и переработка исходного сырья), а на вершине — заключительные (торговля и финансы). Продукция более низких слоев технологических переделов является исходным полуфабрикатом для последующих, более высоких слоев. При этом стоимость товара на каждом технологическом переделе формируется по затратному принципу: величина прибыли, включаемой в цену произведенного товара, определяется по нормативу ко всем произведенным затратам, к полной себестоимости продукции, т. е. включает затраты «прошлого тру-

■ Таблица 3. Расчет эффективности занятости по отраслям промышленности России

Наименование отрасли	$Z_{ог}/Z_{ж}^*$		E_i
	1995	1998	
Электроэнергетика	7,3	6,6	0,92
Нефтедобывающая промышленность	7,3	6,5	0,90
Нефтеперерабатывающая промышленность	17,9	5,2	0,33
Черная металлургия	8,1	5,2	0,68
Химическая и нефтехимическая промышленность	7,0	4,9	0,74
Топливная промышленность	6,2	4,6	0,78
Пищевая промышленность	5,7	4,7	0,85
Газовая промышленность	5,5	4,2	0,80
Промышленность строительных материалов	3,4	3,1	0,93
Цветная металлургия	8,1	3,1	0,45
Лесная, деревообрабатывающая и целлюлозно-бумажная промышленность	3,2	2,9	0,93
Машиностроение и металлообработка	3,0	2,5	0,88
Легкая промышленность	3,1	2,3	0,80
Угольная промышленность	2,1	2,1	1,0

* Данные приведены из [5].

да». В результате чем выше уровень технологического передела, тем больше в себестоимости продукции затрат прошлого труда, и, следовательно, больше масса прибыли при одних и тех же затратах живого труда, т. е. его эффективность.

Формирование положительной динамики увеличения численности работающих с одновременным прогрессивным изменением структуры занятости обеспечивает рост благосостояния региона и его населения, являясь важнейшей целью управления. Так, например, в США за 20 лет, начиная с 1969 г., «экономика работала подобно гигантской машине, создающей новые рабочие места. В абсолютных показателях численность работающих увеличилась за этот период с 78 до 117 млн человек. Рост количества рабочих мест в США в рассматриваемый период затронул как обрабатывающие отрасли промышленности, так и сферу услуг, хотя в последний уровень занятости повышался более высокими темпами» [2].

В богатых странах численность занятых в добывающих отраслях ничтожно мала. Ресурсодобывающие производства, несмотря на наличие природных запасов полезных ископаемых, остановлены или значительно сокращены (нефть в США, каменный уголь в Великобритании). Приобретая ресурсы за рубежом, эти страны создают необходимые условия для создания эффективных рабочих мест у себя.

Россия вовлекается в мировое разделение труда с явно выраженной тенденцией сырьевой страны, что противоречит ее национальным интересам, поскольку ведет к отсталости и бедности. Российские базовые отрасли переключаются с внутреннего рынка на внешний, формируя тем самым внут-

ри страны неэффективную структуру занятости. Однако уровень технологического передела дает качественную оценку рабочих мест, дополняемую количественной оценкой — численностью рабочих мест, создаваемых как в месте приложения инвестиций, так и по всей технологической цепочке (вторичная занятость).

Концепция «разумного роста»

Как показал анализ зарубежного опыта управления СЭС, важнейшей задачей является выравнивание уровней социально-экономического развития регионов и территорий. Без этого становится невозможным поддержание ее динамической устойчивости. Для России эта проблема стоит наиболее остро.

Для создания привлекательных условий приема инвестиций в регионе, городе, области или районе необходимо, в первую очередь, организовать четкую систему привлечения капитала и определить приоритеты. Для этого необходимо:

- обеспечить прозрачность действующего законодательства по инвестициям, чтобы оно было понятно любому инвестору (на сегодняшний день оно запутано);

- сформировать «инвестиционное» управление, при котором все участвующие подразделения работали бы как единая команда на конечный результат — привлечение инвестора. На практике инвестор просто теряется из-за трудностей, связанных с многочисленными бюрократическими согласованиями между ответственными за это подразделениями, пока определяет должное развитие своего проекта. При этом сложилось и отношение к потенциальному инвестору, как к прожитому ресурсу, а не как к их источнику. Мировой опыт свидетельствует о том, что успех реализации инвестиционных программ и проектов обеспечивает именно система «инвестиционного» управления, предусматривающая согласованность действий всех участников проекта.

- сформировать механизм пропорционального участия в инвестициях частных, смешанных и общественных организаций. Такие пропорции должны быть увязаны с приоритетами развития региона (территории).

С учетом вышеизложенного, концепция «разумного роста» представляет собой совокупность следующих принципов:

- поддержания динамической устойчивости СЭС;

- единства инновационно-инвестиционной деятельности;

- перехода от концепции «тяжелого роста» с инвестициями в реализацию крупных проектов в тяжелой промышленности к концепции, нацеленной на реализацию высокоэффективных проектов в отраслях высоких технологий, освоения новых технологических процессов, обновления основных фондов;

- использования показателя занятости с оценкой эффективности ее структуры, как индикатора и цели экономических преобразований на переходном периоде;

- выравнивания социально-экономического уровня развития регионов и территорий;

- формирования четкой системы привлечения капитала и определения приоритетов инвестирования.

Практический пример

В качестве примера приведем результаты, полученные при оценке эффективности занятости по г. Санкт-Петербургу как субъекту федерации и его территориальной составляющей — Адмиралтейскому административному району.

Для оценки влияния изменения структуры занятости на эффективность функционирования города рассчитаем изменение соотношения долей живого и овеществленного труда в валовом региональном продукте (ВРП), используя принятые в отчетности обозначения. Результаты расчета приведены в табл. 4.

Анализ приведенных данных показывает, что относительная оценка эффективности занятости E_i в городе с 1991 по 1996 год снизилась на 2 %, и это означает, что неблагоприятные структурные сдвиги в экономике Санкт-Петербурга равносильны сокращению занятости на 2 %. К 2000 году E_i снизилась на 22 %, что потребовало вмешательства в процесс управления занятостью через инвестиции на уровне района.

В соответствии со сформулированной выше концепцией «разумного роста» администрация Адмиралтейского административного района г. Санкт-Петербурга в 2001–2002 гг. сосредоточила свои усилия на решении следующих задач:

- увеличение абсолютной численности занятых в экономике района;

- определение приоритетов экономического развития и на их основе улучшение существующей структуры занятости и эффективности рабочих мест;

- инновационно-инвестиционным подходе в привлечении инвестиций в экономику района.

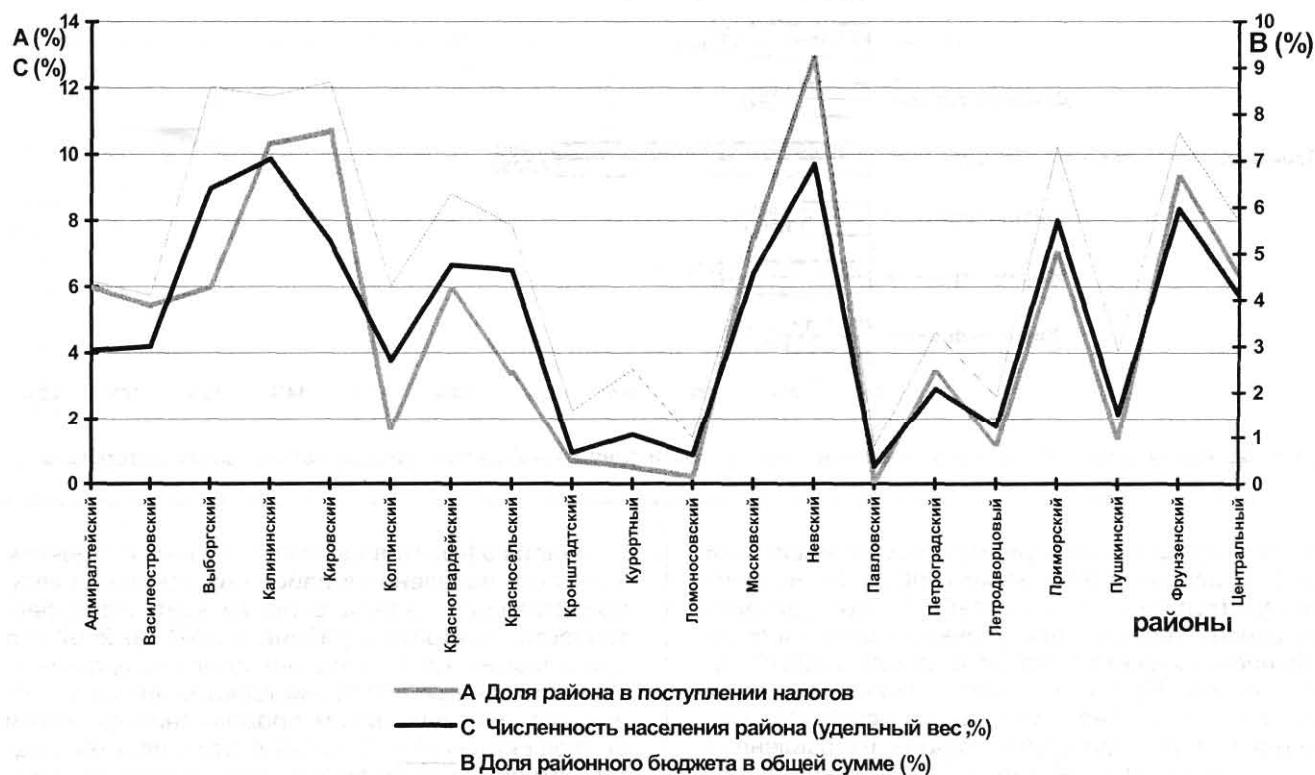
Анализ структуры экономики района, приведенной на диаграммах (см. рис. 1 и 2) и в табл. 5 и 6, показал ее разнообразие и по составу и по численности занятых. Из представленного материала видно, что в экономике района транспорт занимает приоритетное значение (25,3 % по объему и 18,8 % по численности занятых), на втором месте пищевая промышленность (20,6 % и 5,8 % соответственно) и на третьем месте ЖКХ (17,9 % и 3,4 % соответственно). При анализе экономики района предложен подход, увязывающий бюджет района с доходами и численностью проживающего населения.

Анализ данных, приведенных в табл. 5, демонстрирует тесную связь между численностью населения, долей доходов, долей районного бюджета и

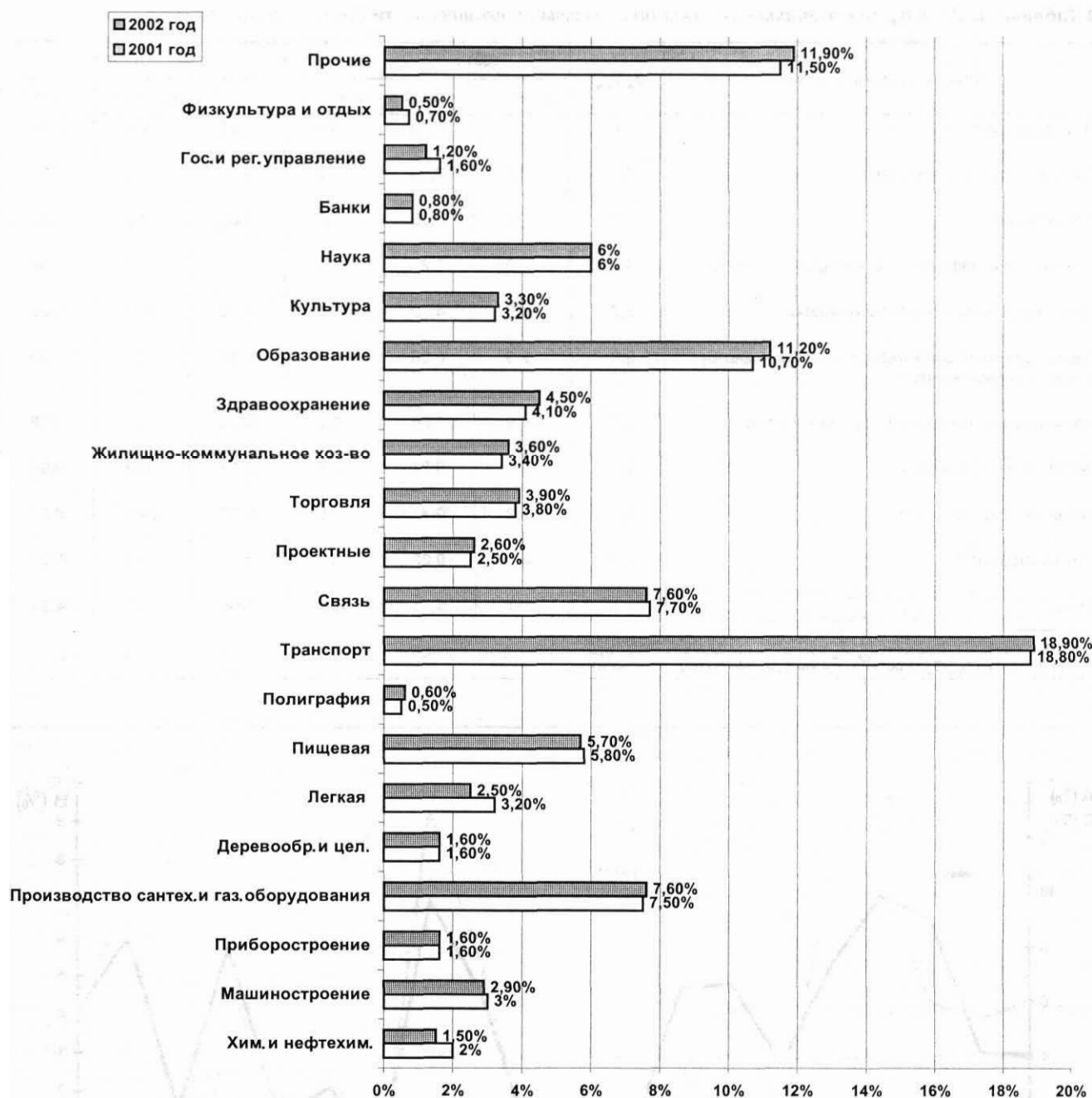
■ Таблица 4. Структурные соотношения труда по отраслям промышленности Санкт-Петербурга

Отрасли промышленности	$Z_{ж}/Z_{ов}^*$	1991 г.		1996 г.		2000 г.	
		уд. вес, %	гр2хгр3	уд. вес, %	гр2хгр5	уд. вес, %	гр2хгр7
Электроэнергетика	6,6	1,3	0,09	3,2	0,21	7,8	0,51
Топливная промышленность	4,6	0,1	0,00	0,2	0,1	0,0	0,0
Металлургия	3,1	1,0	0,03	2,8	0,09	4,5	0,14
Химическая и нефтехимическая промышленность	4,9	4,0	0,2	4,0	0,2	1,7	0,08
Машиностроение и металлообработка	6,7	70,5	4,72	62,1	4,16	35,1	2,35
Лесная, деревообрабатывающая, целлюлозно-бумажная промышленность	2,9	2,9	0,08	4,0	0,12	3,1	0,09
Промышленность строительных материалов	3,1	1,9	0,06	2,0	0,06	2,6	0,08
Легкая промышленность	2,3	9,1	0,21	7,8	0,18	2,3	0,05
Пищевая промышленность	2,0	5,2	0,1	7,6	0,15	34,5	0,69
Прочие отрасли**	5,49	4,0	0,22	6,3	0,33	8,4	0,22
Итого		100	5,71	100	5,60	100	4,21

* При расчете отношения $Z_{ов}/Z_{ж}$ в затраты овецественного труда не включается импорт.
 ** Отношение $Z_{ов}/Z_{ж}$ определяется как среднее по году по всем отраслям.



■ Рис. 1. Расчет расходной части бюджета по модели (2001 г.)



■ Рис. 2. Удельный вес отраслей в совокупной численности занятых в Адмиралтейском районе Санкт-Петербурга

наглядно представляет преимущества и недостатки каждого района в экономике города. По численности, находясь в середине табл. 5, Адмиралтейский район обеспечил поступление налогов и долю районного бюджета в бюджете города выше среднего уровня. При этом в рассматриваемые годы видна положительная динамика: в результате мероприятий районной администрации, направленных на создание новых рабочих мест, численность занятых в районе за год возросла на 11,6 %, а доходы бюджета — на 45,3 %.

Анализ эффективности структурных изменений занятости населения в районе на основе данных, представленных в табл. 6, показывает, что эффективность занятости в районе с 2001 по 2002 год снизилась на 4,5 %, а это значит, что прирост численности занятых в отраслях промышленности происходил при небольшом продолжении снижения ее эффективности. В связи с этим действия администрации реализовались в двух направлениях: стимулировании роста эффективности отраслевой занятости за счет лучшей кооперации и техни-

■ **Таблица 5.** Анализ расходной части бюджета районов Санкт-Петербурга в общей величине расходов бюджета города, %

Наименование района	Численность населения (С)		Доход района, тыс. руб.		Доля дохода (А)		Бюджет района, тыс. руб.		Доля районного бюджета (В)	
	2000 г.	2001 г.	2000 г.	2001 г.	2000 г.	2001 г.	2000 г.	2001 г.	2000 г.	2001 г.
Адмиралтейский	191,2	190,0	13 616	23 602	4,7	6,0	276 753,0	395 506,4	4,5	4,4
Василеостровский	197,4	195,1	19 387	21 644	6,7	5,4	236 503,0	368 480,2	3,9	4,1
Выборгский	419,6	417,8	25 662	23 525	8,9	6,0	535 474,0	767 158,6	8,7	8,6
Калининский	462,7	459,5	21 830	40 601	7,6	10,3	545 221,0	747 232,0	8,9	8,4
Кировский	344,9	343,3	23 760	42 139	8,2	10,7	494 979,5	772 986,7	8,1	8,7
Колпинский	176,7	175,5	5 051	6 835	1,7	1,7	272 119,0	386 852,5	4,4	4,3
Красногвардейский	315,4	310,3	32 034	23 438	11,1	3,9	397 054,0	564 907,1	6,5	6,3
Красносельский	303,8	302,9	13 242	13 572	4,6	3,4	354 041,0	498 262,3	5,8	5,6
Кронштадтский	44,1	43,7	2 205	2 743	0,8	0,7	92 892,4	142 852,5	1,5	1,6
Курортный	69,3	69,2	3 180	1 903	1,1	0,5	164 646,0	227 182,5	2,7	2,5
Ломоносовский	40,6	40,2	1 636	904	0,6	0,2	56 352,0	92 803,1	0,9	1,0
Московский	295,1	291,5	19 065	29 747	6,6	7,5	349 554,0	493 992,4	5,7	5,5
Невский	453,9	452,4	29 444	51 223	10,2	12,9	570 046,0	829 316,0	9,3	9,3
Павловский	24,8	24,4	187	174	0,1	0,1	52 096,0	69 587,3	0,8	0,8
Петроградский	138,2	135,7	8 410	13 476	2,9	3,4	215 695,0	292 863,1	3,5	3,3
Петродворцовый	81,1	81,0	2 564	4 665	0,9	1,2	113 208,0	173 192,4	1,8	1,9
Приморский	369,1	373,0	12 230	27 793	4,2	7,0	445 669,0	653 603,5	7,3	7,3
Пушкинский	100,4	99,9	8 473	5 703	2,9	1,4	180 188,0	258 445,3	2,9	2,9
Фрунзенский	392,3	389,1	25 536	36 709	8,8	9,3	420 113,0	678 457,0	6,8	7,6
Центральный	273,4	266,1	21 319	25 255	7,4	6,4	369 375,0	508 014,1	6,0	5,7
	4694,0	4660,6	288 831	395 651	100	100	6 141 978,9	8 921 695,0	100	100

■ **Таблица 6.** Структура занятости населения по отраслям промышленности

Отрасли промышленности	$Z_{ж}/Z_{ов}^*$	2001 год		2002 год	
		Уд. вес, %	гр2хгр3	Уд. вес, %	гр2хгр5
Химическая и нефтехимическая	6,6	2,0	0,13	1,5	0,10
Машиностроение	6,7	3,0	0,2	2,9	0,19
Приборостроение	6,7	1,6	0,11	1,6	0,11
Производство сантехнического и газового оборудования	4,2	7,5	0,32	7,6	0,32
Деревообрабатывающая и целлюлозно-бумажная	2,9	1,6	0,05	1,6	0,05
Легкая	2,3	3,2	0,07	2,5	0,06
Пищевая	2,0	5,8	0,12	5,7	0,11
Полиграфическая	4,0	0,5	0,02	0,6	0,02
Транспорт	3,4	18,8	0,64	18,9	0,64
Связь	8,0	7,7	0,62	7,6	0,61
Торговля	4,6	3,8	0,17	3,9	0,18
Жилищно-коммунальное хозяйство	5,3	3,4	0,18	3,6	0,19
Здравоохранение	3,1	4,1	0,13	4,5	0,14
Образование	3,1	4,1	0,13	4,5	0,14
Культура	2,0	3,2	0,06	3,3	0,07
Наука	7,0	6,0	0,42	6,0	0,42
Банки	9,0	0,8	0,07	0,8	0,07
Государственное и региональное управление	4,0	1,6	0,06	1,2	0,05
Проектные	7,0	2,5	0,18	2,6	0,18
Туризм, физкультура и спорт	6,5	0,7	0,05	0,5	0,03
Прочие отрасли**	4,15	11,5	0,48	11,9	0,49
Итого		100	4,63	100	4,38

* При расчете отношения $Z_{ов}$ к $Z_{ж}$ в затраты овещественного труда не включается импорт.
** Отношение $Z_{ов}$ к $Z_{ж}$ определяется как среднее по году по всем отраслям.

ческого оснащения (рост отношения $Z_{ов}/Z_{ж}$); обеспечении роста занятости в наиболее эффективных сферах приложения труда. По данным за 2001–2002 гг., это привело к росту числа занятых на 1 % и доходов бюджета на 3,9 %. Соотношение занятого и проживающего в районе населения при этом составило 59 %.

Заключение

Анализ эффективности занятости населения в режиме мониторинга позволяет осуществлять контроль и анализ экономической ситуации в реальном режиме времени и своевременно корректировать управленческие решения в интересах динамичного развития СЭС.

Литература

1. Яковец Ю. Инновационное инвестирование: новые подходы // Экономист. — 1995 — № 11. — С. 351.
2. Клавдиенко В. Инвестиции и экономический рост // Инвестиции в России. — 2002. — № 7. — С. 40–45.
3. Родионов В. Организация эффективного управления регионами России на основе активного подхода. — СПб: Изд-во «Невский фонд», 2002. — 112 с.
4. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики. — М: ЮНИТИ, 1998. — 1022 с.
5. Российский статистический ежегодник. 1999. — М.: Статистика, 2000.

**ЯКОВЕНКО
Максим
Константинович**



Закончил Санкт-Петербургский университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети» в 2000 году. Аспирант кафедры вычислительных систем и сетей ГУАП. Автор семи научных публикаций. Области научных интересов: анализ и обработка биологических изображений; статистическое и эвклид-описание объектов распознавания.

**СЕРГЕЕВ
Михаил
Борисович**



Профессор кафедры вычислительных систем и сетей СПбГУАП. Окончил Ленинградский электротехнический институт (ЛЭТИ) в 1980 г. по специальности «Электронные вычислительные машины». Диссертацию на соискание ученой степени доктора технических наук защитил в 2001 году. Автор более 80 научных работ, в том числе двух монографий. Под его редакцией издано пять сборников научных трудов и монографий. Главный редактор журнала «Информационно-управляющие системы». Области научных интересов: теория разрядных вычислений; методы проектирования спецпроцессоров для систем контроля и управления.

**РЫЖИКОВ
Юрий
Иванович**



Профессор кафедры математического обеспечения ЭВМ Военно-космической академии им. А. Ф. Можайского. Закончил Черноморское высшее военно-морское училище в 1958 г. Докторскую диссертацию защитил в 1969 г. по математической теории управления запасами. Заслуженный деятель науки РФ. Автор более 180 публикаций, в том числе учебников и монографий. Области научных и педагогических интересов: управление запасами; теория очередей; имитационное моделирование; вычислительные методы и прикладное программирование; разработка, эксплуатация и эффективность программного обеспечения; науковедение и подготовка научных кадров.

**СМИРНОВ
Юрий
Михайлович**



Заместитель директора Института интеллектуальных систем и технологий, заведующий кафедрой «Интеллектуальные системы управления» СПбГПУ. Закончил математико-механический факультет Ленинградского государственного университета по специальности «Теоретическая механика». Автор более 250 научных работ. Научные интересы — проблемы управления сложными системами, организационное управление, стратегическое планирование. Член ученого и диссертационного советов СПбГПУ. Действительный член Российской академии естественных наук и Международной академии информационных процессов и технологий.

**ПОЛЯКОВ
Александр
Олегович**



Профессор кафедры интеллектуальных систем управления СПбГПУ, ведущий научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН. Выпускник ЛЭТИ 1972 года. Диссертацию на соискание ученой степени доктора технических наук защитил в 1995 г. Автор более 120 публикаций, в том числе пяти монографий. Научные интересы — информатика, управление в сложных системах. Член ученого и диссертационного советов СПИИРАН. Действительный член Академии инженерных наук РФ.

**ОДНОБОКОВ
Вячеслав
Владимирович**



Старший преподаватель кафедры физики и химии Псковского политехнического института СПбГПУ. Заместитель декана подготовительного факультета. Выпускник Радиофизического факультета Ленинградского политехнического института 1988г. по специальности «Физическая электроника». Автор более 20 публикаций. Научные интересы — управление сложными процессами, компьютерная физика.

**ЧЫОНГ
Динь
Тяу**



Аспирант Санкт-Петербургского государственного политехнического университета (СПбГПУ).
Выпускник кафедры автоматизации и вычислительной техники СПбГТУ 2001 г.
Автор двух научных публикаций.
Области научных интересов: распределенные информационно-управляющие системы; SCADA-системы; OPC-технология.

**КОЛБАНЕВ
Михаил
Олегович**



Доцент СПб ГУТ им. проф. М. А. Бонч-Бруевича, докторант кафедры «Информационные управляющие системы». Окончил СПб ГУТ в 1977 г. Диссертацию на соискание ученой степени кандидата технических наук защитил в 1987 г.
Автор более 80 печатных научных и научно-методических работ.
Область научных интересов — методы моделирования центров коммутации сетей связи.

**ЕРОШ
Игорь
Львович**



Профессор кафедры вычислительных систем и сетей СПбГУАП.
Окончил в 1960 г. Ленинградский электротехнический институт (ЛЭТИ).
Диссертацию на соискание ученой степени доктора технических наук защитил в 1979 г.
Автор более 150 научных трудов, в том числе соавтор двух учебников и трех монографий.
Член-корреспондент Международной академии информатизации.
Области научных интересов: дискретная математика; распознавание образов; защита информации.

**ТУКАБАЕВ
Павел
Тамьянович**



Директор Новороссийского филиала Современного гуманитарного института.
В 1975 году закончил Военно-медицинскую академию им. С. М. Кирова и в 1980 году — Северо-западный политехнический институт.
Кандидат медицинских наук.
Автор 15 научных работ, в том числе — монографии.
Научные интересы — проблемы самосовершенствования и саморазвития человека, взаимодействия биологических и информационных систем.

**ЛУКМАНОВ
Юнис
Халитович**



Глава территориального управления Адмиралтейского административного района. Доцент экономического факультета Санкт-Петербургского государственного университета.
Имеет три высших образования по специальностям: строительство, юриспруденция, государственное и муниципальное управление.
Диссертацию на соискание ученой степени кандидата экономических наук защитил в 1998 г.
Почетный строитель РФ.
Область научных интересов — управление инвестиционными процессами.

Редакция журнала напоминает, что ответственность за подбор, достоверность и точность фактов, экономико-статистических и технических показателей, собственных имен и прочих сведений, а также за то, что в материалах не содержится сведений, не подлежащих открытой публикации, несут авторы публикуемых в журнале материалов и рекламодатели.

УДК 621.865.8

Применение спектральных характеристик для распознавания изображений, полученных при дистанционном зондировании земной поверхности.

Соловьев Н. В. — Информационно-управляющие системы, 2003. — № 2–3. — С. 2–7.

В работе рассмотрены некоторые проблемы распознавания изображений, полученных при дистанционном зондировании земной поверхности с помощью многополосных спектрометров. Основное внимание уделено вопросам снижения размерности пространства признаков при классификации аномалий и определению их геометрических параметров при наличии проективных искажений. Предложен метод идентификации реперных точек по их спектральным характеристикам и рассмотрены особенности его компьютерной реализации.

Список лит.: 9 назв.

УДК 621.865.8

Application of spectral characteristics for the identification of the images obtained during remote sensing of a terrestrial surface

Solovjov N. V. — IUS, 2003. — № 2–3. — P. 2–7.

Some problems of the identification of the images obtained during remote sensing of a terrestrial surface with using of multiline spectrometers are considered in this work. The main attention is given to the problems of reduction of dimensionality of space of indications for classifications of anomalies and calculation of their geometric parameters with availability of projective distortions. The method of identification of reference points through their spectral reflectance is offered and the features of its computer realization are considered.

Refs: 9 titles.

УДК 681.007:61.1

Распознавание объектов с динамическими характеристиками

Яковенко М. К. — Информационно-управляющие системы, 2003. — № 2–3. — С. 8–15.

Распознавание объектов с динамическими характеристиками — сложная задача, обусловленная высокой вариабельностью их ключевых параметров. Основная проблема — это выявление эффективно-го множества признаков, способных описать объект в достаточной мере для корректного распознавания с приемлемым уровнем. В статье в качестве примера объектов с динамическими характеристиками рассмотрены лейкоциты человека (пять основных классов).

Список лит.: 19 назв.

УДК 681.007:61.1

Recognition method for objects with dynamic characteristics

Yakovenko M. K. — IUS, 2003. — № 2–3. — P. 8–15.

Recognition of objects with dynamic characteristics is hard task because of very high variability of their key parameters. Main problem is extraction of effective set of features that will describe target objects enough for classification with acceptable recognition rate. In this paper is considered good example of objects with dynamic characteristics — human white blood cell (five main classes).

Refs: 19 titles.

УДК 681.3:518.5

Гибридный разрядный метод решения систем уравнений в целочисленной арифметике

Сергеев М. Б. — Информационно-управляющие системы, 2003. — № 2–3. — С. 16–18.

Предложен гибридный разрядный метод решения целочисленных систем линейных алгебраических уравнений, гарантирующий абсолютную точность и реализуемый простейшими арифметическими операциями. Рассмотрен класс решаемых систем уравнений, оценена эффективность программной реализации в микропроцессорных системах.

Список лит.: 6 назв.

УДК 681.3:518.5

A hybrid bit-wise method for solution of systems of equations on integer arithmetic

Sergeev M. B. — IUS, 2003. — № 2–3. — P. 16–18.

A hybrid bit-wise method for solution of integer systems of linear algebraic equations is presented. The method guarantees absolute accuracy of the decision and uses the elementary arithmetic operations. The class of solved systems of the equations is considered and efficiency of program solution in microprocessor systems is evaluated.

Refs: 6 titles.

УДК 681.324

Комплексное моделирование процессов управления структурной динамикой информационной системы

Соколов Б. В., Малюгин К. А. — Информационно-управляющие системы, 2003. — № 2–3. — С. 19–29.

Рассматриваются методологические и методические основы организации и проведения комплексного моделирования процессов управления структурной динамикой информационной системы. Предложен оригинальный полимодельный комплекс, описывающий процессы функционирования и развития информационной системы и включающий в себя динамические аналитико-имитационные модели управления операциями, потоками, ресурсами, структурами информационной системы. Рассмотрен пример использования указанного комплекса

Список лит.: 19 назв.

УДК 519.872+519.876.5

Оценка системы моделирования GPSS WORLD
Рыжиков Ю. И. — Информационно-управляющие системы, 2003. — № 2–3. — С. 30–38.

Дан обзор проблемы математического моделирования систем и сетей с очередями. Описаны новые возможности, предоставляемые имитационной системой общего назначения GPSS World. Приводятся результаты ее тестирования. Предлагается комбинированное (в зависимости от сложности задачи, временных ресурсов и квалификации разработчика) использование численных методов теории очередей, программирования на языках типа Фортран и упомянутой системы.

Список лит.: 13 назв.

УДК 577.4

Математические методы внешнего проектирования сложных систем

Смирнов Ю. М., Поляков А. О., Однобокоев В. В. — Информационно-управляющие системы, 2003. — № 2–3. — С. 39–44.

Предлагается рассмотрение особенностей постановки и решения задач оптимального распределения ресурсов. Предложены и обоснованы алгоритмы многошагового улучшения составляющих интегрального показателя. Определены требования к характеру зависимостей интегрального показателя от частных и частных показателей от затрат на их улучшения, обеспечивающие сходимость и оптимальность магистрального решения. Дается схема оптимального распределения ресурсов.

Список лит.: 6 назв.

УДК 681.324

Integrated modeling of information system structure dynamics control

Sokolov B. V., Malyugin K. A. — IUS, 2003. — № 2–3. — P. 19–29.

Within the proposed study the methodological and methodical bases for the problem of integrated modeling of information system structure dynamics control. An original multiple-model complex for simulation of information system (IS) was proposed. The complex included analytical-simulation models of control processes for IS operation, flows, resources and structures. The utilization example of multiple-model complex was outlined

Refs: 19 titles.

УДК 519.872+519.876.5

Evaluating of modelling system of GPSS WORLD
Ryzhikov Y. I. — IUS, 2003. — № 2–3. — P. 30–38.

It is presented a review of the mathematical modeling of the queuing systems and networks, and the description of new possibilities of the GPSS World. The results of its testing are discussed. The combined (depending on the problem complexity, needed precision, time resources and explorer qualification) using of numerical queuing theory methods, programming in Fortran and GPSS modeling is proposed.

Refs: 13 titles.

УДК 577.4

Mathematical Methods of external design of Complex Systems

Smirnov Y. M., Polyakov A. O., Odobokov V. V. — IUS, 2003. — № 2–3. — P. 39–44.

The particularities of formulation and solving the problems of optimum resource distribution are marked out. The algorithms of multi-steps improvement of integral index's components are proposed and grounded. There are determined requirements to the nature of dependencies of integral index from particular factors and particular ones from costs for their improvement, ensuring convergence and optimum of a main solution. The scheme of optimum distribution is given.

Refs: 6 titles.

UDK 621-52:004.52; 629.78; 681.3

Многослойная перцептронная нейронная сеть в задаче моделирования речевых сигналов
Изилов Я. Ю. — Информационно-управляющие системы, 2003. — № 2–3. — С. 44–50.

Рассматриваются возможности применения искусственной нейронной сети перцептронного типа в задаче моделирования речевых сигналов. Приводятся структуры и алгоритм, позволяющий осуществить функционирование многослойной перцептронной нейросети на персональном компьютере. Обсуждаются проблемы, возникающие при практической реализации, а также достоинства и недостатки использования данного подхода.

Список лит.: 20 назв.

UDK 621-52:004.52; 629.78; 681.3

Multilayer perceptron neural network for task of speech signals modelling
Izilov J. Y. — IUS, 2003. — № 2–3. — P. 44–50.

This article considers the opportunities of application artificial neural network based on perceptron type for speech signals modeling task. Structures and the algorithm, which allows to carry out functioning of multilayer perceptron neural network on personal computer, are resulted. The problems arising at practical realization, and also limits and advantages of the given approach usage are discussed.

Refs: 20 titles.

UDK 681.325.5

Взаимодействие открытых систем промышленной автоматизации — состояние и проблемы
Чьонг Динь Тяу — Информационно-управляющие системы, 2003. — № 2–3. — С. 52–57.

Рассматривается технология OPC (OLE for Process Control) — основной стандарт взаимодействия средств современных систем сбора данных и диспетчерского управления — Supervisory Control and Data Acquisition (SCADA), затрагивается механизм обмена данными между OPC-приложениями вопрос производительности OPC-серверов, сравнивается OPC с другими стандартами, перечисляются преимущества и недостатки стандарта. В качестве задачи, подлежащей решению, отмечается создание универсального OPC-сервера.

Список лит.: 9 назв.

UDK 681.325.5

Open-system interconnection in industrial automation — condition and problems
Truong Dinh Chau — IUS, 2003. — № 2–3. — P. 52–57.

This article examines the OLE for Process Control (OPC) technology — the basic standard of systems interaction in modern Supervisory Control and Data Acquisition (SCADA) system. The mechanism of data communication between OPC-applications and performance of OPC servers are mentioned. The article compares OPC by other standards. The advantages and disadvantages of the standard are listed. As a task subject to the decision, the creation universal OPC server is marked.

Refs: 9 titles.

UDK 621.391.28

Формализованное описание процесса функционирования центров обработки информации и управления интеллектуальных сетей для целей оценки вероятностно-временных характеристик
Колбанёв М. О. — Информационно-управляющие системы, 2003. — № 2–3. — С. 58–62.

Приводятся требования к методам описания процессов функционирования центров обработки информации и управления, ориентированным на построение математических моделей оценки вероятностно-временных характеристик. Излагаются основы метода составления описания, отвечающего этим требованиям.

Список лит.: 5 назв.

UDK 621.391.28

Formalized description of process of functioning of centres of information processing and switching of intellectual networks for estimation of time and probabilistic characteristics
Kolbanev M. O. — IUS, 2003. — № 2–3. — P. 58–62.

The methods of the description of the processes functioning of centers switching and information processing focused on construction of mathematical models to an estimation of time and probabilistic characteristics are considered. The bases of a method of drawing up of the description are stated.

Refs: 5 titles.

УДК 681.391.1

Разграничение доступа к ресурсам в системах коллективного пользования

Ерош И. Л. — Информационно-управляющие системы, 2003. — № 2–3. — С. 63–66.

В статье рассмотрена задача на разграничение доступа в системах коллективного пользования, основанная на булевых преобразованиях двоичных последовательностей. Для обеспечения доступа используется метод сравнения запроса с ответом. Приведены примеры построения булевых функций, применение которых позволяет гибко менять доступ для каждой группы пользователей в зависимости от выбираемых запросов.

Приведенные примеры иллюстрируют метод распределенного доступа в системах коллективного пользования, который может быть применен для управления доступом к вычислительным ресурсам компьютерной сети, для обеспечения доступа с ограничениями в различные помещения и др.

Список лит.: 3 назв.

УДК 681.391.1

Differentiation of access to resources in shared systems
Erosh I. L. — IUS, 2003. — № 2–3. — P. 63–66.

In the article task on differentiation of access in shared systems based on Boolean transformations of binary sequences is considered. For providing of access the method of comparison of inquiry with the answer is used. The examples of construction Boolean of functions are given, in which access for each group of the users flexibly varies depending on chosen inquiries.

The given examples illustrate a method of the distributed access in shared systems, which can be applied for management of access to computing resources of a computer network, for providing of access with restrictions in various premises etc.

Refs: 3 titles.

УДК 612.821

Информационные проблемы организации обратной связи при взаимодействии биологических и технических систем

Поляков А. О., Тукабаев П. Т. — Информационно-управляющие системы, 2003. — № 2–3. — С. 67–72.

Предлагается рассмотрение возможности создания новых подходов к организации обратной связи при взаимодействии биологических («био-социальных») и технических компонент высокого уровня организационной сложности. В основу исследования положена необходимость контекстно-зависимого и структурного представления систем для повышения организации уровня управления. Указывается возможность и целесообразность введения понятия биотехнической обратной связи как совокупности сигнального и структурного обмена информацией между компонентами сложной системы.

Список лит.: 11 назв.

УДК 612.821

Informational Problems in Arranging for a Feedback at Interaction between Biological and Technical Systems

Polyakov A. O., Tukabaev P. T. — IUS, 2003. — № 2–3. — P. 67–72.

It is proposed to consider a possibility of developing new approaches to arranging for a feedback at the interaction between biological («bio-social») and technical components of high organizational complexity. The research is based upon the necessity to implement context-dependent and structural representations of the systems to receive a new higher level in control organization. Possibility and advantage of introducing new notions for biotechnical feedback as an aggregate of signal and structural information exchange between the complex system components are indicated.

Refs: 11 titles.

УДК 612(075)

Автоматизированная система измерения динамических характеристик параметров изображения газоразрядного свечения

Коротков К. Г., Крыжановский Э. В., Муромцев Д. И., Бабицкий М. А., Борисова М. Б. — Информационно-управляющие системы, 2003. — № 2-3. — С. 73-79.

Статья посвящена изучению изменения во времени характеристик изображений газового разряда, получаемых в электромагнитном поле высокой напряженности вокруг капель жидкости. Для описания газоразрядных изображений (ГРВ-грамм) были предложены несколько параметров: площадь свечения, средняя интенсивность свечения, средний радиус, нормализованное среднеквадратическое отклонение радиуса и коэффициент формы. Данный метод исследования объектов получил название динамической ГРВ-графики. Выявлены статистически значимые различия между образцами жидкостей, различия между которыми не обнаруживались при изучении их одиночных (статических) ГРВ-грамм. Построена математическая модель для описания поведения трендов параметров ГРВ-грамм.

Список лит.: 5 назв.

УДК 612(075)

The Automated system of measurement of dynamic characteristics of parameters of the image of a gas discharge glow

Korotkov K. G., Krizanovski E. V., Mouromtsev D. I., Babitski M. A., Borisova M. B. — IUS, 2003. — № 2-3. — P. 73-79.

This article devoted to the study of changing of the characteristics of gas discharge glow images (GDV images) around the drop of liquid in the strong electromagnetic field. Several parameters to describe the GDV images were proposed: Glow Area, Average Intensity, etc. This approach was named "Dynamic GDV-graphy". The statistically significant differences were revealed between different samples of liquids in the case when conventional "static" GDV-graphy was unable to reveal ones. The mathematical model of behavior of the time series trends of the GDV-parameters was suggested.

Refs: 5 titles.

УДК 338.24

Управление инвестициями в российских регионах на основе концепции «разумного роста»

Лукманов Ю. Х. — Информационно-управляющие системы, 2003. — № 2-3. — С. 80-88.

Статья посвящена проблеме инновационного управления экономическими процессами в регионах. На основе анализа статистического и методического материала проведен анализ факторов, влияющих на экономический рост промышленности. Автором предлагаются к рассмотрению основные принципы создания системы привлечения капитала в регион на основе концепции «разумного роста».

Список лит.: 5 назв.

УДК 338.24

Management of investments on the basis of the concept of «reasonable growth» in regions of Russia
Lukmanov U. H. — IUS, 2003. — № 2-3. — P. 80-88.

The issue of the article is innovation of control over the economic processes at the local level with using of statistics and methodology material of the analysis of factors affecting economic growth of the industry. The author suggests basic principles of investments into the local economy, basing on the concept of «reasonable growth».

Refs: 5 titles.

Уважаемые авторы журнала

«ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ»

*При подготовке рукописей статей
редакция просит вас руководствоваться следующими рекомендациями*

Редакция принимает к публикации статьи в следующие тематические разделы: обработка информации и управление; информационно-управляющие системы; моделирование систем и процессов; программные и аппаратные средства; защита информации; системный анализ; стохастическая динамика и хаос; управление в социально-экономических системах; управление в медицине и биологии; краткие научные сообщения; рецензии (на книги, журналы, статьи, диссертации); хроника и информация (о семинарах, конференциях, выставках, юбилеях, а также очерки по истории науки и техники).

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word (не более 70 знаков в строке) шрифтом Times New Roman размером 13.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, полное название организации, аннотация (5–10 строк).

В редакцию предоставляются:

- отпечатанный (формат А4) текст статьи, подписанный всеми авторами с указанием даты представления, и иллюстрации, пронумерованные с подрисуночными подписями (в двух экземплярах);
- статья в виде файла Microsoft Word или Adobe PageMaker 6.0 (шрифт Times New Roman размер 12) на дискетах 1,44Mb или CD (иллюстрации в текст не заверстываются и предоставляются отдельными файлами *.tif, *.bmp, *.jpg с максимальным разрешением);
- аннотация (5–10 строк) на русском и английском языках;
- сведения об авторах (фамилия, имя, отчество, ученая степень, ученое звание, место работы, должность, области научных интересов, домашний и служебный адреса и телефоны, факс, e-mail), фото авторов (можно в электронном виде с максимальным разрешением);
- экспертное заключение (при необходимости).

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

- для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее число страниц;
- для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;
- ссылки на иностранную литературу следует давать на языке оригинала без сокращений.

Адрес редакции:

191011, Санкт-Петербург, Инженерная ул., 6
ФГУП «Издательство „Политехника“».
Редакция журнала «Информационно-управляющие системы».
Тел./факс 312-53-90,
E-mail: asklab@aanet.ru, gfm@polytechics.spb.ru