ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 23
2017
№ 2

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

Информация о журнале доступна по сети Internet по адресу **http://novtex.ru/IT.**
Журнал включен в систему Российского индекса научного цитирования и базу данных RSCI на платформе Web of Science.
Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

INFORMATION TECHNOLOGIES
INFORMACIONNYE TEHNOLOGII

Vol. 23
2017
No. 2

THEORETICAL AND APPLIED SCIENTIFIC AND TECHNICAL JOURNAL

Published since November 1995

ISSN 1684-6400

CONTENTS

Complete Internet version of the journal at site: **http://novtex.ru/IT.**
According to the decision of the Higher Certifying Commission of the Ministry of Education of Russian Federation, the journal is inscribed in "The List of the Leading Scientific Journals and Editions wherein Main Scientific Results of Theses for Doctor's or Candidate's Degrees Should Be Published"

УДК 004.942 DOI:10.17587/it.23.83-89

О. М. Полещук, д-р техн. наук, проф., e-mail: olga.m.pol@yandex.ru, poleshchuk@mgul.ac.ru,
Мытищинский филиал МГТУ им. Н. Э. Баумана

Использование нечетких логических функций для поддержки принятия решений по результатам рейтингового оценивания

Для поддержки принятия решений по результатам рейтингового оценивания объектов разработан подход на основе нечетких логических функций. Этот подход позволяет разбить рейтинговые оценки на кластеры, каждому из которых соответствует управляющее воздействие, направленное на успешное функционирование объектов в будущем.

Ключевые слова: *рейтинговые оценки, лингвистические переменные, нечеткие логические функции, поддержка принятия решений*

Введение

Рейтинговые оценки, широко используемые в различных областях деятельности человека (образование, техника, экономика, экология и т.д.), позволяют получать доступную и своевременную информацию в виде некого интегрального показателя, который используется для принятия управленческих решений. Существует ряд сложностей получения рейтинговых оценок, который подробно рассмотрен в работе [1]. Эти сложности связаны с разнородностью характеристик, неустойчивостью конечных результатов при использовании определенных шкал, а также с распознаванием полученных результатов в целях выработки управляющих воздействий. Для распознавания рейтинговых оценок определяют некоторые пороговые значения, которые разбивают всю область значений этих оценок на интервалы (полуинтервалы, отрезки и т.д.). При попадании значений рейтинговых оценок в определенные интервалы применяют соответствующие управляющие воздействия. Задача получения пороговых значений решается, как правило, опытным путем или прямым опросом экспертов, что не всегда доступно. Возможно отсутствие апостериорной статистической информации или отсутствие четкости в процедуре оценивания, что является причиной существенных трудностей экспертов и приводит к ошибочным результатам. Кроме этого, всегда существуют зоны неопределенности значений рейтинговых оценок, которые усложняют принятие решений о выборе управляющего воздействия. Как правило, это зоны, расположенные вблизи пороговых значений, или зоны так называемых "средних значений", для которых сложно выбрать управляющее воздействие ввиду неоднозначности ситуации.

Использование лингвистических переменных позволило определять рейтинговые оценки в условиях разнородных характеристик и избегать некорректных арифметических операций, свойственных традиционным рейтинговым моделям [2]. Однако в решении задачи распознавания рейтинговых оценок в целях выработки управляющих воздействий существуют пробелы, о которых речь шла выше. Причиной этих пробелов является отсутствие формализованного подхода, который позволил бы снизить ошибки экспертов вследствие неполноты или нечеткости информации. В связи с этим актуально дальнейшее исследование, направленное на ликвидацию существующих пробелов. В работе предложен подход к распознаванию рейтинговых оценок и поддержке принятия управляющих решений на основе нечетких логических функций, которые были разработаны и адаптированы для сред с нечеткими условиями и целями [3—5].

1. Построение нечетких функций *k*-значной логики

Рассмотрим характеристики X_j, $j = \overline{1, m}$, каждой из которых поставлено в соответствие некоторое множество значений X_{lj}, $l = \overline{1, m_j}$, $j = \overline{1, m}$, характеризующих их состояние. Будем считать эти характеристики подчиненными характеристике Y с множеством значений Y_l, $l = \overline{1, k}$, если Y приписан некоторый оператор агрегирования информации, позволяющий на основе значений характеристик X_j, $j = \overline{1, m}$ вычислять значения Y. Оператор агрегирования информации (ОАИ) O_Y есть функция, определенная на множестве всех возможных значений X_j, $j = \overline{1, m}$, и принимающая значения на множестве Y_l, $l = \overline{1, k}$:

$$O_Y : X_{1l_1} \times X_{2l_2} \times \ldots \times X_{ml_m} \to Y_l.$$

Исторически первым подходом к выбору ОАИ является геометрический подход, который основан

на представлении оператора как некоторой поверхности в $(m + 1)$-мерном пространстве. Недостатками этого подхода являются необходимость знать значение оператора агрегирования, хотя бы на $(m + 1)$ наборе значений характеристик X_j, $j = \overline{1, m}$, и невозможность использования дополнительной информации от экспертов о его поведении.

Логический подход к выбору ОАИ применим, когда возможна формулировка некоторых условий на поведение оператора O_Y. Если имеется k значений характеристики Y, то мы можем представить оператор агрегирования информации как некоторую функцию k-значной логики. Если число подчиненных характеристик X_j равно m, то в качестве оператора агрегирования может быть использована одна из функций k-значной логики от m переменных. Если эксперт может сформулировать нечеткие условия на поведение искомой функции типа "При сильном возрастании первого аргумента значение функции слегка убывает", "При совместном возрастании аргументов 3 и 5 значение функции сильно возрастает" и т.п., мы можем говорить о нечетких функциях k-значной логики.

Нечеткая логическая функция является расширением понятия обычной функции k-значной логики от m переменных и определяется на нечетких переменных с использованием нечетких условий на ее поведение.

Нечеткой переменной называется тройка

$$\{X, U, \widetilde{A}\},$$

где X — название переменной; U — область ее определения (универсальное множество); $\widetilde{A}$ — нечеткое множество универсального множества, описывающее возможные значения нечеткой переменной.

Согласно [3] нечеткие условия S представляются в виде некоторого нечеткого отношения $\widetilde{S}$.

Нечетким бинарным отношением $\widetilde{S}$ между множествами X, Y называется нечеткое множество $\widetilde{S}$ такое, что $\forall(x, y) \in X \times Y$, $\mu_{\widetilde{S}}(x, y) \in [0, 1]$, $X = \{x\}$, $Y = \{y\}$ — обычные множества.

Если множества X, Y конечны $X = \{x_1, x_2, ..., x_n\}$, $Y = \{y_1, y_2, ..., y_m\}$, то нечеткое бинарное отношение $\widetilde{S}$ может быть задано с помощью его матрицы, строкам и столбцам которой ставят в соответствие элементы множеств, а на пересечении i-й строки и j-го столбца помещается элемент $\mu_{\widetilde{S}}(x_i, y_j)$. Таким образом

$$\widetilde{S} = \begin{pmatrix} \mu_{\widetilde{S}}(x_1, y_1) & \mu_{\widetilde{S}}(x_1, y_2) & \dots & \mu_{\widetilde{S}}(x_1, y_m) \\ \mu_{\widetilde{S}}(x_2, y_1) & \mu_{\widetilde{S}}(x_2, y_2) & \dots & \mu_{\widetilde{S}}(x_2, y_m) \\ \dots & \dots & \dots & \dots \\ \mu_{\widetilde{S}}(x_n, y_1) & \mu_{\widetilde{S}}(x_n, y_2) & \dots & \mu_{\widetilde{S}}(x_n, y_m) \end{pmatrix}.$$

Нечетким бинарным отношением $\widetilde{S}$ на множестве X называется нечеткое множество $\widetilde{S}$ такое, что $\forall(x, y) \in X \times X$ $\mu_{\widetilde{S}}(x, y) \in [0, 1]$.

Рассмотрим одно нечеткое условие S на поведение функции f от одной переменной. Нечеткое отношение $\widetilde{S}$, которое соответствует нечеткому условию S, описывает принадлежность функции к определенному классу (например, слегка возрастающие или слегка убывающие функции) на основе значений функции в точках i и $i + 1$, $0 \leqslant i \leqslant k - 1$. Значение $\mu_{\widetilde{S}}(p, q)$ есть степень принадлежности функции к данному классу при условии, что $f(i) = p$, $f(i + 1) = q$, $0 \leqslant p, q \leqslant k - 1$. Матрица нечеткого отношения $\widetilde{S}$, соответствующего нечеткому условию S, имеет вид

$$\widetilde{S} = (\mu_{\widetilde{S}}(p, q)).$$

Пусть на поведение функции наложено несколько нечетких условий — S^r, $r = \overline{1, s}$. Каждому из этих условий соответствует матрица нечеткого отношения $\widetilde{S}^r$, $r = \overline{1, s}$. Матрица, которая обобщает все условия получается на основе T-нормы:

$$\widetilde{\widetilde{S}}^s = \mathop{T}_{r=1}^{s} \widetilde{S}^r.$$

Треугольной нормой (T-нормой) называется действительная двухместная функция T: $[0,1] \times [0,1] \to \to [0,1]$, удовлетворяющая следующим условиям:

1) $T(0,0) = 0$, $T(\mu_{\widetilde{A}}, 1) = T(1, \mu_{\widetilde{A}}) = \mu_{\widetilde{A}}$ (ограниченность);

2) $T(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}) \leqslant T(\mu_{\widetilde{C}}, \mu_{\widetilde{D}})$, если $\mu_{\widetilde{A}} \leqslant \mu_{\widetilde{C}}$, $\mu_{\widetilde{B}} \leqslant \mu_{\widetilde{D}}$ (монотонность);

3) $T(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}) = T(\mu_{\widetilde{B}}, \mu_{\widetilde{A}})$ (коммутативность);

4) $T(\mu_{\widetilde{A}}, T(\mu_{\widetilde{B}}, \mu_{\widetilde{C}})) = T(T(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}), \mu_{\widetilde{C}})$ (ассоциативность).

Если матрица отношения $\widetilde{\widetilde{S}}^s$ имеет хотя бы одну нулевую строку, то множество условий S^r, $r = \overline{1, s}$, является противоречивым, так как функция f удовлетворяет им с нулевой степенью. Если система условий S^r, $r = \overline{1, s}$ является противоречивой, то предлагается выделить непротиворечивые подсистемы согласно следующему алгоритму. На противоречивость проверяют все пары нечетких условий системы. Найденные противоречивые пары удаляют из рассмотрения. Далее на противоречивость проверяют все тройки нечетких условий (в которые не входят противоречивые пары). Найденные противоречивые тройки удаляют из рассмотрения. Эта операция повторяется, пока на шаге l, $1 \leqslant l \leqslant s$ не окажется, что все подсистемы, состоящие из $l + 1$ нечетких условий, противоречивы. Тогда непротиворечивыми подсистемами окажутся подсистемы, выписанные на шаге $l - 1$ и состоящие из l нечетких условий. Таким образом, любое количество нечетких условий по одной переменной формально легко сводится к одному нечеткому условию по этой переменной.

Рассмотрим функции k-значной логики от m переменных и множество нечетких условий S. Для простоты изложения положим $m = 2$ и $|S| = 2$. Пусть первое нечеткое условие S_1 определено по первой переменной, второе условие S_2 — по второй перемен-

ной. Построим матрицы отношений $\widetilde{S}_1$ и $\widetilde{S}_2$. Удовлетворение условиям S означает одновременное удовлетворение условиям S_1 и S_2. Это, в свою очередь означает, что на наборе (i_1, i_2) значений переменных x_1 и x_2 ($i_1, i_2 \in \{0, 1, ..., k-1\}$) мы в качестве значения строки матрицы отношения $\widetilde{\widetilde{S}}$ должны взять T-норму $(i_1 + 1)$-й строки матрицы отношения $\widetilde{S}^1$ и $(i_2 + 1)$ строки $\widetilde{S}^2$. Полученная таким образом матрица будет искомой. Данная матрица обладает всеми свойствами матриц отношений для одной переменной.

Будем предполагать, что на поведение функции накладывается нечеткое условие S и дополнительно некоторое начальное условие (например, $f(0) = 0$). Нечеткое отношение $\widetilde{S}$, формализующее нечеткое условие S, берется за основу для построения нечеткого отношения $\widetilde{\widetilde{S}}$, формализующего оба условия на поведение функции. Согласно [3, 4]

$$\mu_{\widetilde{\widetilde{S}}}(l, j) = \underset{i=1}{\overset{k}{K}} \mu_{\widetilde{\widetilde{S}}}(l-1, i) \times \mu_{\widetilde{S}}(i, j),$$
$$(1 \leqslant j \leqslant k,\ 2 \leqslant l \leqslant k)$$

при начальном условии $f(0) = 0$. Первой строкой матрицы нечеткого отношения $\widetilde{\widetilde{S}}$, соответствующего значению аргумента 0, является строка (1, 0, ..., 0) — как выражение начального условия. Элементы $\widetilde{\widetilde{S}}$ получаются своеобразным умножением предыдущей строки матрицы отношения $\widetilde{\widetilde{S}}$ на столбцы матрицы отношения $\widetilde{\widetilde{S}}$, где вместо операции сложения используется операция взятия треугольной конормы K.

Треугольной конормой K называется действительная двухместная функция K: $[0,1] \times [0,1] \to [0,1]$, удовлетворяющая следующим условиям:

1) $K(1,1) = 1$, $K(\mu_{\widetilde{A}}, 0) = K(0, \mu_{\widetilde{A}}) = \mu_{\widetilde{A}}$ (ограниченность);

2) $K(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}) \geqslant K(\mu_{\widetilde{C}}, \mu_{\widetilde{D}})$, если $\mu_{\widetilde{A}} \geqslant \mu_{\widetilde{C}}$, $\mu_{\widetilde{B}} \geqslant \mu_{\widetilde{D}}$ (монотонность);

3) $K(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}) = K(\mu_{\widetilde{B}}, \mu_{\widetilde{A}})$ (коммутативность);

4) $K(\mu_{\widetilde{A}}, K(\mu_{\widetilde{B}}, \mu_{\widetilde{C}})) = K(K(\mu_{\widetilde{A}}, \mu_{\widetilde{B}}), \mu_{\widetilde{C}})$ (ассоциативность).

Предположим, что начальное условие сформулировано не для $f(0)$, а для $f(k-1)$. Пусть, для определенности, это будет условие $f(k-1) = q$, $(1 \leqslant q \leqslant k-1)$, которое определяет k-ю строку матрицы отношения $\widetilde{\widetilde{S}}$. Строка содержит одни нули за исключением 1 в q-м столбце. Для остальных строк матрицы $\widetilde{\widetilde{S}}$

$$\mu_{\widetilde{\widetilde{S}}}(l, j) = \underset{i=1}{\overset{k}{K}} \mu_{\widetilde{\widetilde{S}}}(l+1, i) \times \mu_{\widetilde{S}}(j, i),$$
$$(1 \leqslant j \leqslant k,\ 2 \leqslant l \leqslant k-2).$$

Предположим, что начальное условие сформулировано для некоторого промежуточного значения $f(l^*)$, $(1 < l^* < k-1)$. Пусть, для определенности, это будет условие $f(l^*) = q$, $(1 < q < k-1)$, которое определяет l — строку матрицы отношения $\widetilde{\widetilde{S}}$. Строка содержит одни нули за исключением 1 в q-м столбце. Общий метод построения матрицы отношения $\widetilde{\widetilde{S}}$ определяется следующей формулой:

$$\mu_{\widetilde{\widetilde{S}}}(l, j) =$$
$$= \begin{cases} \underset{i=1}{\overset{k}{K}} \mu_{\widetilde{\widetilde{S}}}(l+1, i) \times \mu_{\widetilde{S}}(j, i) \text{ при } 1 \leqslant l \leqslant l^*, \\ \underset{i=1}{\overset{k}{K}} \mu_{\widetilde{\widetilde{S}}}(l-1, i) \times \mu_{\widetilde{S}}(i, j) \text{ при } l^* < l \leqslant k, \end{cases}$$
$$(1 \leqslant l, j \leqslant k).$$

Пусть для функции одной переменной задано одно нечеткое условие и t начальных условий. Таким образом, если начальных условий несколько, то матрица $\widetilde{\widetilde{S}}_i$ строится для каждого условия i в отдельности $(1 < i \leqslant t)$, а итоговая матрица $\widetilde{\widetilde{S}}^1$ получается следующим образом:

$$\widetilde{\widetilde{S}}^1 = \underset{i=1}{\overset{t}{T}} \widetilde{\widetilde{S}}_i.$$

В результате формализации всех условий, налагаемых на поведение функций k-значной логики, получается матрица нечеткого отношения. Функции, определяемые построенным нечетким отношением, применяют к реальной задаче. Полученную апостериорную информацию о результатах работы каждой функции сравнивают с априорной информацией, полученной от экспертов. Функции, для которых сравнительный анализ привел к противоречию, отбрасывают. Оставшиеся функции и есть искомые. Если таких функций нет, то с экспертами уточняют налагаемые на поведение функций условия и повторяется вся процедура построения.

2. Рейтинговые оценки и нечеткие логические функции

Рассмотрим N объектов, у которых оцениваются качественные характеристики X_j, $j = \overline{1, m}$, соответственно со значениями X_{lj}, $l = \overline{1, m_j}$, $j = \overline{1, m}$. Будем считать, что характеристики X_j, $j = \overline{1, m}$, оказывают существенное влияние на характеристику Y (со значениями Y_l, $l = \overline{1, k}$) — успешное функционирование объектов в будущем.

По результатам рейтингового оценивания объектов в рамках характеристик X_j, $j = \overline{1, m}$, необходимо разработать управляющие воздействия на характеристику Y. Для определения рейтинговых оценок объектов результаты оценивания характеристик формализуют на основе полных ортогональных семантических пространств согласно методу работы [2].

Лингвистической переменной называется пятерка [6]

$$\{X,\ T(X),\ U,\ V,\ S\},$$

где X — название переменной; $T(X) = \{X_i, i = \overline{1, m}\}$ — терм-множество переменной X, т.е. множество термов или названий лингвистических значений переменной X (каждое из этих значений — нечеткая переменная со значениями из универсального

множества U); V — синтаксическое правило, порождающее названия значений лингвистической переменной X; S — семантическое правило, которое ставит в соответствие каждой нечеткой переменной с названием из $T(X)$ нечеткое подмножество универсального множества U.

Семантическим пространством называется лингвистическая переменная с фиксированным термножеством $\{X, T(X), U, S\}$.

Согласно [7] семантическое пространство называется полным ортогональным семантическим пространством (ПОСП), если функции принадлежности $\mu_l(x)$, $l = \overline{1, m}$, его термов удовлетворяют следующим требованиям

1. Для каждого понятия X_l, $l = \overline{1, m}$, существует $\widehat{U}_l \neq \varnothing$, где $\widehat{U}_l = \{x \in U: \mu_l(x) = 1\}$ есть точка или отрезок.

2. Пусть $\widehat{U}_l = \{x \in U: \mu_l(x) = 1\}$, тогда $\mu_l(x)$, $l = \overline{1, m}$, не убывает слева от $\widehat{U}_l$ и не возрастает справа от $\widehat{U}_l$.

3. $\mu_l(x)$, $l = \overline{1, m}$, имеют не более двух точек разрыва первого рода.

4. Для каждого $x \in U$ $\sum\limits_{l=1}^{m} \mu_l(x) = 1$.

Опираясь на метод работы [2], построим m ПОСП $X_j, j = \overline{1, m}$, с терм-множествами X_{lj}, $l = \overline{1, m_j}$, $j = \overline{1, m}$. Обозначим через $\mu_{lj}(x)$ функцию принадлежности нечеткого числа $\tilde{X}_{lj}$, соответствующего l-му терму j-го ПОСП, $l = \overline{1, m_j}$, $j = \overline{1, m}$. Обозначим через $\tilde{X}_j^n$ и $\mu_j^n(x) \equiv (a_{j1}^n, a_{j2}^n, a_{jL}^n, a_{jR}^n)$, $n = \overline{1, N}$, $j = \overline{1, m}$, оценку n-го объекта в рамках характеристики X_j. Нечеткое число $\tilde{X}_j^n$ с функцией принадлежности $\mu_j^n(x)$ равно одному из нечетких чисел $\tilde{X}_{lj}$, $l = \overline{1, m_j}$, $j = \overline{1, m}$. Первые два параметра в скобках — абсциссы соответственно левых и правых концов верхних оснований трапеций, которые являются графиками функций принадлежности, а вторые два параметра — длины соответственно левых и правых крыльев трапеций. Обозначим весовые коэффициенты оцениваемых характеристик через $\omega_j, j = \overline{1, m}$, $\sum\limits_{j=1}^{m} \omega_j = 1$.

Нечеткая рейтинговая оценка n-го объекта [8], $n = \overline{1, N}$, в рамках характеристик X_j, $j = \overline{1, m}$, определяется в виде нечеткого числа

$$\tilde{A}_n = \omega_1 \otimes \tilde{X}_1^n \otimes ... \otimes \omega_k \otimes \tilde{X}_m^n$$

с функцией принадлежности

$$\mu_n(x) \equiv \left(\sum_{j=1}^{m} \omega_j a_{j1}^n, \sum_{j=1}^{m} \omega_j a_{j2}^n, \sum_{j=1}^{m} \omega_j a_{jL}^n, \sum_{j=1}^{m} \omega_j a_{jR}^n \right),$$

$$n = \overline{1, N}.$$

Дефазифицируем нечеткие числа $\tilde{A}_n$, $n = \overline{1, N}$, $\tilde{B}_1 = \omega_1 \otimes \tilde{X}_{11} \oplus ... \oplus \omega_k \otimes \tilde{X}_{1m}$, $\tilde{B}_m = \omega_1 \otimes \tilde{X}_{m_1 1} \oplus \oplus ... \oplus \omega_k \otimes \tilde{X}_{m_m m}$ по методу центра тяжести [9] и полученные четкие числа обозначим через A_n, $n = \overline{1, N}$, B_1, B_m. Число A_n, $n = \overline{1, N}$, называется точечной рейтинговой оценкой проявления качественных характеристик X_j, $j = \overline{1, m}$, у n-го объекта $n = \overline{1, N}$.

Нормированную рейтинговую оценку n-го объекта, $n = \overline{1, N}$, найдем по формуле

$$E_n = \frac{A_n - B_1}{B_m - B_1}, \; n = \overline{1, N}.$$

Область изменения рейтинговой оценки E_n, $n = \overline{1, N}$, есть отрезок [0, 1].

Для выработки управляющих воздействий по результатам рейтингового оценивания будем использовать нечеткие логические функции. Для этого рассмотрим m переменных X_j, $j = \overline{1, m}$, а искомая функция будет принимать k значений (в соответствии с числом значений характеристики Y_l, $l = \overline{1, k}$). Таким образом, задача сводится к построению нечеткой функции k-значной логики от m переменных. Построенная функция позволит разбить рейтинговые оценки на k кластеров по числу значений характеристики Y. Каждому кластеру в соответствие ставится управляющее воздействие, направленное на успешное функционирование объектов в будущем.

На искомую функцию накладываются начальные условия и нечеткие условия на ее поведение. Построение таких функций изложено выше и на конкретном примере будет продемонстрировано в следующем разделе.

3. Выработка управляющих рекомендаций, направленных на обеспечение коммерческой успешности программных продуктов

Для исследования были отобраны двенадцать программных продуктов, разрабатываемых для автоматизации розничного бизнеса, банковской деятельности, страхового бизнеса и внутрихозяйственного учета. Разрабатываемые продукты были использованы потребителями в тестовом режиме. В качестве входных характеристик программных продуктов были взяты три характеристики: X_1 — модифицируемость; X_2 — изучаемость; X_3 — функциональность.

Модифицируемость — это характеристика программного продукта, которая упрощает внесение в него необходимых изменений и доработок и включает в себя характеристики расширяемости, структурированности и модульности. *Изучаемость* — это характеристика, которая позволяет минимизировать усилия по изучению и пониманию программ и документации программных продуктов и включает в себя характеристики информативности, понят-

ности, структурированности и удобочитаемости. *Функциональность* — это характеристика, которая показывает способность программного продукта выполнять набор функций, определенных в его внешнем описании и удовлетворяющих заданным или подразумеваемым потребностям пользователей.

В качестве выходной характеристики рассматривалась успешность программных продуктов — Y, которая включает в себя популярность этих продуктов у потребителей, их продаваемость и признание специалистами-экспертами.

Все характеристики оценивали в рамках трех лингвистических значений — "низкая", "средняя", "высокая", которым в соответствие были поставлены баллы 0, 1, 2. По результатам тестового использования программных продуктов и их рейтингового оценивания была поставлена задача выработки управляющих воздействий, направленных на обеспечение успешности этих продуктов в будущем.

Результаты оценивания программных продуктов экспертами занесены в табл. 1.

Полученные данные были формализованы по методу работы [2] с помощью ПОСП. Функции принадлежности лингвистических значений "низкая", "средняя", "высокая" занесены в табл. 2. Если графиком функции принадлежности является трапеция, то функция определяется четырьмя параметрами. Первые два параметра — абсциссы соответственно левого и правого концов верхнего основания трапеции, а вторые два параметра — длины соответственно левого и правого крыльев трапеции. Если графиком функции принадлежности является треугольник, то функция определяется тремя параметрами. Первый параметр — абсцисса вершины треугольника, а вторые два параметра — длины соответственно левого и правого крыльев треугольника.

Для программных продуктов были вычислены рейтинговые оценки и полученные результаты занесены в табл. 3. Весовые коэффициенты $\omega_j, j = \overline{1, 3}$, по согласованию с экспертами взяты равными 1/3.

Полученные результаты оценивания и рейтинговые оценки программных продуктов были использованы для разработки управляющих рекомендаций, направленных на достижение успешности этих продуктов.

Обычно значения рейтинговых оценок по согласованию с экспертами разбиваются на несколько интервалов. При попадании рейтинговой оценки в определенный интервал применяется соответствующее управляющее воздействие. Будем считать, что таких интервалов три по числу значений выходной характеристики и разрабатываемых управляющих воздействий. Если рейтинговая оценка программного продукта попадает в первый интервал $[0, x)$, то успешность программного продукта Y низкая, и программный продукт нуждается в серьезных доработках. Если рейтинговая оценка программного продукта попадает в средний интервал $[x, y)$, то успешность программного продукта Y средняя, и программный продукт нуждается в незначительных доработках. Если рейтинговая оценка программного продукта попадает в последний интервал $[y, 1]$, то успешность программного продукта Y высокая, и программный продукт выпускается на рынок. Для выработки управляющих рекомендаций и определения границ интервалов рейтинговых оценок будем использовать нечеткую логическую функцию.

Нечеткая логическая функция F, зависящая от переменных X_1, X_2, X_3, принимает три значения: "успешность программного продукта низкая"; "успешность программного продукта средняя"; "успешность программного продукта высокая". Этим значениям в соответствие были поставлены значе-

Таблица 1

Результаты оценивания программных продуктов

n	X_1	X_2	X_3
1	0	1	0
2	0	0	0
3	1	0	2
4	1	1	1
5	2	0	2
6	2	1	1
7	0	1	1
8	1	0	1
9	1	2	0
10	1	2	0
11	0	0	1
12	1	1	0

Таблица 2

Формализованные данные оценивания программных продуктов

n	X_1	X_2	X_3
1	(0,0.15,0,0.3)	(0.375,0.425,0.25,0.35)	(0,0.125,0,0.25)
2	(0,0.15,0,0.3)	(0,0.125,0,0.25)	(0,0.125,0,0.25)
3	(0.45,0.55,0.3,0.3)	(0,0.125,0,0.25)	(0.85,1,0.3,0)
4	(0.45,0.55,0.3,0.3)	(0.375,0.425,0.25,0.35)	(0.375,0.55,0.25,0.3)
5	(0.85,1,0.3,0)	(0,0.125,0,0.25)	(0.85,1,0.3,0)
6	(0.85,1,0.3,0)	(0.375,0.425,0.25,0.35)	(0.375,0.55,0.25,0.3)
7	(0,0.15,0,0.3)	(0.375,0.425,0.25,0.35)	(0.375,0.55,0.25,0.3)
8	(0.45,0.55,0.3,0.3)	(0,0.125,0,0.25)	(0.375,0.55,0.25,0.3)
9	(0.45,0.55,0.3,0.3)	(0.775,1,0.35,0)	(0,0.125,0,0.25)
10	(0.45,0.55,0.3,0.3)	(0.775,1,0.35,0)	(0,0.125,0,0.25)
11	(0,0.15,0,0.3)	(0,0.125,0,0.25)	(0.375,0.55,0.25,0.3)
12	(0.45,0.55,0.3,0.3)	(0.375,0.425,0.25,0.35)	(0,0.125,0,0.25)

Таблица 3

Рейтинговые оценки и рейтинг программных продуктов

n	Рейтинговые оценки	Рейтинг
1	0,248	11
2	0	12
3	0,746	2
4	0,542	6
5	0,816	1
6	0,676	3
7	0,457	8
8	0,433	9
9	0,613	4, 5
10	0,613	4, 5
11	0,329	10
12	0,462	7

ния 0, 1 и 2 и соответственно управляющие воздействия — "программный продукт нуждается в серьезных доработках", "программный продукт нуждается в незначительных доработках", "программный продукт выпускается на рынок". Лингвистическим значениям X_1, X_2, X_3 — "низкая", "средняя", "высокая" в соответствие были поставлены значения 0, 1 и 2. Эксперты сформулировали следующие начальные условия: $F(X_1 = 2) = 2$, $F(X_2 = 2) = 2$, $F(X_3 = 2) = 2$ и нечеткие условия "слегка-возрастание" на поведение функции по каждому из аргументов. Эти условия были формализованы с помощью нечетких отношений, матрицы которых представлены соответственно в табл. 4—6.

В результате формализации условий на поведение функции и начальных условий были получены следующие матрицы нечетких отношений, элементы которых занесены соответственно в табл. 7—9.

На пересечении $(i + 1)$-й строки и $(j + 1)$-го столбца табл. 7—9 стоят значения степени уверенности в том, что функция F примет значение j при значении аргументов соответственно X_1, X_2, X_3, равных i, $i = \overline{0, 2}$, $j = \overline{0, 2}$.

В результате формализации всех условий получено отношение, матрица которого имеет 27 строк (по числу возможных значений аргументов) и 3 столбца (по числу значений функции). Элементами матрицы являются значения степени уверенности в том, что функция F примет то или иное значение в зависимости от значений аргументов X_1, X_2, X_3. Например, чтобы получить степень уверенности в том, что функция F примет значение 1 при $X_1 = 0$, $X_2 = 1$, $X_3 = 0$, нужно взять минимум из элемента матрицы табл. 7, стоящего на пересечении первой строки и второго столбца, элемента матрицы табл. 8, стоящего на пересечении второй строки и второго столбца и элемента матрицы табл. 9, стоящего на пересечении первой строки и второго столбца.

После всех операций получено нечеткое отношение, описывающее работу нечеткой логической функции F. Элементы матрицы этого отношения занесены в табл. 10.

Обсуждение с экспертами полученных результатов, позволило получить нечеткую функцию 3-значной логики от трех переменных. Результаты работы этой функции занесены в табл. 11.

Используя построенную нечеткую логическую функцию, можем заключить, что программные продукты № 1, 2, 8, 11, рейтинговые оценки которых попадают в первый интервал [0, 0,45), нуждаются в серьезных доработках. Программные продукты № 4, 7, 12, рейтинговые оценки которых попадают во второй интервал [0,45, 0,55), нуждаются в незначительных доработках. Программные продукты № 3, 5, 6, 9, 10, рейтинговые оценки которых попадают в последний интервал [0,55, 1], выпускаются на рынок. Таким образом, построенная нечеткая логическая функция позволила разбить рейтинговые оценки на три интервала и каждому из интервалов поставить в соответствие управляющее воздействие, направленное на достижение успешности программных продуктов в будущем. Полученные решения полностью согласуются с мнением экспертов, которые оценили успешность представленных программных продуктов исходя из своего опыта и знаний.

Таблица 4

Матрица нечеткого отношения, описывающего "слегка-возрастание" логической функции F по аргументу X_1

	0	1	2
0	0,9	1	0,9
1	0	0,9	1
2	0	0	0,9

Таблица 5

Матрица нечеткого отношения, описывающего "слегка-возрастание" логической функции F по аргументу X_2

	0	1	2
0	0,7	1	0,7
1	0	0,7	1
2	0	0	0,7

Таблица 6

Матрица нечеткого отношения, описывающего "слегка-возрастание" логической функции F по аргументу X_3

	0	1	2
0	0,8	1	0,8
1	0	0,8	1
2	0	0	0,8

Таблица 7

Матрица нечеткого отношения, описывающего значения логической функции F по аргументу X_1

	0	1	2
0	1	0,9	0,9
1	0,9	0,9	0,9
2	0	0	1

Таблица 8

Матрица нечеткого отношения, описывающего значения логической функции F по аргументу X_2

	0	1	2
0	1	0,7	0,7
1	0,7	0,7	0,7
2	0	0	1

Таблица 9

Матрица нечеткого отношения, описывающего значения логической функции F по аргументу X_3

	0	1	2
0	1	0,8	0,8
1	0,8	0,8	0,8
2	0	0	1

Таблица 10

Нечеткое отношение, описывающее работу функции F

	0	1	2		0	1	2		0	1	2
000	1	0,7	0,7	100	0,9	0,7	0,7	200	0	0	0,7
001	0,8	0,7	0,7	101	0,8	0,7	0,7	201	0	0	0,7
002	0	0	0,7	102	0	0	0,7	202	0	0	0,7
010	0,7	0,7	0,7	110	0,7	0,7	0,7	210	0	0	0,7
011	0,7	0,7	0,7	111	0,7	0,7	0,7	211	0	0	0,7
012	0	0	0,7	112	0	0	0,7	121	0	0	0,7
020	0	0	0,8	120	0	0	0,8	220	0	0	0,8
021	0	0	0,8	121	0	0	0,8	221	0	0	0,8
022	0	0	0,9	122	0	0	0,9	222	0	0	1

Таблица 11

Функция 3-значной логики от трех переменных

Значения аргументов	Значения функции	Значения аргументов	Значения функции	Значения аргументов	Значения функции
000	0	100	0	200	2
001	0	101	0	201	2
002	2	102	2	202	2
010	0	110	1	210	2
011	1	111	1	211	2
012	2	112	2	121	2
020	2	120	2	220	2
021	2	121	2	221	2
022	2	122	2	222	2

Заключение

Рейтинговое оценивание применяется во многих областях деятельности человека и используется для выработки управляющих воздействий, направленных на эффективное функционирование объектов оценивания. Сложность принятия решений о выборе соответствующего управляющего воздействия состоит в том, что вследствие неполноты информации или сложности оценочных процедур возможны ошибки экспертов. Кроме этого, всегда существуют зоны неопределенности значений рейтинговых оценок, которые делают этот выбор неоднозначным.

В работе предлагается подход к поддержке принятия решений по результатам рейтингового оценивания на основе нечетких логических функций. Построение таких функций осуществляется на основе начальных условий и нечетких условий на поведение функций. Построенные функции позволяют отнести рейтинговые оценки объектов к одному из кластеров с соответствующим управляющим воздействием, направленным на обеспечение успешности их функционирования.

В работе приведен практический пример, который подтверждает адекватность и состоятельность разработанного подхода.

Список литературы

1. **Домрачев В. Г., Комаров Е. Г., Полещук О. М.** Мониторинг функционирования объектов на основе нечеткого описания их состояний // Информационные технологии. 2007. № 11. С. 46—52.
2. **Poleshchuk O.** The determination of students' fuzzy rating points and qualification levels // International Journal of Industrial and Systems Engineering. 2011. Vol. 9, N. 1. P. 3—20.
3. **Ryjov A.** Fuzzy data bases: description of objects and retrieval of information // Proceeding of the First European Congress in Intelligent Technologies. Aachen. Germany, 1993. Vol. 3. P. 1557—1562.
4. **Ryjov A.** Description of Objects in Human-Machine Information Systems // Application of Fuzzy Systems — Proceeding of the International Conference on Application of Fuzzy Systems. ICAES — 94. Iran, 1994. P. 246—249.
5. **Полещук О. М.** О развитии систем обработки нечеткой информации на базе полных ортогональных семантических пространств // Вестник Московского государственного университета леса. Лесной вестник. 2003. № 1 (26). С. 112—117.
6. **Заде Л. А.** Понятие лингвистической переменной и его применение к принятию приблизительных решений. М.: Мир, 1976. 165 с.
7. **Poleshchuk O., Komarov E.** Expert Fuzzy Information Processing. Berlin Heidelberg: Springer-Verlag, 2011. 237 p.
8. **Poleshchuk O., Darwish A.** New models for monitoring and clustering of the state of plant species based on sematic spaces // Journal of Intelligent and Fuzzy Systems. 2014. Vol. 26, N. 3. P. 1089—1094.
9. **Yager R., Filev D. P.** On the issue of defuzzification and selection based on a fuzzy set // Fuzzy Sets and Systems. 1993. Vol. 55. N. 3. P. 255—272.

O. M. Poleshchuk, D. Sc., Professor,
Professor of Higher Mathematics Department, olga.m.pol@yandex.ru, poleshchuk@mgul.ac.ru,
Mytishchi Branch of the Moscow State Technical University named after N. E. Bauman

Using Fuzzy Logic Functions for Decision Making Based on the Rating Assessment

The complication of rating points obtaining is associated with the heterogeneity of characteristics, sustainability of results when used ordinal scales, as well as recognition of the results in order to develop control actions. Using the linguistic variables has allowed to determine the rating points for different characteristics and to avoid incorrect arithmetic operations in the traditional rating models. However, there are some gaps in the decision of ratings recognition problem. The reason for these gaps is the lack of a formalized approach that would enable experts to reduce the errors that arise due to incomplete or fuzzy information. In this connection the urgency of further research aimed at eliminating existing gaps. An approach based on fuzzy logic functions was developed in the paper for decision making based on the rating assessment. This approach allows to cluster the rating points. Each cluster corresponds to the control action aimed at the successful operation of objects in the future. The paper presents a practical example, which confirms the adequacy and viability of the developed approach.

Keywords: *rating assessment, linguistic variables, fuzzy logic functions, decision making*

References

1. **Domrachev V. G., Komarov E. G., Poleshchuk O. M.** Monitoring funktsionirovaniya ob"ektov na osnove nechetkogo opisaniya ikh sostoyaniy, [Performance monitoring objects based on fuzzy descriptions of their states], *Informatsionnye tehnologii*, 2007, no. 11, pp. 46—52.
2. **Poleshchuk O.** The determination of students' fuzzy rating points and qualification levels, *International Journal of Industrial and Systems Engineering*, 2011. Vol. 9, no. 1, pp. 3—20.
3. **Ryjov A.** Fuzzy data bases: description of objects and retrieval of information, *Proceeding of the First European Congress in Intelligent Technologies, Aachen, Germany, 1993,* vol. 3, pp. 1557—1562.
4. **Ryjov A.** Description of Objects in Human-Machine Information Systems, Application of Fuzzy Systems, *Proceeding of the International Conference on Application of Fuzzy Systems, ICAES — 94, Iran, 1994*, pp. 246—249.
5. **Poleshchuk O. M.** O razvitii sistem obrabotki nechetkoj informacii na baze polnyh ortogonal'nyh semanticheskih prostranstv [On the development of fuzzy information processing systems on the basis of complete orthogonal semantic spaces], *Vestnik Moskovskogo gosudarstvennogo universiteta lesa — Lesnoj vestnik*, 2003, no. 1 (26), pp. 112—117.
6. **Zadeh L. A.** *Ponjatie lingvisticheskoj peremennoj i ego primenenie k prinjatiju priblizitel'nyh reshenij* [Concept of a linguistic variable and its application to adoption of approximate decisions], Moscow, Mir, 1976. 165 p.
7. **Poleshchuk O., Komarov E.** *Expert Fuzzy Information Processing*, Berlin Heidelberg, Springer-Verlag, 2011. 237 p.
8. **Poleshchuk O., Darwish A.** New models for monitoring and clustering of the state of plant species based on sematic spaces, *Journal of Intelligent and Fuzzy Systems*, 2014, vol. 26, no. 3, pp. 1089—1094.
9. **Yager R., Filev D. P.** On the issue of defuzzification and selection based on a fuzzy set, *Fuzzy Sets and Systems*, 1993, vol. 55, no. 3, pp. 255—272.

УДК 004.8 DOI:10.17587/it.23.90-96

В. Я. Цветков, д-р техн. наук, проф., заместитель руководителя
центра перспективных фундаментальных и прикладных исследований, e-mail: cvj2@mail.ru,
Научно-исследовательский и проектно-конструкторский институт информатизации,
автоматизации и связи на железнодорожном транспорте (НИИАС), г. Москва

Когнитивные технологии

Проводится анализ содержания когнитивных технологий. Показаны связь и различие между информационными технологиями и моделями с когнитивными технологиями, связь и различие между информационными и когнитивными технологиями сбора информации на примере рецепции информации, а также связь и различие между информационными и когнитивными конструкциями. Описаны информационно определенные и информационно недоопределенные параметры. Описаны методы интерпретации, основанные на когнитивных технологиях.

Ключевые слова: *когнитология, информационные технологии, когнитивные технологии, систематика, моделирование, интерпретация, когнитивные конструкции, информационные модели, информационная определенность, информационная неопределенность, точечные значения параметров, интервальные значения параметров, рецепция информации*

Введение

В последние десятилетия в ряде областей (искусственный интеллект, качественный и количественный анализ, системный анализы, принятие решений) широко применяется термин "когнитивный" (*cognitive*) и производные от него понятия: технология, модель. Существует понятие *cognitive science* [1] — когнитивистика, когнитивная наука, исследующая и моделирующая принципы организации и работы естественных и искусственных интеллектуальных систем". Латинский корень cognito ("co" — вместе и "gnoscere" — знаю) обозначает познание чего-либо. Когнитивная наука опирается на искусственный интеллект, психологию, лингвистику, визуальное моделирование, образование и включает следующие компоненты: представление знаний, информационное и когнитивное взаимодействие, информирование, мышление и восприятие [2]. Естественные и искусственные системы связаны с естественным и искусственным информационными полями [3], из которых человек получает информацию и знания. Поэтому одно из главных назначений когнитивных технологий — получение знаний. Когнитивные технологии широко применяют в разных областях от медицины [4] до управления [5]. Когнитивные технологии являются основой анализа слабоструктурированных информационных коллекций и ситуаций. Однако они не существуют сами по себе, а связаны с информационными и интеллектуальными технологиями.

Информационная определенность и информационная недоопределенность параметров моделей

Когнитивные технологии связаны с информационными технологиями, но между ними есть различия. Одно из различий связано с параметрами моделей, которые в этих технологиях применяют. Информационные технологии и модели применяют в ситуациях, для которых параметры информационной модели или информационной ситуации являются информационно определенными (*information certainty*) или имеют информационно полную определенность (*complete certainty*).

Информационная определенность в аспекте измерений означает, что эти параметры могут быть измерены с помощью существующих измерительных средств или вычислены на основе измерений других параметров. Информационная определенность в аспекте описаний означает, что эти параметры соответствуют точечным четким величинам, а не интервальным или нечетким величинам. На плоскости такому параметру соответствует точка, т.е. элемент множества.

Когнитивные технологии и модели применяют в ситуациях, для которых параметры информационной модели или информационной ситуации являются информационно недоопределенными или информационно частично определенными. Информационная недоопределенность (*underdetermined*) отличается от информационной неопределенности (*information uncertainty*) тем, что частично содержит информационно определенные параметры, т. е. может информировать, но давать не полную информацию, а частичную. На рис. 1 приведена триада отношений между понятиями: определенность, недоопределенность и неопределенность.

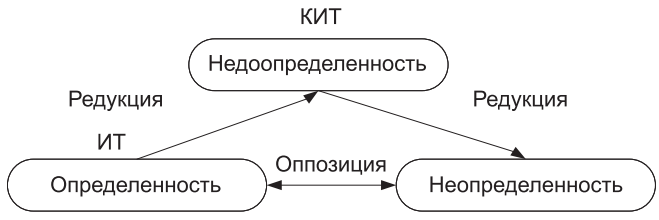


Рис. 1. Триада отношений между понятиями "определенность", "недоопределенность" и "неопределенность"

Определенность и неопределенность связаны двухсторонними отношениями оппозиции. Определенность и недоопределенность, а также недоопределенность и неопределенность связаны отношениями редукции. В условиях определенности применяют информационные технологии (ИТ), в условиях недоопределенности — когнитивные информационные технологии (КИТ).

Информационная недоопределенность в аспекте измерений означает, что не все параметры могут быть измерены с помощью существующих измерительных средств или вычислены на основе измерений других параметров. Информационная недоопределенность в аспекте описаний означает, что эти параметры соответствуют не точечным четким величинам, а интервальным или нечетким величинам. На плоскости такому параметру соответствует прямая, т.е. множество. В интервале значений такому параметру ставят в соответствие так называемые трапециевидные, прямоугольные или треугольные числа. Различие между теорией нечетких множеств и когнитивным анализом в том, что в когнитивном анализе включается когнитивное пространство человека, механизм которого не всегда понятен как неявное знание [6].

Когнитивные и информационные технологии

Следует напомнить, что после четвертой информационной революции, которую также называют цифровой революцией, возникли термины "Новые информационные технологии" (НИТ) и синонимический ему термин "Информационные и коммуникационные технологии" (ИКТ). Информационные технологии (НИТ или ИКТ), связанные с компьютерной или цифровой обработкой, отличаются от других информационных технологий, например, почтовой связи или записи лекций. С течением времени НИТ и ИКТ стали заменять на ИТ и это привело к тому, что не всегда просматривается грань между "компьютерными" ИТ и "не компьютерными" ИТ. Когнитивные технологии связаны с компьютерными ИТ.

Если использовать аспект коммуникации, то можно отметить следующее различие информационных и когнитивных технологий. Информационные технологии используют две характеристики при передаче или обработке информации: информационная емкость сообщения; семантика, или смысловое содержание информационного сообщения. В соответствии с этими характеристиками информационные модели используют структурные информационные единицы и семантические информационные единицы [7].

Когнитивные технологии используют дополнительные характеристики. Две из них являются информационными и свойственны информационным технологиям. В этом состоит связь между ИТ и КИТ. КИТ имеют дополнительные характеристики: логическую, предикативную, ассоциативную. Соответственно, информационных единиц в когнитивных технологиях и моделях больше. Это логические информационные единицы, предикативные информационные единицы [8] и ассоциативные информационные единицы. Ассоциативная характеристика, или ассоциации, — это способ нахождения связи на основе уже существующего опыта или образов (прецедентов) [9] вместо кропотливого анализа информации.

Информационные и когнитивные конструкции

В настоящее время накоплен опыт моделирования и существует много информационных и когнитивных технологий и моделей. Наряду с понятием информационной модели часто употребляют термин "информационная конструкция" [10, 11]. Его применяют для концептуального описания информационных моделей и технологий. Термин "когнитивная информационная конструкция" соответствует понятию "сложная система".

Когнитивные информационные конструкции имеют связь с информационными конструкциями (две характеристики информационные) и имеют свои собственные характеристики. Когнитивную информационную конструкцию можно представить как развитие понятия информационная конструкция.

Как модели концептуализации, информационные конструкции (ИК) и когнитивные информационные конструкции (КИК) выполняют две основные функции: дескриптивную и процессуальную. Дескриптивные (описательные) функции ИК и КИК состоят в том, что они являются средством описания объекта исследования и информируют об этом объекте, играя роль носителя информации, т.е. ИК и КИК являются средствами описания картины мира.

Процессуальная функция ИК и КИК состоит в том, что эти конструкции формируют динамическое описание объектов во внешней среде. Это отражает динамику развития, организацию, самоорганизацию или деградацию объектов.

При описании объектов в условиях недоопределенности применяют КИК. КИК — обобщенное понятие, которое объединяет когнитивные модели, элементы моделей и допускает количественную и качественную вариабельность. КИК подразумевает наличие структуры, является более широким понятием, чем когнитивная карта, и допускает больше разновидностей когнитивного моделирования. В аспекте структурной сложности выделяют сложные и простые когнитивные конструкции. Сложные конструкции включают в свой состав другие конструкции, а простые конструкции — только элементы. КИК имеют информационные (И) и когнитивные (К) признаки. Эти признаки следующие:

- системность, означающая, что компоненты конструкции в совокупности образуют систему (И, К);

- структурированность, означающая наличие структуры и структурных элементов конструкции (И, К);
- интерпретируемость, означающая возможность интерпретации конструкций или ее частей (И, К);
- предикативное соответствие, означающее наличие свойства предикативности для всей конструкции (К);
- полисемия, означающая, что конструкция может содержать много смысловых значений (И, К);
- когнитивность, означающая участие когнитивного пространства субъекта в формировании конструкции (К).

Когнитивная конструкция имеет определенный смысл, если существует какая-либо ее интерпретация. Интерпретировать когнитивную конструкцию — это значит связать с ней семантическую область, называемую также областью интерпретации. С точки зрения когнитивной лингвистики интерпретация может включать субъективные когнитивные процедуры [12].

Структурированность информационных конструкций позволяет осуществлять их морфологический анализ. Морфологический анализ включает идентификацию формы и структуры информационной конструкции. Таким образом, информационная и когнитивная конструкции являются обобщением ряда понятий: информационный объект, информационная модель, информационная система. Структурность конструкции отличает ее от информационной совокупности или информационной коллекции. Информационная совокупность может содержать не связанные или слабосвязанные между собой объекты и элементы. Структура в такой совокупности может отсутствовать.

Информационная конструкция включает связанные компоненты и связанные элементы. Для нее применим системный, морфологический и структурный анализы. Когнитивные конструкции становятся объективным фактором интерпретации и описания явлений во многих областях. Совокупности конструкций дают возможность оценки морфологической и смысловой сложности явлений окружающего мира.

Интерпретация когнитивной информационной конструкции

Интерпретация КИК представляет собой процесс нахождения информационного соответствия между конструкцией и ее значениями в информационном интерпретационном поле [3]. Для интерпретации конструкции вводят понятия: объект интерпретации, семантическое окружение объекта интерпретации, набор интерпретационных признаков и метод интерпретации.

Важное значение в методах интерпретации играет когнитивная информационная семантика [13], которая является частью когнитивной информационной лингвистики. Обе науки используют междисциплинарный перенос понятий наук об информации в семантику и лингвистику. Это определяет тесную связь интерпретации с когнитивистикой. Интерпретация — это типичная когнитивная технология.

Интерпретационное информационное поле содержит семантическое окружение объекта интерпретации. Интерпретационное поле информационной конструкции включает многочисленные признаки, характеризующие данную область исследования. Интерпретационное поле неоднородно, и в нем вычленяются несколько зон интерпретационного поля, которые обладают тематической однородностью.

Информацию, применяемую в информационных и когнитивных технологиях, можно разделить на две группы, одна из которых описывает факты, а другая является интерпретацией фактов [14]. Интерпретация осуществляется по правилам, среди которых выделяют три вида: аксиоматические, эмпирические, продуктивные [15].

Аксиоматические правила интерпретации основаны на системе аксиом и сводят новое знание к системе базисных положений. Доминирующей в этом виде интерпретации является система аксиом, на основе которой строятся интерпретационные цепочки. В результате новые факты исследователь интерпретирует с помощью известных базисных положений (стереотипов). Этот подход можно назвать стереотипным.

Эмпирические правила интерпретации определяются опытным путем, применительно к конкретной ситуации и могут быть не пригодны для иных условий. Они применяются при использовании набора моделей и дают интерпретацию некоей модельной ситуации, при которой исследователь видит начало и результат. Доминирующим является выявление взаимосвязей и тенденций между фактами, видимыми исследователю. Этот подход можно назвать причинно-следственным [16].

Продуктивные правила интерпретации получают на основе механизма правил, который допускает самообучение и коррекцию. Доминирующим в этом подходе является использование механизма вывода, который может привести к результатам и выявить факты, ранее не известные исследователю. Все подходы дополняют друг друга. С позиций познания интерпретация КИК представляет собой извлечение неявного знания [17], которое плохо формализовано и структурировано.

Рецепция информации как когнитивная технология

Современные технологии сбора информации основаны на применении информационно-измерительных систем. Однако во многих случаях информацию собирает субъект или эксперт. Эксперт не просто накапливает информацию, а осуществляет рецепцию информации [18, 19] с использованием всех сенсорных систем на уровне сознания и под-

сознания. Рецепция осуществляется определенными структурными образованиями — сенсорными системами. Причем чем больше опыт эксперта, тем выше результат рецепции. Рецепция информации применяет когнитивные методы анализа информации и дополнительные каналы анализа.

На практике сбор информации в сложных ситуациях сталкивается со следующими проблемами [20, 21]: большой объем информации; избыточная информативность о ситуации. Это мотивирует применение когнитивного подхода для анализа информации. Рецепция информации может быть определена как совокупность процессов декомпозиции информации по разным каналам человеческого восприятия, проведение качественного и количественного анализа информации, применение метода прецедентов и ассоциативного анализа, когнитивного синтеза информации в единую систему моделей.

Когнитивный фильтр. Одним из механизмов рецепции является когнитивный фильтр [22] (рис. 2), который можно рассматривать как когнитивную информационную конструкцию, состоящую из четырех слоев.

Комбинации слоев фильтра формируют различные модели: коммуникационную (Ком М), информационную (Инф М), когнитивную (Когн М). Базисным является формальный, или коммуникационный слой. Формальность заключается в кодировании информации, т.е. в преобразовании внешней информации в кодифицированную. Эта процедура задает информационный объем кодированной информации, а слой содержит все остальные слои. Формализация присутствует на каждом слое и на каждом слое она разная, т.е. соответствует типу слоя. Первый слой позволяет формировать коммуникационную модель (Ком М), которую рассматривает К. Э. Шеннон в своей известной работе по математической теории коммуникации.

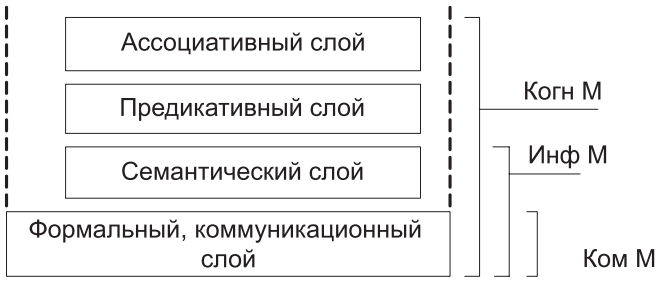


Рис. 2. Когнитивный фильтр при рецепции информации

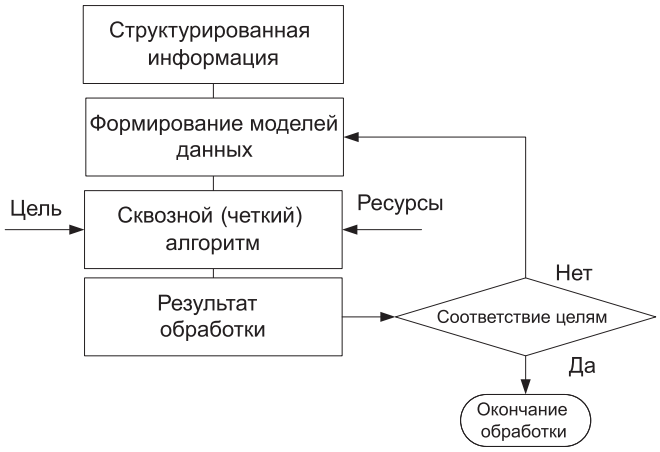


Рис. 3. Алгоритмическая обработка информации

Второй слой является семантическим. Он отвечает за смысловое наполнение модели. Первый и второй слои позволяют формировать информационную модель (Инф М) и создают условия для информационного взаимодействия и информирования. Применение этих двух слоев достаточно для сбора информации и для трансформации информации. Если информация структурированная, а параметры моделей информационно определены, то этих слоев достаточно для обработки информации и принятия решений. Поэтому можно считать, что первые два слоя создают информационный фильтр, который полностью решает задачи информационного анализа и информационного моделирования.

В качестве альтернативы рецепции и когнитивной обработке целесообразно рассмотреть алгоритмическую обработку информации. Классическая алгоритмическая обработка информации (рис. 3) осуществляется с использованием информационного фильтра.

Особенностью данной схемы является наличие структурированной информации и применение сквозного алгоритма. Сквозным называют алгоритм, который позволяет решать задачу или проводить обработку от начала до конца, без итераций или промежуточных этапов.

Возвращаясь к когнитивному фильтру (см. рис. 2), следует отметить, что третий и четвертый слои определяют специфику когнитивного анализа и рецепции информации. Эти слои не входят в схему на рис. 3. Третий слой является предикативным [8]. Он соотносит содержание входной информации или анализируемой модели с реальностью и позволяет определять область истинности для них. Четвертый слой является ассоциативным. Он связывает анализируемую информацию или модель или ее характеристики с тезаурусом, с базой данных, с семантической сетью, с базой стереотипов, с базой прецедентов или с базой знаний. Все четыре слоя позволяют формировать когнитивную модель (Когн М) и совместно осуществляют рецепцию информации. Причем следует подчеркнуть, что эта рецепция распространяется именно на неструктурированную информацию. Рецепция информации основана на информационном и когнитивном взаимодействии. Когнитивный фильтр создает возможность рецепции информации и когнитивной обработки. На рис. 4 приведена схема когнитивной обработки информации с использованием рецепции информации. Центральная ветвь схемы на рис. 4 является аналогом схемы на рис. 3. Две дополнительные ветви обработки на рис. 4 являются

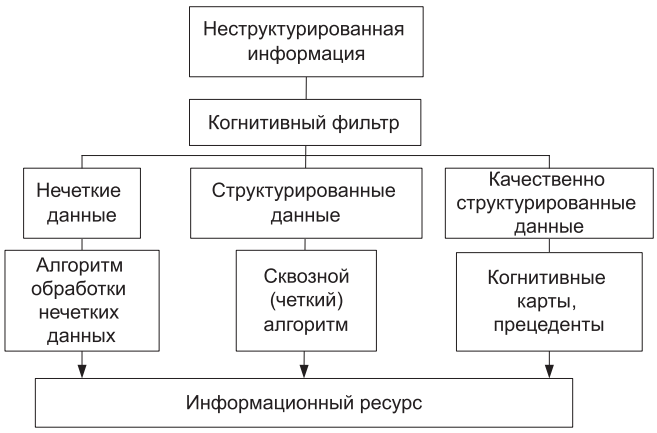


Рис. 4. Когнитивная обработка информации с применением рецепции информации

расширением схемы алгоритмической обработки информации (см. рис. 3) и включают дополнительные возможности, которые информационный метод исключает.

Особенностью когнитивной обработки является то, что в уровне "Информационный ресурс" (рис. 4) могут присутствовать два или три результата обработки информации, полученные по разным ветвям обработки. Эти результаты сравниваются и анализируются с помощью рецепции информации. Рецепция информации позволяет не только на входе (неструктурированная информация), но и на выходе (информационный ресурс) осуществлять анализ и повышать обоснованность принятия решения.

Когнитивное взаимодействие. Взаимодействие в информационном поле осуществляется между объектами и субъектами, между объектами, между субъектами. По этому критерию оно разделяется на объектное (формальное) и субъектное (когнитивное) [15]. Объектное взаимодействие основано на полностью формализованных моделях, для которых достаточно двух уровней когнитивного фильтра. Когнитивное взаимодействие основано на включении всех четырех уровней когнитивного фильтра.

Для осуществления когнитивного взаимодействия субъект (эксперт) или интеллектуальная система должны обладать следующими признаками:

- коммуникационной лингвистической способностью. Эксперт и интеллектуальная система должны владеть несколькими языками (лингвистическим, топологическим, схемным, унифицированным и др.) для описания ситуации и процессов;
- наличием рецепторов информации, а именно: зрением (видеоканал), слухом (аудиоканал), обонянием (канал качественного восприятия), вкусом (канал качественного восприятия), осязанием (канал качественного восприятия) и вестибулярного аппарата (канал качественного восприятия). Это характерная особенность рецепции информации;
- наличием механизма самоорганизации. Самоорганизация — это свойства субъекта и интеллектуальной системы, проявляющиеся в наличии у них способностей к накоплению опыта и неявных знаний для последующей трансформации их в новое знание, направленное на улучшение характеристик системы и достижение цели с учетом динамики внешней среды и противоборства других субъектов и объектов;
- наличием механизма ассоциативного анализа. Информационные системы применяют только количественный анализ. Когнитивные системы позволяют осуществлять качественный и количественно-качественный анализ;
- наличием организованной ассоциативной памяти. Память — одна из психических функций и видов умственной деятельности, предназначенная сохранять, накапливать и воспроизводить информацию. Оперативная память человека содержит до семи чанков. Долговременная память эксперта позволяет длительно хранить информацию о событиях внешнего мира и реакциях организма и многократно использовать ее в сфере сознания для организации последующей деятельности.

Следует отметить, что базы данных хранят только структурированную информацию.

Выделяют следующие типы взаимодействия "субъкт — объект" в рамках рецепции информации.

Паралингвистическое взаимодействие — в рамках этого взаимодействия [23] используются символы и сигналы, не входящие в лингвистические языки. Информационные признаки паралингвистических моделей лежат вне лингвистического языка. Такое когнитивное взаимодействие активно используется в процессе коммуникации между людьми, особенно в сфере образования и в театральной деятельности. Производными невербального информационного взаимодействия являются: музыка, танец. Это когнитивное взаимодействие осуществляется в системе "субъект — субъект".

Вербальное когнитивное взаимодействие осуществляется сущностями, обладающими речевой способностью, и подразумевает использование коммуникации на основе естественного языка. Это когнитивное взаимодействие осуществляется в системе "субъект — субъект" [24].

Иконическое когнитивное взаимодействие осуществляется сущностями с использованием знаков и изображений, не входящих в состав алфавита, и слов естественного языка. Это когнитивное взаимодействие осуществляется в системах "субъект — субъект", "объект — субъект", "субъект — объект".

Лингвистическое когнитивное взаимодействие осуществляется сущностями с использованием лингвистики — с помощью языковых единиц естественного и искусственного языка. Это когнитивное взаимодействие осуществляется в системах "субъект — субъект", "объект — субъект", "субъект — объект".

Рецепция в когнитивной области. Когнитивная область субъектов и объектов (только интеллектуальные системы) представляет собой не только область индивидуального сознания (индивидуальная), но и область коллективного сознания групп индивидов (групповая), объединенных общей целью. Примером коллективного сознания являются мультиагентные системы. Важным для когнитивной области является возникновение синергетического эффекта в рамках коллективного сознания, не сводящегося к простой сумме индивидуальных сознаний [25]. В когнитивной области осуществляется коллективное понимание и осознание текущей ситуации.

В когнитивной области можно выделить следующие уровни рецепции информации:

- на уровне понятий, суждений и умозаключений;
- на уровне гипотез, теорий и знаний;
- на уровне осведомления о текущей ситуации;
- на уровне концепций, целей, задач, замыслов, решений, планов;
- на уровне корпоративного проектирования;
- на уровне мозгового штурма.

На каждом из уровней специфицируются свои информационные ресурсы. Информационное взаимодействие в когнитивной области позволяет обеспечить коллективное понимание и осознание текущей ситуации исходя из стандартизованных терминов, терминологических отношений, общей базы данных, общей базы прецедентов, согласованных стереотипов задач, общей базы данных. При этом включаются ассоциативные и предикативные методы анализа информации. Когнитивное взаимодействие существенно влияет на преодоление проблем "нечеткости" и "диссипации" информации.

Применение модели рецепции информации при анализе сложной и неструктурированной информации позволяет расширить виды обрабатываемой и анализируемой информации. Рецепция информации позволяет на входе (неструктурированная информация) и на выходе (информационный ресурс, рис. 4) осуществлять дополнительный анализ, что повышает обоснованность принятия решения. Рецепция информации расширяет виды исходной информации, применяемой в управлении или анализе. Применение рецепции информации в сочетании с информационными технологиями позволяет строить сложную структурную модель и создает синергетический эффект.

Заключение

Существует различие между информационными технологиями и когнитивными технологиями. Когнитивные технологии позволяют обрабатывать большее число информационных коллекций, которые в информационных технологиях не обрабатывают. Особенностью когнитивных технологий является включение когнитивного взаимодействия между познающим субъектом и объектом исследования. Когнитивное взаимодействие отличается от информационного взаимодействия тем, что в когнитивной области осуществляется не передача информации, а рецепция информации, которая подключает дополнительные каналы взаимодействия к техническому каналу. Когнитивные технологии нельзя сводить только к применению когнитивных карт. Введение понятия "когнитивные информационные конструкции" расширяет возможности когнитивных технологий. Оно создает условия для применения методов интерпретации в когнитивной области на основе когнитивной информационной лингвистики. Понятие рецепции информации не тождественно сбору информации. Принципиальным отличием рецепции в технологическом плане является подключение предикативных и ассоциативных параметров модели и когнитивного пространства к анализу информации на качественном и количественном уровнях. Когнитивные технологии решают задачи преобразования неявного знания в явное знание. Однако область когнитивных технологий требует широкого исследования.

Список литературы

1. **Першиков В. И., Савинков В. М.** Толковый словарь по информатике. М.: Финансы и статистика, 1995. 544 с.
2. **Eysenk M. W.** ed. The Blackwell Dictionary of Cognitive Psychology. Cambridge. Massa-chusetts: Basil Blackwell Ltd., 1990.
3. **Цветков В. Я.** Естественное и искусственное информационное поле // Международный журнал прикладных и фундаментальных исследований. Ч. 2. 2014. № 5. С. 178—180.
4. **Номоконова О. Ю.** Опыт врача как когнитивный информационный ресурс // Славянский форум, 2015. № 3 (9). С. 200—209.
5. **Болбаков Р. Г.** Основы когнитивного управления // Государственный советник. 2015. № 1. С. 45—49.
6. **Полани М.** Личностное знание. На пути к посткритической философии: Пер. с англ. / Под ред. В. А. Лекторского и В. И. Аршинова. М.: Прогресс, 1985. 343 с.
7. **Ozhereleva T. A.** Systematics for information units // European Researcher, 2014. Vol. (86), N. 11/1. P. 1894—1900.
8. **Цветков В. Я.** Информационные единицы сообщений // Фундаментальные исследования. 2007. № 12. С. 123—124.
9. **Варшавский П. Р., Еремеев А. П.** Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений // Искусственный интеллект и принятие решений. 2009. № 2. С. 45—57.
10. **Тесленко П. А.** Информационная конструкция и атрибуты ее исследования // Проблемы техники. Научно-производственный журнал. 2008. № 3. С. 22—31.
11. **Бондур В. Г.** Информационные конструкции в космических исследованиях // Образовательные ресурсы и технологии. 2016. № 3 (15). С. 79—88.
12. **Чехарин Е. Е.** Интерпретируемость информационных единиц // Славянский форум. 2014. №2 (6). С. 151—155.
13. **Nomokonov I. B.** The Semantic Informativeness // European Journal of Medicine. Series B. 2015. Vol. (4). Iss. 3. P. 141—147.
14. **Искусственный** интеллект. В 3-х кн. Кн. 2. Модели и методы: Справочник / Под ред. Д. А. Поспелова. М.: Радио и связь, 1990. 340 с.
15. **Соловьев И. В., Мордвинов В. А., Жигалов О. С.** Информационное и когнитивное взаимодействие. М.: МаксПресс, 2015.
16. **Номоконова О. Ю.** Импакт анализ в диагностике. М.: МаксПресс, 2016. 56 с.
17. **Сигов А. С., Цветков В. Я.** Неявное знание: оппозиционный логический анализ и типологизация // Вестник Российской Академии Наук, 2015. Т. 85, № 9. С. 800—804.
18. **Номоконова О. Ю.** Рецепция информации при медицинской диагностике // Славянский форум. 2015. № 4 (10). С. 238—243.

19. **Цветков В. Я.** Рецепция информации // Образовательные ресурсы и технологии. 2016. № 1 (13). С. 121—129.
20. **Чехарин Е. Е.** Большие данные: большие проблемы // Перспективы науки и образования. 2016. № 3. С. 7—11.
21. **Майер-Шенбергер В., Кукьер К.** Большие данные: Революция, которая изменит то, как мы живем, работаем и мыслим. М.: Манн, Иванов и Фербер, 2014. 240 с.
22. **Tsvetkov V. Ya.** Intelligent control technology // Russian Journal of Sociology. 2015. Vol. (2), Iss. 2. P. 97—104.
23. **Цветков В. Я.** Паралингвистические информационные единицы в образовании // Перспективы науки и образования. 2013. № 4. С. 30—38.
24. **Duval R.** A cognitive analysis of problems of comprehension in a learning of mathematics // Educational studies in mathematics. 2006. Vol. 61, N. 1—2. P. 103—131.
25. **Кулинич А. А.** Когнитивная система поддержки принятия решений "Канва" // Программные продукты и системы. 2002. № 3. С. 25—28.

V. Ya. Tsvetkov, Professor, Deputy Head, e-mail: cvj2@mail.ru,
Center fundamental and advanced research,
Research and Design Institute of design information, automation
and communication on railway transport, Moscow

Cognitive Technologies

The article analyzes the content of cognitive technologies. This article describes the relationship and the difference between information technology and models with cognitive technologies and models. This article describes the relationship and the difference between information technology and cognitive information-gathering techniques to collect information example of the reception of information. This article describes the relationship and differences between cognitive information structures and information structures. This article describes the information certainty parameters s and information underdetermined parameters. This article describes interpretation techniques based on cognitive technologies.

Keywords: *cognitive science, information technology, cognitive technology, taxonomy, modeling, interpretation, cognitive design, information models, information certainty, information uncertainty, spot parameter values, interval settings, information reception*

References

1. **Pershikov V. I., Savinkov V. M.** *Tolkovyj slovar' po informatike*, Moscow: Finansy i statistika, 1995. 544 p.
2. **Eysenk M. W.** ed. *The Blackwell Dictionary of Cognitive Psychology*. Cambridge. Massa-chusetts: Basil Blackwell Ltd, 1990.
3. **Tsvetkov V. Ya.** Estestvennoe i iskusstvennoe informacionnoe pole, *Mezhdunarodnyj zhurnal prikladnyh i fundamental'nyh issledovanij*, Ch. 2, 2014, no. 5, pp. 178—180.
4. **Nomokonova O. Ju.** Opyt vracha kak kognitivnyj informacionnyj resurs, *Slavjanskij forum*, 2015, no. 3 (9), pp. 200—209.
5. **Bolbakov R. G.** Osnovy kognitivnogo upravlenija, *Gosudarstvennyj sovetnik*, 2015, no. 1, pp. 45—49.
6. **Polani M.** *Lichnostnoe znanie. Na puti k postkriticheskoj filosofii.* Per. s angl. Pod red. V. A. Lektorskogo i V. I. Arshinova, Moscow, Progress, 1985, 343 p.
7. **Ozhereleva T. A.** Systematics for information units, *European Researcher*, 2014, vol. (86), no. 11/1, pp. 1894—1900.
8. **Tsvetkov V. Ya.** Informacionnye edinicy soobshhenij, *Fundamental'nye issledovanija*, 2007, no. 12, pp. 123—124.
9. **Varshavskij P. R., Eremeev A. P.** Modelirovanie rassuzhdenij na osnove precedentov v intellektual'nyh sistemah podderzhki prinjatija reshenij, *Iskusstvennyj intellekt i prinjatie reshenij*, 2009, no. 2, pp. 45—57.
10. **Teslenko P. A.** Informacionnaja konstrukcija i atributy ee issledovanija, *Problemy tehniki. Nauchno-proizvodstvennyj zhurnal*, 2008, no. 3, pp. 22—31.
11. **Bondur V. G.** Informacionnye konstrukcii v kosmicheskih issledovanijah, *Obrazovatel'nye resursy i tehnologii*, 2016, no. 3 (15), pp. 79—88.
12. **Cheharin E. E.** Interpretiruemost' informacionnyh edinic, *Slavjanskij forum*, 2014, no. 2 (6), pp. 151—155.
13. **Nomokonov I. B.** The Semantic Informativeness, *European Journal of Medicine. Series B*, 2015, vol. (4), Is. 3, pp. 141—147.
14. **Iskusstvennyj** intellekt. V 3-h kn. Kn. 2. Modeli i metody: Spravochnik. Pod red. D. A. Pospelova, Moscow, Radio i svjaz', 1990. 340 p.
15. **Solov'ev I. V., Mordvinov V. A., Zhigalov O. S.** *Informacionnoe i kognitivnoe vzaimodejstvie*. Moscow, MaksPress, 2015.
16. **Nomokonova O. Ju.** *Impakt analiz v diagnostike*, Moscow, MaksPress, 2016, 56 p.
17. **Sigov A. S., Tsvetkov V. Ya.** Nejavnoe znanie: oppozicionnyj logicheskij analiz i tipologizacija, *Vestnik Rossijskoj Akademii Nauk*, 2015, vol. 85, no. 9, pp. 800—804.
18. **Nomokonova O. Ju.** Recepcija informacii pri medicinskoj diagnostike, *Slavjanskij forum*, 2015, no. 4 (10), pp. 238—243.
19. **Tsvetkov V. Ya.** Recepcija informacii, *Obrazovatel'nye resursy i tehnologii*, 2016, no. 1 (13), pp. 121—129.
20. **Cheharin E. E.** Bol'shie dannye: bol'shie problemy, *Perspektivy nauki i obrazovanija*, 2016, no. 3, pp. 7—11.
21. **Majer-Shenberger V., Kuk'er K.** *Bol'shie dannye: Revoljucija, kotoraja izmenit to, kak my zhivem, rabotaem i myslim*. Moscow, Mann, Ivanov i Ferber, 2014, 240 p.
22. **Tsvetkov V. Ya.** Intelligent control technology, *Russian Journal of Sociology*, 2015, vol. (2), Is. 2, pp. 97—104.
23. **Tsvetkov V. Ya.** Paralingvisticheskie informacionnye edinicy v obrazovanii, *Perspektivy nauki i obrazovanija*, 2013, no. 4, pp. 30—38.
24. **Duval R.** A cognitive analysis of problems of comprehension in a learning of mathematics, *Educational studies in mathematics*, 2006, vol. 61, no. 1—2, pp. 103—131.
25. **Kulinich A. A.** Kognitivnaja sistema podderzhki prinjatija reshenij "Kanva", *Programmnye produkty i sistemy*, 2002, no. 3, pp. 25—28.

УДК 004.827 DOI:10.17587/it.23.97-106

Г. Э. Яхъяева, канд. физ.-мат. наук, доц., e-mail: gul_nara@mail.ru,
А. А. Карманова, магистрант, e-mail: anast.karmy.aa@gmail.com,
А. А. Ершов, магистрант, e-mail: alaershov@gmail.com,
Н. П. Савин, магистрант, e-mail: npsavin@rambler.ru,
Новосибирский государственный университет, г. Новосибирск

Вопросно-ответная система для управления информационными рисками на основе теоретико-модельной формализации предметных областей[1]

Работа посвящена описанию вероятностной вопросно-ответной системы QA-RiskPanel, позволяющей пользователю в диалоговом режиме проводить анализ различных рисков, связанных с компьютерными атаками. QA-RiskPanel является вопросно-ответной системой, основанной на знаниях (knowledge based QA-system). В качестве источника знаний в системе QA-RiskPanel используется постоянно пополняемая база прецедентов компьютерных атак, что обеспечивает актуализацию прогнозирования рисков. Онтологический подход к формализации предметной области позволяет проводить анализ рисков на различных уровнях конкретизации/обобщения.

Приводится теоретико-модельная формализация базы знаний описываемой предметной области, описана классификация вопросных типов, которые в данной системе носят вероятностный характер. Представлены алгоритмы поиска ответов на вопросные типы данной классификации.

Ключевые слова: *информационная безопасность, компьютерная атака, прецедент компьютерной атаки, база знаний, вопросно-ответная система, теория нечетких моделей, обобщенная нечеткая модель*

Введение

Компьютер является неотъемлемой частью жизни большинства людей. С развитием компьютерных технологий возросло и число желающих получить прибыль или другие привилегии за счет изъянов в безопасности компьютерного программного обеспечения (ПО) или уязвимости железной части системы. По статистике МВД РФ [1] число совершаемых как в России, так и в мире, киберпреступлений увеличивается с каждым годом. Так, в 2015 г. по итогам расследования преступлений в сфере информационных технологий в судебные органы было направлено 11 223 материала, что на 137 % превышает показатели 2014 г.

На сегодняшний день существует множество различных программных систем для обеспечения информационной безопасности. По способу подхода к решению проблемы компьютерной безопасности все существующие программные продукты можно разделить на три основные категории:

- *программы, предназначенные для проведения профилактических работ*: проведение резервного копирования данных, проверка прочности паролей, сканирование операционной системы и установленного ПО и т.д.;
- *программы, предназначенные для обнаружения и предотвращения компьютерных атак*, т.е. аппаратно-программные решения, направленные на автоматизацию контроля событий, протекающих в компьютерной системе или сети, и анализ этих событий в целях поиска признаков проблем безопасности;
- *программы, предназначенные для анализа и оценки рисков неблагоприятных событий*, т.е. возможность анализировать текущую безопасность функционирования информационной системы, оценивать и прогнозировать риски, управлять их влиянием на бизнес-процессы организации, корректно и обоснованно подходить к вопросу поддержания безопасности ее активов.

Управление информационными рисками становится одним из наиболее актуальных и динамично развивающихся направлений стратегического и оперативного менеджмента в области защиты информации [2]. Компьютерные сети играют важную роль во многих областях. Растущие требования к информационной инфраструктуре предприятий приводят к усложнению структуры и увеличению размера сетей. Это, в свою очередь, ведет к росту сложности анализа их безопасности и запаздыванию применения контрмер [3]. При возникновении угрозы необходимо знать симптомы заражения, возможные потери, предлагаемые решения проблемы и любую информацию, которая могла бы помочь своевременно и правильно среагировать и справиться с проблемой.

Первым шагом любой методики управления информационными рисками является идентификация рисков. Типовым подходом к решению данной задачи является использование различных стандартных списков классов рисков. Компании либо разрабатывают свои списки рисков информа-

[1] Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 14-07-00903_а.

ционной безопасности, либо используют некоторые общепринятые стандарты, либо покупают такие списки у крупных производителей ПО. Трудности реализации такого подхода обусловлены излишней конкретизацией. Заказчик, используя подробные каталоги, может "потонуть" в море рисков, которые система для него идентифицирует.

Еще одна трудность, с которой сталкиваются разработчики систем информационной безопасности, — это очень динамичное развитие информационных технологий и, как следствие, непрерывное появление новых угроз и уязвимостей. Разработчики стараются отслеживать появление новых рисков, выпускают новые версии программ, которые очень быстро устаревают.

Для решения этих проблем в Новосибирском государственном университете была разработана программная система управления информационными рисками RiskPanel[2] [4]. Проблема излишней конкретизации рисков в данной системе решается средствами онтологического моделирования предметной области. Иерархическая структура понятий предметной области позволяет проводить анализ рисков на различных уровнях конкретизации/обобщения. Проблема постоянного появления новых угроз и уязвимостей решается с помощью модуля полуавтоматического пополнения базы данных прецедентов компьютерных атак. Пополнение осуществляется за счет анализа оперативной информации, представленной в Интернет-источниках.

Программный комплекс RiskPanel имеет модульную структуру, позволяющую в дальнейшем подключать к нему новые модули. Например, к основным модулям был подключен модуль ранней диагностики компьютерного нападения [5, 6], основанный на применении ДСМ-метода [7] и анализа формальных понятий [8].

В данной работе мы описываем вопросно-ответную систему QA-RiskPanel[3], разработанную в рамках программного комплекса RiskPanel. Вопросы, которые пользователь может задавать вопросно-ответной системе QA-RiskPanel, носят вероятностный характер и направлены на прогнозирование компьютерных атак и исследование рисков, связанных с ними [9]. В данном случае мы имеем в виду фреквентистский подход к определению вероятности, предложенный Джоном Веном в 1876 г. [10, 11]. Согласно этому подходу под вероятностью понимается частота появления события в общем числе наблюдений. Использование данного подхода обусловлено тем, что мы имеем дело с динамично меняющейся предметной областью. Постоянно появляются новые виды атак, угроз и уязвимостей, которые в любой момент могут кардинально изменить статистику. Поэтому мы вынуждены каждый раз пересчитывать частоту появления события и в общем числе (актуальных на данный момент времени) наблюдений. Только таким образом мы обладаем "объективной" информацией о предметной области.

Система QA-RiskPanel состоит из трех модулей: модуль безусловных вопросов, модуль условных вопросов и модуль модальных вопросов. При формализации условных вопросов кроме понятия объективной (фреквентистской) вероятности в работе используются так называемые субъективные вероятности. Понятие "субъективной вероятности" было введено в 30-х годах прошлого века Фрэнком Рамсеем [11] и подразумевает степень уверенности эксперта в наступлении того или иного события.

Следующий раздел статьи посвящен теоретико-модельному описанию базы знаний рассматриваемой предметной области. Во втором разделе рассматриваются шаблоны безусловных, условных и модальных вопросов, реализованные в системе. Третий раздел посвящен описанию алгоритмов поиска ответов для каждого рассмотренного выше вопросного шаблона.

1. База знаний по компьютерной безопасности

1.1. Описание структуры базы знаний

Формализация предметной области "Компьютерные атаки" проводилась на языке теории нечетких моделей [12, 13] с применением методов логики описаний [14]. На первом этапе вводится множество P атомарных понятий предметной области Δ. Все атомарные понятия разделены на шесть классов:

P_1: "Симптомы"; P_4: "Последствия";
P_2: "Угрозы"; P_5: "Потери";
P_3: "Уязвимости"; P_6: "Контрмеры".

Каждый из этих классов понятий иерархически упорядочен. Множество понятий для каждого класса формируется в результате анализа базы данных прецедентов компьютерной безопасности National Vulnerability Database (NVD, the agency NIST[4]). NVD — это государственный проект США, постоянно пополняемая база компьютерных уязвимостей, открытых в различных программных системах и компонентах. В описаниях уязвимостей содержится информация о том, в какой версии какого продукта эта уязвимость присутствует, как ей могут воспользоваться злоумышленники, какие меры можно предпринять для устранения этой уязвимости, и другая информация.

Множество всех понятий CON предметной области "Компьютерные атаки" строится согласно стандартному синтаксису Логики Описаний. Каждое понятие $\varphi \in CON$ является булевой комбинацией атомарных понятий. Онтология предметной области (которую, согласно традициям Логики Описаний

[2] Свидетельство о государственной регистрации программ для ЭВМ № 2011617412 от 23.09.2011.

[3] Свидетельство о государственной регистрации программ для ЭВМ № 2016611258 от 28.01.2016.

[4] http://www.nist.gov/

будем называть *TBox*) состоит из множества всех аксиом специализации, которые отражают иерархическую структуру классов понятий.

TBox является первым компонентом Базы Знаний рассматриваемой предметной области. Вторым компонентом является описание прецедентов компьютерных атак. Основным источником информации о компьютерных атаках для системы QA-Risk-Panel служат базы данных корпораций NIST и MITRE[5]. Каждая атака e характеризуется наличием/отсутствием тех или иных понятий из каждого класса понятий P_i. Таким образом, на данном этапе каждое атомарное понятие мы воспринимаем как одноместный предикат, т.е. $P(x) \in P$.

Далее, используя аксиомы *TBox*, мы пополняем наши знания об истинности атомарных понятий на различных прецедентах компьютерных атак. Введем обозначение:

$$ABox = \{P(e) \mid \text{атомарное понятие } P(x) \text{ истинно на прецеденте } e\}.$$

Далее пару $KB = \langle TBox, ABox\rangle$ и будем считать Базой Знаний предметной области "Компьютерные атаки". По мере появления в данной предметной области новых понятий (т.е. новых угроз, уязвимостей и т.д.) или новых прецедентов компьютерных атак база знаний будет расширяться. Однако общая структура Базы Знаний будет оставаться неизменной.

1.2. Теоретико-модельная формализация базы знаний

Для решения задач статистической обработки данных нам понадобятся ***прецедентная*** и ***нечеткая*** модели рассматриваемой предметной области [13]. Эти модели строятся на основе класса ***интерпретаций*** Базы Знаний *KB*.

Рассмотрим конечное множество компьютерных атак $E = \{e_1, ..., e_n\}$, которые были использованы для описания *ABox*, и класс P одноместных предикатов, используемых для описания *TBox*.

Определение 1. *Алгебраическую систему* $A_E = \langle E, P\rangle$ *будем называть интерпретацией Базы Знаний KB, если* $A_E \models ABox$ *(т.е. для каждого предложения* $\varphi(e_i) \in ABox$ *мы имеем* $A_E \models \varphi(e_i)$*).*

Определение 2. *Упорядоченную тройку* $Case(A_E) = \langle\{a\}, P, \tau\rangle$ *назовем* ***прецедентной моделью****, порожденной интерпретацией* $A_E = \langle E, P\rangle$*, если для любого понятия* $\varphi(x) \in CON$ *имеем* $\tau(\varphi(a)) = \{e \in E \mid A_E \models \varphi(e)\}$.

В прецедентной модели каждому понятию ставится в соответствие множество прецедентов компьютерных атак, обладающих этим понятием. Заметим, что с теоретико-модельной точки зрения модель $Case(A_E)$ является булевозначной моделью. В этой булевозначной модели каждому предложению сигнатуры $P \cup \{c_a\}$ (где c_a — константа) ставится в соответствие элемент булевой алгебры $\rho(E)$, порожденный множеством атак E [15].

[5] http://www.mitre.org/

В большинстве методик статистической обработки данных используются объективные и/или субъективные вероятности. Под объективной вероятностью понимается относительная частота появления какого-либо события в общем объеме наблюдений или отношение числа благоприятных исходов к общему числу наблюдений. Под субъективной вероятностью имеется в виду мера уверенности некоторого эксперта или группы экспертов в том, что данное событие в действительности будет иметь место [16].

В рассматриваемом подходе для описания объективных вероятностей используется понятие нечеткой модели.

Определение 3. *Упорядоченную тройку* $Fuz(A_E) = \langle\{a\}, P, \mu\rangle$ *назовем* ***нечеткой моделью*** *предметной области* Δ*, порожденной интерпретацией* $A_E = \langle E, P\rangle$*, если для любого понятия* $\varphi(x) \in CON$ *имеем*

$$\mu(\varphi(a)) = \frac{\left\|\{e \in E | A_E \models \varphi(e)\}\right\|}{\|E\|}.$$

Значениями истинности предложений (понятий) в нечеткой модели являются числа из интервала [0, 1], которые отражают объективную вероятность наличия того или иного понятия у потенциальной компьютерной атаки. Более подробное описание свойств прецедентных и нечетких моделей можно найти в работах [13, 15].

Заметим, что получаемая из Интернета информация о прецедентах компьютерных атак в подавляющем большинстве случаев является неполной. Следовательно, согласно парадигме "open-world semantics", мы имеем класс различных интерпретаций базы знаний *KB*. Обозначим этот класс I_E, т.е.

$$I_E = \{A_E = \langle E, P\rangle \mid A_E \models ABox\}.$$

Определение 4. *Упорядоченную тройку* $Fuz(E) = \langle\{a\}, P, \xi_E\rangle$ *назовем* ***обобщенной нечеткой моделью****, порожденной классом интерпретаций* I_E*, если для любого понятия* $\varphi(x) \in CON$ *имеем*

$$\xi_E(\varphi(a)) = \{\mu(\varphi(a)) | Fuz(A_E) = \langle\{a\}, \sigma_\Delta, \mu\rangle \text{ и } A_E \in I_E\}.$$

Таким образом, значениями истинности предложений на обобщенной нечеткой модели являются различные подмножества рациональных чисел из интервала [0, 1]. В работе [17] показано, что значениями истинности на модели $Fuz(E)$ являются интервалы, определенные на множестве

$$Q^n = \left\{0, \frac{1}{n}, ..., \frac{n-1}{n}, 1\right\},$$

где $n = \|E\|$ — число прецедентов, занесенных в базу знаний.

Заметим, что в строго математическом смысле мы не будем получать интервальную модель. Однако при $n \to \infty$ значения истинности предложений на модели $Fuz(E)$ будут стремиться к интервалам на множестве $[0, ..., 1] \cap Q$. Таким образом, на практике, имея дело с достаточно большим множеством

прецедентов, мы можем воспринимать значения истинности на модели $Fuz(E)$ как интервалы рациональных чисел. Исходя из этого будем обозначать

$$Fuz(E) \models_{[\alpha,\ \beta]} \varphi(a),$$

если $\alpha = \inf(\xi_E(\varphi(a)))$ и $\beta = \sup(\xi_E(\varphi(a)))$. В частном случае, когда $\alpha = \beta$, будем обозначать $Fuz(E) \models_\alpha \varphi(a)$. Будем говорить, что предложение $\varphi(a)$ ***истинно*** на модели $Fuz(E)$, если $Fuz(E) \models_1 \varphi(a)$, и ***ложно*** на модели $Fuz(E)$, если $Fuz(E) \models_0 \varphi(a)$.

Рассмотрим следующие подмножества множества прецедентов E:

$$T(E, \varphi) = \{e \in E \mid \forall A_E \in I_E\colon A_E \models \varphi(e)\};$$

$$F(E, \varphi) = \{e \in E \mid \forall A_E \in I_E\colon A_E \not\models \varphi(e)\}; \qquad (1)$$

$$N(E, \varphi) = E \setminus (T(E, \varphi) \cup F(E, \varphi)).$$

Пусть $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$. Тогда, согласно Определению 3, имеем

$$\|T(E, \varphi)\| = \beta_1\|E\|,$$

$$\|N(E, \varphi)\| = (\beta_2 - \beta_1)\|E\|, \qquad (2)$$

$$\|F(E, \varphi)\| = (1 - \beta_2)\|E\|.$$

Заметим, что $\inf(\xi_E(\varphi(a))) = \sup(\xi_E(\varphi(a)))$ тогда и только тогда, когда $N(E_1, \varphi) = \varnothing$. Более того, предложение $\varphi(a)$ истинно (ложно) на модели $Fuz(E)$ тогда и только тогда, когда $\|E\| = \|T(E_1, \varphi)\|$ ($\|E\| = \|F(E_1, \varphi)\|$).

1.3. Оптимальные сужения базы знаний

Для поиска ответов на условные вопросы (см. раздел 2.4) нам будет необходимо сужать базу знаний *KB*, оставляя столько прецедентов компьютерных атак из множества E, чтобы полученная при этом база знаний максимально подходила под заданное условие. Более того, мы будем стремиться к тому, чтобы сужение базы знаний было минимальным, т.е. "выкидывалось" как можно меньшее число прецедентов.

Пусть $E_1 \subseteq E (E_1 \neq \varnothing)$ и модель $A_E = \langle E, P\rangle$ является интерпретацией базы знаний *KB*. Подмодель $A_{E_1} = \langle E_1, P\rangle$ модели A_E будем называть ***сужением интерпретации*** A_E базы знаний *KB* на множество прецедентов E_1 (и обозначать $A_{E_1} \subseteq A_E$). Тогда класс моделей

$$I_{E_1} = \{A_{E_1} = \langle E_1, P\rangle \mid \exists A_E \in I_E\colon A_{E_1} \subseteq A_E\}$$

будем называть ***сужением*** класса интерпретаций I_E на множество E_1.

Определение 5. *Рассмотрим множество прецедентов $E_1 \subseteq E (E_1 \neq \varnothing)$. Обобщенную нечеткую модель $Fuz(E_1)$, порожденную классом I_{E_1}, будем называть* ***сужением*** *модели $Fuz(E)$ и обозначать $Fuz(E_1) \preccurlyeq Fuz(E)$.*

Обозначим $\rho(Fuz(E))$ множество всех сужений модели $Fuz(E)$. Так как сигнатура P обобщенной нечеткой модели $Fuz(E)$ является чисто предикатной, каждая модель $Fuz(E_1) \preccurlyeq Fuz(E)$ однозначно определяется множеством прецедентов $E_1 \subseteq E$. Нетрудно показать, что частично упорядоченное множество $\langle \rho(Fuz(E)), \preccurlyeq\rangle$ образует булеву решетку.

Пусть $\varphi(x) \in CON$ и $\alpha \in [0, 1]$. Рассмотрим следующие подмножества множества $\rho(Fuz(E))$:

$$M(\varphi \geqslant \alpha) = \{Fuz(E_1)) \preccurlyeq Fuz(E) \mid \xi_{E_1}(\varphi(a)) \subseteq [\alpha, 1]\};$$

$$M(\varphi \leqslant \alpha) = \{Fuz(E_1)) \preccurlyeq Fuz(E) \mid \xi_{E_1}(\varphi(a)) \subseteq [0, \alpha]\}.$$

Заметим, что множество $M(\varphi \geqslant \alpha)$ ($M(\varphi \leqslant \alpha)$) пусто тогда и только тогда, когда $\alpha \neq 0$ ($\alpha \neq 1$) и предложение $\varphi(a)$ ложно (истинно) на модели $Fuz(E)$.

Пусть $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$. Тогда, если $\alpha \leqslant \beta_1$, то $Fuz(E) \in M(\varphi \geqslant \alpha)$. Следовательно, эта модель является наибольшей в упорядоченном множестве $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$.

Вместе с тем, если $\alpha = 1$, то $M(\varphi \geqslant \alpha) = \{Fuz(E_1) \mid E_1 \subseteq T(E, \varphi)\}$. Следовательно, модель $Fuz(T(E, \varphi))$ является наибольшей в упорядоченном множестве $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$.

Аналогично, если $\alpha \geqslant \beta_2$, то модель $Fuz(E)$ является наибольшей в упорядоченном множестве $\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle$. И если $\alpha = 0$, то модель $Fuz(F(E, \varphi))$ является наибольшей в упорядоченном множестве $\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle$.

В остальных случаях, если множество $M(\varphi \geqslant \alpha)$ ($M(\varphi \leqslant \alpha)$) не пусто, то частично упорядоченное множество $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$ ($\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle$) имеет больше одного максимального элемента.

Предложение 1. *а) Пусть модель $Fuz(E_1)$ является максимальной в упорядоченном множестве $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$. Тогда $T(E_1, \varphi) = T(E, \varphi)$.*

б) Пусть модель $Fuz(E_1)$ является максимальной в упорядоченном множестве $\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle$. Тогда $F(E_1, \varphi) = F(E, \varphi)$.

Доказательство. Докажем пункт а), пункт б) доказывается аналогично.

Рассмотрим модель $Fuz(E_1) \in \rho\{Fuz(E))$ такую, что $Fuz(E_1) \in M(\varphi \geqslant \alpha)$ и $T(E, \varphi) \setminus T(E_1, \varphi) \neq \varnothing$. Покажем, что модель $Fuz(E_1)$ не является максимальной в $\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle$.

Из того, что $Fuz(E_1) \in M(\varphi \geqslant \alpha)$ следует, что $\alpha \leqslant \dfrac{\|T(E_1, \varphi)\|}{\|E_1\|}$. А из того, что $T(E, \varphi) \setminus T(E_1, \varphi) \neq \varnothing$, следует, что существует хотя бы один прецедент e такой, что $e \in T(E, \varphi) \setminus T(E_1, \varphi)$.

Рассмотрим модель $Fuz(E_1 \cup \{e\})$. Очевидно, что

$$Fuz(E_1) \preccurlyeq Fuz(E_1 \cup \{e\}).$$

Из равенств $\|T(E_1 \cup \{e\}, \varphi)\| = \|T(E_1, \varphi)\| + 1$ и $\|E_1 \cup \{e\}\| = \|E_1\| + 1$ получим $\alpha \leqslant \dfrac{\|T(E_1, \varphi)\|}{\|E_1\|} \leqslant \dfrac{\|T(E_1 \cup \{e\}, \varphi)\|}{\|E_1 \cup \{e\}\|}$. А это означает, что $Fuz(E_1 \cup \{e\}) \in M(\varphi \geqslant \alpha)$, т.е. модель $Fuz(E_1)$ не является максимальной в $\langle M(\varphi \geqslant \alpha), \subseteq\rangle$. ■

Определение 6. *Пусть* $\varphi(x) \in CON$ *и* $\alpha \in [0, 1]$. *Модель* $Fuz(E_1)$ *будем называть* $(\varphi \geqslant \alpha)$***-оптимальным (****$(\varphi \leqslant \alpha)$***-оптимальным) сужением*** *модели* $Fuz(E)$, *если она удовлетворяет следующим условиям:*

1. Модель $Fuz(E_1)$ *является максимальной в частично упорядоченном множестве* $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$ $(\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle)$;

2. Для любой модели $Fuz(E_2)$, *являющейся максимальной в* $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$ $(\langle M(\varphi \leqslant \alpha), \preccurlyeq\rangle)$, *выполняется условие* $\xi_{E_2}(\varphi(a)) \subseteq \xi_{E_1}(\varphi(a))$.

Теорема 1. *Рассмотрим* $\varphi(x) \in CON$ *и* $\alpha \in [0, 1]$. *Пусть* $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$. *Тогда*

а) если $\beta_1 < \alpha$ *и модель* $Fuz(E_1)$ *является* $(\varphi \geqslant \alpha)$*-оптимальным сужением модели* $Fuz(E)$, *то выполняются следующие условия:*

1. $\|E_1\| = \left[\!\left[\frac{\beta_1}{\alpha}\|E\|\right]\!\right]$, *где* $[\![x]\!]$ *— целая часть числа x;*

2. $Fuz(E_1) \models_{[k\beta_1; \min(1, k\beta_2)]} \varphi(a)$, *где* $k = \frac{\|E\|}{\|E_1\|}$;

б) если $\alpha < \beta_2$ *и модель* $Fuz(E_1)$ *является* $(\varphi \leqslant \alpha)$*-оптимальным сужением модели* $Fuz(E)$, *то выполняются следующие условия:*

1. $\|E_1\| = \left[\!\left[\frac{1-\beta_2}{1-\alpha}\|E\|\right]\!\right]$;

2. $Fuz(E_1) \models_{[\max\{0; 1-k(1-\beta_1)\}; 1-k(1-\beta_2)]} \varphi(a)$, *где* $k = \frac{\|E\|}{\|E_1\|}$.

Доказательство. Докажем пункт а), пункт б) доказывается аналогично.

Так как модель $Fuz(E_1)$ является $(\varphi \geqslant \alpha)$-оптимальным сужением модели $Fuz(E)$, то по Предложению 1 должно выполняться условие $T(E_1, \varphi) = T(E, \varphi)$. А в силу того, что $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$, получим $\|T(E_1, \varphi)\| = \|T(E, \varphi)\| = \beta_1\|E\|$.

Так как $Fuz(E_1) \in M(\varphi \geqslant \alpha)$, то $\alpha \leqslant \frac{\|T(E_1, \varphi)\|}{\|E_1\|}$.

Следовательно, получим $\|E_1\| \leqslant \frac{\|T(E_1, \varphi)\|}{\alpha} = \frac{\beta_1}{\alpha}\|E\|$.

И в силу того, что модель $Fuz(E_1)$ является максимальной в упорядоченном множестве $\langle M(\varphi \geqslant \alpha), \preccurlyeq\rangle$, мощность множества E_1 будет равна наибольшему целому числу n, удовлетворяющему условиям:

$$n \leqslant \frac{\beta_1}{\alpha}\|E\| \text{ и } n - \|T(E, \varphi)\| \leqslant \|F(E, \varphi)\| + \|N(E, \varphi)\|.$$

Покажем, что этим условиям удовлетворяет число $\left[\!\left[\frac{\beta_1}{\alpha}\|E\|\right]\!\right]$.

Очевидно, что число $\left[\!\left[\frac{\beta_1}{\alpha}\|E\|\right]\!\right]$ является наибольшим числом, удовлетворяющим первому неравенству. Допустим, что $\left[\!\left[\frac{\beta_1}{\alpha}\|E\|\right]\!\right] - \|T(E, \varphi)\| > \|F(E, \varphi)\| + \|N(E, \varphi)\|$. Тогда получим

$$\frac{\beta_1}{\alpha}\|E\| - \beta_1\|E\| > (1 - \beta_2)\|E\| + (\beta_2 - \beta_1)\|E\|.$$

Разделив данное неравенство на $\|E\|$, получим $\frac{\beta_1 - \alpha\beta_1}{\alpha} > 1 - \beta_1$. Следовательно, $\beta_1 - \alpha\beta_1 > \alpha - \alpha\beta_1$, значит, $\alpha < \beta_1$. А это противоречит условию теоремы. Таким образом, пункт а) доказан.

Далее, из $\|T(E_1, \varphi)\| = \beta_1\|E\|$ следует, что

$$\inf(\xi_{E_1}(\varphi(a))) = \frac{\|T(E_1, \varphi)\|}{\|E_1\|} = \frac{\beta_1\|E\|}{\|E_1\|} = k\beta_1.$$

Вместе с тем, так как модель $Fuz(E_1)$ является $(\varphi \geqslant \alpha)$-оптимальным сужением модели $Fuz(E)$, то

$$\sup(\xi_{E_1}(\varphi(a))) = \min\left\{1, \frac{\|T(E_1, \varphi)\| + \|N(E, \varphi)\|}{\|E_1\|}\right\} =$$

$$= \min\left\{1, \frac{\beta_2\|E\|}{\|E_1\|}\right\} = \min(1, k\beta_2). \blacksquare$$

Следствие 1. *Рассмотрим* $\varphi(x) \in CON$ *и* $\alpha \in [0, 1]$. *Пусть* $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$, $\beta_1 < \alpha$, *модель* $Fuz(E_1)$ *является* $(\varphi \geqslant \alpha)$*-оптимальным сужением модели* $Fuz(E)$ *и* $k = \frac{\|E\|}{\|E_1\|}$. *Тогда*

а) если $\alpha \geqslant \frac{\beta_1}{\beta_2}$, *то* $Fuz(E_1) \models_{[\beta_1, \beta_2]} \varphi(a)$;

б) если $\alpha < \frac{\beta_1}{\beta_2}$, *то* $Fuz(E_1) \models_{[k\beta_1; k\beta_2]} \varphi(a)$.

Доказательство. По Теореме 1 имеем, что $\sup(\xi_{E_1}(\varphi(a))) = \min\{1, k\beta_2\}$.

Рассмотрим случай, когда $\alpha \geqslant \frac{\beta_1}{\beta_2}$. Тогда $\|E_1\| = \left[\!\left[\frac{\beta_1}{\alpha}\|E\|\right]\!\right] \leqslant [\![\beta_2\|E\|]\!]$. Из равенств (2) вытекает, что $[\![\beta_2\|E\|]\!] = \beta_2\|E\|$, т.е. $\|E_1\| \leqslant \beta_2\|E\|$. Тогда $k = \frac{\|E\|}{\|E_1\|} \geqslant \frac{1}{\beta_2}$. Следовательно, $k\beta_2 \geqslant 1$. Таким образом, получаем, что $\sup(\xi_{E_1}(\varphi(a))) = 1$.

Вместе с тем, если $\alpha < \frac{\beta_1}{\beta_2}$, то $k < \frac{1}{\beta_2}$. Тогда получаем, что $\sup(\xi_{E_1}(\varphi(a))) = k\beta_2$. $\blacksquare$

Следствие 2. *Рассмотрим* $\varphi(x) \in CON$ *и* $\alpha \in [0, 1]$. *Пусть* $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$, $\alpha < \beta_2$, *модель* $Fuz(E_1)$

является $(\varphi \leqslant \alpha)$*-оптимальным сужением модели* $Fuz(E)$ *и* $k = \frac{\|E\|}{\|E_1\|}$. *Тогда*

а) если $\alpha \leqslant 1 - \frac{1-\beta_2}{1-\beta_1}$,

то $Fuz(E_1) \models_{[0;1-k(1-\beta_2)]} \varphi(a)$;

б) если $\alpha > 1 - \frac{1-\beta_2}{1-\beta_1}$,

то $Fuz(E_1) \models_{[1-k(1-\beta_1);1-k(1-\beta_2)]} \varphi(a)$.

Доказательство аналогично доказательству Следствия 1.

2. Формализация и классификация вопросных типов

Вопросно-ответные системы направлены на удовлетворение информационных потребностей пользователя. Однако анализ, понимание и удовлетворение этих потребностей — непростая задача даже для человека, не говоря уже о программных системах. Поэтому важной задачей для вопросно-ответных систем является формализация вопросно-ответных отношений. Понятно, что от качества проведенной формализации будет зависеть качество и эффективность взаимодействия пользователя с QA-системой.

Множество всех вопросов, которые пользователь может задать системе на естественном языке, бесконечно, даже в рамках довольно узкой предметной области, и формализация этого множества не представляется возможной. Мы можем формально описать лишь некоторые типовые шаблоны вопросов. Выбор формализации базы знаний, очевидно, накладывает определенные ограничения на типы вопросов, которые может обрабатывать вопросно-ответная система. В данной работе база знаний вопросно-ответной системы формализуется в виде алгебраической системы $Fuz(E)$. Это позволяет проводить формализацию и классификацию типов вопросов в духе эротетической логики, основополагающие идеи которой можно найти в работе [18].

Традиционно в эротетической логике рассматриваются два типа вопросов: "ли"-вопросы и "какой"-вопросы. Так как наши суждения носят вероятностный характер, то мы будем рассматривать еще и третий тип вопросов — "вероятностные вопросы".

2.1. "Ли"-вопросы

Вопросы "ли"-типа направлены на выяснение истинности некоторого суждения. Обычно такой вопрос начинается со слов *"Верно ли, что ..."*. В качестве ответа на "ли"-вопрос пользователь ожидает ответ "да" или "нет". С теоретико-модальной точки зрения "ли"-вопрос формализуется как "запрос": истинно ли заданное предложение на данной алгебраической системе.

В системе QA-RiskPanel "ли"-вопросы модифицируются, так как, с одной стороны, цель системы в определении вероятности наступления тех или иных рисков, а с другой стороны, база знаний системы формализуется в виде обобщенной нечеткой модели $Fuz(I_E)$. Поэтому при формулировке "ли"-вопроса помимо суждения мы будем указывать вероятностную характеристику данного суждения. Приведем примеры такого вопроса:

"Верно ли, что вероятность использования ненадежного пароля в атаке менее 0,3?"

"Верно ли, что вероятность использования в атаке SQL-инъекции равна 0,8?"

2.2. Вероятностные вопросы

Вероятностный вопрос есть требование вычислить вероятность некоторого суждения. К примеру, информационная потребность пользователя может состоять в том, чтобы узнать, как часто в компьютерных атаках используется уязвимость "ненадежный пароль". В этом случае ответом на вопрос

"Какова вероятность, что в компьютерной атаке будет использован ненадежный пароль?"

будет интервал рациональных чисел из отрезка [0; 1].

В более общем случае, рассматривая интерпретацию вопросов на моделях с истинностными функциями различной природы (например, булевозначные или прецедентные модели [15]), вместо вероятностных вопросов мы получим *оценочные вопросы*. Общая схема таких вопросов следующая:

Каково значение истинности того, что ⟨*суждение*⟩*?*

В случае интерпретации вопроса на классической модели оценочный вопрос будет отождествляться с "ли"-вопросом.

2.3. "Какой"-вопросы

В эротетической логике "какой"-вопрос трактуется как запрос на выявление объема субъекта данного вопроса. Каждый "какой"-вопрос рассматривается как множество (возможно, бесконечное) различных "ли"-вопросов. Ответом на такой вопрос является список субъектов тех "ли"-вопросов, на которые получен положительный ответ.

В нашем подходе каждый "какой"-вопрос будет разбиваться на *конечное* множество "ли"-вопросов. Приведем пример "какой"-вопроса:

Какие скрытые атаки могут произойти с вероятностью более 0,8?

Под алгоритмом поиска ответа на "какой"-вопрос будем понимать процедуру, состоящую из следующих трех шагов:

(1) выделение подмножества C множества всех атомарных понятий P;

(2) вычисление значений истинности всех понятий из C;

(3) выделение подмножества понятий $C' \subseteq C$, значения истинности которых удовлетворяют вероятностной характеристике, заданной в вопросе. Ответом на такой вопрос будет являться список понятий из C'.

2.4. Условные вопросы

Условные вопросы — это вопросы, содержащие условие и требующие ответа в случае, когда известно выполнение этого условия. Условные вопросы играют важную роль в формализации вопросно-ответных систем. Дело в том, что для вопросов реального мира обычно подразумевается, что спрашивающий человек обладает некоторым предварительным знанием о предмете, и игнорирование этого знания может привести к снижению релевантности извлеченного системой ответа ожиданиям пользователя; первоочередная же задача вопросно-ответных систем — выдавать релевантные ответы.

В разработанной вопросно-ответной системе реализуется возможность добавления условия к вопросу любого типа. Таким образом, имеется возможность задавать условные "ли"-вопросы, условные вероятностные вопросы, условные "какой"-вопросы и условные модальные вопросы (см. раздел 2.5). Вопросы, на которые не наложено условие, будем называть *безусловными*. Рассмотрим пример условного вероятностного вопроса:

"*Если шифрование сети отсутствует, то какова вероятность утечки информации в результате атаки?*"

Этот вопрос можно понимать, как "запрос" отыскать условную вероятность события "происходит утечка информации", при условии, что событие "шифрование сети отсутствует" достоверно. С теоретико-модельной точки зрения для ответа на такой вопрос требуется найти значение истинности на обобщенной нечеткой модели, являющейся ($\varphi = 1$)-оптимальной подмоделью модели $Fuz(E)$, где $\varphi =$ = "*шифрование сети отсутствует*" (см. раздел 1.3).

Заметим, что в нашем случае суждения носят вероятностный характер, а значит, и условия могут быть выполнимыми с некоторой вероятностью. Пример такого вопроса:

"Если шифрование сети отсутствует как минимум в 30 % случаев, то какова вероятность утечки информации в данной атаке?"

Для отыскания ответа на такой вопрос мы должны найти ($\varphi \geqslant 0{,}3$)-оптимальную подмодель, на которой и будем считать значение истинности предложения "*существует утечка информации*". Заметим, что в общем случае мы будем получать целый класс моделей с заданным условием.

2.5. Модальные вопросы

Компьютерные атаки можно разделить на два класса: одношаговые и многошаговые [19]. В одношаговой атаке злоумышленник сразу использует уязвимость для достижения конечной цели. В многошаговой атаке злоумышленник может использовать существующую уязвимость для открытия новой уязвимости, которую можно использовать для проведения другой атаки.

Для анализа многошаговых атак в системе QA-RiskPanel строятся графы атак. Множеством вершин графа является множество одношаговых атак $E = \{e_1, ..., e_n\}$. Ребро между двумя атаками строится, если последствия одной атаки приводят к созданию условий, необходимых для совершения второй атаки. Многошаговая атака является путем в этом направленном графе, каждая следующая вершина которого достижима из предыдущей однократным переходом по направленному ребру.

Таким образом, мы имеем возможность получать статистическую информацию о многошаговых атаках, добавляя различные модальности к любому типу вопроса. На сегодняшний день в системе QA-RiskPanel реализована возможность задавать модальные вопросы двух типов: "возможно"-вопросы и "рано или поздно"-вопросы. Приведем пример безусловного вероятностного вопроса типа *"возможно":*

Какова вероятность того, что в результате многошаговой атаки возможна порча базы данных?

и условного вероятностного вопроса типа *"рано или поздно":*

Если вероятность переполнения буфера не менее 0,8, то какова вероятность того, что в результате многошаговой атаки рано или поздно произойдет порча базы данных?

Для формализации вопросов и поиска ответов о многошаговых атаках мы будем использовать методы модальной логики и алгоритмы model checking.

3. Алгоритмы поиска ответов

3.1. Безусловные вопросы

Алгоритмы поиска ответов на безусловный "ли"-вопрос и безусловный вероятностный вопрос схожи между собой. В обоих случаях необходимо найти значение истинности некоторого предложения $\varphi(a)$ (порожденного понятием $\varphi \in CON$) на модели $Fuz(E)$. Это значение истинности и является ответом на вероятностный вопрос. Ответом на "ли"-вопрос является результат сравнения данного значения истинности с заданным в вопросе ограничением.

"Какой"-вопрос задает конечное множество предложений $\varphi_1(a)$, ..., $\varphi_n(a)$, для которых необходимо вычислить значения истинности на модели $Fuz(E)$. Ответом на такой вопрос будет список предложений, значения истинности которых удовлетворяют требованиям, описанным в вопросе.

Таким образом, ядром алгоритма на любой безусловный вопрос является процедура поиска значения истинности заданного предложения на обобщенной нечеткой модели.

Заметим, что значение истинности предложения $\varphi(a)$ на обобщенной нечеткой модели $Fuz(E)$ строится как объединение значений истинности этого предложения на всех интерпретациях, входящих в класс I_E. Каждая интерпретация $A_\Delta \in I_E$ является конечной моделью, сигнатура которой состоит из конечного числа одноместных предикатов. Поэтому процедура вычисления значения истинности предложения на модели A_Δ разрешима и может быть решена методами логики высказываний.

Класс I_E также является конечным. Однако его мощность экспоненциально зависит от степени неопределенности базы знаний KB. Допустим, база знаний KB содержит 10 000 полностью описанных прецедентов компьютерных атак, и только в 10 случаях неизвестно, был ли в атаке использован вирус. Эта незначительная неопределенность порождает 2^{10} различных интерпретаций. Очевидно, что перебор всех интерпретаций слишком трудоемок и не реализуем программно.

В работах [17, 20] рассмотрен алгоритм, вычисляющий значение истинности бесквантворного предложения на обобщенной нечеткой модели, сигнатура которой состоит из конечного числа одноместных предикатов. Данный алгоритм основан на идее разложения обобщенной нечеткой модели в прямое произведение обобщенных прецедентов и имеет полиномиальную сложность.

3.2. Модальные вопросы

Опишем теоретико-модельную формализацию вопросов типа "возможно" и "рано или поздно". На множестве E введем бинарное отношение R: атаки e_1 и e_2 находятся в отношении R, если последствия атаки e_1 открывают уязвимости, характерные для атаки e_2. В данной работе нас будет интересовать транзитивное замыкание R^T отношения R.

Пополним терминологию $TBox$ базы знаний KB одним ролевым понятием R^T. Введем обозначение: $TBox' = TBox \cup \{R^T\}$. Это, в свою очередь, повлечет расширение множества истинных атомарных понятий $ABox$ до множества $ABox'$ и, как следствие, расширение базы знаний KB' и обобщенной нечеткой модели $Fuz(E)'$, формализующей эту базу знаний.

Пусть $\varphi \in CON$, т.е. не содержит в своей сигнатуре ролевого понятия R^T. Тогда (согласно синтаксису логики описания) "возможно"-вопросы будут формализоваться с помощью понятия $\varphi_\Diamond = \varphi \vee \exists R^T.\varphi$, а "рано или поздно"-вопросы — с помощью понятия $\varphi_\Box = \varphi \vee (\forall R^T.(\exists R^T.\varphi))$. А ответы на эти вопросы будут зависеть от значений истинности соответствующих предложений на обобщенной нечеткой модели $Fuz(E)'$.

Для реализации вычислений значений истинности предложений $\varphi_\Diamond$ и $\varphi_\Box$ рассматривается граф $G = \langle E, R^T \rangle$. Проводится постфиксный обход графа G. В результате обхода графа каждой вершине графа присваивается одно из трех значений: TRUE, UNKNOWN или FALSE по следующему принципу:

$$e \in T(E, \varphi_{\Diamond/\Box}) \Rightarrow e := TRUE;$$

$$e \in F(E, \varphi_{\Diamond/\Box}) \Rightarrow e := FALSE;$$

$$e \in N(E, \varphi_{\Diamond/\Box}) \Rightarrow e := UNKNOWE.$$

Данная разметка графа и позволяет определить inf и sup значений истинности предложений.

В работах [21, 22] приведено описание алгоритмов *FuzGLEMP* и *FuzGLEMN* вычисления значений истинности предложений типа $\varphi_\Diamond$ и $\varphi_\Box$ на обобщенной нечеткой модели $Fuz(E)'$. За основу был взят алгоритм пометки графа из [23], модифицированный для работы с неполной базой знаний.

3.3. Условные вопросы

На сегодняшний день в системе QA-RiskPanel реализованы алгоритмы поиска ответов на условные вопросы следующих двух типов:

Если $\varphi \geqslant \alpha$, то (безусловный модальный/немодальный вопрос)?

Если $\varphi \leqslant \alpha$, то (безусловный модальный/немодальный вопрос)?

Таким образом, алгоритм поиска ответа на условный вопрос сводится к нахождению $(\varphi \geqslant \alpha)$-оптимальной или $(\varphi \geqslant \alpha)$-оптимальной модели $Fuz(E_1)$, где $E_1 \subseteq E$. Далее, на модели $Fuz(E_1)$ запускается один из алгоритмов, описанных в разделах 3.1 и 3.2.

Пусть $Fuz(E) \models_{[\beta_1, \beta_2]} \varphi(a)$. Тогда при $\alpha \leqslant \beta_1$ и при $\alpha = 1$ мы имеем единственную $(\varphi \geqslant \alpha)$-оптимальную модель, а при $\alpha \geqslant \beta_2$ и при $\alpha = 0$ мы имеем единственную $(\varphi \leqslant \alpha)$-оптимальную модель. На этой оптимальной модели и запускается алгоритм поиска значения истинности предложения $\psi(a)$, формализующего соответствующий безусловный вопрос.

В остальных случаях мы имеем целый класс $K_{\varphi \geqslant \alpha}$ различных $(\varphi \geqslant \alpha)$-оптимальных и класс $K_{\varphi \leqslant \alpha}$ различных $(\varphi \leqslant \alpha)$-оптимальных моделей. Наша задача найти такие модели $Fuz(E_{\min})$, $Fuz(E_{\max}) \in K_{\varphi \geqslant \alpha/\varphi \leqslant \alpha}$, чтобы выполнялись условия:

$$\inf(\xi_{E_{\min}}(\psi(a))) =$$
$$= \min\{\inf(\xi_{E'}(\psi(a))) \mid Fuz(E') \in K_{\varphi \geqslant \alpha/\varphi \leqslant \alpha}\};$$

$$\sup(\xi_{E_{\max}}(\psi(a))) =$$
$$= \max\{\sup(\xi_{E'}(\psi(a))) \mid Fuz(E') \in K_{\varphi \geqslant \alpha/\varphi \leqslant \alpha}\}.$$

Рассмотрим алгоритм выбора $Fuz(E_1)$ и $Fuz(E_2)$ для класса $K_{\varphi \geqslant \alpha}$. В этом случае выполняется условие: $\beta_1 < \alpha < 1$. Согласно Следствию 1 имеем два случая.

Случай 1. Выполняется условие $\alpha \geqslant \dfrac{\beta_1}{\beta_2}$. Тогда для любой модели $Fuz(E') \in K_{\varphi \geqslant \alpha}$ имеем $E' = T(E, \varphi) \cup N$, где $N \subseteq N(E, \varphi)$. Таким образом, разброс значений истинности предложения $\psi(a)$ на моделях из класса $K_{\varphi \geqslant \alpha}$ полностью зависит от того, какие именно пре-

цеденты из множества прецедентов $N(E, \varphi)$ были выбраны. Разделим множество $N(E, \varphi)$ на три подмножества (согласно формулам, определенным в (1)):

$$T(N(E, \varphi), \psi),\ F(N(E, \varphi), \psi),\ N(N(E, \varphi), \psi).$$

Для построения модели $Fuz(E_{\min})$ в первую очередь выбираем прецеденты из множества $F(N(E, \varphi), \psi)$. Затем, если $\|F(N(E, \varphi), \psi)\| < \|N\|$, то выбираем прецеденты из множества $N(N(E, \varphi), \psi)$. И в последнюю очередь, если $\|F(N(E, \varphi), \psi)\| + \|N(N(E, \varphi), \psi)\| < \|N\|$, выбираем прецеденты из множества $T(N(E, \varphi), \psi)$.

Для построения модели $Fuz(E_{\max})$ приоритет выбора прецедентов следующий: $T(N(E, \varphi), \psi)$, $N(N(E, \varphi), \psi)$, $F(N(E, \varphi), \psi)$.

Случай 2. Выполняется условие $\alpha < \dfrac{\beta_1}{\beta_2}$. Тогда для любой модели $Fuz(E') \in K_{\varphi \geqslant \alpha}$ имеем $E' = T(E, \varphi) \cup F$, где $F \subseteq F(E, \varphi)$. Тогда, по аналогии со случаем 1, разделим множество $F(E, \varphi)$ на три подмножества: $T(F(E, \varphi), \psi)$, $F(F(E, \varphi), \psi)$, $N(F(E, \varphi), \psi)$.

Для построения модели $Fuz(E_{\min})$ приоритет выбора прецедентов следующий: $F(F(E, \varphi), \psi)$, $N(F(E, \varphi), \psi)$, $N(F(E, \varphi), \psi)$.

Для построения модели $Fuz(E_{\max})$ приоритет выбора прецедентов следующий: $T(F(E, \varphi), \psi)$, $N(F(E, \varphi), \psi)$, $F(F(E, \varphi), \psi)$.

Алгоритм выбора $Fuz(E_1)$ и $Fuz(E_2)$ для класса $K_{\varphi \leqslant \alpha}$, согласно Следствию 2, также разбивается на два случая: $\alpha \leqslant 1 - \dfrac{1-\beta_2}{1-\beta_1}$ и $\alpha > 1 - \dfrac{1-\beta_2}{1-\beta_1}$.

Заключение

Статья посвящена описанию математической формализации и алгоритмической части вопросно-ответной системы с ограниченным доменом QA-RiskPanel. Данная вопросно-ответная система основана на прецедентном подходе к моделированию предметных областей и позволяет пользователю задавать вероятностные вопросы в целях определения и прогнозирования различных рисков, связанных с компьютерными атаками.

База знаний системы QA-RiskPanel состоит из множества прецедентов компьютерных атак. Исходя из этих прецедентов оценивается вероятность различных утверждений, имеющих отношение к безопасности корпоративной информационной системы.

На сегодняшний день в вопросно-ответной системе QA-RiskPanel реализованы модули обработки трех видов вопросов: безусловных, условных и модальных. Первый модуль ориентирован на обработку информации в ситуации, когда о начавшемся компьютерном нападении ничего не известно. Второй модуль работает в условиях, когда уже имеется некоторая вероятностная информация о начавшейся компьютерной атаке. Целью работы третьего модуля является обеспечение возможности получения информации о многошаговых атаках.

Для каждого модуля разработаны шаблоны вопросов и алгоритмы поиска ответов. Все алгоритмы разработаны на основе методологии теории нечетких моделей и имеют полиномиальную сложность.

Список литературы

1. **Официальный сайт МВД России.** URL: https://xn--b1aew.xn--p1ai/news/item/7693833. Доступ: 24.10.2016.
2. **Васенин В. А.** К созданию международной системы мониторинга и анализа информационного пространства для предотвращения и прекращения военно-политических киберконфликтов // Информационные технологии. 2012. № 9. С. 2—10.
3. **Михайлов В. Ю., Гридин В. Н., Мазепа Р. Б.** Безопасное информационное взаимодействие. Проблемы и решения // Информационные технологии. 2014. № 10. С. 72—77.
4. **Пальчунов Д. Е., Яхъяева Г. Э., Хамутская А. А.** Программная система управления информационными рисками RiskPanel // Программная инженерия. 2011. № 7. С. 29—36.
5. **Яхъяева Г. Э., Ясинская О. В.** Применение методологии прецедентных моделей в системе риск-менеджмента, направленного на раннюю диагностику компьютерного нападения // Вестник НГУ. Серия: Информационные технологии. 2012. Т. 10, вып. 2. С. 106—115.
6. **Yakhyaeva G. E., Yasinskya O. V.** Application of Case-based Methodology for Early Diagnosis of Computer Attacks // Journal of Computing and Information Technology — CIT 22, 2014. Vol. 3. P. 145—150.
7. **ДСМ-метод** автоматического порождения гипотез: Логические и эпистемологические основания. Сост. Аншаков О. М., Фабрикантова Е. Ф.; Под общ. ред. Аншакова О. М. М.: Книжный дом "ЛИБРОКОМ", 2009. 432 с.
8. **Ganter B., Stumme G., Wille R.** Formal Concept Analysis. Foundations and Applications. Berlin Heidelberg: Springer-Verlag. 2005.
9. **Яхъяева Г. Э., Ясинская О. В., Карманова А. А.** Вероятностная вопросно-ответная система в области компьютерной безопасности // Вестник НГУ. Серия: Информационные технологии. 2014. Т. 12, вып. 3. С. 132—145.
10. **Рассел С., Норвиг П.** Искусственный интеллект: современный подход. 3-е изд. М.: Вильямс, 2015. 1408 с.
11. **Hajek A.** Interpretation of probability // The Stanford Encyclopedia of Philosophy (Winter 2012 Edition), Edward N. Zalta (ed.), 2007. URL: http://plato.stanford.edu/archives/win2012/entries/probability-interpret/ [Электронный ресурс] Доступ: 24.10.2016.
12. **Palchunov D. E., Yakhyaeva G. E.** Interval fuzzy algebraic systems // Mathematical Logic in Asia. Proceedings of the 9-th Asian Logic Conference. World Scientific Publishers. 2006. P. 191—202.
13. **Пальчунов Д. Е., Яхъяева Г. Э.** Нечеткие логики и теория нечетких моделей // Алгебра и логика. 2015. Т. 54, № 1. С. 109—118.
14. **Baader F., Calvanese D., Mcguinness D. L., Nardi D., Patel-Schneider P. F.** The description logic handbook: Theory, implementation, and applications // Cambridge: Cambridge University Press, 2007.
15. **Пальчунов Д. Е., Яхъяева Г. Э**. Нечеткие алгебраические системы // Вестник НГУ. Серия: Математика, механика, информатика. 2010. Т. 10, вып. 3. С. 75—92.
16. **Gillies D.** Philosophical Theories of Probability. London: Routledge, 2012, 240 p.
17. **Яхъяева Г. Э., Ясинская О. В.** Методы согласования знаний по компьютерной безопасности, извлеченных из различных документов // Вестник НГУ. Серия: Информационные технологии. 2013. Т. 11, вып. 3. С. 63—73.
18. **Белнап Н., Стил Т.** Логика вопросов и ответов. М.: Прогресс. 1981. 288 с.
19. **Alhomidi M., Reed M.** Attack graph-based risk assessment and optimization approach // International Journal of Network Security & Its Applications. 2014, Vol. 6, N. 3. P. 31—43.
20. **Yakhyaeva G. E., Yasinskya O. V.** An Algorithm to Compare Computer-Security Knowledge from Different Sources // Proceedings

of the 17th International Conference on Enterprise Information Systems. 2015. P. 565—572.

21. **Яхъяева Г. Э., Ершов А. А.** О применении прецедентного подхода к анализу многошаговых компьютерных атак // Известия Юго-Западного государственного университета. Серия Управление, вычислительная техника, информатика. Медицинское приборостроение. 2016. Т. 18. № 1. С. 33—36.

22. **Yakhyaeva G. E., Ershov A. A.** Knowledge Base System for Risk Analysis of the Multi-Step Computer Attacks // Proceedings of the 18th International Conference on Enterprise Information Systems. 2016. Vol. 2. P. 143—150.

23. **Blackburn P., Van Benthem J., Wolter F.** Handbook of Modal Logic // Amsterdam: Elsevier. 2007.

G. E. Yakhyaeva, Associate Professor, e-mail: gul_nara@mail.ru,
A. A. Karmanova, Master Student, e-mail: anast.karmy.aa@gmail.com,
A. A. Ershov, Master Student, e-mail: alaershov@gmail.com,
N. P. Savin, Master Student, e-mail: npsavin@rambler.ru,
Novosibirsk State University, Novosibirsk

Question-Answering System for Managing of the Information Risks Based on Model-Theoretic Formalization of the Object Domains

The work is devoted to describing the probabilistic question-answer system QA-RiskPanel which let user to interactively analyze the various risks related to computer attacks. QA-RiskPanel is a knowledge based QA-system. QA-RiskPanel system is using a constantly updated database of precedents of computer attacks that enables actualization of the risk prediction. The ontological approach to the formalization of the object domains allows the analysis of risks at various levels of specification/generalization.

The article is provided a model-theory formalization of the knowledge base considered by object domain. The classification of types of questions is described. All types of questions have probabilistic specificity in this system. We present algorithms for finding the answers to the questions of all types.

Keywords: *information security, computer attacks, case of the computer attack, knowledge base, question-answering system, theory of the fuzzy models, generalized fuzzy model*

References

1. **Officialni** *siit MVD Rossii*. https://xn--b1aew.xn--p1ai/news/item/7693833 [Electronic resource]

2. **Vasenin V. A.** K sozdaniu megdunarodnoi sistemi monitoring i analiza informacionnogo prostranstva dly predotvrashenia i prekrachenia voenno-poluticheskih kiberkonfliktov, *Informacionnye tehnologii*, 2012, no. 9, pp. 2—10.

3. **Mikhaylov V. Yu., Gridin V. N., Mazepa R. B**. Bezopasnoe informacionnoe vzaimodeistvie. Problemi i reshenia, *Informacionnye tehnologii*, 2014, no. 10, pp. 72—77.

4. **Palchunov D. E., Yakhjyaeva G. E., Hamutskya A. A.** Programnaiy sistema upravleniy informacionnimi riskami RiskPanel, *Programmnaya ingeneriya*, 2011, no. 7, pp. 29—36.

5. **Yakhyaeva G. E., Yasinskaya O. V.** Primenenie metodologii precedentnih modelei v sisteme risk-menedgementa, napravlenogo na rannuudiagnostiku komputernogo napadenia, *Vestnik Novosib. gos. un-ta*. Seriya: *Informacionnye tehnologii*, 2012, vol. 10, no. 2, pp. 106—115.

6. **Yakhyaeva G. E., Yasinskya O. V.** Application of Case-based Methodology for Early Diagnosis of Computer Attacks, *Journal of Computing and Information Technology — CIT 22*, 2014, vol. 3, pp. 145—150.

7. **DSM-metod** *avtomaticheskogo porogdenia giptez: Logicheskie i epistemologicheskie osnocania*. Ed.: Anshakov O. M. Moscow. Knigni dom "LIBRIKOM", 2009. 432 p.

8. **Ganter B., Stumme G. и Wille R.** *Formal Concept Analysis. Foundations and Applications.* Berlin Heidelberg: Springer-Verlag. 2005.

9. **Yakhyaeva G. E., Yasinskya O. V., Karmanova A. A.** Verоyatnostnaya voprosno-otvetnaya sistema v oblasti komputernoi, *Vestnik Novosib. gos. un-ta. Seriya: Informacionnye tehnologii*, 2014, vol. 12, no. 3, pp. 132—145.

10. **Russell S. J., Norvig P.** *Artificial Intelligence: A Modern Approach*. 3nd Edition, Moscow, Vilyams, 2015, 1408 p.

11. **Hajek A.** *Interpretation of probability*. The Stanford Encyclopedia of Philosophy (Winter 2012 Edition), Edward N. Zalta (ed.), URL: http://plato.stanford.edu/archives/win2012/entries/probability-interpret/

12. **Palchunov D. E., Yakhyaeva G. E.** Interval fuzzy algebraic systems, *Mathematical Logic in Asia. Proceedings of the 9-th Asian Logic Conference*. World Scientific Publishers. 2006, pp. 191—202.

13. **Palchunov D. E., Yakhyaeva G. E.** Nechetkie logiki i teoria nechetkih modelei, *Algebra i Logika*, 2015, vol. 54, no. 1, pp. 74—80.

14. **Baader F., Calvanese D., Mcguinness D. L., Nardi D., Patel-Schneider P. F**. *The description logic handbook: Theory, implementation, and applications*, Cambridge: Cambridge University Press, 2007.

15. **Palchunov D. E., Yakhyaeva G. E.** Nechetkie algebraicheskie sistemi. *Vestnik Novosib. gos. un-ta. Seriya: Matematica, mehanica, informatica*, 2010, vol. 10, no. 3, pp. 75—92.

16. **Gillies D.** *Philosophical Theories of Probability*, London: Routledge, 2012, 240 p.

17. **Yakhyaeva G. E., Yasinskaya O. V.** Metodi soglasovania znanij po komputernoj bezopasnosti, izvlechennih iz razlichnih dokumentov, *Vestnik Novosib. gos. un-ta. Seriya: Informacionnye tehnologii*, 2013, vol. 11, no. 3, pp. 63—73.

18. **Belnap N. D., Steel T. B.** *The Logic of Questions and Answers*, London: Yale University Press, 1976, 176 p.

19. **Alhomidi M., Reed M.** Attack graph-based risk assessment and optimization approach, *International Journal of Network Security & Its Applications*, 2014, vol. 6, no. 3, pp. 31—43.

20. **Yakhyaeva G. E., Yasinskya O. V.** An Algorithm to Compare Computer-Security Knowledge from Different Sources, *Proceedings of the 17th International Conference on Enterprise Information Systems*, 2015, pp. 565—572.

21. **Yakhyaeva G. E., Ershov A. A.** O primenenii precedentnogo podhoda k analizu mnogishagovih komputernih atak, *Izvestia Ugo-Zapadnogo gos. un-ta. Seriya: Upravlenie, vichislitelnay tehnika, informatika. Medicinskoe priborostroenie*, 2016, vol. 18, no. 1, pp. 33—36.

22. **Yakhyaeva G. E., Ershov A. A.** Knowledge Base System for Risk Analysis of the Multi-Step Computer Attacks, *Proceedings of the 18th International Conference on Enterprise Information Systems*, 2016, vol. 2. pp. 143—150.

23. **Blackburn P., Van Benthem J., Wolter F.** *Handbook of Modal Logic*, Amsterdam: Elsevier. 2007.

УДК 539.3 DOI:10.17587/it.23.107-112

Г. Г. Булычев, д-р физ.-мат. наук, проф., e-mail: geo-bulychev@mail.ru,
Московский технологический университет (МИРЭА)

Метод пространственных характеристик в задачах механики деформируемого твердого тела. Часть 2*

Методом численного моделирования на ПК решается задача исследования разрушения двухэтажного дома при сейсмической нагрузке. В качестве математической модели динамики и динамического разрушения дома выбрана характеристическая форма соответствующих уравнений и используется метод пространственных характеристик. Приводятся все этапы выполнения проекта и анализируются результаты моделирования. Программирование проводится на языке ФОРТРАН FPS-11/2012, при этом используется ПК с процессором FX8350 и оперативной памятью 8 Гбайт и авторская графика.

Ключевые слова: *метод пространственных характеристик, численное моделирование, динамика и разрушение строительных сооружений и конструкций*

Моделирование динамического разрушения дома при действии сейсмической нагрузки

План моделируемого двухэтажного сооружения показан на рис. 1, *а*, *б*, где цифрами обозначены оконные дверные и лючный проемы. Оконные проемы на обоих рисунках одинаковы, также одинаковы и дверные проемы. Дверные проемы обозначены цифрами 1—3 и 9, 10; оконные проемы — цифрами 4—7 и 11—15. Лючный проем обозначен штриховой линией и цифрой 8. Рис. 1, *а* соответствует плану первого этажа сооружения, а рис. 2, *б* — плану второго этажа. Все проемы и внутренние границы сооружений предполагаются свободными от напряжений. Здание рассматривается как двухслойный упруговязкопластический материал (фундамент и здание) с разными характеристиками слоев. Предполагается, что сейсмическая нагрузка действует на нижнюю поверхность фундамента.

Модель и метод. Математическая модель динамики материала здания и фундамента состоит из уравнений движения, условий аддитивности упругих и вязкопластических деформаций, уравнений Коши, закона пластичности и закона Гука для упругих деформаций изотропного упруговязкопластического материала. Эти уравнения, соответственно, имеют следующий вид:

$$\partial_j \sigma_{ij}^k = \rho^k \partial_t V_i^k;\ \varepsilon_{ij}^k = (\varepsilon_{ij}^k)^e + (\varepsilon_{ij}^k)^{vp};$$

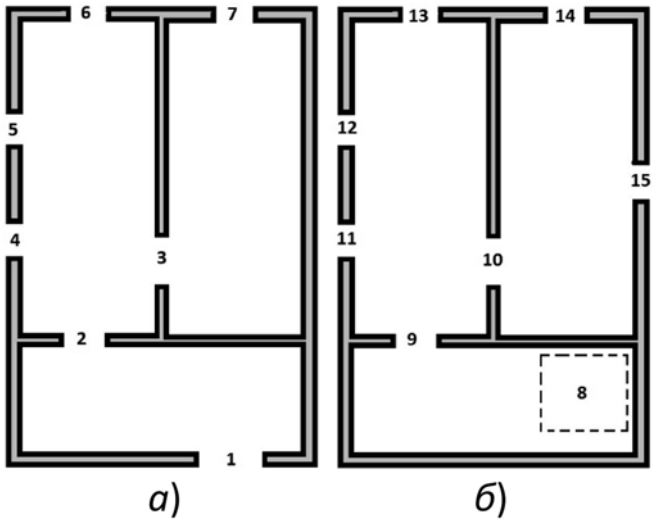


Рис. 1

Рис. 2

* Часть 1 статьи опубликована в № 1, 2017 г.

$$\sigma_{ij}^k = 2\mu^k(\varepsilon_{ij}^k)^e + \lambda^k\delta_{ij}(\varepsilon_{mm}^k)^e;$$
$$\partial_t\varepsilon_{ij}^k = (\partial_j V_j^k + \partial_i V_j^k)/2;$$
$$(\varepsilon_{ij}^k)^{vp} = \Phi^k(\sigma_{ij}^k);\ i, j = 1, 2, 3,\ k = 1, 2, \qquad (1)$$

где σ_{ij}^k, ε_{ij}^k, V_i^k, ρ^k — напряжения, деформации в k-м слое, скорости частиц и плотность материала в k-м слое, μ^k и λ^k — параметры Ламе; $\Phi^k(\sigma_{ij}^k) = \Phi^k(S^k - k_S^k)$ — функция пластичности, в которой $S^k = \sqrt{s_{ij}^k s_{ij}^k/2}$ — интенсивность напряжений в материале, $s_{ij}^k = \sigma_{ij}^k - \delta_{ij}p^k$ — девиатор напряжений, $p^k = \sigma_{ii}^k/3$ — гидростатическое давление, k_S^k — статический предел текучести материала при сдвиге.

Видно, что модель (1) является частным случаем модели динамики ортотропного тела [1], поскольку:

1) второе уравнение системы (1) получается из общего вида закона Гука при следующих упрощениях в матрице жесткости C: $c_{11}^k = c_{22}^k = c_{33}^k = 2\mu^k + \lambda^k$, $c_{44}^k = c_{55}^k = c_{66}^k = \mu^k$, $c_{12}^k = c_{13}^k = c_{23}^k = \lambda$;

2) внутренние силы и геометрические члены отсутствуют;

3) функция пластичности выбирается в простейшем виде:

$$\Phi_{\alpha j}^k = \gamma^k\sqrt{S^k - k_S^k}\cdot s_{\alpha j}^k/S \text{ при } S^k \geqslant k_S^k$$
$$\text{и } \tilde{\Phi}_{\alpha j}^k = 0 \text{ при } S^k < k_S^k.$$

Следовательно, характеристическая форма (1) получается из соответствующей характеристической формы при указанных выше упрощениях.

Проведем обезразмеривание системы (1) с помощью соотношений

$$\tilde{x}_i = x_i/x_0,\ \tau^k = c_0^k t/x_0,\ \tilde{\sigma}_{ij}^k = \sigma_{ij}^k/k_S^k,$$
$$\tilde{S}^k = S^k/k_S^k,\ \tilde{V}_i^k = V_i^k/V_S^k,\ k_S^k = \rho^k c_0^k V_S^k,$$
$$\gamma^k = x_0/(V_S^k\tau_0^k),$$

в которых x_0 — нормировочная константа, за которую принят размер сооружения по оси x_1 ($x_0 = l$, где l — максимальный линейный размер строения); V_S^k — скорость частиц материала, при которой начинается пластическое течение, $\alpha, i, j = 1, 2, 3,\ k = 1, 2$; по повторяющимся греческим индексам суммирование не производится, $c_0^k = \sqrt{(2\mu^k + \lambda^k)/\rho^k}$ — скорость продольной волны в слое, τ_0^k — время задержки текучести.

После проведения указанных упрощений и обезразмеривания характеристическая форма системы (1) примет вид: для продольной волны, движущейся со скоростью c_0^k в обе стороны вдоль оси x_α

$$(\partial/\partial\tau^k \pm \partial/\partial x_\alpha)(\sigma_{\alpha\alpha}^k \mp V_\alpha^k) = \partial/\partial x_\beta(\nu_1^k V_\beta^k \pm \sigma_{\alpha\beta}^k) +$$
$$+ \partial/\partial x_\gamma(\nu_1^k V_\gamma^k \pm \sigma_{\alpha\beta}^k) - \Phi_{\alpha\alpha}^k - \nu_1^k(\Phi_{\beta\beta}^k + \Phi_{\gamma\gamma}^k), \qquad (2)$$

для поперечных волн, распространяющихся со скоростью $c_\perp^k = \sqrt{\mu^k/\rho^k}$ вдоль той же оси

$$(\partial/\partial\tau^k \pm \xi^k\partial/\partial x_\alpha)(\sigma_{\alpha\beta}^k \mp \xi^k V_\beta^k) =$$
$$= \partial/\partial x_\beta(\nu_2^k V_\alpha^k/2 \pm \xi^k\sigma_{\beta\beta}^k) \pm \partial/\partial x_\gamma(\xi^k\sigma_{\beta\gamma}^k) - \Phi_{\alpha\beta}^k; \qquad (3)$$

$$(\partial/\partial\tau^k \pm \xi^k\partial/\partial x_\alpha)(\sigma_{\alpha\gamma}^k \mp \xi^k V_\gamma^k) =$$
$$= \partial/\partial x_\gamma(\nu_2^k V_\alpha^k/2 \pm \xi^k\sigma_{\gamma\gamma}^k) \pm \partial/\partial x_\beta(\xi^k\sigma_{\beta\gamma}^k) - \Phi_{\alpha\gamma}^k \qquad (4)$$

для неподвижных разрывов с нормалью x_α

$$\partial/\partial\tau^k(\sigma_{\beta\beta}^k - \nu_1^k\sigma_{\alpha\alpha}^k) = \nu_2^k(1 - \nu^k)^{-1}(\partial V_\beta^k/\partial x_\beta +$$
$$+ \nu^k\partial V_\gamma^k/\partial x_\gamma - \Phi_{\beta\beta}^k - \nu^k\Phi_{\gamma\gamma}^k); \qquad (5)$$

$$\partial/\partial\tau^k(\sigma_{\gamma\gamma}^k - \nu_1^k\sigma_{\alpha\alpha}^k) = \nu_2^k(1 - \nu^k)^{-1}(\partial V_\gamma^k/\partial x_\gamma +$$
$$+ \nu^k\partial V_\beta^k/\partial x_\beta - \Phi_{\gamma\gamma}^k - \nu^k\Phi_{\beta\beta}^k); \qquad (6)$$

$$\partial/\partial\tau^k\sigma_{\beta\gamma}^k = \nu_2^k[(\partial V_\beta^k/\partial x_\gamma + \nu^k\partial V_\gamma^k/\partial x_\beta)/2 - \Phi_{\beta\gamma}^k], \qquad (7)$$

где ν^k — коэффициент Пуассона, $\nu_1^k = \nu^k/(1 - \nu^k)$, $\nu_2^k = 1 - \nu_1^k$, $\xi^k = c_\perp^k/c_0^k$ — отношение скорости поперечной волны $c_\perp^k = \sqrt{\mu^k/\rho^k}$ к скорости продольной волны c_0^k. Значок "волна", поставленный над функциями при обезразмеривании, снят, поскольку в дальнейшем, если не оговорено, будут использоваться эти нормировки.

Заметим, однако, что формулы (2)—(7) не могут быть использованы непосредственно, если (как в данном случае) конструкции имеют кусочно-однородное строение. Это связано с тем, что нормировка времени в каждом слое различна и зависит от скорости звука c_0^k в нем. Для того чтобы исключить эту зависимость проведем следующие преобразования.

1. Выберем $c_0 = \max_k c_0^k$ и разделим все c_0^k на c_0, обозначим $\bar{c}_0^k \equiv c_0^k/c_0$.

2. Обозначим $\tau = c_0 t/x_0$, тогда $\tau^k = \bar{c}_0^k\tau$.

3. Умножим все уравнения (2)—(7) на $\bar{c}_0^k$ и обозначим $x_\eta^k = x_\eta/\bar{c}_0^k$, $\eta = \alpha, \beta, \gamma$. Обозначим $\bar{\gamma}^\alpha \equiv \bar{c}_0^\alpha\gamma^\alpha$.

Рассмотрим получившиеся уравнения, используя введенные соотношения. Так, например, уравнение (2) в новых обозначениях примет вид

$$(\partial/\partial\tau \pm \partial/\partial x_\alpha^k)(\sigma_{\alpha\alpha}^k \mp V_\alpha^k) = \partial/\partial x_\beta^k(\nu_1^k V_\beta^k \pm \sigma_{\alpha\beta}^k) +$$
$$+ \partial/\partial x_\gamma^k(\nu_1^k V_\gamma^k \pm \sigma_{\alpha\beta}^k) - \bar{\Phi}_{\alpha\alpha}^k - \nu_1^k(\bar{\Phi}_{\beta\beta}^k + \bar{\Phi}_{\gamma\gamma}^k), \quad (8)$$

где $\bar{\Phi}_{\eta\varsigma}^k = \bar{\gamma}^k \dfrac{\sqrt{S^k - 1}}{S^k} s_{\eta\varsigma}^k$, $\eta, \varsigma = \alpha, \beta, \gamma$.

Вид остальных уравнений изменится аналогичным образом.

Зададим изменение времени $\Delta\tau$, тогда положение фронта волны в k-м слое изменится на величину $\Delta x^\alpha = \pm\Delta\tau \cdot \bar{c}_0^\alpha$. Это соотношение определяет связь между временной и пространственными сетками, используемыми при численном моделировании динамики и динамического разрушения кусочно однородных изотропных тел методом пространственных характеристик. Заметим, что все величины, стоящие под знаком дифференциала в уравнениях (2)—(7) при этом не меняются, а алгебраические члены просто умножаются на константу ($\bar{c}_0^\alpha$); в силу изотропии тела эта константа оказывается одинаковой для всех направлений распространения волн.

Приведенные рассуждения показывают, что алгоритмы, разработанные для однородных изотропных тел, могут быть использованы и во внутренних точках кусочно-однородных тел, меняется только пространственная сетка и некоторые константы, задаваемые как входные данные.

В точках границы раздела слоев расчетные схемы могут быть получены модернизацией схем, описывающих возникновение и распространение трещин вдоль этой границы раздела. При этом на границе раздела k-го и $k+1$-го слоев с единичным вектором нормали n и касательной τ в зависимости от напряженно-деформированного состояния в ее окрестности и наличия или отсутствия трещин на ней выполняются либо условия непрерывности

$$\bar{V}_i^k = a\bar{V}_i^{k+1},\ \bar{\sigma}_{in}^k = b\bar{\sigma}_{in}^{k+1},$$
$$\text{где } a = V_S^{k+1}/V_S^k,\ b = k_S^{k+1}/k_S^k; \quad (9)$$

либо условия свободной поверхности

$$\sigma_{in}^k = \sigma_{in}^{k+1} = 0; \quad (10)$$

либо условия кулоновского трения

$$\bar{V}_n^k = a\bar{V}_n^{k+1},\ \bar{\sigma}_{nn}^k = b\bar{\sigma}_{nn}^{k+1},\ \bar{\sigma}_{n\tau}^k = -\bar{\sigma}_{n\tau}^{k+1},$$
$$\bar{\sigma}_{n\tau}^{k+1} = \omega^k|\bar{\sigma}_{nn}^{k+1}|\mathrm{sgn}(V_\tau^{k+1} - V_\tau^k), \quad (11)$$

где $\tau \neq n$ и $\omega^k < 1$ — коэффициент сухого трения. Черточка над функциями показывает, что нормировка проводится на k-й слой.

Другая важная проблема, возникающая при моделировании динамики на границе раздела слоев, заключается в несовпадении узлов сеток с разных сторон границы раздела. Эта проблема решается с помощью схем линейной интерполяции опорных точек: вначале опорные точки $k+1$-го слоя интерполируют таким образом, чтобы получить расчетную схему для контактной поверхности со стороны k-го слоя, и с их помощью рассчитывают σ_{ij} и V_i на k-м слое (эти значения используют в дальнейшем для расчетов на k-м слое), затем интерполируют опорные точки на k-м слое и с их помощью рассчитывают σ_{ij} и V_i на $k+1$-м слое и их используют для расчетов на $k+1$-м слое. В самих расчетных схемах проводят анализ напряжений и перемещений на внутренних границах контакта и на основе этого анализа принимается решение о применении того или иного из соотношений (9)—(11).

Критерии образования трещин во внутренних точках конструкции выбирают с учетом накопления повреждений в виде $J_n = A$ или $J_\tau = B$, где

$$J_n = \int_{\tau_0}^{\tau} <\sigma_n(\tau)/\sqrt{3} - 1> d\tau,\ J_{\tau i} = \int_{\tau_1}^{\tau} <\sigma_{\tau i}(\tau) - 1> d\tau,$$

$i = 1, 2$, константы A и B определяют экспериментально из опытов, а направление нормали n или касательных τ — по результатам моделирования динамики.

Здесь $<f> = f$ при $f > 0$ и $<f> = 0$ при $f \leqslant 0$; $A = B = 1$; τ_0 и τ_1 — моменты времени, в которые соответствующие напряжения впервые выходят в пластическую область.

До выполнения критериев разрушения во всех внутренних точках сооружений и фундаментов предполагается непрерывность скоростей частиц и напряжений. Далее предполагается, что трещина развивается в соответствии с условиями (9)—(11), в которых $a = b = 1$.

Граничные и начальные условия. Нагрузка состоит из двух частей: собственно сейсмической волны $P_l(t)$, имеющей треугольную форму, и упругой реакции грунта на основании фундамента, выражающейся формулой Винклера, в которой, в силу специфики принятой формы уравнений модели, коэффициент k, связывающий нормальные напряжения в фундаменте с деформациями грунта, оказывается нормированным на $\rho^1(c_0^1)^2$, где верхний индекс единица соответствует фундаменту. Вес сооружения и фундамента не учитывается, поскольку считается, что он уравновешен статической реакцией грунта и в процессе моделирования не меняется.

Общая формула нагружения в каждой точке $\{x_2, x_3\}$ основания фундамента ($x_1 = 0$) имеет вид

$$P(t, x_2, x_3) = ku^1(t, x_2, x_3) - P_l(t) \leqslant 0 \text{ и}$$
$$P(t, x_2, x_3) = 0 \text{ при } ku(t, x_2, x_3) - P_l(t) > 0, \quad (12)$$

где $u^1(t, 0, x_2, x_3) = \int_0^t V(t, 0, x_2, x_3)dt$ — перемещение нижней точки $\{x_2, x_3\}$ фундамента вниз по оси x_1. Первая формула в (12) относится к такому нагружению, при котором сохраняется контакт между

грунтом и подошвой основания сооружения, вторая формула — когда такой контакт нарушен и между подошвой и грунтом образовалась полость. В начальный момент здание находится в состоянии покоя.

Общая схема проекта представлена на рис. 2. С ее помощью решаются две задачи моделирования:

1. Исследование динамики здания под действием сейсмической нагрузки.

2. Исследование его динамического разрушения при нагрузке того же типа.

Решение указанных задач подразделяют на шесть последовательных этапов.

На первом этапе проводят моделирование динамики сооружений и фундаментов, разрушение их не предусматривается. Нагрузкой являются нормальные сжимающие напряжения, действующие на фундамент и моделирующие сейсмическую нагрузку [3].

Форма нагружающего импульса имеет вид напряжений $P_f = P_0 f(t)$, где $f(t)$ — треугольный импульс, амплитуда которого равна 1, длительность соответствует 30 временам пробега волны по максимальному размеру здания (что равно времени моделирования), отношение проекций сторон на основание — 1 к 3; амплитуда P_0 имеет такое значение, что разрушения, даже локального, не происходит. Массивы данных, полученные в результате моделирования, на этом этапе используют для графической обработки в целях определения максимальных напряжений и областей их локализации.

На втором этапе проводят указанную обработку. При этом используют примитивы OpenGL и авторские подпрограммы построения контуров и изолиний, разработанные на базе пакета ГРАФОР.

На третьем этапе выбирают критерии образования трещин и условия, возникающие на них в процессе нагружения. В качестве критериев начала разрушения используют условия накопления неупругих нормальных или касательных напряжений до заранее заданного значения. Дальнейшее поведение образовавшейся при этом трещины определяют с помощью соотношений (9)—(11).

На четвертом этапе проводят моделирование динамического разрушения конструкции при различных амплитудах нагружающих напряжений P_0. В качестве критерия полного разрушения конструкции принимается полное разрушение какого-либо ее сечения. Анализ разрушения проводится автоматически, а выполнение указанного критерия используется для окончания счета. Предполагается также, что разрушение происходит за время, не превышающее 30 времен пробега продольной волны по максимальному размеру сооружения.

На пятом этапе массивы, полученные в момент полного разрушения конструкции, подвергают анализу в целях определения областей катастрофического разрушения и характера разрушений в них.

И, наконец, *на шестом,* последнем, этапе решается задача определения минимальных нагрузок, приводящих к полному разрушению конструкций.

Параметры моделирования. В силу выбора нормировок изменяемыми физико-механическими параметрами являются: коэффициенты Пуассона ν^1 и ν^2 соответственно фундамента и строения; относительные продольные скорости $\bar{c}_0^1$ и $\bar{c}_0^2$; относительный предел текучести на сдвиг b; относительный предел скоростей начала текучести a; коэффициенты, определяющие начало текучести, γ^1 и γ^2; коэффициенты сухого трения — d^1 и d^2 и шаг по времени — h. Величина h определяется из нормировки на максимальный размер сооружения, который принимается за 1, и максимальной скорости, которая также принимается за 1. Остальные параметры геометрические и определяют (в ячейках) габаритные размеры и пространственную структуру строений и фундаментов; размер ячеек определяется пространственной сеткой в соответствующих слоях.

Строение имеет размеры ($x \times y \times z$) 80 × 144 × × 80 ячеек, толщина основания 4 ячейки, а толщина боковых и верхней стенок меняется и составляет 2, 3 или 4 ячейки, размеры окон 10 × 10 ячеек, двери — 20 × 10 ячеек. Размеры люка 24 × 32 ячеек. Расположение окон, дверей и люка показаны на планах рис. 1. Максимальный размер строения принимается равным единице.

На первом этапе моделирования критерии разрушения выбирают столь большими, что разрушения, даже локального, не происходит. При различном порядке следования слоев в зависимости от толщины фундамента или стен строения определяли напряжения в средней точке фундамента в течение всего времени моделирования. На рис. 3—6 показана эволюция нормального напряжения σ_{11} в срединной точке основания строения, $P = -0{,}2$, что соответствует сжатию. Время τ измеряется в пробегах продольных волн по наибольшему размеру строения (здесь по y). Штриховой линией на рис. 5 показана сейсмическая нагрузка. Коэффициент k в соответствии с результатами [4] выбран равным 0,0005.

Рис. 3 и 4 соответствуют случаю, когда фундамент изготовлен из более жесткого материала, при этом

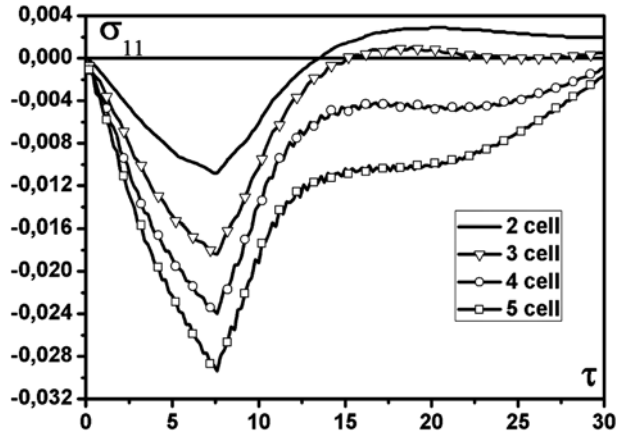


Рис. 3

$c_0^1 = 1$; $c_0^2 = 0{,}5$; $\nu^1 = 0{,}1$; $\nu^2 = 0{,}2$; $a = 1$; $b = 0{,}5$; $\omega^1 = 0{,}3$; $\omega^2 = 0{,}3$, $\bar{\gamma}^1 = 0{,}2$; $\bar{\gamma}^2 = 0{,}1$. Аргументом графиков является безразмерное время τ, определяемой величиной — нормальное напряжение σ_{11}, действующее в том же направлении, что и нагрузка. Это напряжение пронормировано на тот слой, где приложена нагрузка. Изменяемым параметром на рис. 3 является толщина здания, которая составляет, соответственно, 2, 3, 4 и 5 ячеек, фиксированным параметром является толщина фундамента, которая составляет 3 ячейки. На рис. 4 изменяемым параметром является толщина фундамента, которая изменяется от 1 до 5 ячеек, фиксированный параметр — толщина стен здания, которая составляет 3 ячейки.

На рис. 5 и 6 показаны те же напряжения σ_{11} в той же точке того же строения и при той же нагрузке, но при обратной последовательности материалов: здесь $c_0^1 = 0{,}5$; $c_0^2 = 1$; $\nu^1 = 0{,}2$; $\nu^2 = 0{,}1$; $\omega^1 = 0{,}3$; $\omega^2 = 0{,}3$, $\bar{\gamma}^1 = 0{,}1$; $\bar{\gamma}^2 = 0{,}2$; $a = 1$; $b = 2$. Рис. 5, так же как и рис. 3, соответствует случаю, когда изменяемым параметром является толщина стенок здания (от 2 до 5 ячеек), а неизменяемым — толщина фундамента (3 ячейки). На рис. 6 изменяемым параметром является толщина фундамента: число ячеек фундамента по его толщине меняется от 1 до 10 с учетом того, что размер ячейки сетки в фундаменте в этом случае в 2 раза меньше, чем в случае, показанном на рис. 3; толщина здания остается неизменной и составляет 3 ячейки.

Анализ графиков, приведенных на рис. 3—6, показывает, что в том случае, когда материал фундамента прочнее материала строения, напряжения σ_{11}, создаваемые нагрузкой в основании строения, в 2 раза меньше.

Во второй постановке определяются амплитуды минимальных сейсмических нагрузок, приводящих к разрушению сооружений. При моделировании сооружение считается разрушенным, если полностью разрушенным оказывается одно из его поперечных сечений. Это требование является достаточным для разрушения, но не является необходимым. Поэтому полученные значения нагрузок оказываются завышенными. В критериальных величинах $A = 1$ и $B = 1$; поскольку разрушение может иметь сложный характер рассматриваются 9 критериев разрушения: 3 по нормальным напряжениям и 6 по касательным. Анализ механизмов разрушения показывает, что возможно как пересечение трещин (с частичным локальным залечиванием материала в области пересечения), так и слияние трещин одного направления с образованием магистральной трещины.

При моделировании изменялась как толщина фундамента, так и толщина всех стен сооружения (толщина стен предполагается одинаковой). В соответствии с анализом рис. 3—6, предполагалось, что жесткость фундамента выше жесткости строения ($b = 0{,}5$).

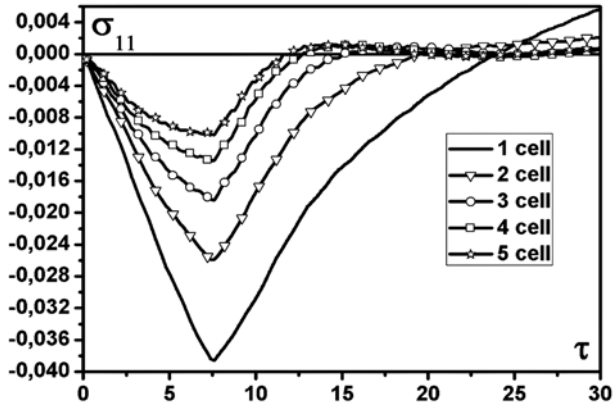


Рис. 4

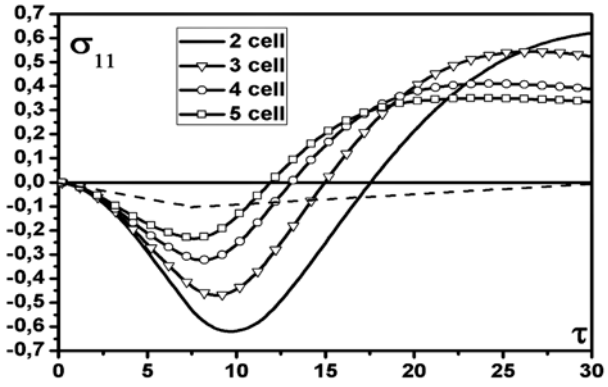


Рис. 5

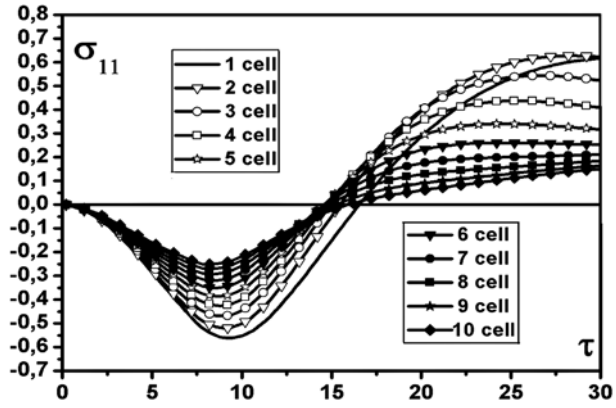


Рис. 6

Таблица

	2	3	4	5	6
1	−0,01	−0,35	−0,41	−0,45	−0,41
2	−0,02	−0,38	−0,52	−0,46	−0,46
3	−0,02	−0,37	−0,54	−0,55	−0,45
4	−0,02	−0,35	−0,56	−0,56	−0,45
5	−0,02	−0,34	−0,54	−0,54	−0,41
6	−0,02	−0,33	−0,52	−0,54	−0,42

В таблице приведены значения минимальных амплитуд сейсмических нагрузок P_f, приводящих к полному разрушению строений. Нагрузка P_f имеет треугольную форму, показанную штриховой линией на рис. 5. В таблице цифрами 1—6 по вертикали указана выраженная в количестве ячеек

толщина фундамента; цифрами 2—6 по горизонтали — толщина стен и перегородок сооружений (они приняты одинаковыми), выраженная также в ячейках. Знак минус соответствует сжатию.

Анализ данных, приведенных в таблице, показывает, что при минимальной толщине стен — 2 ячейки, независимо от толщины фундамента разрушающие напряжения крайне малы. Далее, с ростом толщины стен, эти напряжения быстро нарастают; зависимость от толщины фундамента при одной и той же толщине стен оказывается более слабой. Многообразие возможных механизмов разрушения и движение трещин в различных направлениях может приводить как к их частичному залечиванию — когда трещины различных направлений, пересекаясь, останавливают друг друга, или ускоряют разрушение, когда трещины одного направления, двигающиеся по одной прямой, сливаются. Эти эффекты проявляются тогда, когда толщина стен или фундамента оказывается достаточно большой, — при этом минимальные разрушающие напряжения P_f могут колебаться в значительных пределах. Максимум разрушающей нагрузки равен −0,56 и достигается при толщине стен в 4 или 5 ячеек и толщине фундамента 4 ячейки, при больших толщинах стен вследствие слияния трещин одного направления в фундаменте минимальные разрушающие нагрузки несколько падают.

Заключение

В работе с помощью метода пространственных характеристик проводится моделирование практически важной и достаточно сложной задачи — динамики и динамического разрушения малоэтажных строений под действием сейсмической нагрузки. Использование характеристической формы представления математической модели процессов позволило решить эту задачу на ПК.

При моделировании динамики были определены наиболее уязвимые точки строения и наиболее опасные напряжения в них. В соответствии с заданными критериями разрушения были найдены точки начала локального разрушения. С использованием двойной пространственной сетки и заданной в зависимости от напряжений в окрестностях трещины последовательности условий на трещине рассмотрено распространение, слияние и залечивание трещин в стенах и фундаменте здания вплоть до его полного разрушения.

Проанализированы зависимости напряжений, возникающих в фундаменте, при различных свойствах фундамента и различной его толщине. Определены минимальные амплитуды сейсмических нагрузок, при которых происходит полное разрушение здания, в зависимости от толщины фундамента и толщины стен сооружения.

Список литературы

1. **Булычев Г. Г.** Построение матричной характеристической формы уравнений динамики анизотропных упруговязкопластических сред // Известия РАН. Механика твердого тела. 1995. № 1. С. 91—95.
2. **Булычев Г. Г., Кукуджанов В. Н.** Динамическое разрушение предварительно напряженного волокнистого композита, вызванное обрывом волокна // Известия РАН. Механика твердого тела. 1993. № 3. С. 207—214.
3. **Тяпин А. Г.** Расчет динамического отрыва фундаментной плиты от основания при сейсмическом воздействии. Часть 2: Простейшее сочетание отрыва с запаздывающими жесткостями основания // Строительная механика и расчет сооружений. 2013. № 3. С. 39—43.
4. **Булычев Г. Г.** Динамическое разрушение малоэтажных строений под действием сейсмической нагрузки // Строительная механика и расчет сооружений. 2015. № 4. С. 40—45.

G. G. Bulychev, D.Sc., Professor, Moscow Technological University, Moscow

Method of Spatial Characteristics in Problems of a Mechanics of a Deformable Solid Body. Part 2*

By the method of numerical simulation on PC the research problem of destruction of the two-storeyed house is solved at seismic loading. As mathematical model of dynamics and dynamic destruction of the house the characteristic form of the corresponding equations is chosen and the method of spatial characteristics is used. All stages of execution of the project are resulted and results of modelling are analyzed. Programming execute in language FORTRAN FPS-11/2012, thus was used PC with processor FX8350 and with operative memory of 8 Gb and the author's drawing.

Keywords: *method of spatial characteristics, numerical modelling, dynamics and destruction of building constructions and designs*

References

1. **Bulihchev G. G.** Postroenie matrichnoyj kharakteristicheskoyj formih uravneniyj dinamiki anizotropnihkh uprugovyazkoplasticheskikh sred, *Izvestiya RAN. Mekhanika tverdogo tela,* 1995, no. 1, pp. 91—95.
2. **Bulihchev G. G., Kukudzhanov V. N.** Dinamicheskoe razruhshenie predvariteljno napryazhennogo voloknistogo kompozita, vihzvannoe obrih-vom volokna, *Izvestiya RAN, Mekhanika tverdogo tela,* 1993, no. 3, pp. 207—214.
3. **Tyapin A. G.** Raschet dinamicheskogo otrihva fundamentnoyj plitih ot osnovaniya pri seyjsmich-eskom vozdeyjstvii. Chastj 2: prosteyjshee sochetanie otrihva s zapazdihvayuthimi zhestkostyami osnovaniya, *Stroiteljnaya mekhanika i raschet sooruzheniyj*, 2013, no. 3, pp. 39—43.
4. **Bulihchev G. G.** Dinamicheskoe razrushenie maloehtazhnihkh stroeniyj pod deyjstviem seyjsmicheskoyj nagruzki, *Stroiteljnaya mekhanika i raschet sooruzheniyj*, 2015, no. 4, pp. 40—45.

*Part 1 was published in N. 1, 2017.

УДК 004.94 DOI:10.17587/it.23.113-119

Д. Л. Кишлаков[1], студент, e-mail: daniel_kish@outlook.com,
П. В. Тараканов[1, 2], аспирант, e-mail: pashabeetle@yandex.ru,
Г. В. Шашурин[1, 2], канд. техн. наук, декан ф-та, e-mail: shashuring@mail.ru,
Ю. В. Берчун[1], ст. преподаватель, e-mail: y_berchun@mail.ru,
[1]Московский государственный технический университет им. Н. Э. Баумана, г. Москва
[2]Институт машиноведения им. А. А. Благонравова Российской академии наук, г. Москва

Эффективность облачных вычислений в моделировании кинетики трещин в наводороженных элементах конструкций

Рассматриваем задачу исследования роста трещин в наводороженных элементах конструкций. Представляем математическую модель задачи, и на этой основе — алгоритм определения предельной длины трещины. Разработанное математическое и алгоритмическое обеспечение реализовано в виде так называемых локального (последовательного) и параллельного (распределенного) программных приложений. Поставлена задача оценки эффективности последнего приложения при использовании веб-службы Microsoft Azure Batch. С помощью вычислительного эксперимента показана зависимость эффективности приложения от параметров модели роста трещин. Получены характеристики роста трещин в зависимости от свойств исследуемых материалов.

Ключевые слова: *автоматизированное проектирование, долговечность, охрупчивание, параллельные вычисления, облачные технологии, эффективность*

Введение

Актуальность автоматизации проектирования долговечных элементов конструкций, подверженных статическому нагружению в условиях длительного воздействия агрессивных сред, обусловлена развитием нефтехимической, нефтеперерабатывающей промышленности и водородной энергетики, что во многом объясняется высоким уровнем ущерба, который преждевременный отказ элементов конструкций может нанести экологии, экономике и безопасности населения [1, 2].

При проектировании указанных элементов конструкций применяют системы автоматизированного проектирования (САПР) с модулями прочностного анализа, позволяющими проводить оценку ресурса (долговечности) проектируемых изделий. При этом в качестве основы расчета долговечности используют классические, хорошо зарекомендовавшие себя модели накопления повреждений [3, 4], не учитывающие влияния агрессивных сред на ресурс, что может привести к получению его завышенного значения. Для уточненного расчета ресурса разрабатывают специализированные модели накопления повреждений [5, 6], в которых оценку влияния нагружения, свойств материала и агрессивной среды на долговечность проводят путем введения соответствующих параметров в уравнения кинетики растущих трещин. Такой подход позволяет оценить влияние агрессивных сред на весь процесс роста трещин: от начальных производственно-технологических дефектов до предельного состояния элементов конструкций, когда, собственно, и происходит исчерпание их ресурса.

Комплекс проектных задач требует многократного вычисления ресурса при различных значениях параметров модели накопления повреждений, что необходимо для исследования чувствительности модели к значениям ее параметров, а также для статистического моделирования роста трещин. Последнее особенно важно ввиду того, что при прогнозировании разрушения в проектируемом объекте трудно предсказать расположение и типы возможных дефектов, поэтому приходится прибегать к статистическому моделированию разрушения тел, ослабленных случайными дефектами, как, например, описано в работе [7]. Это требование приводит к необходимости построения программного решения, которое бы осуществляло последовательно-параллельное решение множества однотипных задач определения ресурса изделия.

Организация последовательно-параллельной обработки обсуждаемых задач подразумевает использование многопроцессорной ЭВМ с распределенной памятью, в связи с чем встает вопрос о балансировке ее загрузки. Используем динамический метод балансировки загрузки, основанный на наличии централизованной очереди задач, из которой процессоры ЭВМ последовательно извлекают задачи и решают их [8].

Существуют различные способы организации параллельных вычислений с названным методом балансировки загрузки. Наиболее распространенной технологией параллельных вычислений является технология *MPI*. Однако при использовании *MPI* возникает необходимость самостоятельно реализовывать механизм очереди задач. Существуют другие технологии организации параллельных вы-

числений, где данная задача уже решена. Так, в настоящее время одним из возможных способов организации параллельных вычислений является использование пакетной службы *Azure* [9], позволяющей строить распределенные (сервис-ориентированные [10, 11]) приложения. Такие приложения используют облачные вычислительные мощности для большого числа вычислительных процессов в режиме пакетной обработки. Служба хранения данных *Azure Storage* используется для реализации очереди задач, к которой открыт доступ процессорам ЭВМ посредством *REST*-протокола.

В работе рассмотрена модель роста трещин в элементах конструкций при статическом нагружении в агрессивных водородсодержащих средах. На основе модели разработаны два альтернативных программных решения задачи многократного вычисления долговечности: локальное приложение, осуществляющее последовательные вычисления, и распределенное приложение на основе пакетной службы *Azure*. С помощью этих программных решений получены кривые роста трещин в элементах конструкций, а также оценки долговечности рассматриваемых элементов конструкций. Для количественной оценки эффективности указанных программных решений в работе предложена формальная зависимость времени вычислений от значений факторов, характеризующих задачу, и используемые вычислительные средства. Эта зависимость позволяет ввести критерий эффективности способа организации вычислений. На основе данного критерия поставлена и решена задача определения области в пространстве параметров модели, в которой параллельное решение задач определения ресурса изделия является наиболее эффективным.

1. Модель роста трещин в элементах конструкций при статическом нагружении в агрессивных водородсодержащих средах

Водородное охрупчивание является частой причиной отказа элементов конструкций, эксплуатируемых в условиях воздействия статической нагрузки и агрессивной среды [12]. Охрупчивание заключается в проникновении водорода в металл и, как следствие, в локальном снижении его прочностных характеристик вблизи вершины трещины по мере увеличения в металле локальной концентрации водорода. Такой процесс, снижая трещиностойкость металла, приводит к последовательному подрастанию трещины и в итоге к разрушению элемента конструкции.

Ограничимся рассмотрением элементов конструкций, характерные размеры которых существенно больше характерных размеров, находящихся в них трещин.

В машиностроении при рассмотрении задачи определения ресурса элементов конструкций традиционно выделяют два этапа процесса разрушения: инкубационный период и стадия роста трещины. На первом этапе трещина не растет, но вблизи ее вершины постепенно реализуются условия последующего роста. На втором этапе наблюдается устойчивое увеличение размеров трещины вплоть до начала динамического разрушения всего исследуемого объекта.

Будем считать, что в плоском образце в начальный момент времени имеется дефект типа трещины в виде прямолинейного разреза длины l_0 пренебрежимо малой ширины. Предположим, что напряжения в образце на достаточно большом расстоянии от разреза не зависят от пространственных координат и равны $\sigma_x = 0$, $\sigma_y = \sigma^\infty$, $\tau_{yx} = 0$. Свяжем с вершиной начальной трещины неподвижную "глобальную" декартову систему координат XOY, ось OX которой направлена вдоль прямой линии роста трещины, а ось OY — перпендикулярно ей. "Глобальное время" обозначим t. Также введем подвижную систему координат xOy, в каждый момент времени t связанную текущим положением вершины трещины. Через вершину в материал проникает водород [6, рис. 1]. В качестве превалирующего направления проникновения водорода в металл выберем направление предполагаемого роста трещины — вдоль оси OX [14].

Для количественной оценки изменения локальной концентрации водорода $C(x, \tau)$ используем уравнение однонаправленной диффузии, заданное в локальной системе координат xOy, где τ — локальное время. Задачу диффузии решаем как одномерную краевую задачу в форме

$$\frac{\partial C(x,\tau)}{\partial \tau} = D\frac{\partial^2 C(x,\tau)}{\partial x^2} + \zeta\frac{\partial C(x,\tau)}{\partial x}\frac{\partial \tilde{\sigma}(x)}{\partial x},$$

$$C(0,\tau) = C^0,\ C(x,\tau) = C_0(x) \qquad (1)$$

где $\tilde{\sigma} = \tilde{\sigma}(x)$ — локальное растягивающее напряжение вблизи вершины трещины; D — коэффициент диффузии; ζ — параметр, определяемый свойствами среды и материала; C^0 — концентрация водорода в среде [12].

Для упрощения расчетов используем полученное в рамках линейной механики разрушения [4] распределение напряжений $\tilde{\sigma}(x)$ вблизи вершины трещины вида $\tilde{\sigma}(x) = \dfrac{\tilde{K}_I}{\sqrt{\pi x}}$. Здесь $\tilde{K}_I = \sigma^\infty\sqrt{\pi l}$ — коэффициент интенсивности напряжений; l — текущая длина трещины.

Первый этап (инкубационный период) характеризуется постепенным повышением средней концентрации $\overline{C}_a(\tau)$ водорода вблизи вершины трещины:

$$\overline{C}_a(\tau) = \frac{1}{a_0^d}\int_0^{a_0^d} C(x,\tau)dx.$$

Здесь отрезок $x \in [0;\ a_0^d]$ — область предразрушения. Полагаем, что с течением времени по мере увеличения $\overline{C}_a(\tau)$ в указанной области происходит снижение локальной трещиностойкости металла $\widetilde{K}_{Ic}$ по зависимости

$$\left(\frac{\widetilde{K}_{Ic} - \widetilde{K}^*_{Ic}}{\widetilde{K}^0_{Ic} - \widetilde{K}^*_{Ic}}\right)^{\alpha_1} + \left(\Omega\frac{\overline{C}_a(\tau)}{C^0}\right)^{\beta_1} = 1,\ \Omega > 0, \qquad (2)$$

где α_1, β_1 — константы материала; $\widetilde{K}^0_{Ic}$, $\widetilde{K}^*_{Ic}$ — трещиностойкость металла при $C_a = 0$ и $C_a = C^*$ соответственно. Здесь величина C^* равна предельной растворимости водорода в металле, величина $\Omega = C^0/C^*$ является мерой интенсивности водородного охрупчивания для исследуемой пары среда—металл как принято в работе [6].

Пусть в момент времени τ_0 — инкубационный период трещиностойкость снизилась до значения коэффициента интенсивности напряжений. Это условие называем локальным критерием разрушения и представляем в форме

$$\widetilde{K}_I(\sigma^\infty,\ l) = \widetilde{K}_{Ic}(C_a(\tau_0)). \qquad (3)$$

Одно из главных допущений рассматриваемой модели роста трещин заключается в том, что именно в этот момент трещина подрастает, причем ее длина увеличивается на длину области предразрушения a_0^d. Преобразовав зависимость (2) с учетом условия (3), можем получить выражение для критического значения концентрации $\overline{C}^*_a$:

$$\overline{C}^*_a = \sqrt[\beta_1]{1 - \left(\frac{\widetilde{K}_I(\sigma^\infty,\ l) - \widetilde{K}^*_{Ic}}{\widetilde{K}^0_{Ic} - \widetilde{K}^*_{Ic}}\right)^{\alpha_1}}\ \frac{C^0}{\Omega}.$$

Для оценки продолжительности инкубационного периода τ_0 используем приближенное решение уравнения диффузии (1) вида $C(x, \tau) \approx T(\tau)\varphi(x)$, где $\varphi(x)$ — координатная функция, удовлетворяющая граничным условиям задачи (1), а $T(\tau)$ — функция, зависящая только от времени. Примем $\varphi(x)$ равной $Ae^{-\lambda x}$; A, $\lambda \in \mathbb{R}$. По методу Галеркина (путем ортогонализации невязки к координатной функции) получим приближенное решение вида

$$C(x,\ \tau) \approx C^0\varphi(x)\exp\left(\frac{D(\varphi;\ \varphi_{xx}) + \zeta(\varphi;\ \varphi_x\widetilde{\sigma}_x)}{(\varphi;\ \varphi)}\tau\right). \qquad (4)$$

Здесь скалярные произведения вида $(f;\ g)$ определяются как $\int_\Phi f(x)g(x)dx$; x и xx в нижнем индексе — частная производная по x и вторая частная производная по x соответственно; область интегрирования $\Phi = [a_0^d;\ +\infty)$.

Для значения τ_0 выполняется равенство средней локальной концентрации ее критическому значению:

$$\overline{C}_a(\tau_0) = \overline{C}^*_a. \qquad (5)$$

Отсюда, обозначив $\dfrac{\int_0^{a_0^d}\varphi(x)dx}{a_0^d} = \overline{\varphi}_0$, получим следующее выражение для τ_0:

$$\tau_0 = \Upsilon(l_0,\ \sigma^\infty)\ln\left(\frac{\overline{\varphi}_0}{\Omega}\left[1 - \left(\frac{\widetilde{K}_I(\sigma^\infty,\ l_0) - \widetilde{K}^*_{Ic}}{\widetilde{K}^0_{Ic} - \widetilde{K}^*_{Ic}}\right)^{\alpha_1}\right]^{\frac{1}{\beta_1}}\right),$$

$$\Upsilon(l_0,\ \sigma^\infty) = \frac{(\varphi;\ \varphi)}{D(\varphi;\ \varphi_{xx}) + \zeta(\varphi;\ \varphi_x\widetilde{\sigma}_x)};$$

$$(\varphi;\ \varphi_x\widetilde{\sigma}_x) = -\frac{\widetilde{K}_I(\sigma^\infty,\ l_0)}{2\sqrt{\pi}}\left(\varphi;\frac{\varphi_x}{x\sqrt{x}}\right).$$

Итак, инкубационный период заканчивается в момент времени $t = \tau_0$ (в глобальной системе отсчета) при выполнении равенства (5) подростом трещины на значение a_0^d с распределением концентрации $C(x,\ \tau_0) = C_0(x)$, которая играет роль начального условия аналогичной задачи диффузии для следующих этапов роста трещины.

Стадия роста трещины. Полагаем, что данная стадия (продолжительностью t^*) представляет собой серию из n последовательных (через время Δt_i) скачкообразных подростов трещины на значение a_i^d каждый; $i \in [1:\ n]$. При этом длина трещины увеличивается от $l_0 + a_0^d$ до исчерпания ресурса исследуемого образца.

Для определения значений Δt_i решим задачу, аналогичную рассмотренной выше для инкубационного периода. При этом допустим, что размер области предразрушения a_i^d разный для каждого из подростов и зависит от длины трещины:

$$a^d(l) = a_0^d + (a_*^d - a_0^d)\cdot\sqrt[\alpha_2]{1 - \left(\frac{L^*_{fr} - l}{L^*_{fr} - l_0}\right)^{\beta_2}},$$

где L^*_{fr} — предельная длина трещины; a_*^d — максимальная длина области предразрушения; α_2, β_2 — произвольные положительные константы.

Итак, решая рассмотренную задачу многократно получим последовательность состояний трещины $\{l_i,\ t_i\}$, каждое из которых задается текущей длиной трещины и текущим значением глобального времени:

$$\begin{cases} l_i = l_0 + \sum\limits_{k=0}^{i-1} a_k^d(l_k), \\ t_i = \sum\limits_{k=0}^{i-1}\Delta t_k, \end{cases} \quad i \in [1:\ n]. \qquad (6)$$

Моделирование проводится до наступления стадии динамического разрушения.

Критерий предельного состояния. Используем критерий предельного состояния образца с трещиной в форме

$$l_n = L^*_{fr}, \qquad (7)$$

где n — номер последнего подроста; соответствующее время равно $t_n = t^*$.

Чтобы определить предельную длину трещины воспользуемся силовым критерием прочности [3]:

$$\widetilde{K}^0_{Ic} = \sigma^\infty \sqrt{\pi L^*_{fr}} \Leftrightarrow L^*_{fr} = \frac{1}{\pi}\left(\frac{\widetilde{K}^0_{Ic}}{\sigma^\infty}\right)^2.$$

При этом предполагаем, что трещиностойкость металла $\widetilde{K}_{Ic}(t)$ по мере стремления длины трещины к ее предельному значению приближается к значению трещиностойкости ненаводороженного металла $\widetilde{K}^0_{Ic}$. Последнее предположение обоснуем следующими рассуждениями. Рассмотрим трещину в моменты времени $t_0 = t_0$ и $t_1 = t_0 + \Delta t_1$, т.е. в состояниях, непосредственно предшествующих первому (инкубационный период) и второму подростам трещины. Сравним соответствующие критические концентрации $(\bar{C}^*_a)_0$, $(\bar{C}^*_a)_1$, используя формулу (2):

$$\frac{C_0}{\Omega}\left[1 - \left(\frac{\widetilde{K}_I(\sigma^\infty, l_0) - \widetilde{K}^*_{Ic}}{\widetilde{K}^0_{Ic} - \widetilde{K}^*_{Ic}}\right)^{\alpha_1}\right]^{\frac{1}{\beta_1}} >$$

$$> \frac{C_0}{\Omega}\left[1 - \left(\frac{\widetilde{K}_I(\sigma^\infty, l_0 + a^d_0) - \widetilde{K}^*_{Ic}}{\widetilde{K}^0_{Ic} - \widetilde{K}^*_{Ic}}\right)^{\alpha_1}\right]^{\frac{1}{\beta_1}} \Leftrightarrow$$

$$\Leftrightarrow \widetilde{K}_I(\sigma^\infty, l_0) < \widetilde{K}_I(\sigma^\infty, l_0 + a^d_0) \Leftrightarrow l_0 < l_0 + a^d_0. \qquad (8)$$

Таким образом, пользуясь тем, что $\bar{C}^*_a$ — убывающая функция по l, получили верное неравенство. Отметим, что неравенство (8) верно при сравнении критических концентраций для любых двух состояний, предшествующих i-му и j-му подросту трещины; $i < j$.

Во всяком i-м состоянии трещины выполняется условие

$$(\bar{C}^*_a)_i = \max_{\tau \in [0;\, \tau_i]} C^i_a(\tau),$$

поскольку $C^i_a(\tau)$ — монотонно возрастающая величина и $C^i_a(0) < (\bar{C}^*_a)_i$. Другими словами, критическую величину $(\bar{C}^*_a)_i$ допустимо считать верхней оценкой для средней концентрации в течение процесса диффузии, протекающего между двумя последовательными подростами трещины.

Теперь легко показать, что

$$\widetilde{K}_{Ic}(C_a)\big|_{C_a = 0} = \widetilde{K}^0_{Ic},$$

$$(\bar{C}^*_a)_i = \max_{\tau \in [0;\, \tau_i]} C^i_a(\tau), \ \lim_{i \to \infty} (\bar{C}^*_a)_i = 0 \Rightarrow \widetilde{K}_{Ic} \to \widetilde{K}^0_{Ic}.$$

Другими словами, с ростом трещины соответствующие значения величины $(\bar{C}^*_a)_i$ образуют убывающую последовательность и, поскольку минимальным значением концентрации является ноль, на определенном этапе разрушения (и при определенном значении длины трещины) локальная средняя концентрация станет близка к нулю и, как видно из выражения (2), локальная условная трещиностойкость станет равной трещиностойкости ненаводороженного материала $\widetilde{K}^0_{Ic}$. Это обстоятельство позволяет сформулировать критерий в форме (7).

Из сказанного также следует, что по мере приближения $(\bar{C}^*_a)_i$ к нулю последовательность Δt_i также сходится к нулю (поскольку для достижения критической концентрации перед каждым подростом требуется все меньше времени), а это означает стремление скорости роста трещины к бесконечности, т.е. лавинообразное разрушение образца.

2. Локальное приложение для моделирования роста трещин

Последовательность $\{l_i\}$ можно рассматривать как значения некоторой функции вида $l = l(t)$, $t \in [0; t^*]$ на неоднородной сетке, с узлами t_i. Все параметры модели в совокупности зададим вектором

$$\Psi = \{l_0, \sigma^\infty, \Omega, \widetilde{K}^0_{Ic}, \widetilde{K}^*_{Ic}, \alpha_1, \beta_1, \alpha_2, \beta_2, \zeta, D, \delta\}.$$

Тогда процесс роста трещины можно представить в виде

$$\begin{cases} l = l(t, \Psi), \\ l(t^*, \ldots) = L^*_{fr} = \dfrac{(1-\delta)}{Y\pi}\left(\dfrac{\widetilde{K}^0_{Ic}}{\sigma^\infty}\right)^2, \end{cases} \qquad (9)$$

где δ — параметр, близкий к единице. Формула (9) позволяет выразить ресурс изделия в виде $t^* = t^*(\Psi)$.

Представленная модель роста трещин реализована программно в виде локального приложения (ЛП). С помощью ЛП было проведено моделирование кинетики трещин с параметрами из табл. 1 и

Таблица 1

Исходные данные для моделирования роста трещин

l_0, м	σ^∞, МПа	$\widetilde{K}^*_I$, МПа$\sqrt{\text{м}}$	α_1	β_1	a^d_0, мм	α_2	β_2	D	δ	Ω	ζ, lg(кг · $\frac{\text{м}}{\text{с}^2}$)
0,005	140	10	2	2	10^{-5}	2	2	10^{-10}	0,05	2,5	−19,094

Таблица 2

Результаты моделирования для различных значений K_{Ic}^0

K_{Ic}^0, МПа$\sqrt{\text{м}}$	Ресурс t^*, с	Число итераций n	Время расчета $t_{\text{расч}}$, мс	$t_{\text{расч.ср}}$
80	204 500	1373	2285	1,6
90	214 051	1756	2992	1,7
110	225 540	2657	7094	2,6

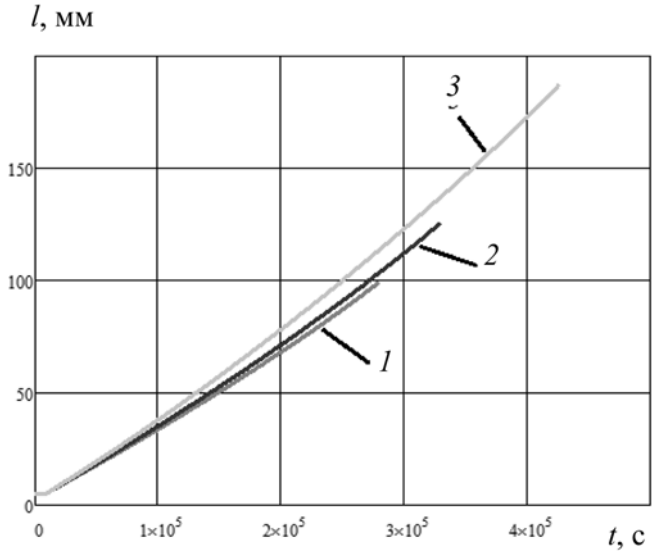


Рис. 1. Кривые роста трещин для различных значений $\tilde{K}_{Ic}^0$:

1 — $\tilde{K}_{Ic}^0 = 80$ МПа$\sqrt{\text{м}}$; *2* — $\tilde{K}_{Ic}^0 = 90$ МПа$\sqrt{\text{м}}$;

3 — $\tilde{K}_{Ic}^0 = 110$ МПа$\sqrt{\text{м}}$

каждым из трех значений параметра K_{Ic}^0: 80, 90, 110 МПа$\sqrt{\text{м}}$. Соответствующие кривые роста трещин представлены на рис. 1. Для каждой кривой было зафиксировано число итераций, потребовавшихся для расчета кинетики дефекта, оценка ресурса модельного образца и время расчета (табл. 2).

Представленные результаты вычислительных экспериментов свидетельствуют о существенном росте времени расчета с ростом значения величины K_{Ic}^0. Поэтому при любой организации параллельного решения набора задач моделирования кинетики трещин необходимо учитывать, что время решения одной задачи, в общем случае, существенно зависит от параметров моделирования.

3. Распределенное приложение для моделирования роста

Определим термин *параметрический анализ* как проведение многофакторного испытания описанной в разд. 1 модели, заключающегося в многократном моделировании роста трещин (9) с Z различными наборами параметров Ψ^i, $i \in [1{:}Z]$.

Пусть при использовании ЛП решение i-й задачи моделирования роста трещин при наборе параметров Ψ^i занимает время T_i и суммарное время расчетов при решении всех Z задач моделирования равно T. С целью уменьшить это время реализуем программное приложение, которое решает задачи моделирования последовательно-параллельно, используя N независимых вычислительных ресурсов.

При реализации указанного программного приложения учтем следующие требования.

1. Выделение в приложении программы ЛП, реализующей непосредственно моделирование, и программы ОП, которая выполняет множественные запуски ЛП с различными наборами исходных данных. Это позволяет обеспечить большую гибкость приложения, поскольку программу ЛП можно использовать и традиционным способом на локальном компьютере.

2. Реализация централизованной очереди задач, когда их число существенно превышает число независимых вычислительных ресурсов N. Группировать задачи в более крупные не представляется рациональным в силу разного времени выполнения вычислений в зависимости от параметров моделирования.

3. Реализация вычислений за счет внешних вычислительных мощностей, получаемых в краткосрочную "аренду" по требованию.

Для разработки приложения, вообще говоря, можно было бы использовать и локальные ресурсы. Однако при этом неизбежным станет применение таких технологий, как *MPI* или *ZeroMQ*. Технология *MPI* не отвечает перечисленным требованиям, поскольку, во-первых, в этом случае программный код, реализующий непосредственно моделирование, должен быть интегрирован в код, осуществляющий параллельное решение множества задач. Во-вторых, в этой технологии отсутствует реализация общей очереди задач (сообщений). В технологии *ZeroMQ* выполнено лишь требование 2.

Для выполнения требования 3 необходимо использовать облачные ресурсы [14]. На основе анализа известных облачных высокопроизводительных ресурсов было принято решение использовать пакетную службу *Azure* (*Azure Batch*), развиваемую в рамках платформы *Microsoft Azure* с осени 2014 г.

Пакетная служба позволяет разграничить программу ЛП и облачный программный комплекс (ОП) для использования первой одновременно на N виртуальных машинах. Главным отличием пакетной службы от других подобных технологий является то, что организация централизованной очереди задач и решение всех технических вопросов, связанных с выделением затребованных пользователем ресурсов, служба реализует как свою основную функциональность. Необходимо только предоставить этой службе исполняемый файл программы ЛП и библиотеку стандартизованной структуры на

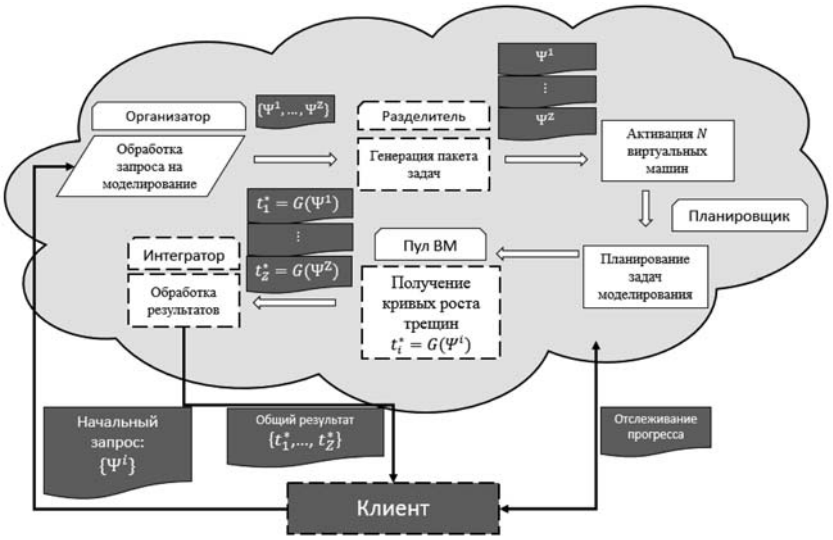


Рис. 2. Структура распределенного приложения. Штриховой линией обведены компоненты, реализуемые пользователем самостоятельно. Остальные составляющие облачной части приложения реализованы в пакетной службе

языке *C#*, которая описывает, как именно к задаче, полученной из очереди, следует применить программу ЛП. Взаимодействие между всеми частями распределенной системы осуществлялось посредством *REST*-запросов. Архитектура приложения (рис. 2) продиктована устройством и принципами функционирования пакетной службы *Azure*.

Совокупность всех задач, которые решаются в облачной части приложения, называется заданием. Каждую задачу определяет вектор Ψ. Под решением задачи понимаем вызов ЛП с вектором Ψ в качестве входных данных.

Облачное приложение работает в режиме вебслужбы и доступно в любое время. Оно активируется, когда от клиента приходит *REST*-запрос, содержащий информацию, необходимую для генерации последовательности векторов $\{\Psi^i\}$. Модуль под названием "Разделитель" ("*Job Splitter*") генерирует пакет задач моделирования. Далее планировщик пакетной службы запускает необходимое число виртуальных машин и с помощью службы хранения данных *Azure Storage* создает централизованную очередь задач. На каждой виртуальной машине запускается процесс, который также через протокол *REST* обращается к "службе хранилища" и извлекает очередную задачу. После этого из предоставленной пользователем библиотеки вызывается компонент, который решает извлеченную задачу. Результат решения передается в модуль "Интегратор". После того как "Интегратор" получает решения всех задач, пул виртуальных машин деактивируется, решения архивируются и становятся доступными для загрузки клиентом.

Во время работы облачной части ОП клиент может отслеживать процессы, происходящие в пуле, и получать информацию о сбоях, ошибках и этапах работы приложения.

4. Исследование эффективности распределенного приложения

Параметрический анализ заключается в формировании Z различных наборов исходных параметров Ψ^i и вычислении для каждого из них ресурса по формуле (9):

$$t_i^* = t^*(\Psi^i),\ i = [1{:}Z].$$

Объединим множество наборов Ψ^i в матрицу

$$\mathbf{\Psi} = (\Psi^1\ \ldots\ \Psi^Z).$$

Определим время T выполнения всего объема $\mathbf{\Psi}$ заданий моделирования роста трещин с использованием ОП как функцию вида

$$T = T(\mathbf{\Psi},\ N), \qquad (10)$$

где N — число доступных одинаковых вычислительных ресурсов, которые в случае облачных вычислений являются виртуальными машинами.

Коэффициент эффективности параллельных вычислений имеет вид

$$E(\mathbf{\Psi},\ N) = \frac{T(\mathbf{\Psi},\ 1)}{T(\mathbf{\Psi},\ N)N} \in [0;\ 1]. \qquad (11)$$

На практике число доступных виртуальных машин всегда ограничено. Обозначим максимальное число доступных ресурсов N^*. Также зададим некоторое значение коэффициента эффективности E^*, при достижении которого приложение будет считаться эффективным.

Сформулируем критерий эффективности программы ОП при решении задач моделирования из набора $\mathbf{\Psi}$ в виде

$$\begin{cases} E(\mathbf{\Psi},\ N) \geqslant E^* \\ N \leqslant N^*. \end{cases} \qquad (12)$$

Условия (12) формируют область D_Ψ в пространстве параметров модели, в которой критерий (12) выполняется, назовем эту область зоной эффективности ОП применительно к задаче (10)—(12). Если эта задача не имеет решения, констатируем недостижимость требуемой эффективности в рамках заданных ограничений.

Для анализа эффективности ОП поставим вычислительный эксперимент, заключающийся в проведении серии вычислений с матрицей

$$\mathbf{\Psi}(K_{Ic}^0) = \begin{pmatrix} l_0^1\ \sigma^\infty\ \ldots\ \tilde{K}_{Ic}^0\ \ldots\ D\ \ \delta \\ \vdots \\ l_0^Z\ \sigma^\infty\ \ldots\ \tilde{K}_{Ic}^0\ \ldots\ D\ \ \delta \end{pmatrix}^{\mathrm{T}}, \qquad (13)$$

где все параметры принимают значения из табл. 1; величина K_{Ic}^0 является переменной; l_0^i есть некоторая последовательность допустимых начальных длин трещины.

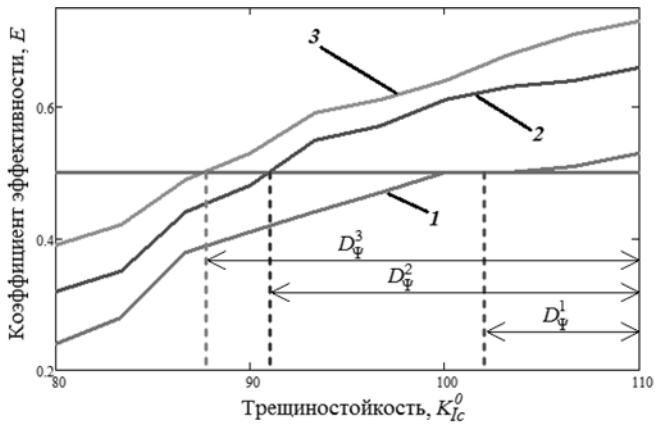


Рис. 3. Кривые *1*, *2* и *3* — $E(\tilde{K}_{Ic}^0, N)$ при $N \in \{10, 15, 20\}$. Области D_Ψ^i — соответствующие зоны эффективности, $i \in [1: 3]$. Серая линия — $E^* = 0{,}5$

Пакетная служба *Azure* по умолчанию предоставляет не более 20 виртуальных машин, поэтому $N^* = 20$. Установим $E^* = 0{,}5$. Таким образом, критерий эффективности принимает вид

$$\begin{cases} E(\Psi, N) \geqslant 0{,}5, \\ N \leqslant 20. \end{cases} \qquad (14)$$

Для исследования эффективности выделим некоторый набор значений $\tilde{K}_{Ic}^0$ в диапазоне [80; 110]. Для каждого из этих значений проведем расчет с помощью ОП, используя в качестве входных данных матрицу Ψ вида (13), где $l_0^i = 0{,}005 + 0{,}005i$, $i \in [0; 50]$. Полученные таким образом графики приведены на рис. 3.

Из результатов исследования видно, что с ростом значения $\tilde{K}_{Ic}^0$ (с ростом вычислительной сложности каждой из задач набора) эффективность ОП повышается и в определенном диапазоне значений $\tilde{K}_{Ic}^0$ выполняется критерий эффективности. Таким образом, найдена зона эффективности комплекса ОП при входных данных вида (13).

Заключение

В работе предложена модель роста трещин в наводороженном металле. На ее основе разработано локальное приложение ЛП, с помощью которого получены кривые роста трещин.

Рассмотрен один из возможных подходов к организации параллельных вычислений в задаче моделирования роста трещин с помощью облачных технологий. С помощью вычислительного эксперимента показано, что разработанный облачный программный комплекс ОП выполняет вычисления с разной эффективностью в зависимости от параметров моделирования. Для исследования эффективности комплекса предложен обобщенный критерий эффективности.

Выявлены ситуации, когда ЛП быстрее, чем ОП, справляется даже с большим объемом работы. Чаще всего причиной этого эффекта является низкая вычислительная сложность большинства (или всех) задач, которая, в свою очередь, определяется конкретными параметрами модели. Эффективность приложения ОП повышается с ростом вычислительной сложности задач, поскольку это приводит к относительному уменьшению коммуникационных затрат.

Причиной низкой эффективности приложения ОП может являться также неравномерная загрузка процессоров системы. Используемый в работе подход не позволяет непосредственно исследовать это явление.

Введенный в работе критерий эффективности вычислений позволяет обнаружить зоны эффективности приложения ОП в области допустимых значений параметров моделирования, что дает возможность более экономно расходовать облачные ресурсы.

Список литературы

1. **Александров А. А., Ларионов В. И., Сущев С. П., Идрисова Я. Р.** Условия транспорта нефти и оценка безопасности трубопроводов при аварийном выходе из грунта в низкотемпературную атмосферу // Проблемы сбора, подготовки и транспорта нефти и нефтепродуктов. 2011. № 4. С. 113—119.
2. **Александров А. А., Павлихин Г. П.** МГТУ им. Н. Э. Баумана: решение проблем промышленной экологии // Экология и промышленность России. 2011. № 4. С. 24—25.
3. **Матвиенко Ю. Г.** Моделирование и критерии разрушения в современных проблемах прочности, живучести и безопасности машин // Проблемы машиностроения и надежности машин. 2014. № 3. С. 80—89.
4. **Матвиенко Ю. Г.** Двухпараметрическая механика разрушения в современных проблемах прочности // Проблемы машиностроения и надежности машин. 2013. № 5. С. 37—46.
5. **Романов А. Н., Тараканов П. В., Шашурин Г. В., Берчун Ю. В., Резчикова Л. А., Сокольников П. С.** Моделирование роста трещин в наводороживаемых высокопрочных сталях при циклическом нагружении // Проблемы машиностроения и автоматизации. 2014. № 4. С. 87—93.
6. **Tarakanov P. V., Romanov A., Shashurin G.** Numerical life estimation of structure components subjected to hydrogen embrittlement and cycling // Key Engineering Materials. 2014. Vol. 592—593. С. 117—120.
7. **Витвицкий П. М., Попина С. Ю.** Прочность и критерии хрупкого разрушения стохастически дефектных тел. Киев: Наук. думка, 1980. 187 с.
8. **Gupta A., Sarood O., Kale L. V., Milojicic D.** Improving HPC application performance in cloud through dynamic load balancing // Proc. of 13th IEEE/ACM International Symposium On Cluster, Cloud, And Grid Computing. 2013. P. 402—409.
9. **Пакетная** служба Azure Batch. URL: http://azure.microsoft.com/ru-ru/services/batch/. (Дата обращения: 15.08.2016).
10. **Гридин В. Н., Дмитревич Г. Д., Анисимов Д. А.** Архитектура распределенных сервис-ориентированных систем автоматизированного проектирования // Известия ЮФУ. Технические науки. 2014. № 7 (156). С. 51—58.
11. **Селиванов Е. В.** Сервис-ориентированная архитектура в облачных технологиях // Современные тенденции в образовании и науке: сб. науч. тр. по материалам Международной научно-практической конференции 31 октября 2013 г. Тамбов: Изд-во ТРОО "Бизнес—Наука—Общество", 2013. С. 140—141.
12. **Tomohiko O.** Hydrogen embrittlement of steel in corrosive environments and high-pressure gaseous hydrogen environments // Corrosion Engineering. 2009. N 4. P. 151—167.
13. **Yokobori A. T., Chinda Y., Nemoto T., Satoh K., Yamada T.** The characteristics of hydrogen diffusion and concentration around a crack tip concerned with hydrogen embrittlement // Corrosion Science. 2002. N 3. P. 407—424.
14. **Алмаметов В. Б.** Возможности повышения эффективности использования САПР на основе облачных технологий // Труды международного симпозиума "Надежность и Качество". Пенза. 2012. Т. 2. С. 62—63.

D. L. Kishlakov[1], Student, e-mail: daniel_kish@outlook.com,
P. V. Tarakanov[1, 2], Postgraduate Student, e-mail: pashabeetle@yandex.ru,
G. V. Shashurin[2], Dean, e-mail: shashuring@mail.ru,
Yu. V. Berchun[1], Senior Lecturer, e-mail: y_berchun@mail.ru
[1]Moskovskij Gosudarstvennyj Tehniceskij Universitet im. N. E. Baumana,
[2]Mechanical Engineering Research Institute of the Russian Academy of Sciences, Moscow

Cloud Computing Efficiency in Crack Growth Simulation in Hydrogenated Structure Components

The design of durable structure components requires durability analysis in CAE systems. Such analysis often requires batch processing of problem with two-level tree data stream topology. The leaves of the tree are multiple same-type calculation routines and cloud computations use is a favorable choice for such a task.

To develop CAE durability modules analytical approach to different batch processing systems efficiency estimation is desirable. Such modules are designed for durability analysis of hydrogenated and loaded structure components with initial defects. Crack kinetics model was used to simulate fracture processes; crack length curves had been obtained.

An approach to parallel batch processing organization is described: a distributed cloud application, built on top of Microsoft Azure services, which engages multiple computational resources from a distant cloud server to perform parallel execution of simulations.

An efficiency criterion of parallel simulation tasks batch processing is suggested. The criterion can be applied to estimate efficiency taking modeling parameters into consideration. Based on the criterion a problem can be stated of finding such a subset in the modeling parameters domain, for which any given application will be efficient while processing simulation tasks.

The criterion was applied for efficiency analysis of the developed cloud application to find a subset of one modelling parameter in which the application is efficient being constrained in number of computational resources available and in the desirable level of performance.

Keywords: *automated design, durability, computer simulation, embrittlement, parallel computing, cloud technologies, efficiency, task parallelism, web-services, service oriented architecture*

References

1. **Aleksandrov A. A., Larionov V. I., Sushhev S. P., Idrisova Ya. R.** Uslovija transporta nefti i ocenka bezopasnosti truboprovodov pri avarijnom vyhode iz grunta v nizkotemperaturnuju atmosferu.(Oil transport conditions and evaluation of the underground pipeline safety in the event of pipeline emergency exposure to low-temperature environment), *Problemy sbora, podgotovki i transporta nefti i nefteproduktov*, 2011, no. 4, pp. 113—119.

2. **Aleksandrov A. A., Pavlihin G. P.** MGTU im. N. E. Baumana: reshenie problem promyshlennoj jekologii (Bauman Moscow State Technical Univeristy: solving problems of industrial ecology), *Jekologija i promyshlennost' Rossii*, 2011, no. 4, pp. 24—25.

3. **Matvienko Yu. G.** Modelirovanie i kriterii razrushenija v sovremennyh problemah prochnosti, zhivuchesti i bezopasnosti mashin (Modeling and destruction criteria in modern problems of strength, vitality and machine safety), *Problemy mashinostroenija i nadezhnosti mashin,* 2014, no. 3, pp. 80—89.

4. **Matvienko Yu. G.** Dvuhparametricheskaja mehanika razrushenija v sovremennyh problemah prochnosti (Two-parameter fracture mechanics in modern strength problems), *Problemy mashinostroenija i nadezhnosti mashin*, 2013, no. 5, pp. 37—46.

5. **Romanov A. N., Tarakanov P. V., Shashurin G. V., Berchun Yu. V., Rezchikova L. A., Sokol'nikov P. S.** Fatigue crack propagation modeling of hydrogenating high-strength steels, *Problemy mashinostroenija i avtomatizacii*, 2014, no. 4, pp. 87—93.

6. **Tarakanov P. V., Romanov A., Shashurin G.** Numerical life estimation of structure components subjected to hydrogen embrittlement and cycling, *Key Engineering Materials*, 2014, vol. 592—593, pp. 117—120.

7. **Vitvickij P. M., Popina S. Yu.** *Prochnost' i kriterii hrupkogo razrushenija stohasticheski defektnyh tel* (Strength and brittle failure criteria of stochastically damaged bodies), Kiev: Nauk. dumka, 1980, 187 p.

8. **Gupta A., Sarood O., Kale L. V., Milojicic D.** Improving HPC application performance in cloud through dynamic load balancing, *Proc. of 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing*, 2013, pp. 402—409.

9. **Microsoft** Azure Batch (15.08.2016), available at: http://azure.microsoft.com/ru-ru/services/batch/

10. **Gridin V. N., Dmitrevich G. D., Anisimov D. A.** Arhitektura raspredelennyh servis-orientirovannyh sistem avtomatizirovannogo proektirovanija (Architecture of distributed service-oriented system computer-aided design), *Izvestija JuFU. Tehnicheskie nauki*, 2014, no. 7 (156), pp. 51—58.

11. **Selivanov E. V.** Servis-orientirovannaja arhitektura v oblachnyh tehnologijah (Service oriented architecture in cloud technologies), *Sovremennye tendencii v obrazovanii i nauke*, 2013, pp. 140—141.

12. **Tomohiko O.** Hydrogen embrittlement of steel in corrosive environments and high-pressure gaseous hydrogen environments, *Corrosion Engineering*, 2009, no. 4. pp. 151—167.

13. **Yokobori A. T., Chinda Y., Nemoto T., Satoh K., Yamada T.** The characteristics of hydrogen diffusion and concentration around a crack tip concerned with hydrogen embrittlement, *Corrosion Science*, 2002, no. 3, pp. 407—424.

14. **Almametov V. B.** Vozmozhnosti povyshenija jeffektivnosti ispol'zovanija SAPR na osnove oblachnyh tehnologij (Possibilities of increasing CAD/CAE systems efficiency using cloud technologies), *Proceedings of international conference "Nadezhnost' i Kachestvo",* Penza, 2012, vol. 2, pp. 62—63.

УДК 621.391 DOI:10.17587/it.23.121-129

А. А. Устинов, д-р техн. наук, проф., ст. науч. сотр.,
С. В. Дворников, д-р техн. наук, проф., e-mail: practicdsv@yandex.ru,
Н. С. Агеева, аспирант,
Военная академия связи им. Маршала Советского Союза С. М. Буденного, г. Санкт-Петербург

Научно-методический аппарат адаптивного ортогонального преобразования видеоданных

Представлены результаты исследования по сжатию видеоданных на основе принципов адаптивного ортогонального преобразования и анализ существующих подходов к сжатию изображений. Показаны ограничения существующих методов вследствие их относительно низких значений коэффициентов сжатия по отношению к адаптивному ортогональному преобразованию. Анализируются результаты численного моделирования, подтверждающие наибольшую устойчивость метода адаптивного ортогонального преобразования к канальным ошибкам. Демонстрируются типичные артефакты в виде блочности, характерные для дискретного косинусного преобразования, и в виде выделенных фоновых участков, характерных для вейвлет-преобразования. Обосновываются преимущества разработанного метода адаптивного ортогонального преобразования по отношению к известным методам ортогональных преобразований.

Ключевые слова: *адаптивное ортогональное преобразование видеоданных, сжатие изображений, статистическое кодирование без потерь*

Введение

В настоящее время на практике широко используют разнообразные методы сжатия изображений [1—7], большинство из которых предполагает выполнение следующих процедур над исходным видеорядом:

- ортогональные преобразования дискретных отсчетов изображения;
- кодирование дискретных отсчетов видеоизображения или коэффициентов их преобразований;
- статистическое кодирование двоичных последовательностей.

Заметим, что преобразование дискретных отсчетов видеоданных проводится в целях минимизации числа битов, используемых при кодировании. Операции кодирования преобразованных отсчетов видеоизображения и статистического кодирования выполняют в целях приведения их к двоичному виду для передачи по цифровым каналам связи (ЦКС).

Алгоритмы, основанные на непосредственном преобразовании дискретных отсчетов видеоизображения в двоичные последовательности, образуют группу методов пространственного кодирования [8]. К данной группе относятся методы импульсно-кодовой модуляции и ее разновидности (дифференциальная импульсно-кодовая модуляция, адаптивная дифференциальная импульсно-кодовая модуляция и др.). Однако на практике применение данных методов является ограниченным вследствие малых коэффициентов сжатия, которые они обеспечивают.

В основе методов ортогонального преобразования лежит процедура декорреляции, в общем случае коррелированных дискретных отсчетов видеоданных. При этом энергия коэффициентов указанных преобразований перераспределяется таким образом, что при кодировании существенной их части используют лишь незначительное число битов, либо вообще отказываются от их кодирования, приравнивая эти коэффициенты к нулю. На этом явлении основаны практически все методы сжатия с преобразованием.

Между тем анализ методов статистического кодирования без потерь показывает [7, 8], что реализуемые ими коэффициенты сжатия в основном достигли своего предела, определяемого избыточностью источника сообщений. В то же время использование адаптивного ортогонального преобразования (АОП), учитывающего статистические особенности кодируемого изоизображения, может привести либо к увеличению коэффициента сжатия при сохранении заданного качества восстановленного изображения, либо к улучшению качества восстановленного изображения при сохранении коэффициента сжатия. В связи с этим в настоящей работе представлены результаты исследования возможностей АОП.

Адаптивное ортогональное преобразование видеоданных

Из теории факторного анализа [9] известно, что какова бы ни была прямоугольная матрица **A** размером $P \times N$, всегда существует разложение вида

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}^{\mathrm{T}}, \qquad (1)$$

где **U**, $\mathbf{V}^{\mathrm{T}}$ — унитарные матрицы; Λ — прямоугольная диагональная матрица размером $P \times N$ с невозрастающими неотрицательными элементами на диагонали.

Между тем, согласно [10], если задано сингулярное разложение (1), то диагональные элементы Λ матрицы **A** являются ее сингулярными числами. А столбцы матрицы **U** образуют ортонормированный базис из собственных векторов матрицы $\mathbf{AA}^{\mathrm{T}}$. При этом столбцы $\mathbf{V}^{\mathrm{T}}$ образуют ортонормированный базис из собственных векторов матрицы $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, а столбцы матриц $\mathbf{V}^{\mathrm{T}}$ и **U** в своей совокупности образуют сингулярные базисы матрицы **A**.

Из выражения (1) следует, что

$$\Lambda = \mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{V}. \qquad (2)$$

Следовательно, Λ можно рассматривать как набор коэффициентов преобразования при использовании сингулярных базисов матрицы **A**. Пример результата выполнения преобразования (2) показан на рис. 1.

Поскольку матрица Λ является диагональной, то при известных на приеме матрицах **U** и **V**, для восстановления изображения необходимо располагать данными о диагональных элементах Λ. Однако условие того, что матрицы **U** и **V** должны быть известны в точке восстановления исходного изображения, не позволяет использовать преобразование (2) при практической реализации алгоритмов сжатия.

Действительно, если изображение, представленное матрицей **A**, имеет размер $P \times N$ пикселей, то **U** и **V** будут иметь размеры $P \times P$ и $N \times N$, соответственно, поэтому передача матриц **U** и **V** на приемную сторону потребует расхода канального ресурса значительно большего, чем непосредственная передача **A**. В связи с этим определим следующий набор коэффициентов преобразования исходного изображения:

$$\tilde{\Lambda} = \tilde{\mathbf{U}}^{\mathrm{T}}\mathbf{A}\tilde{\mathbf{V}}, \qquad (3)$$

где $\tilde{\mathbf{U}}^{\mathrm{T}}$, $\tilde{\mathbf{V}}$ — квадратные матрицы размером $P \times P$ и $N \times N$, соответственно, с ограничениями вида:

$$\tilde{u}(i, j) = 0,\ \forall i, j = \overline{1, P};$$
$$i \neq 1 + (k-1)m_1,\ 2 + (k-1)m_1,\ \ldots,\ km_1; \qquad (4)$$
$$k = 1,\ 2,\ \ldots,\ P/m_1,$$

$$\tilde{v}(i, j) = 0,\ \forall i, j = \overline{1, N};$$
$$j \neq 1 + (k-1)m_2,\ 2 + (k-1)m_2,\ \ldots,\ km_2; \qquad (5)$$
$$k = 1,\ 2,\ \ldots,\ N/m_2,$$

$$\tilde{\mathbf{U}}^{\mathrm{T}}\tilde{\mathbf{U}} = \mathbf{I};\ \tilde{\mathbf{V}}^{\mathrm{T}}\tilde{\mathbf{V}} = \mathbf{I}. \qquad (6)$$

Здесь $\tilde{\Lambda}$ — прямоугольная матрица размером $P \times N$ элементов с минимальным числом ненулевых элементов; m_1, m_2 — номера строк, k — номера столбцов матриц $\tilde{\mathbf{U}}^{\mathrm{T}}$, $\tilde{\mathbf{V}}$.

Ограничения (4) и (5) означают, что в каждой строке матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$ и в каждом столбце матрицы $\tilde{\mathbf{V}}$ число ненулевых элементов составляет ровно m_1 и m_2 (значения m_1 и m_2 должны быть кратны P и N соответственно). Пример структуры $\tilde{\mathbf{U}}^{\mathrm{T}}$, $\tilde{\mathbf{V}}$ показан на рис. 2 и 3 соответственно.

Если m_1 и m_2 удовлетворяют неравенствам $m_1 \ll P$ и $m_2 \ll N$, то введение ограничений (5) и (6) позволит существенно снизить требования к ресурсу цифрового канала, затрачиваемого на передачу $\tilde{\mathbf{U}}^{\mathrm{T}}$ и $\tilde{\mathbf{V}}$.

Ограничение (6) обеспечивает ортогональность разложения в

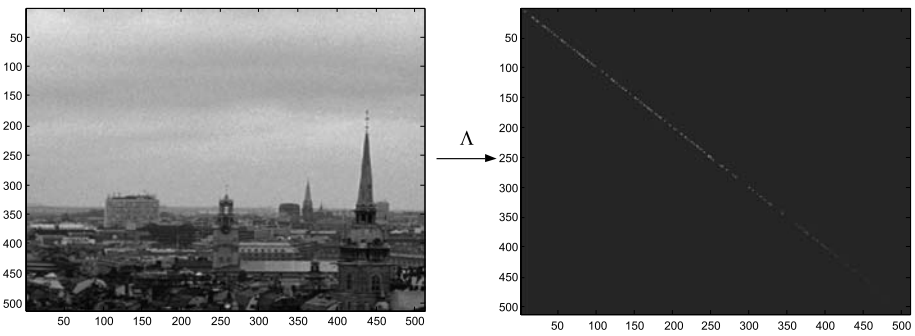


Рис. 1. Пример исходного изображения и матрицы коэффициентов сингулярного преобразования

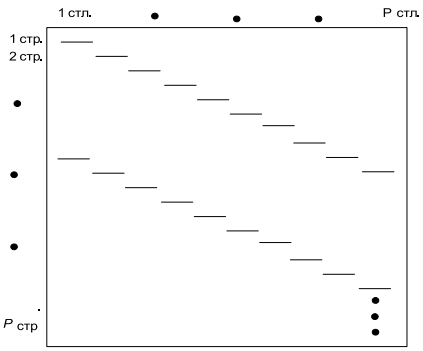


Рис. 2. Визуализация структуры матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$:
___ — последовательность из m_1 ненулевых элементов

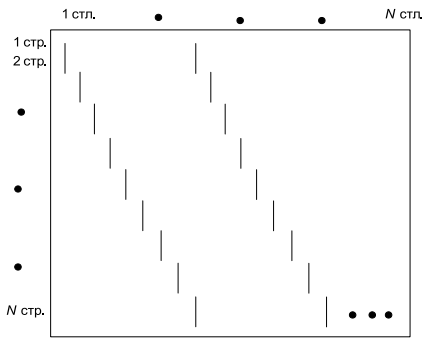


Рис. 3. Визуализация структуры матрицы $\tilde{\mathbf{V}}$:
| — последовательность из m_2 ненулевых элементов

соответствии с выражением (3). Ниже рассмотрим решение задачи поиска $\tilde{\mathbf{U}}^{\mathrm{T}}$ и $\tilde{\mathbf{V}}$ при ограничениях (4)—(6).

Вначале рассмотрим последовательность операций поиска матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$ для заданной матрицы $\tilde{\mathbf{V}}$. Для простоты рассуждений примем, что $\tilde{\mathbf{V}}$ является единичной матрицей. В этом случае, с учетом выражения (3) имеет место равенство

$$\tilde{\mathbf{\Lambda}} = \tilde{\mathbf{U}}^{\mathrm{T}}\mathbf{A}. \qquad (7)$$

Заметим, что для любого j-го столбца матрицы $\tilde{\mathbf{\Lambda}}$ справедливо равенство

$$\tilde{\mathbf{\Lambda}}_j = \tilde{\mathbf{U}}^{\mathrm{T}}\mathbf{A}_j,\ j = 1, 2, \ldots, N. \qquad (8)$$

С учетом размерности матрицы $\mathbf{A}$ ($P \times N$ элементов) число элементов в векторе $\tilde{\mathbf{\Lambda}}_j$ равно P.

Как было отмечено выше, матрица $\tilde{\mathbf{\Lambda}}$ должна содержать минимальное число ненулевых элементов. Определим выражение для средней квадратической ошибки (СКО), определяемой обнулением последних $P - M$ элементов вектора $\tilde{\mathbf{\Lambda}}_j$.

Любую i-ю компоненту вектор-столбца $\tilde{\mathbf{\Lambda}}_j$ с учетом правила матричного умножения [10], запишем в виде:

$$\tilde{\mathbf{\Lambda}}_{ij} = \tilde{\mathbf{U}}_i^{\mathrm{T}}\mathbf{A}_j, \qquad (9)$$

где $\tilde{\mathbf{U}}_i^{\mathrm{T}}$ — i-я строка матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$.

Тогда СКО, обусловленную обнулением $P - M$ элементов вектора $\tilde{\mathbf{\Lambda}}_j$, запишем следующим образом:

$$\varepsilon_j = \sum_{i=M+1}^{P} \mathbf{\Lambda}_{ij}^2 = \sum_{i=M+1}^{N} \tilde{\mathbf{U}}_i^{\mathrm{T}}\mathbf{A}_j\mathbf{A}_j^{\mathrm{T}}\tilde{\mathbf{U}}_i. \qquad (10)$$

В случае если обнуление $P - M$ компонент осуществляется во всех N столбцах матрицы $\tilde{\mathbf{\Lambda}}$, то результирующая ошибка будет иметь следующий вид:

$$\varepsilon_N = \sum_{j=1}^{N}\varepsilon_j = \sum_{j=1}^{N}\sum_{i=M+1}^{P}\mathbf{\Lambda}_{ij}^2 = \sum_{j=1}^{N}\sum_{i=M+1}^{P}\tilde{\mathbf{U}}_i^{\mathrm{T}}\mathbf{A}_j\mathbf{A}_j^{\mathrm{T}}\tilde{\mathbf{U}}_i. \qquad (11)$$

С учетом выражения (11) задачу поиска $\tilde{\mathbf{U}}^{\mathrm{T}}$ формально запишем следующим образом:

$$\sum_{j=1}^{N}\sum_{i=M+1}^{P}\tilde{\mathbf{U}}_i^{\mathrm{T}}\mathbf{A}_j\mathbf{A}_j^{\mathrm{T}}\tilde{\mathbf{U}}_i \to \min_{\tilde{\mathbf{U}}_i,\ i = M+1,\ M+2,\ \ldots,\ P}, \qquad (12)$$

при ограничениях (4) и (6).

Рассмотрим произведение $\tilde{\mathbf{U}}_i^{\mathrm{T}}\mathbf{A}_j$, $\forall i = \overline{1, P}$; $\forall j = \overline{1, N}$ с учетом структуры $\tilde{\mathbf{U}}^{\mathrm{T}}$ (см. рис. 4, здесь элементы матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$ представлены как $\tilde{\mathbf{u}}_{i,n}^{\mathrm{T}}$, вектор $\mathbf{A}_j$ как $\mathbf{A}_j$, а его элементы как $\mathbf{a}_{i,j}$).

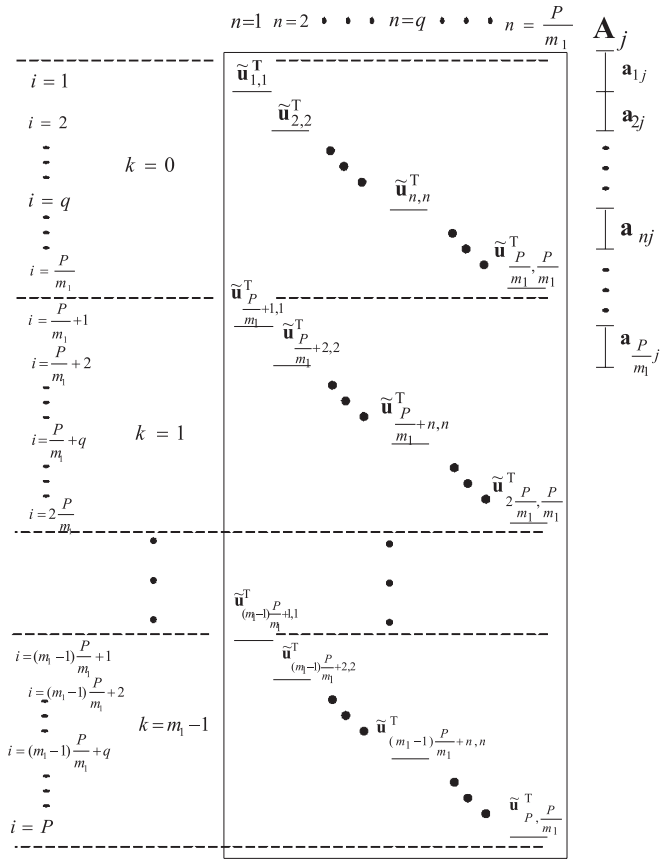


Рис. 4. Структура матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$ и вектора $\mathbf{A}_j$

Ограничение (4) определяет блочную структуру $\tilde{\mathbf{U}}^{\mathrm{T}}$.

Матрица состоит из m_1 блоков, с нулевого по $(m_1 - 1)$-й. В свою очередь, каждый блок содержит P/m_1 строк, каждая из которых включает только m_1 ненулевых компонент.

Пусть n есть текущий номер ненулевого фрагмента из m_1 компонент. В этом случае начальный и конечный индексы ненулевого фрагмента строки определяются, как $1 + (n - 1)m_1$ и $n \times m_1$. В соответствии со структурой матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$, приведенной на рис. 4, строки, в которых ненулевые элементы, расположенные с $1 + (n - 1)m_1$ по $n \times m_1$ — индексы, имеют номер $k\dfrac{P}{m_1} + n$, $k = 0, 1, \ldots, m_1 - 1$. Ненулевые фрагменты строк $\tilde{\mathbf{U}}_i^{\mathrm{T}}$, $i = 1, 2, \ldots, P$, обозначим в виде векторов: $\tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n}$, $k = 0, 1, \ldots, m_1 - 1$; $n = 1, 2, \ldots, \dfrac{P}{m_1}$.

В соответствии с рис. 4 представим j-й столбец $\mathbf{A}$ в виде совокупности P/m_1 фрагментов, при этом номера n-го фрагмента, расположенные с $1 + (n - 1)m_1$

по $n \times m_1$-индексы. Обозначим n-й фрагмент j-го столбца матрицы $\mathbf{A}$, как $\mathbf{a}_{n,\,j}$. Тогда справедливы равенства

$$\tilde{\mathbf{U}}_i^{\mathrm{T}} \mathbf{A}_j = \tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n} \mathbf{a}_{n,\,j} \qquad (13)$$

для индексов i, удовлетворяющих равенствам $i = k\dfrac{P}{m_1} + n$ и $k = 0, 1, ..., m_1 - 1$.

Представим число необнуляемых компонент в каждом столбце $\mathbf{A}$ в виде $M = q\dfrac{P}{m_1}$, где q — целое число, удовлетворяющее неравенствам $0 \leqslant q\dfrac{P}{m_1} < m_1 - 1$.

В этом случае с учетом равенства (13) и перестановки порядка суммирования выражение для результирующей ошибки будет иметь вид:

$$\varepsilon_N = \sum_{k=q+1}^{m_{11}-1} \sum_{n=1}^{\frac{P}{m_1}} \tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n} \left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right] \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n}. \qquad (14)$$

С учетом выражения (14) задачу поиска матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$ аналитически можно представить в следующем виде:

$$\sum_{k=q+1}^{m_{11}-1} \sum_{n=1}^{\frac{P}{m_1}} \tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n} \left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right] \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n} \rightarrow \min$$

$$\rightarrow \mathbf{u}_{k\frac{P}{m_1}+n,\,n},\ k = q+1, q+2, ..., m_1-1;\ n = 1, 2, ..., \frac{P}{m_1}. \qquad (15)$$

В отличие от задачи (12) искомые вектора в задаче (15) имеют размерность не P, а m_1 элементов.

При решении задачи (15) необходимо учесть ограничение (6) на ортонормальность строк матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$. Поскольку места расположения векторов $\tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n}$, $k = 0, 1, ..., m_1 - 1$; $n = 1, 2, ..., \dfrac{P}{m_1}$ в соответствующих строках матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$ являются непересекающимися, то для выполнения ограничения $\tilde{\mathbf{U}}^{\mathrm{T}}\tilde{\mathbf{U}} = \mathbf{I}$ достаточно выполнить ограничения

$$\tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n} \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n} = 1,$$

$$k = 0, 1, ..., m_1 - 1;\ n = 1, 2, ..., \frac{P}{m_1}. \qquad (16)$$

Для решения задачи (15) с учетом ограничений (16) в виде равенств воспользуемся методом множителей Лагранжа [1]. С учетом равенств (16) определим функцию Лагранжа для результирующей ошибки:

$$L_\varepsilon = \sum_{k=q+1}^{m_{11}-1} \sum_{n=1}^{\frac{P}{m_1}} \tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n} \left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right] \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n} -$$

$$- \sum_{k=q+1}^{m_1-1} \sum_{n=1}^{\frac{P}{m_1}} \beta_{k\frac{P}{m_1}+n,\,n} \left(\mathbf{u}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n} \mathbf{u}_{k\frac{P}{m_1}+n,\,n} - 1\right), \qquad (17)$$

где $\beta_{k\frac{P}{m_1}+n,\,n}$ — множители Лагранжа, $k = 0, 1, ..., m_1 - 1$; $n = 1, 2, ..., \dfrac{P}{m_1}$.

Вычислим производную функции Лагранжа по искомой векторной переменной:

$$\frac{\partial L_\varepsilon}{\partial \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n}} = 2\left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right] \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n} - 2\beta_{k\frac{P}{m_1}+n,\,n} \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n}. \qquad (18)$$

Приравняв левую часть равенства (18) к нулю, получаем

$$\left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right] \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n} = \beta_{k\frac{P}{m_1}+n,\,n} \tilde{\mathbf{u}}_{k\frac{P}{m_1}+n,\,n}. \qquad (19)$$

Выражение (19) означает, что $\tilde{\mathbf{u}}^{\mathrm{T}}_{k\frac{P}{m_1}+n,\,n}$, $k = 0, 1, ..., m_1 - 1$; $n = 1, 2, ..., \dfrac{P}{m_1}$ есть собственные вектора матрицы $\left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right]$, а $\beta_{k\frac{P}{m_1}+n,\,n}$ — соответствующие им собственные значения.

Таким образом, для вычисления $\tilde{\mathbf{U}}^{\mathrm{T}}$ при ограничениях (4), (5) и (6) необходимо вычислить собственные вектора корреляционных матриц $\left[\sum_{j=1}^{N} \mathbf{a}_{n,j}\mathbf{a}^{\mathrm{T}}_{n,j}\right]$, $n = 1, 2, ..., \dfrac{P}{m_1}$ и расположить их в строках матрицы $\tilde{\mathbf{U}}^{\mathrm{T}}$, как показано на рис. 4. При этом для выполнения обратного преобразования декодеру необходимо передать P/m_1 матриц размером $m_1 \times m_1$ элементов каждая. В этом случае общее число элементов, которые необходимо передать декодеру, составит $P \times m_1$.

Отметим, что при использовании разложения (1) без учета ограничений (4), (5) и (6), для выпол-

нения обратного преобразования декодеру необходимо передать ровно $P \times P$ элементов.

По аналогии с приведенными рассуждениями рассмотрим поиск матрицы $\tilde{\mathbf{V}}$ для найденной $\tilde{\mathbf{U}}^{\mathrm{T}}$. Если $\tilde{\mathbf{V}}$ не единичная, то имеет место равенство

$$\tilde{\mathbf{\Lambda}} = \tilde{\mathbf{U}}^{\mathrm{T}} \mathbf{A} \tilde{\mathbf{V}}. \qquad (20)$$

Определим матрицу $\mathbf{B}^{\mathrm{T}} = \tilde{\mathbf{U}}^{\mathrm{T}} \mathbf{A}$. Тогда для любой j-й строки матрицы $\tilde{\mathbf{\Lambda}}$ справедливо равенство

$$\tilde{\mathbf{\Lambda}}_j^{\mathrm{T}} = \mathbf{B}_j^{\mathrm{T}} \tilde{\mathbf{V}}, \; j = 1, 2, ..., P. \qquad (21)$$

С учетом размерности $\mathbf{B}$ ($P \times N$ элементов) число элементов в вектор-строке $\tilde{\mathbf{\Lambda}}_j^{\mathrm{T}}$ равно N. Исходя из требований к сжатию, матрица $\tilde{\mathbf{\Lambda}}$ должна содержать минимальное число ненулевых элементов. Минимальное значение СКО при обнулении последних элементов каждого столбца $\tilde{\mathbf{\Lambda}}$ обеспечивается выбором оптимальной $\tilde{\mathbf{U}}^{\mathrm{T}}$.

Ниже вычислим $\tilde{\mathbf{V}}$, для которой СКО, определяемая обнулением последних $N - M$ элементов вектора-строки $\tilde{\mathbf{\Lambda}}_j^{\mathrm{T}}$, будет минимальна.

Запишем произвольную i-ю компоненту вектор-строки $\tilde{\mathbf{\Lambda}}_{ji}^{\mathrm{T}}$ с учетом правила матричного умножения в следующем виде:

$$\tilde{\mathbf{\Lambda}}_{ji}^{\mathrm{T}} = \mathbf{B}_j^{\mathrm{T}} \tilde{\mathbf{V}}_i, \qquad (22)$$

где $\tilde{\mathbf{V}}_i$ — i-й столбец матрицы $\tilde{\mathbf{V}}$.

Тогда СКО, обусловленную обнулением $N - M$ элементов вектора $\tilde{\mathbf{\Lambda}}_j^{\mathrm{T}}$, запишем следующим образом:

$$\varepsilon_j = \sum_{i = M+1}^{N} \mathbf{\Lambda}_{ji}^2 = \sum_{i = M+1}^{N} \tilde{\mathbf{V}}_i^{\mathrm{T}} \mathbf{B}_j \mathbf{B}_j^{\mathrm{T}} \tilde{\mathbf{V}}_i. \qquad (23)$$

С учетом выражения (23) задачу поиска матрицы $\tilde{\mathbf{V}}$ можно формально представить следующим аналитическим выражением:

$$\sum_{j=1}^{P} \sum_{i=M+1}^{N} \tilde{\mathbf{V}}_i^{\mathrm{T}} \mathbf{B}_j \mathbf{B}_j^{\mathrm{T}} \tilde{\mathbf{V}}_i \rightarrow \min_{\tilde{\mathbf{U}}_i,\; i = M+1,\; M+2,\; ...,\; N} \qquad (24)$$

при ограничениях (4) и (6).

Решение задачи (24) аналогично решению задачи (12). Учитывая, что структура $\tilde{\mathbf{V}}$ и структура $\tilde{\mathbf{U}}^{\mathrm{T}}$ с учетом транспонирования совпадают, то пропуская промежуточные выкладки, по аналогии с выражением (19) можно записать

$$\left[\sum_{j=1}^{P} \mathbf{b}_{j,\,n} \mathbf{b}_{j,\,n}^{\mathrm{T}} \right] \tilde{\mathbf{v}}_{n,\,k\frac{N}{m_2}+n} = \alpha_{n,\,k\frac{N}{m_2}+n} \tilde{\mathbf{v}}_{n,\,k\frac{N}{m_2}+n}, \qquad (25)$$

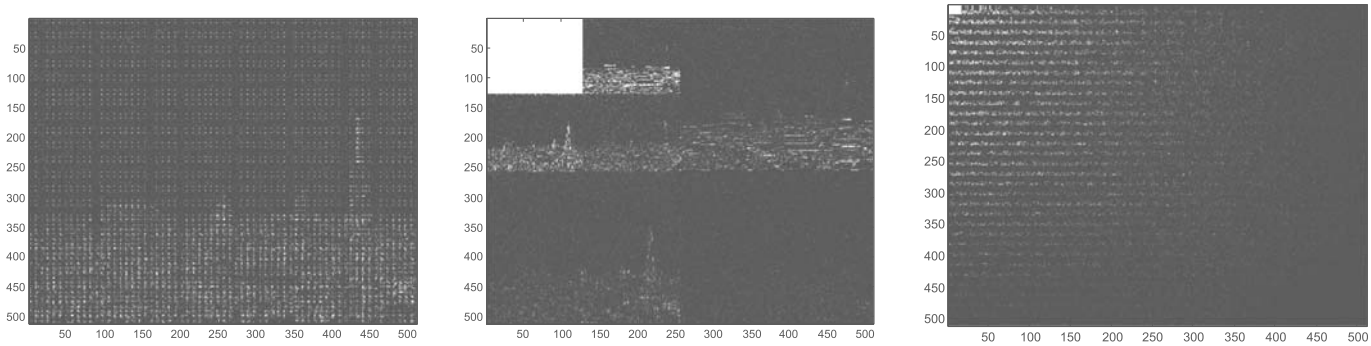
где $\mathbf{b}_{j,\,n}$ — n-й ненулевой фрагмент длиной m_2 символов в j-й строке матрицы $\mathbf{B}^{\mathrm{T}}$; $\tilde{\mathbf{v}}_{n,\,k\frac{N}{m_2}+n}$ — ненулевые фрагменты столбцов матрицы $\tilde{\mathbf{V}}$ с индексами $k\frac{N}{m_2} + n$; $\alpha_{n,\,k\frac{N}{m_2}+n}$ — собственные значения корреляционной матрицы $\left[\sum_{j=1}^{P} \mathbf{b}_{j,\,n} \mathbf{b}_{j,\,n}^{\mathrm{T}} \right]$, $k = 0, 1, ..., m_1 - 1$; $n = 1, 2, ..., \frac{P}{m_1}$.

Таким образом, $\tilde{\mathbf{U}}^{\mathrm{T}}$ и $\tilde{\mathbf{V}}$ полностью определяются собственными векторами корреляционных матриц $\left[\sum_{j=1}^{N} \mathbf{a}_{n,\,j} \mathbf{a}_{n,\,j}^{\mathrm{T}} \right]$, $n = 1, 2, ..., \frac{P}{m_1}$ и $\left[\sum_{j=1}^{P} \mathbf{b}_{j,\,n} \mathbf{b}_{j,\,n}^{\mathrm{T}} \right]$, где $n = 1, 2, ..., \frac{N}{m_2}$ соответственно.

Преобразование, описываемое выражением (20) и основанное на поиске ортогональных матриц специального вида $\tilde{\mathbf{U}}^{\mathrm{T}}$ и $\tilde{\mathbf{V}}$ для изображения $\mathbf{A}$, будем называть адаптивным преобразованием.

Результаты численного моделирования

Примеры набора коэффициентов преобразования для адаптивного и известных преобразований показаны на рис. 5.

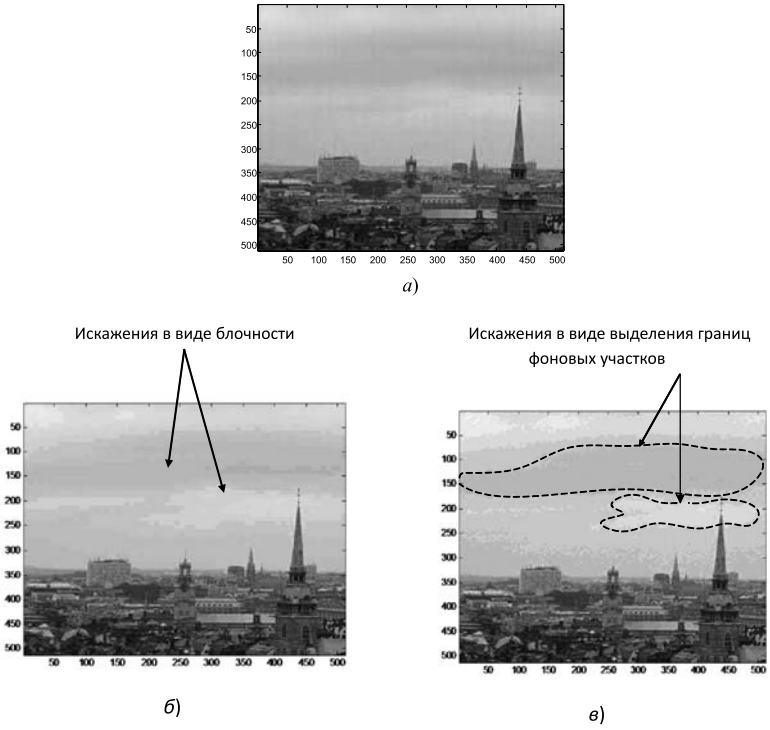


а) *б*) *в*)

Рис. 5. Визуализация матриц коэффициентов преобразования:
а — ДКП (размер матрицы 8×8); *б* — вейвлет-преобразования; *в* — адаптивного преобразования ($m_1 = 32$, $m_2 = 32$)

Результаты имитационного моделирования

Вид преобразования	Оцениваемый параметр							
	PSNR, дБ				*NZ*			
	$kv = 4$	$kv = 8$	$kv = 16$	$kv = 32$	$kv = 4$	$kv = 8$	$kv = 16$	$kv = 32$
Адаптивное преобразование $m_1 = 8$, $m_2 = 8$	46,88	41,10	36,84	34,27	161 688	92 088	34 838	15 820
Адаптивное преобразование $m_1 = 16$, $m_2 = 16$	46,89	41,12	36,96	34,44	159 176	88 561	32 005	12 989
Адаптивное преобразование $m_1 = 32$, $m_2 = 32$	46,89	41,17	37,05	34,54	156 814	86 154	30 966	12 003
ДКП	46,88	41,07	36,84	34,23	162 828	92 770	35 274	16 331
Вейвлет-преобразование	46,89	40,96	36,62	33,72	171 135	103 176	45 249	26 745



а) *б*) *в*)

Рис. 6. Примеры восстановленных изображений:
а — при использовании адаптивного преобразования; *б* — при использовании ДКП; *в* — при использовании вейвлет-преобразования

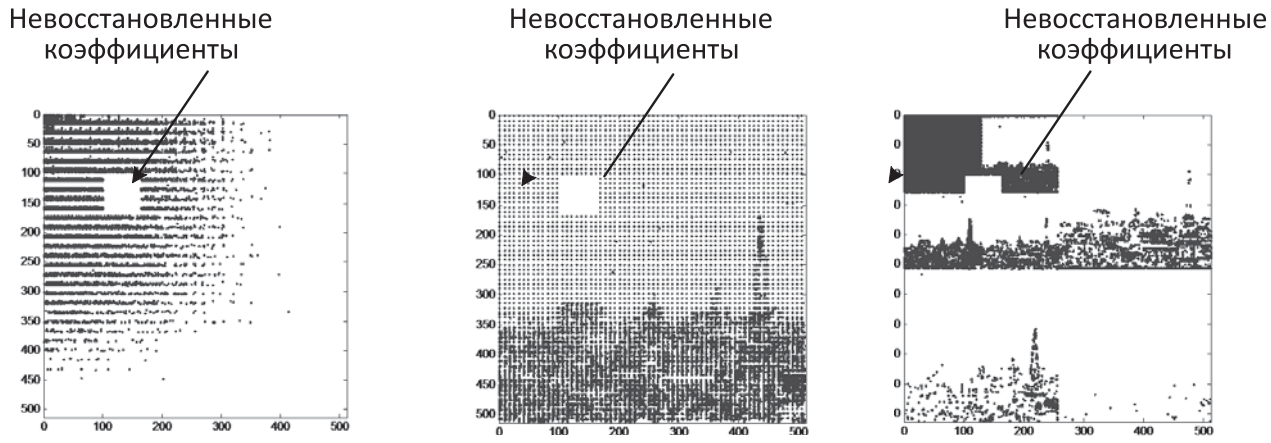
Результаты имитационного моделирования полученного преобразования представлены в таблице.

В ходе моделирования осуществлялось квантование полученных коэффициентов преобразования с различным шагом: $kv = 4$, $kv = 8$, $kv = 16$, $kv = 32$. Квантование матрицы коэффициентов преобразования $\tilde{\mathbf{\Lambda}}$ выполнялось в соответствии с выражением

$$\tilde{\mathbf{\Lambda}}_{kv} = \text{round}\left(\frac{\tilde{\mathbf{\Lambda}}}{kv}\right), \qquad (26)$$

где $\tilde{\mathbf{\Lambda}}_{kv}$ — матрица квантованных коэффициентов преобразования; kv — шаг квантования; round(*) — операция арифметического округления.

После вычисления квантованных коэффициентов оценивали два параметра: пиковое отношение сигнал/шум (*PSNR*) и число ненулевых коэффициентов преобразования (*NZ*). Величину *PSNR* использовали для объективной оценки качества восста-



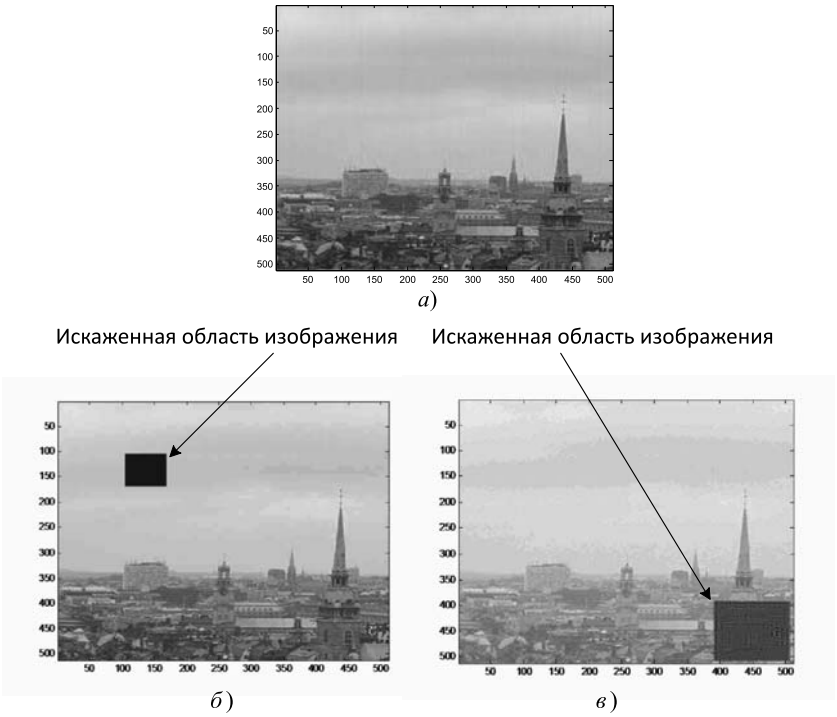
NZ = 11 416 *NZ* = 16 267 *NZ* = 25 313

а) *б*) *в*)

Рис. 7. Визуализация невосстановленных коэффициентов:
а — адаптивного преобразования; *б* — ДКП; *в* — вейвлет-преобразования

новленного изображения. Параметр *NZ* косвенно характеризует реализуемый коэффициент сжатия изображения. Известно [2], что чем меньше число ненулевых квантованных коэффициентов преобразования, тем больше может быть обеспечен коэффициент сжатия.

Имитационное моделирование показывает преимущество адаптивного преобразования по сравнению с дискретно-косинусным преобразованием (ДКП) и вейвлет-преобразованием. Так, при $m_1 = 32$ и $m_2 = 32$ число ненулевых коэффициентов при адаптивном преобразовании на 30 % меньше, чем при ДКП, и более чем на 100 % меньше, чем при вейвлет-преобразовании. Кроме того, восстановленное на основе адаптивного преобразования изображение не имеет артефактов в виде блочности (рис. 6, *б*), характерных для ДКП, и артефактов в виде выделения границ фоновых участков, характерных для вейвлет-преобразования (см. рис. 6, *в*).



а) *б*) *в*)

Рис. 8. Восстановленные изображении при обнулении части коэффициентов на основе:
а — адаптивного преобразования; *б* — ДКП; *в* — вейвлет-преобразования

В реальных каналах связи вследствие воздействия помех возникают искажения передаваемых данных. В связи с этим проанализируем устойчивость ортогональных преобразований к искажению части коэффициентов. Предположим, что ошибки под воздействием канала связи возникли при передаче одной и той же области коэффициентов размером 64 × 64 элементов. Коэффициенты преобразования данной области приравняем к нулю, как показано на рис. 7.

На рис. 8 показаны восстановленные изображения при обнулении части коэффициентов преобразования.

Представленные на рис. 8 восстановленные изображения получены в условиях обнуления квадратной области коэффициентов изображения размером 64 × 64 элемента. Верхний левый угол области имеет координаты (100, 100).

Заключение

Использование АОП обеспечивает получение меньшего числа ненулевых коэффициентов по сравнению с неадаптивными преобразованиями (ДКП, вейвлет-преобразование). Данное обстоятельство свидетельствует о том, что при использовании АОП может быть достигнуто большее значение коэффициента сжатия по сравнению с использованием неадаптивных методов преобразования. Кроме того, восстановленное по АОП изображение не имеет искажений в виде блочности и в виде выделения границ фоновых участков, характерных для известных методов ортогональных преобразований.

Анализ полученных изображений показывает, что наиболее устойчивым к канальным ошибкам является адаптивное преобразование. Визуально искажения почти незаметны при обнулении фиксированной области размером 64 × 64 элементов.

При использовании ДКП искажения проявляются в том, что не восстанавливается область изображения, координаты которой соответствуют координатам искаженных коэффициентов (см. рис. 7, *б*). Снижение *PSNR* для ДКП составляет около 12,5 дБ.

При использовании вейвлет-преобразования ошибки проявляются в виде искаженных областей большего размера (см. рис. 7, *в*), при этом снижение *PSNR* составило 12,6 дБ. Размеры искаженной области определяются числом уровней вейвлет-разложения. В рассматриваемом примере использовано двухуровневое разложение. В связи с этим вертикальный и горизонтальный размеры искаженной области увеличены ровно в 2 раза и составляют 128 × 128 пикселей.

Таким образом, разработанное адаптивное ортогональное преобразование имеет явные преимущества по сравнению с известными не адаптивными преобразованиями, как по числу ненулевых коэффициентов преобразования при заданном шаге квантования, так и по устойчивости к канальным ошибкам. Однако более детального рассмотрения требуют вопросы разработки алгоритмов быстрого вычисления собственных векторов матриц малого размера (32 × 32 элемента) и передачи данных матриц по каналу связи. Авторам видится интересным

использование методов частотно-временной обработки, рассмотренных в работах [11—13]. Данные вопросы, являющиеся направлением дальнейших исследований, будут рассмотрены в последующих статьях.

Список литературы

1. **Гонсалес Р., Вудс Р.** Цифровая обработка изображений: пер. с англ. М.: Техносфера, 2006. 1072 с.
2. **Ричардсон Я.** Видеокодирование H.264 и MPEG-4 — стандарты нового поколения. М.: Техносфера, 2005. 368 с.
3. **Сэломон Д.** Сжатие данных, изображений и звука. М.: Техносфера, 2004. 368 с.
4. **Тропченко А. Ю., Курносенков И. Н.** Анализ современных стандартов сжатия видеоданных // Научно-технический вестник СПбГУ ИТМО. 2006. Вып. 32. СПб.: Изд-во СПбГУ ИТМО. С. 17—21.
5. **Умбиталиев А. А., Дворников С. В., Оков И. Н., Устинов А. А.** Способ сжатия графических файлов методами вейвлет-преобразований // Вопросы радиоэлектроники. Серия: Техника телевидения. 2015. № 3 (21). С. 100—106.
6. **Lee M., Chan R., Adjieroh D.** Quantization of 3D-DCT coefficients and scan order for video compression // J. Vis. Commun. Image Represent. Dec. 1997. N. 8. P. 405—422.
7. **Дворников С. В., Устинов А. А., Цветков В. В.** Устранение межкадровой избыточности подвижных изображений на основе поиска оптимального базиса преобразования во временной области // Вопросы радиоэлектроники. Серия: Техника телевидения. 2012. № 2. С. 45—54.
8. **Дворников С. В., Устинов А. А., Цветков В. В.** Компенсация движения в видеокодеках, использующих трехмерные ортогональные преобразования, на основе оптимальных разбиений кодируемых блоков во временной области // Вопросы радиоэлектроники. Серия: Техника телевидения. 2013. № 2. С. 98—111.
9. **Лоули Д., Максвелл А.** Факторный анализ как статистический метод. М.: Мир, 1967.
10. **Воеводин В. В., Кузнецов Ю. А.** Матрицы и вычисления. М.: Наука. Главная редакция физико-математической литературы, 1984. 320 с.
11. **Дворников С. В., Железняк В. К., Комарович В. Ф., Храмов Р. Н.** Метод обнаружения радиосигналов на основе обработки их частотно-временных распределений плотности энергии // Информация и космос. 2005. № 4. С. 13—16.
12. **Дворников С. В., Сауков А. М.** Модификация частотно-временных описаний нестационарных процессов на основе показательных и степенных функций // Научное приборостроение. 2004. Т. 14, № 3. С. 76—85.
13. **Алексеев А. А., Дворников С. В., Железняк В. К.** и др. Применение методов частотно-временной обработки акустических сигналов для анализа параметров реверберации // Научное приборостроение. 2001. Т. 11, № 1. С. 65—76.

A. A. Ustinov, D. Sc., Professor, **S. V. Dvornikov,** D. Sc., Professor, e-mail: practicdsv@yandex.ru,
N. S. Ageeva, Postgraduate Student
Military Communications Academy, St. Petersburg

Adaptive Orthogonal Transformation of Video Image Method

In the article the important research and technology task of development and research methods and video data compression algorithms. In many social spheres of human activity means of objective remote control and monitoring are applied. Such means on the basis of analysis of data received in real time. The same social spheres the following aero video monitoring of traffic situation in big megalopolis, environmental resources, prevention and operational of emergency recovery, etc.

Generally, in such video monitoring systems video data image detecting data are transmitted to point data analysis by broadband heterogeneous networks. The heterogeneous is caused by resort employment of combination different communication technologies. This and broadband networking technology, which are used, for example, information transmission is provided from an unmanned aerial vehicles to a ground control station. Also batch high-speed communication systems, which are provided information transmission for subscribers of the terrestrial infrastructure. Accordingly an important aspect of similar systems is providing its high-capacity, minimum data loss during transmission and also providing minimum data garbling. In such a way the development task of adoptive data compression methods is the actual research and technology task. As well as quality requirement such video data are placed increased demands on video data quality from different user application. Besides, there is a need to parameter optimization of communication link by capacity and possible timing delay during compressed video data transfer.

The article includes synthesis of adaptive algorithm for video data compressed on the basis of discrete cosine transformation methods and wavelet transform. Analytical techniques is shown that proposed adaptive procedures video data compressed allow to reduce structural and parametric by communicated information about 25—50 fold with no loss in quality. On the basis of existing methods and algorithms video data compressed by analytical approach is shown that appliance of existing methods is limited, in consequence of low data-compression ratio (2—4 fold), which methods and algorithms provide. Application of the offered procedures allows reducing such limitation. As received compression ratios are large in the same procedures in comparison with known methods. From the analysis of attained result by computational modeling research and technology results are received about the adoptive orthogonal transformation is most resistant to link errors recovery video data by the proposed method has not artifacts as blocks, which are typical for discrete cosine transform and artifacts as boundary detection which are typical for wavelet transform. In such away the developed method of adaptive orthogonal transformation has a number of adaptive before known methods by orthogonal transformation.

Keywords: *adaptive orthogonal transformation, video image compression, lossless entropy coding*

References

1. **Gonsales R., Vuds R.** *Cifrovaya obrabotka izobrazhenij* // Per. s angl. Moscow, Tekhnosfera, 2006. 1072 p.
2. **Richardson Ya.** *Videokodirovanie N.264 i MPEG-4 — standarty novogo pokoleniya*, Moscow, Tekhnosfera, 2005, 368 p.
3. **Sehlomon D.** *Szhatie dannyh, izobrazhenij i zvuka*, Moscow, Tekhnosfera, 2004, 368 p.
4. **Tropchenko A. Yu., Kurnosenkov I. N.** Analiz sovremennyh standartov szhatiya videodannyh // *Nauchno-tekhnicheskij vestnik SPbGU ITMO*. 2006. Issue 32. P. 17—21.
5. **Umbitaliev A. A., Dvornikov S. V., Okov I. N., Ustinov A. A.** Sposob szhatiya graficheskih fajlov metodami vejvlet-preobrazovanij, *Voprosy radioehlektroniki. Seriya: Tekhnika televideniya*, 2015, no. 3 (21), pp. 100—106.
6. **Lee M., Chan R., Adjieroh D.** Quantization of 3D-DCT coefficients and scan order for video compression, *J. Vis. Commun. Image Represent*, Dec. 1997, no. 8, pp. 405—422.
7. **Dvornikov S. V., Ustinov A. A., Cvetkov V. V.** Ustranenie mezhkadrovoj izbytochnosti podvizhnyh izobrazhenij na osnove poiska optimal'nogo bazisa preobrazovaniya vo vremennoj oblasti, *Voprosy radioehlektroniki. Seriya: Tekhnika televideniya*, 2012, no. 2, pp. 45—54.
8. **Dvornikov S. V., Ustinov A. A., Cvetkov V. V.** Kompensaciya dvizheniya v videokodekah, ispol'zuyushchih tryohmernye ortogonal'nye preobrazovaniya, na osnove optimal'nyh razbienij kodiruemyh blokov vo vremennoj oblasti, *Voprosy radioehlektroniki. Seriya: Tekhnika televideniya*, 2013, no. 2, pp. 98—111.
9. **Louli D., Maksvell A.** *Faktornyj analiz kak statisticheskij metod*, Moscow, Mir, 1967.
10. **Voevodin V. V., Kuznecov Yu. A.** *Matricy i vychisleniya*, Moscow, Nauka. Glavnaya redakciya fiziko-matematicheskoj literatury, 1984, 320 p.
11. **Dvornikov S. V., Zheleznyak V. K., Komarovich V. F., Hramov R. N.** Metod obnaruzheniya radiosignalov na osnove obrabotki ih chastotno-vremennyh raspredelenij plotnosti ehnergii, *Informaciya i kosmos*, 2005, no. 4, pp. 13—16.
12. **Dvornikov S. V., Saukov A. M.** Modifikaciya chastotno-vremennyh opisanij nestacionarnyh processov na osnove pokazatel'nyh i stepennyh funkcij, *Nauchnoe priborostroenie*, 2004, vol. 14, no. 3, pp. 76—85.
13. **Alekseev A. A., Dvornikov S. V., Zheleznyak V. K.** et al. Primenenie metodov chastotno-vremennoj obrabotki akusticheskih signalov dlya analiza parametrov reverberacii, *Nauchnoe priborostroenie*, 2001, vol. 11, no. 1, pp. 65—76.

УДК 519.688 DOI:10.17587/it.23.129-134

А. С. Потехин, программист, andrew.potekhin94@gmail.com,
А. В. Стрельников, аспирант, ведущий программист, ArkadiyS@RD-Science.com,
ООО "Ар Ди Сайнс", Красноярск

Методика оценки транспортного потока на перекрестке по данным видеонаблюдения

Исследуется задача оценки транспортного потока на перекрестке по данным видеонаблюдения. Для распознавания объектов на кадре использована методика разности кадров. Разработанные алгоритмы реализованы в виде программного модуля. Приведены результаты статистических экспериментов, на основании которых можно сделать вывод о качестве работы алгоритмов.

Ключевые слова: *компьютерное зрение, транспортный поток, распознавание, движущиеся объекты, видеопоток, обработка видео*

Введение

В данной работе рассматривается задача оценки транспортного потока на перекрестках. Особенность этой задачи состоит в том, что здесь не просто ведется подсчет объектов, пересекающих особую метку, но и также строится траектория движения; это необходимо для записи маршрута перемещающегося объекта. Исследуемая задача относится к общей проблеме моделирования транспортной системы [1], связанной со сложной структурой системы и нехваткой достоверных данных. Для решения поставленной задачи используют алгоритмы машинного зрения [2].

В результате применения алгоритмов собирается статистика транспортных потоков на сложном участке дорожной системы. Эту статистику можно использовать для проверки качества построенной модели транспортной системы [3]. При этом обновление статистики может происходить в режиме реального времени, что позволяет модели реагировать на определенные события в короткий период времени. Дальнейшая оптимизация модели транспортной системы происходит за счет настройки работы светофоров, изменения типов узлов соединения, изменения числа полос. Также полученная статистика может быть использована для иного рода задач, как например, оценка количества выбросов от выхлопных проезжающих автомобилей, прогнозирование износа дорожного полотна или оценка спроса торговой точки за счет объема клиентов-водителей.

Подсчет проезжающего транспорта с использованием только данных видеонаблюдения является наиболее экономичным с точки зрения затрат на закупку, установку, настройку и обслуживание не-

обходимого оборудования. В данном случае используется только камера видеонаблюдения. А поскольку алгоритм работает и в сложных условиях, то в некоторых случаях можно использовать данные с уже размещенных городских камер.

1. Постановка задачи оценки транспортного потока

Объектом исследования является часть транспортной системы — перекресток. Транспортная система может быть представлена в виде одного или нескольких взвешенных ориентированных графов.

Узлами этого графа являются условно прямые участки дороги без ключевых своротов, а дугами — минимальный по расстоянию дорожный маршрут, проложенный между двумя узлами графа. При этом перекресток является не узлом графа, а небольшим подграфом.

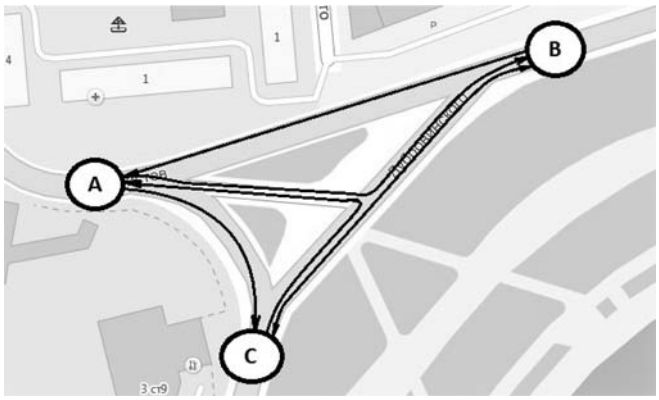


Рис. 1. Пример объекта исследования

Рис. 2. Кадр из видеопотока данных

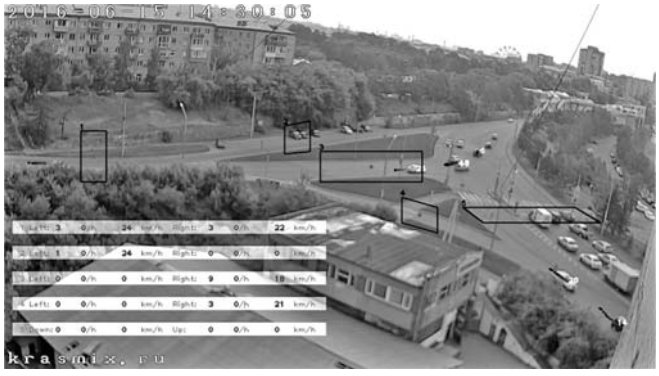


Рис. 3. Кадр из видеопотока данных с элементами обработки

На рис. 1 представлен перекресток с тремя узлами. Все связи между узлами двунаправленные, хотя перекресток содержит односторонние движения на некоторых участках дороги. Таким образом, перекресток содержит шесть маршрутов, или направлений:

$$1.\ A \to B;\ 3.\ A \to C;\ 5.\ B \to C;$$

$$2.\ B \to A;\ 4.\ C \to A;\ 6.\ C \to B.$$

Обозначим транспортный поток через T_{lm} — число автомобилей, проехавших из узла l в узел m за единицу времени. Задача заключается в том, чтобы оценить значение транспортного потока $\hat{T}_{lm}$ для выявленных направлений. Источником информации в задаче являются данные видеонаблюдения (рис. 2).

Видео, кадр которого изображен на рис. 2, записано с ресурса [4], имеющего свободный доступ.

Входными данными для решения задачи является таблица данных с переменными, описывающими характеристики пикселей:

- координаты пикселя (X, Y); $X \in \{0; w\}$; $Y \in \{0; h\}$, где w — ширина кадра; h — высота кадра;
- цвет пикселя $(R, G, B) \in \{0; 255\}$ или интенсивность $I \in \{0; 255\}$ цвета;
- номер кадра N.

Все описанные выше переменные являются целочисленными. Каждую итерацию кадр подается целиком, т.е. на входе имеется весь массив данных: $i_{x, y, j}$; $x = 1, ..., w$; $y = 1, ..., h$, а j — номер текущего кадра. Обозначим через I_j — матрицу значений интенсивности цвета j-го кадра.

На рис. 3 представлен кадр в процессе обработки. По ходу описания решения поставленной задачи в тексте будут поясняться отдельные элементы обработки, представленные на кадре.

Черными пронумерованными четырехугольниками на рис. 3 обозначены области на кадре, которые необходимы для определения маршрутов при подсчете. Каждая область характеризуется рядом параметров:

- координаты четырех углов $q = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)\}$;
- вектор направления d: влево-вправо или вверх-вниз.

Указание направлений необходимо для формализации соответствий между областями и маршрутами. Области находят либо экспертным путем, либо с помощью автоматизированного алгоритма, и этих областей может быть несколько. Обозначим через $S_k = (q_k, d_k)$; $k = 1, ..., K$, k-ю область. Таким образом, будем подразумевать, что значение потока $\hat{T}$ необходимо оценить не по маршруту, а по области.

Области можно располагать согласно узлам транспортного графа, а можно расположить по направлению дуг, как это сделано на примере (см. рис. 1), но в таком случае нужно будет создать соответствие между областями и маршрутами. В текущем примере используются пять областей, у каждой из которых задано по два направления потока (либо вверх-вниз, либо влево-вправо). Соответствия между маршрутами и областями следующие:

1. $A \to B$: поток в области 4, направление вправо;
2. $B \to A$: поток в области 3, направление влево;
3. $A \to C$: поток в области 3, направление вправо;
4. $C \to A$: поток в области 2, направление влево;
5. $B \to C$: поток в области 5, направление вверх;
6. $C \to B$: поток в области 5, направление вниз.

Область 1 уточняет подсчет на маршруте $B \to A$. Это необходимо, так как в данном направлении в области 3 происходит перекрытие объектов, что значительно снижает точность расчетов.

В итоге математическую постановку задачи можно представить следующим образом. Необходимо получить оценку потока:

$$\hat{T}_k = F(I_k, I_{k-1}, \ldots, I_{k-p}, S_k),$$

где F — оператор, описывающий последовательность действий; p — число предыдущих кадров, необходимых для оценки. В данном случае $\hat{T}_k$ является двумерным вектором, так как оценивается значение потока в две стороны: влево и вправо или вверх и вниз.

При обработке данных алгоритм может столкнуться с рядом трудностей:

1. *Острый угол камеры*. Если камера расположена слишком низко к земле, то, во-первых, размеры автомобилей на кадре будут изменяться в процессе движения, во-вторых, машины большего размера будут перекрывать машины меньшего размера.

2. *Препятствия для обзора*. Часть дорожного полотна может быть перекрыта деревьями или инфраструктурой города.

3. *Погодные условия*. Такие погодные ситуации, как снег, дождь и туман могут ухудшать качество распознавания движущихся объектов на кадре.

4. *Пробки и светофоры с длинным периодом*. Медленно двигающийся или находящийся в покое объект может восприниматься алгоритмом как фон на кадре.

Для решения поставленной задачи используется ряд алгоритмов, описание которых представлено в следующем разделе.

2. Методика оценки значения транспортного потока

Существует несколько подходов к решению поставленной задачи. В данной работе рассматривается методика разницы кадров (рис. 4). Любой из подходов состоит из двух этапов:

- распознавание движущихся объектов на кадре;
- построение траектории движения.

Рис. 4. Результат разницы кадров

Для того чтобы выделить движущиеся объекты среди неподвижной среды всякий инструмент распознавания (в том числе и человеческий мозг) сравнивает несколько картинок, зафиксированных с некоторым шагом в течение определенного периода времени. Если человек может это сделать с помощью своего зрения, то вполне вероятно, что алгоритм сможет это повторить с помощью компьютерного зрения.

Алгоритм оперирует информацией интенсивности кадров в черно-белом диапазоне. Поэтому перед непосредственной обработкой происходит перевод кадров из гаммы RGB в серый цвет. В библиотеках OpenCV [5] это преобразование происходит так:

$$I = 0{,}299R + 0{,}587G + 0{,}114B.$$

Если поток считывается уже в черно-белом диапазоне, то этого преобразования не требуется.

В процессе распознавания выполняется ряд повторяющихся шагов. В первую очередь происходит непосредственно вычитание интенсивности цвета кадров:

$$D_j = |I_j - I_{j-1}|,$$

где j — номер текущего кадра. Соответственно алгоритм начинает свою работу лишь со второго кадра. Операция модуля разности происходит поэлементно.

Таким образом, может быть получен серый кадр, на котором будут отображены изменения, произошедшие за время одного кадра. Другим подходом является сравнение текущего кадра не с предыдущим, а с фоном — некоторым статичным эталоном, на котором не отображаются движущиеся объекты.

Как бы хорошо ни была закреплена камера, даже при небольшом ветре камера может испытывать микроколебания. Это вносит небольшой шум на контурах объектов. Помимо этого может происхо-

дить автоматическая настройка яркости видеоизображения, что также отразится на результате в виде шума. Необходимо отфильтровать получившийся результат. Для этого бинаризируем результат разницы двух кадров:

$$b_{x,\,y,\,j} = \begin{cases} 1, \text{ если } i_{x,\,y,\,j} > \delta; \\ 0 \text{ иначе.} \end{cases}$$

Здесь δ — некоторая целочисленная граница в диапазоне [0; 255], которая отражает силу фильтрации.

Как можно заметить на рис. 5, движущиеся машины представляются не цельным объектом, а набором разрозненных точек. Поэтому используем размытие. Для этого сначала происходит обратный перевод из бинарной матрицы в матрицу интенсивности:

$$\hat{i}_{x,\,y,\,j} = \begin{cases} 255, \text{ если } b_{x,\,y,\,j} = 1; \\ 0 \text{ иначе} \end{cases}$$

после происходит размытие:

$$\hat{i}_{x,\,y,\,j} = \frac{1}{M_{blur}} \sum_{x_1 = x-\gamma}^{x+\gamma} \sum_{y_1 = y-\gamma}^{y+\gamma} \hat{i}_{x_1,\,y_1,\,j},$$

при $x_1 \in [0;\, w]$ и $y_1 \in [0;\, h]$,

где $M_{blur} = \sum_{x_1 = x-\gamma}^{x+\gamma} \sum_{y_1 = y-\gamma}^{y+\gamma} 1$; γ — ширина ядра размытия. Далее необходимо провести повторную бинаризацию:

$$\hat{b}_{x,\,y,\,j} = \begin{cases} 1, \text{ если } (\hat{I}_{x,\,y,\,j} > \delta. \\ 0 \text{ иначе.} \end{cases}$$

Результат работы алгоритма размытия представлен на рис. 6.

Сформируем выборку, где каждый светлый пиксель ($\tilde{b}_{x,\,y,\,j} = 1$) будет являться точкой в двумерном пространстве. Эту выборку будем использовать в качестве данных для кластеризации объектов. Алгоритмом кластеризации был выбран метод контуров [6]. Каждый кластер будет движущимся объектом, который можно охарактеризовать размером (средний радиус и число точек).

В результате выполнения всех шагов имеются координаты $c_{i_1} = (x_{i_1}, y_{i_1})$ всех движущихся объектов на кадре. Далее объекты объединяются в траектории по следующему принципу: $c_{i_1,\,j}$ является продолжением точки $c_{i_2,\,j-1}$, если:

- точка $c_{i_2,\,j-1}$ находится ближе всех к точке $c_{i_1,\,j}$ среди всех точек на предыдущем кадре, но расстояние между ними не превышает допустимого радиуса R;
- ни одна из точек на текущем кадре не является уже продолжением точки $c_{i_2,\,j-1}$;
- траектория, в которую входит точка $c_{i_2,\,j-1}$, не является законченной.

Траектория считается законченной, если в течение определенного времени (например, пяти кадров), для нее не было найдено ни одной новой точки. Если траектория еще не закончена, а новой точки для нее не найдено, то последняя точка траектории объявляется текущей.

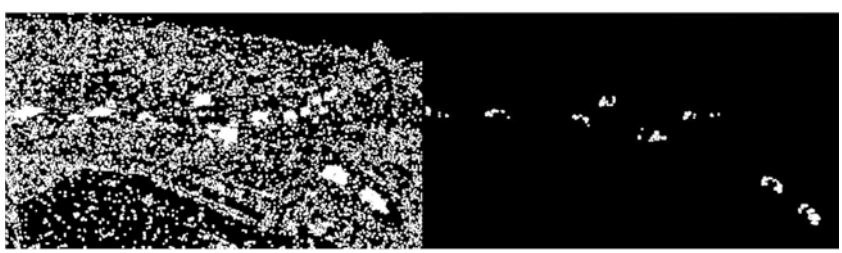
Если для какой-то точки из текущего кадра $c_{i_1,\,j}$ не найдено ни одной точки на предыдущих траекториях, то в этом случае образуется новая траектория.

При входе и выходе в размеченную область для траектории записывается изменение маршрута, а именно координаты текущего положения объекта и номер области, в которую он вошел, либо "0", если объект вышел из области. Таким образом, можно составить маршрут прохождения для каждого объекта.

Согласно записанному маршруту можно вести подсчет объектов в определенном направлении (поток объектов в секунду):

$$\hat{T}_k = \frac{1}{fD} \sum_{j_1 = 0}^{D} \sum_{i_1 = 1}^{S} c_{i_1,\,j-j_1}, \text{ если } c_{i_1,\,j-j_1} \in S_k,$$

где D — длина расчетного периода в кадрах; f — частота кадров/с; s — число распознанных объектов на кадре. Под объектом понимается траектория целиком. Это необходимо для того, чтобы не посчитать один и тот же объект несколько раз. Выражение $c_{i_1,\,j-j_1} \in S_k$ означает, что объект принадлежит области S_k и движется в указанном направлении. Объект принадлежит области S_k, если находится внутри



а) *б*)

Рис. 5. Результат фильтрации при δ = 0 (*а*) и δ = 15 (*б*)

Рис. 6. Результат размытия при γ = 20

четырехугольника $\vec{q}_k$. Объект движется в направлении $\vec{d}_k$, если угол между векторами $\vec{d}_k$ и $\overrightarrow{c_{i_1,j-j_1-1}c_{i_1,j-j_1}}$ меньше либо равен π. Объект движется в обратном направлении, если угол между векторами $\vec{d}_k$ и $\overrightarrow{c_{i_1,j-j_1-1}c_{i_1,j-j_1}}$ больше π.

Выше, на рис. 3, изображен пример подсчета движущихся объектов. При входе в область сравнивается текущее положение объекта и его предыдущие координаты точки, отмеченной в маршруте. В результате определяется направление движения. Объект считается для каждой области отдельно.

3. Исследование качества расчета

В качестве эксперимента будем исследовать перекресток, рассмотренный в предыдущем разделе.

Видео имеют следующие характеристики:

- $f = 25$ кадров/с;
- скорость передачи — 3,5 Мбит/с при разрешении 1280 × 720 пикселей;
- кодек MPEG;
- разрешение 1280 × 720 пикселей.

На рис. 7 изображен кадр из видеопотока.

Камера может быть расположена таким образом, что по мере продвижения размеры машин в пикселях могут изменяться. На рис. 7 можно заметить (подчеркнуто жирными линиями), что размеры в районе области 2 и области 5 отличаются примерно в 1,5 раза.

Таким образом, в описанном алгоритме параметры δ и γ задаются для каждой области свои. Это увеличило количество вычислений в пять раз на этапах фильтрации и размытия, но повысило точность на 3...5 %.

На рис. 7 можно заметить, что в области 3 есть вероятность, что поток $A \to B$ может перекрыть поток $C \to A$. Вследствие этой помехи точность оценки потока $C \to A$ падает до 50 %. Эта ситуация исправляется с помощью ввода дополнительной области под номером 1, которая оценивает проблемный поток $C \to A$ путем вычисления разницы между числом машин, проехавших через область 1, и числом машин через область 2 в левом направлении.

На рис. 8 представлены кадры из видеопотока, записанного во время снегопада.

Как можно заметить по правому изображению на рис. 8, погодные осадки игнорируются с помощью дополнительного ограничения: объекты не воспринимаются, если они меньше порогового значения. Снег все еще виден фильтрацией, но на оригинальном изображении в оценке потока не участвует. Снижения точности при помехах в виде осадков не происходит.

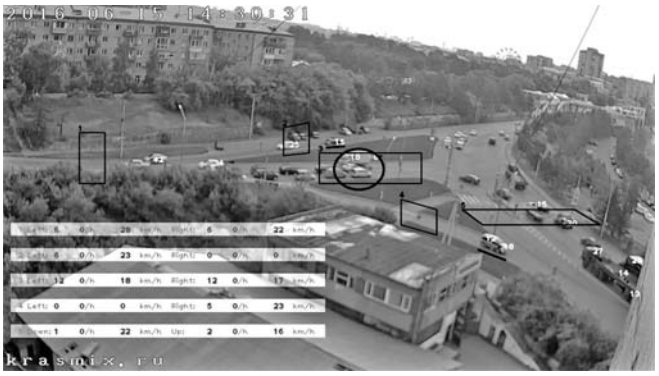


Рис. 7. Изображение кадра с помехами

Запишем одно и то же видео при разном разрешении (см. таблицу). Здесь N — номер эксперимента.

Алгоритм имеет следующие параметры: $\delta = 30$ и $\gamma = 20[1 - 0{,}1(N-1)]$, где N — номер эксперимента.

На каждом выделенном направлении ведется подсчет ошибки распознавания. Абсолютной ошибкой распознавания будем считать абсолютную разницу между оценкой числа проехавших машин и действительным их числом. Относительная ошибка — отношение абсолютной ошибки к действительному числу проехавших машин. Точностью является обратная величина средней относительной ошибки, рассчитанной на каждой области и направлении.

Если за приемлемую точность взять 80 % (средняя точность распознавания на имеющихся камерах видеонаблюдения), то достаточно будет качества изображения формата 800 (рис. 9). Это каче-

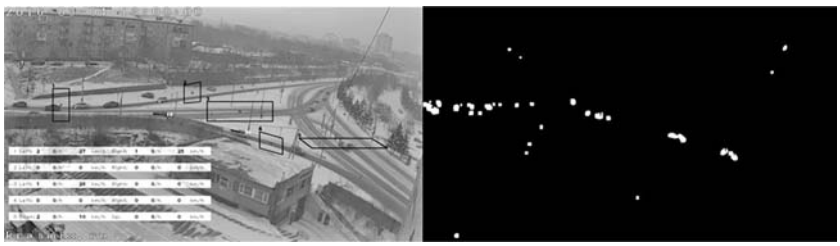


Рис. 8. Изображение и результат фильтрации кадра с осадками

Разрешение видео

Параметры	Номер эксперимента N									
	1	2	3	4	5	6	7	8	9	10
Ширина, пиксели	1280	1152	1024	896	768	640	512	384	256	128
Высота, пиксели	720	648	576	504	432	360	288	216	144	72
Разрешение, мегапиксели	0,92	0,75	0,59	0,45	0,33	0,23	0,15	0,08	0,04	0,01

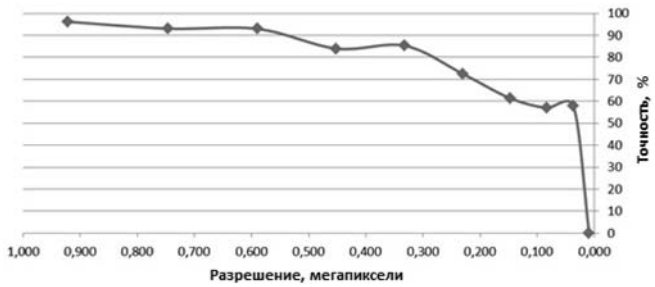


Рис. 9. Результаты работы алгоритма

ство не является очень высоким и не нагружает сильно канал.

Заключение

В ходе решения задачи распознавания объектов в видеопотоке была сформирована методика оценки потока движущихся объектов через размеченную область в указанном направлении. Алгоритм подсчета объектов работает с точностью 80 % при среднем качестве изображения. Методика справляется со своей задачей на таких сложных участках транспортной системы, как перекресток. Также модификации, введенные в методику, позволяют справиться с такими препятствиями, как плохие погодные условия, расположение камеры под острым углом, перекрытия объектов. Открытым вопросом остаются исследования оценки потока в период повышенной интенсивности — пробки. Для решения этого вопроса необходимо комбинировать методы вычитания фона и предыдущих кадров. Все алгоритмы реализованы в виде облачного сервиса для широкого использования под ОС MS Windows и Linux, в настоящее время проходит тестирование бета-версии продукта. В результате реализация позволит осуществлять сбор статистики по оценкам дорожных потоков, что, в свою очередь, позволит вести настройку модели транспортной системы в режиме реального времени.

Список литературы

1. **Шапиро Л., Стокман Д.** Компьютерное зрение. М.: Бином. Лаборатория знаний, 2006. 752 с.
2. **Pulli K., Baksheev A., Kornyakov K., Eruhimov V.** Real-time computer vision with OpenCV // Commun. 2012. Vol. 55, N. 6. P. 61—69.
3. **Еремин С. Н.** Алгоритмическое и программное обеспечение автоматизированной системы контроля параметров автотранспортных потоков: дис. канд. техн. наук: 05.13.14. Череповец, 2000.
4. **Живой** Красноярск. Сайт города Красноярска. Пробки. URL: http://krasmix.ru/cameras (дата обращения: 15.06.2016).
5. **Документация** OpenCV. URL: http://docs.opencv.org (дата обращения: 04.07.2016).
6. **Калмыков В. Г.** Структурный метод описания и распознавания отрезков цифровых прямых в контурах бинарных изображений // Штучний інтелект. 2002. Т. 4. С. 450—457.

A. S. Potekhin, Programmer, andrew.potekhin94@gmail.com,
A. V. Strelnikov, Postgraduate Student, Lead Programmer, ArkadiyS@RD-Science.com,
OOO "RD-Science", Krasnoyarsk

The Method of Traffic Flow Estimation on the Crossroad by Video Control Data

The problem of traffic flow estimation on the crossroads with help video observation data is researched in this article. The purpose of this article is to show decision of formulated problem. This purpose is solved by couple of developed algorithms. The difference frames technique is used for the objects recognition. The intelligence associative search algorithm is used for objects path recovery. The developed algorithms are implemented as software modules. The statistical experiments results are shown. Based on these results we can conclude about the algorithms quality. The algorithm accuracy is 85—90 %. This quality makes developed algorithms competitive. Realized system successfully completes the task with set of complications: poor weather, obstacles on vision lines, camera position at an acute angle. Today the combined algorithm using background estimation is developed. It is necessary to achieve high recognition quality with conditions of the traffic jams. Implemented software modules are part of transport system. The modules of traffic flow estimation allow fitting transport system parameters. For example, it can be used for fitting parameters of traffic lights. The collected statistics are updated every minute. It could be updated every second, but it is not necessary. Such discretization allows fitting transport system in real time. Also collected statistics could be used for solving other problems, such is evaluation of the wear resistance of the roadway and evaluation of the quantity of exhaust gas.

Keywords: *computer vision, traffic flow, recognition, mobile objects, video stream, video process*

References

1. **Shapiro L., Stokman D.** *Komp'yuternoe zrenie*, Moscow, Binom. Laboratoriya znanij, 2006, 752 p.
2. **Pulli K., Baksheev A., Kornyakov K., Eruhimov V.** Real-time computer vision with OpenCV, *Commun.*, 2012, vol. 55, no. 6, pp. 61—69.
3. **Eremin S. N.** Algoritmicheskoe i programmnoe obespechenie avtomatizirovannoj sistemy kontrolya parametrov avtotransportnyh potokov, Dis. ... kand. tekhnicheskim naukam, 05.13.14. Cherepovec, 2000.
4. **Zhivoj Krasnoyarsk**. *Sajt goroda Krasnoyarska*. Probki. URL: http://krasmix.ru/cameras (data of access: 15.06.2016).
5. **Dokumentaciya** *OpenCV*, URL: http://docs.opencv.org (data of access: 04.07.2016).
6. **Kalmykov V. G.** Strukturnyj metod opisaniya i raspoznavaniya otrezkov cifrovyh pryamyh v konturah binarnyh izobrazhenij, *Shtuchnij intelekt*, 2002, vol. 4, pp. 450—457.

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ
CRYPTOSAFETY INFORMATION

УДК 004.9, 004.94, 004.56 DOI:10.17587/it.23.135-142

А. А. Коляда, д-р физ.-мат. наук, доц., e-mail: razan@tut.by,
Н. А. Коляда, науч. сотр., e-mail: razan@tut.by,
С. Ю. Протасеня, мл. науч. сотр., e-mail: Estellita@mail.ru,
Е. В. Шабинская, канд. техн. наук, доц., ст. науч. сотр., e-mail: shabinskaya@rambler.ru,
Научно-исследовательское учреждение "Институт прикладных физических проблем имени А. Н. Севченко" Белорусского государственного университета, Минск, Беларусь

Мультипликативно-субстрактивный метод вычисления денормирующего коэффициента для криптографических RSA-преобразований в модулярном коде

Статья посвящена проблеме денормировки базовых преобразований в криптосистеме RSA с минимально избыточной модулярной кодовой организацией. Для решения поставленной задачи предложен новый метод, основанный на мультипликативно-субстрактивной вычислительной схеме рекурсивного типа. Теоретическую базу применяемого подхода составляет аппарат интервально-модулярных форм чисел и интервально-индексных характеристик. Используемый инструментарий позволяет достичь существенного упрощения немодульных операций, входящих в состав синтезированной процедуры расчета денормирующего коэффициента для криптографических RSA-преобразований. К таким операциям относятся операции умножения на основании модулярной системы счисления и приведение аккумулируемых произведений оснований к остатку по модулю криптосистемы. На множестве модулей криптосистемы разрядностью 1024—2048 бит время работы предложенной процедуры на ПЭВМ с процессором Intel Core i5 (частота 2,27 ГГц) находится в секундном диапазоне.

Ключевые слова: *криптосистема RSA, криптографическое RSA-преобразование, модулярная система счисления, интервально-индексные характеристики, минимально избыточная модулярная арифметика, умножение Монтгомери*

Введение

Как известно [1—4], компоненты математического обеспечения любой криптосистемы распределяются по двум группам, отличающимся друг от друга функциональным назначением. Первую группу составляют средства, работающие в реальном времени. Они осуществляют шифрование и дешифрование данных. Вторую группу образуют средства вспомогательного характера. На них возлагается решение широкого круга задач, выполняемых вне основного вычислительного процесса. В случае криптосистем RSA с модулярной кодовой организацией к таким задачам, в частности, относятся: формирование необходимой базы простых чисел для выбора рабочего системного модуля, а также оснований используемой модулярной системы счисления (МСС), факторизация больших чисел, генерирование ключей шифрования и дешифрования сообщений в требуемом коде, построение комплекса таблиц и системных констант для компьютерной реализации как процедур базовой конфигурации модулярной арифметики (МА), так и синтезированных на ее основе алгоритмов криптографических RSA-преобразований. Задачи вспомогательной группы, как правило, представляют собой так называемые сложные задачи. Поэтому, несмотря на то что их решение осуществляется на этапе предварительных вычислений, к которому, вообще говоря, не предъявляется жестких требований к временным затратам, всемерное их снижение является исключительно важной и актуальной проблемой.

Настоящая работа посвящена созданию методологического и алгоритмического обеспечения криптографических RSA-преобразований, выполняемых с помощью МСС. В частности, предлагается новый мультипликативно-субстрактивный метод вычисления денормирующего коэффициента (ДНК) для процедуры возведения в степень по системному модулю, которая служит основой базовых криптографических преобразований. Синтезированный алгоритм расчета коэффициента денормировки, осуществляя декомпозицию кольца вычетов по модулю криптосистемы на набор колец по модулям МСС, сводит все вычисления к операциям стан-

дартной разрядности в рамках простой рекурсивной схемы. При этом в качестве теоретической базы применяется аппарат так называемых интервально-модулярных форм чисел. Известные подходы к решению задачи нейтрализации влияния нормировки произведений Монтгомери при выполнении криптографических RSA-преобразований в модулярном коде [5—7] базируются на применении алгоритмов общего деления больших чисел [8—12] и операциях вычитания в коде МСС с последующим детектированием знаков получаемых разностей с использованием трудоемких алгоритмов. С точки зрения производительности указанные способы в виду оперирования в диапазоне больших чисел не эффективны.

Постановка задачи и способы ее решения

Ключевым отличительным признаком криптосистем RSA [1—3, 13] является применение процедур умножения типа Монтгомери по большому модулю p, который относится к открытым системным параметрам. В случае криптосистемы RSA с модулярной кодовой организацией базовая мультипликативная операция имеет вид

$$\tilde{\gamma} = |ABM_l^{-1}|_p (A, B \in \mathbf{Z}_p = \{0, 1, ..., p-1\}), \quad (1)$$

где A и B — операнды, заданные соответственно кодами $(\alpha_1, \alpha_2, ..., \alpha_k)$ и $(\beta_1, \beta_2, ..., \beta_k)$, $\alpha_i = |A|_{m_i}$, $\beta_i = |B|_{m_i}$ $(i = \overline{1, k})$ МСС с базисом $\mathbf{M} = \{m_1, m_2, ..., m_k\}$ оснований $m_1, m_2, ..., m_k$; через $|x|_m$ обозначается элемент множества $\mathbf{Z}_m = \{0, 1, ..., m-1\}$, сравнимый с x (в общем случае рациональным числом) по натуральному модулю m; $M_l = \prod_{i=1}^{l} m_i$ — нормирующий коэффициент, используемый в алгоритме умножения Монтгомери по МА-схеме RSA [13]; основания $m_1, m_2, ..., m_l (1 < l < k)$ составляют базис $\mathbf{M_1} = \{m_1, m_2, ..., m_l\}$ усеченной МСС; основания $m_{l+1}, m_{l+2}, ..., m_k$ образуют базис $\mathbf{M_2} = \{m_{l+1}, m_{l+2}, ..., m_k\}$ другой из используемых в алгоритме Монтгомери усеченной МСС.

Базовые преобразования криптосистемы RSA осуществляются согласно правилу

$$Y = |X^E|_p, \quad (2)$$

где X — целое число из диапазона $\mathbf{Z}_p$, отождествляемое с шифруемым или дешифруемым с помощью ключа E сообщением. Процесс вычисления функции (2) в МСС с базисом $\mathbf{M} = \{\mathbf{M_1}, \mathbf{M_2}\}$ сводится к последовательности операций умножения по модулю p вида (1). При этом из-за включения в (1) коэффициента M_l^{-1} нормировки в применяемой процедуре возведения в степень приходится использовать денормирующий коэффициент

$$N = |M_l^2|_p = (\nu_1, \nu_2, ..., \nu_k) (\nu_i = |N|_{m_i} (i = \overline{1, k})), \quad (3)$$

который позволяет исключить из конечного результата коэффициент нормировки произведений. Поскольку p является большим числом, то получение модулярного кода $(\nu_1, \nu_2, ..., \nu_k)$ денормирующего коэффициента (3) относится к разряду сложных задач.

Предлагаемое решение данной задачи базируется на приводимых далее теоретических положениях [14—17].

Согласно Китайской теореме об остатках следует, что целое число X и цифры его модулярного кода $(\chi_1, \chi_2, ..., \chi_l)$ $(\chi_i = |X|_{m_i} (i = \overline{1, l}))$ по базису $\mathbf{M_1}$ связаны между собой соотношением

$$X = \sum_{i=1}^{l-1} M_{i, l-1} \chi_{i, l-1} + M_{l-1} I_l(X)$$
$$(\chi_{i, l-1} = |M_{i, l-1}^{-1} \chi_i|_{m_i}), \quad (4)$$

где $M_{i, l-1} = M_{l-1}/m_i$; $M_{l-1} = \prod_{j=1}^{l-1} m_j$; $I_l(X)$ — интервальный индекс числа X в МСС с базисом $\mathbf{M}_1$. Выражение (4) называется интервально-модулярной формой целого числа X по базису $\mathbf{M}_1$.

Справедлива следующая теорема.

Теорема 1. Для того чтобы в МСС с попарно простыми основаниями $m_1, m_2, ..., m_l$ интервальный индекс $I_l(X)$ каждого элемента $X = (\chi_1, \chi_2, ..., \chi_l)$ диапазона $\mathbf{Z}_{2M}^{-} = \{-M, -M+1, ..., M-1\}$ $(M = m_0 M_{l-1}$; m_0 — вспомогательный модуль) полностью определялся компьютерным интервальным индексом — вычетом $\hat{I}_l(X) = |I_l(X)|_{m_l}$, необходимо и достаточно, чтобы l-е основание удовлетворяло условию: $m_l \geqslant 2m_0 + l - 2$ $(m_0 \geqslant l - 2)$. При этом для $I_l(X)$ верны расчетные соотношения:

$$I_l(X) = \begin{cases} \hat{I}_l(X), \text{ если } \hat{I}_l(X) < m_0, \\ \hat{I}_l(X) - m_l, \text{ если } \hat{I}_l(X) \geqslant m_0; \end{cases} \quad (5)$$

$$\hat{I}_l(X) = \left| \sum_{i=1}^{l} R_{i, l}(\chi_i) \right|_{m_l}; \quad (6)$$

$$R_{i, l}(\chi_i) = |-m_i^{-1} |M_{i, l-1}^{-1} \chi_i|_{m_i}|_{m_l}$$
$$(i \neq l), \; R_{l, l}(\chi_l) = |M_{l-1}^{-} \chi_l|_{m_l}. \quad (7)$$

Определение 1. Набор величин $(\chi_1, \chi_2, ..., \chi_{l-1}, I_l(X))$ или $(\chi_{1, l-1}, \chi_{2, l-1}, ..., \chi_{l-1, l-1}, I_l(X))$ называют интервально-модулярным кодом числа X по базису $\mathbf{M}_1$.

В указанных формах интервально-модулярного кода допускается также использование вместо $I_l(X)$ компьютерного интервального индекса $\hat{I}_l(X)$.

Определение 2. Интегральная характеристика модулярного кода числа X в МСС с базисом $\mathbf{M_1}$ вида $J_l(X) = \lfloor I_l(X)/m_l \rfloor$ называется главным интервальным индексом относительно модуля $\mathbf{M}_l$ (через $\lfloor x \rfloor$ обозначается целая часть вещественной величины x).

Достигаемая в условиях теоремы 1 простота вычисления интервально-индексных характеристик $I_l(X)$, $\hat{I}_l(X)$ и $J_l(X)$ (см. (5)—(7), определение 2) приводит к адекватной оптимизации и немодульных процедур, базирующихся на интервально-модулярной форме (4). В частности, аппарат интервально-модулярных форм может успешно применяться для синтеза процедур расширения кодов, определения знака числа и др. [13—17]. Для детектирования знаков целого числа в классе решаемых задач криптографических МА-приложений высокую эффективность обеспечивает нижеследующая теорема.

Теорема 2. Пусть в МСС с основаниями $\mathbf{M}_1$, $\mathbf{M}_2$, ..., $\mathbf{M}_l \geqslant l - 2$ ($l \geqslant 2$) целому числу X отвечает код $(\chi_1, \chi_2, ..., \chi_l)$ и пусть $J_1(X)$ — главный интервальный индекс данного целого числа. Знаки чисел X и $J_1(X)$ совпадают при $l = 2$, а также при $l > 2$, если $J_1(X) \neq -1$.

Применение минимально-избыточного модулярного кодирования, сущность которого раскрывает теорема 1, снижает сложность расчета интервально-индексных характеристик к предельно низкому уровню. Это открывает принципиально новые возможности для уменьшения объема вычислений в МСС, осуществляемых на базе интервально-модулярных форм чисел. Кроме того, использование теоремы 2, как упрощенного инструментария детектирования знаков чисел, позволяет синтезировать эффективные логические процедуры. В полной мере сказанное относится к разработанному мультипликативно-субстрактивному методу расчета денормирующего коэффициента для криптографических RSA-преобразований.

Мультипликативно-субстрактивный метод вычисления денормирующего коэффициента для криптографических RSA-преобразований

Как показывает равенство (3), базовое число M_l^2, определяющее совместно с модулем p денормирующий коэффициент N, имеет мультипликативную структуру. Благодаря данному обстоятельству для расчета N удалось разработать метод, который весьма прост в реализации. Демонстрируемый далее подход к решению рассматриваемой задачи базируется на мультипликативно-субстрактивном способе приведения целого числа M_l^2 к остатку по модулю p с помощью теоремы 2 и вычислительной схемы, конструируемой согласно сравнению вида

$$((...(((...((m_1 - u_1 p)m_2 - u_2 p)...)m_l - u_l p)m_1 - - v_1 p)...)m_l - v_l p) \equiv \left(\prod_{i=1}^{l} m_i \prod_{j=1}^{l} m_j \right) \equiv N(\mathrm{mod}\, p), \quad (8)$$

где $u_1, u_2, ... u_l, v_1, v_2, ..., v_l$ — адаптивно выбираемые подходящие целочисленные множители.

Запишем (8) в более наглядной и приемлемой для компьютерной реализации форме:

$$\langle N_0^{(1)} = 1,\ N_i^{(1)} = N_{i-1}^{(1)} \cdot m_i - u_i p\ (i = \overline{1,l}\,),$$
$$N_0^{(2)} = N_l^{(1)},\ N_i^{(2)} = N_{i-1}^{(2)} \cdot m_i - v_i p\ (i = \overline{1,l}\,)\rangle. \quad (9)$$

Процесс приведения целого числа M_l^2 к остатку по модулю p, осуществляемый по схеме (9), является рекурсивным. На каждой итерации этого процесса выполняется одна и та же типовая операция, которая состоит в выделении из множества

$$\mathbf{N}(m) = \{\forall n' = nm - fp | f \in \mathbf{Z}_m\}$$
$$(n \in \mathbf{Z}_p;\ m \in \mathbf{M}_1 = \{m_1, m_2, ..., m_l\})$$

элемента, принадлежащего также и к $\mathbf{Z}_p$. Поиск искомого значения параметра f требует детектирования знаков целого числа вида

$$n' = nm - fp (f \in \mathbf{Z}_m). \quad (10)$$

При использовании сравнительно небольших модулей МСС, например $m \in (2^{15}; 2^{16})$, для выполнения указанных операций может быть применен упрощенный алгоритмический инструментарий, основанный на интервально-модулярной форме чисел и теореме 2.

Пусть целые числа n и p заданы своими интервально-модулярными формами в базисе $\mathbf{M}_1$:

$$n = \sum_{i=1}^{l-1} M_{i,\,l-1} \upsilon_{i,\,l-1} + M_{l-1} I_l(n)$$
$$(\upsilon_{i,\,l-1} = |M_{i,\,l-1}^{-1} \upsilon_i|_{m_i},\ \upsilon = |n|_{m_i}), \quad (11)$$

$$p = \sum_{i=1}^{l-1} M_{i,\,l-1} \pi_{i,\,l-1} + M_{l-1} I_l(p)$$
$$(\pi_{i,\,l-1} = |M_{i,\,l-1}^{-1} \pi_i|_{m_i}),\ \pi_i = |p|_{m_i}), \quad (12)$$

где $I_l(n)$ и $I_l(p)$ — интервальные индексы чисел n и p соответственно. Обозначим через $(\upsilon'_1, \upsilon'_2, ..., \upsilon'_l)$ код числа n' в МСС с модулями $m_1, m_2, ..., m_l$ и его интервальный индекс через $I_l(n')$. Для получения интервально-модулярной формы целого числа (10):

$$n' = \sum_{i=1}^{l-1} M_{i,\,l-1} \upsilon'_{i,\,l-1} + M_{l-1} I_l(n')$$
$$(\upsilon'_{i,\,l-1} = |M_{i,\,l-1}^{-1} \upsilon'_i|_{m_i}) \quad (13)$$

достаточно в (10) подставить (11), (12) и затем, применяя лемму Эвклида, выполнить преобразование:

$$n' = m\left(\sum_{i=1}^{l-1} M_{i,\,l-1}\upsilon_{i,\,l-1} + M_{l-1}I_l(n)\right) - f\left(\sum_{i=1}^{l-1} M_{i,\,l-1}\pi_{i,\,l-1} + M_{l-1}I_l(p)\right) =$$
$$= \sum_{i=1}^{l-1} M_{i,\,l-1}(m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1}) + M_{l-1}(mI_l(n) - fI_l(p)) =$$
$$= \sum_{i=1}^{l-1} M_{i,\,l-1}(|m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1}|_{m_i} + \lfloor (m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1})/m_i \rfloor m_i) + M_{l-1}(mI_l(n) - fI_l(p)) =$$
$$= \sum_{i=1}^{l-1} M_{i,\,l-1}|m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1}|_{m_i} + M_{l-1}\left(mI_l(n) - fI_l(p) + \sum_{i=1}^{l-1} \lfloor (m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1})/m_i \rfloor\right). \quad (14)$$

Из (13) и (14) следует, что

$$\upsilon'_{i,\,l-1} = |m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1}|_{m_i} \ (i = \overline{1, l-1}), \quad (15)$$

$$I_l(n') = mI_l(n) - fI_l(p) + \sum_{i=1}^{l-1} \lfloor (m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1})/m_i \rfloor. \quad (16)$$

Согласно теореме 2 число n' и его главный интервальный индекс $J_l(n')$ (определение 2), который в соответствии с (16) вычисляется по формуле

$$J_l(n') = \lfloor I_l(n')/m_l \rfloor = \left\lfloor \left(mI_l(n) - fI_l(p) + \sum_{i=1}^{l-1} \lfloor (m\upsilon_{i,\,l-1} - f\pi_{i,\,l-1})/m_i \rfloor \right) / m_l \right\rfloor, \quad (17)$$

имеют одинаковые знаки, если $J_l(n') \neq -1$, т.е.

$$\mathbf{sn}(n') = \mathbf{sn}(J_l(n')) \ (J_l(n') \neq -1), \quad (18)$$

где **sn** — знаковая функция вида

$$\mathbf{sn}(x) = \begin{cases} 0, \text{ если } x \geqslant 0, \\ 1, \text{ если } x < 0. \end{cases}$$

Случай $J_l(n') = -1$ отвечает неопределенной ситуации при детектировании знака целого числа n' по правилу (18) с использованием (17). Возникновение указанной ситуации на той или иной итерации вычислительной схемы (9) маловероятно и не является критичным. На последующих итерациях возможная неопределенность устраняется с вероятностью, практически близкой к 1.

Цель анализа знаков целого числа (10) состоит в отыскании значения $\tilde{f}$ параметра $f \in \mathbf{Z}_m$, которое обеспечивает выполнение условия:

$$\begin{cases} nm - \tilde{f}\,p \geqslant 0, \\ nm - (\tilde{f} + 1)p < 0. \end{cases} \quad (19)$$

Поскольку в (10) $n \in \mathbf{Z}_p$, то $nm \geqslant 0$, а nm-, $mp = m(n-p) < 0$. Следовательно, в $\mathbf{Z}_m$ всегда существует единственный элемент $\tilde{f}$, удовлетворяющий (19).

Алгоритм расчета денормирующего коэффициента для криптографических RSA-преобразований

На базе изложенных теоретико-методологических положений синтезирован алгоритм расчета денормирующего коэффициента (РДНК), который представлен ниже.

Параметры алгоритма:

- Модуль p криптосистемы.
- Набор $\mathbf{M} = \{m_1, m_2, ..., m_l, m_{l+1}, ..., m_k\}$ из k 16-битовых простых модулей, который объединяет базисы $\mathbf{M}_1 = \{m_1, m_2, ..., m_l\}$ и $\mathbf{M}_2 = \{m_{l+1}, m_{l+2}, ..., m_k\}$ $(1 < l < k)$ минимально избыточной МСС с диапазонами $\mathbf{Z}^-_{2m_0M_{l-1}}$ и $\mathbf{Z}^-_{2m_0M_{k-1}/M_l}$, удовлетворяющие условиям

$$m_l \geqslant 2m_0 + l - 2, \ m_k \geqslant 2m_0 + k - l - 2,$$
$$m_0 \geqslant \max\{l - 2, \ k - l - 2\},$$
$$2p < \min\{m_0M_{l-1}, \ m_0M_{k-1}/M_l\}.$$

Входные данные алгоритма: модулярный код $(\pi_1, \pi_2, ..., \pi_l)$ модуля p в базисе $\mathbf{M}_1(\pi_i = |p|_{m_i} \ (i = \overline{1,l}))$.

Выходные данные: модулярный код $(\nu_1, \nu_2, ..., \nu_l, \nu_{l+1}, ..., \nu_k)$ денормирующего коэффициента $N = |M_l^2|_p$ по полному базису $\mathbf{M}(\upsilon_i = |N|_{m_i} \ (i = \overline{1,k}))$.

Предварительно получаемые данные:

- Коэффициенты нормировки цифр модулярного кода в базисе $\mathbf{M}_1$: $\tilde{C}_i = |M^{-1}_{i,\,l-1}|_{m_i} \ (i = \overline{1, l-1})$.
- Коэффициенты денормировки цифр модулярного кода в базисе $\mathbf{M}_1$: $C_i = |M_{i,\,l-1}|_{m_i} \ (i = \overline{1, l-1})$.
- Коэффициенты для операции расширения интервально-модулярного кода по базису $\mathbf{M}_1$ на модули $m_l, m_{l+1}, ..., m_k$: $C_{i,\,j} = |M_{i,\,l-1}|_{m_j}$ $(i = \overline{1, l-1})$, $C_{l,\,j} = |M_{l-1}|_{m_j}$, где $j = \overline{l, k}$.
- Таблицы TIIi интервального индекса, генерируемые по правилу:

$$TIIi[\chi] = R_{i,\,l}(\chi) = |-m_i^{-1}\,|M^{-1}_{i,\,l-1}\,\chi|_{m_i}|_{m_l}$$
$$(\chi = \overline{0, m_i - 1}, \ i = \overline{1, l-1}),$$
$$TIII[\chi] = R_{l,\,l}(\chi) = |M^{-1}_{l-1}\,\chi|_{m_l} \ (\chi = \overline{0, m_l - 1}).$$

Тело алгоритма расчета денормирующего коэффициента (РДНК)

РДНК 1. Для модуля $p = (\pi_1, \pi_2, ..., \pi_{l-1}, \pi_l)$ криптосистемы сформировать интервально-модулярный код $(\pi_{1,\,l-1}, \pi_{2,\,l-1}, ..., \pi_{l-1,\,l-1}, I_l(p))$ согласно формулам:

$$\pi_{i,\,l-1} = |\widetilde{C}_i \pi_i|_{m_i} \ (i = \overline{1, l-1}),$$

$$I_l(p) = \begin{cases} \hat{I}_l(p), \text{ если } \hat{I}_l(p) < m_0, \\ \hat{I}_l(p) - m_l, \text{ если } \hat{I}_l(p) \geqslant m_0; \end{cases}$$

$$\hat{I}_l(p) = \left| \sum_{i=1}^{l} TIIi[\pi_i] \right|_{m_l} \text{ (см. (5)—(7)).}$$

РДНК 2. Сформировать интервально-модулярный код $(\widetilde{C}_1, \widetilde{C}_2, ..., \widetilde{C}_{l-1}, I_l(1))$ числа 1, где

$$I_l(1) = \begin{cases} \hat{I}_l(1), \text{ если } \hat{I}_l(1) < m_0, \\ \hat{I}_l(1) - m_l, \text{ если } \hat{I}_l(1) \geqslant m_0; \end{cases}$$

$$\hat{I}_l(1) = \left| \sum_{i=1}^{l} TIIi[1] \right|_{m_l}.$$

РДНК 3. Выполнить операцию присвоения: $n = (\upsilon_{1,\,l-1}, \upsilon_{2,\,l-1}, ..., \upsilon_{l-1,\,l-1}, I_l(n)) = (\widetilde{C}_1, \widetilde{C}_2, ..., \widetilde{C}_{l-1}$ и $I_l(1)) = 1$.

РДНК 4. Положить $S = 1$.

РДНК 5. Порядковому номеру текущего модуля базиса $\mathbf{M}_1 = \{m_1, m_2, ..., m_l\}$ присвоить начальное значение: $j = 1$.

РДНК 6. Применяя (15), (16) (при $f = 0$), рассчитать цифры интервально-модулярного кода $(\upsilon'_{1,\,l-1}, \upsilon'_{2,\,l-1}, ..., \upsilon'_{l-1,\,l-1}, I_l(n'))$ целого числа $n' = nm_j$:

$$\upsilon'_{i,\,l-1} = |m_j \upsilon_{i,\,l-1}|_{m_i} \ (i = \overline{1, l-1}),$$

$$I_l(n')) = m_j I_l(n) + \sum_{i=1}^{l-1} \lfloor m_j \upsilon_{i,\,l-1} / m_i \rfloor.$$

РДНК 7. Выполнить операцию присваивания:

$$n = n' \Leftrightarrow (\upsilon_{1,\,l-1}, \upsilon_{2,\,l-1}, ..., \upsilon_{l-1,\,l-1}, I_l(n)) = (\upsilon'_{1,\,l-1}, \upsilon'_{1,\,l-1}, ..., \upsilon'_{l-1,\,l-1}, I_l(n')).$$

РДНК 8. Получить интервально-модулярный код $(\upsilon'_{1,\,l-1}, \upsilon'_{2,\,l-1}, ..., \upsilon'_{l-1,\,l-1}, I_l(n')) = (\upsilon'_{1,\,l-1}, \upsilon'_{2,\,l-1}, ..., \upsilon'_{l-1,\,l-1}, I_l(n')) - (\pi_{1,\,l-1}, \pi_{2,\,l-1}, ..., \pi_{l-1,\,l-1}, I_l(p))$ разности $n' = n' - p$, используя формулы типа (15), (16):

$$\upsilon'_{i,\,l-1} = |\upsilon'_{i,\,l-1} - \pi_{i,\,l-1}|_{m_i} \ (i = \overline{1, l-1}), \ I_l(n') =$$

$$= I_l(n') + \sum_{i=1}^{l-1} \lfloor (\upsilon'_{i,\,l-1} - \pi_{i,\,l-1} - \pi_{i,\,l-1}) / m_i \rfloor.$$

РДНК 9. Вычислить главный интервальный индекс числа n: $J_l(n') = \lfloor I_l(n')/m_l \rfloor$.

РДНК 10. Если $J_l(n') < -1$ (целое число $n' < 0$) или $J_l(n') = -1$ (случай неопределенности знака целого числа n'), то при $j \neq l$ увеличить j на 1 ($j = j + 1$) и перейти к РДНК 6, а по достижении равенства $j = l$ перейти к РДНК 12.

РДНК 11. В случае $J_l(n') \geqslant 0$, указывающем на $n' \geqslant 0$, перейти к РДНК 7.

РДНК 12. При $S = 2$ инкрементировать S ($S = S + 1$) и перейти к РДНК 5.

РДНК 13. Полученный интервально-модулярный код $(\upsilon_{1,\,l-1}, \upsilon_{2,\,l-1}, ..., \upsilon_{l-1,\,l-1}, I_l(n))$ целого числа $n = N = |M_l^2|_p$ расширить на модули $m_l, m_{l+1}, ..., m_k$ согласно правилу:

$$\upsilon_j = \left| \sum_{i=1}^{l-1} |C_{i,\,j} \upsilon_{i,\,l-1}|_{m_j} + |C_{l,\,j} I_l(n)|_{m_j} \right|_{m_j}.$$

РДНК 14. Получить цифры модулярного кода числа $N = n$ по модулям $m_1, m_2, ..., m_{l-1}$: $\upsilon_i = |C_i \upsilon_{i,\,l-1}|_{m_i}$ ($i = \overline{1, l-1}$).

РДНК 15. Зафиксировать модулярный код $(\upsilon_1, \upsilon_2, ..., \upsilon_k)$ по полному базису $\mathbf{M} = \{m_1, m_2, ..., m_k\}$ в качестве искомого кода денормирующего коэффициента $N = |M_l^2|_p$ и завершить работу алгоритма.

Оценка эффективности алгоритма расчета денормирующего коэффициента

Остановимся кратко на реализационных аспектах вычислительной схемы (9) разработанного метода расчета денормирующего коэффициента $N = |M_l^2|_p$. В приведенном алгоритме РДНК 1—РДНК 15 выделение в множестве $\mathbf{N}(m)$ искомого элемента $n' = n'(\widetilde{f}) = nm - \widetilde{f} p$ осуществляется путем простого перебора целых чисел 1, 2, ..., $m - 1$ в указанном порядке до тех пор, пока не будет найдено требуемое $\widetilde{f} \in \mathbf{Z}_m$. Для схемы (9) в общей сложности это потребует времени, ограниченного сверху порогом

$$t_{\text{РДНК, max}} = 2t_{\text{бо}} \sum_{i=1}^{l} (m_i - 1), \qquad (20)$$

где $t_{\text{бо}}$ — длительность базовой операции реализуемого рекурсивного процесса (9), которая заключается в получении по известным n, m и фиксированном $f \in \mathbf{Z}_m$ с помощью (15)—(18) интервально-модулярного кода $\upsilon'_{1,\,l-1}, \upsilon'_{2,\,l-1}, ..., \upsilon'_{l-1,\,l-1}, I_l(n')$, главного интервального индекса $J_l(n')$ и знака $\text{sn}(n')$

целого числа $n' = n'(f) = nm - fp \in \mathbf{N}(m)$. Для проведения всех необходимых расчетов в рамках схемы (9) удобно использовать минимально избыточную МСС, удовлетворяющую теореме 1. При этом полагается $2p < m_0 M_{l-1}$. Из данного требования ввиду условия $M_l \geqslant 2m_0 + l - 2$ вытекает неравенство $4p < 2m_0 M_{l-1} \leqslant M_{l-1}(m_l - l + 2)$. При $2p < m_0 M_{l-1}$ диапазон $\mathbf{Z}^{-}_{2m_0 M_{l-1}}$ базовой минимально избыточной МСС включает множество $\mathbf{Z}^{-}_{4p}$, а значит, и множества $\mathbf{Z}_{2p}$, $\mathbf{Z}_p$.

Поэтому согласно формуле (5) интервально-индексные характеристики $I_l(n)$ и $I_l(p)$ целых чисел n и p из (10) принимают значения, не выходящие за пределы промежутка $[-m_0 - l + 2; m_0 - 1]$. С учетом отмеченного обстоятельства из (16) находим:

$$\min\{I_l(n')\} = -2m_0 - 2(l-2) - (m-1)(l-1) = \\ = -2m_0 - (m+1)(l-1) + 2; \quad (21)$$

$$\max\{I_l(n')\} = 2m_0 + (m-1)(l-1) - 2. \quad (22)$$

Оценки (21) и (22) границ изменения интервального индекса $I_l(n')$ показывают, что в случае применения МСС с 16-битовыми основаниями для проведения в рамках схемы (9) вычислений, связанных с интервально-индексными характеристиками (16) и (17), достаточно использовать арифметику целых чисел разрядностью 32, 64 бит. Сказанное относится и к расчетным соотношениям (15).

Общие временные затраты на выполнение базовой операции схемы (9) определяем оценочным выражением:

$$t_{\text{бо}} = (3l-1)(t_{\text{сл}} + t_{\text{ум}}) + lt_{\text{дел}}, \quad (23)$$

где $t_{\text{сл}}$, $t_{\text{ум}}$, $t_{\text{дел}}$ — длительности операций сложения/вычитания, умножения и деления 32-битовых целых чисел соответственно. Подстановка (23) в (20) дает:

$$t_{\text{РДНК, max}} = \\ = 2\sum_{i=1}^{l}(m_i - 1)((3l-1)(t_{\text{сл}} + t_{\text{ум}}) + lt_{\text{дел}}. \quad (24)$$

Пусть в качестве инструментальной базы для реализации схемы (9) используется ПЭВМ с процессором Intel Core i5, тактовая частота которого составляет 2,27 ГГц. Согласно тестам скоростных характеристик для данного процессора $t_{\text{ум}} > (25/3)\, t_{\text{сл}}$, $t_{\text{дел}} > 2{,}7\, t_{\text{ум}} > 22{,}5\, t_{\text{сл}}$, $t_{\text{сл}} = 5$ нс. С учетом этого при p разрядностью 2462 бит оценка (24) ввиду $l = 155$ позволяет заключить, что $t_{\text{РДНК, max}} < 2 \cdot 155 \cdot 2^{16}$ $(3 \cdot 155 \cdot (28/3) - (28/3) + 155 \cdot 22{,}5)5$ нс $< 13{,}252$ мин.

Время расчета денормирующего коэффициента N по схеме (9) сокращается в $\sum_{i=1}^{l}((m_i - 1)/\lfloor \log_2 m_i \rfloor) = \frac{1}{16}\sum_{i=1}^{l}(m_i - 1)$ раз, если для поиска $\widetilde{f} \in \mathbf{Z}_m$ вместо простого перебора элементов последовательности 1, 2, ..., $m - 1$ применить процедуру направленного поиска, основанную на делении отрезков, начиная с [1; $m - 1$] пополам с последующим переходом к одному из получаемых отрезков.

В настоящее время для вычисления коэффициента N обычно используют алгоритмы умножения по модулю p, основанные на тех или иных процедурах деления [5—9], в частности, на процедурах общего деления, реализующих рекурсивную схему спуска Ферма. В сравнении с данным подходом предложенный мультипликативно-субстрактивный метод позволяет снизить число необходимых модульных операций в процессе денормировки экспонент в криптографических RSA-преобразованиях с $O(l^2)$ до $O(l)$, т.е., как минимум в l раз. В случае 2462-битовых p достигается не менее чем 155-кратное уменьшение временных затрат.

Заключение

Основные результаты представленной разработки по проблеме вычисления денормирующего коэффициента для криптографических RSA-преобразований в модулярном коде состоят в нижеследующем.

1. Изложены базовые теоретические положения аппарата интервально-модулярной формы чисел, который представляет собой эффективный инструментарий для выполнения немодульных операций, используемых в рамках предлагаемого решения поставленной задачи вычисления денормирующего коэффициента для процедуры возведения в степень по системному модулю. Ключевыми составляющими применяемого инструментария являются интервально-индексные характеристики как отдельных исходных данных, так и результатов оперирования на диапазонах больших чисел в интервально-модулярном коде, получаемые по упрощенным схемам.

2. Дано теоретическое обоснование нового метода вычисления денормирующего коэффициента для криптографических RSA-преобразований, реализуемых с применением минимально избыточной МА. Основой предложенного метода служат интервально-модулярные формы чисел, позволяющие с помощью интервально-индексных характеристик, удовлетворяющих условию минимально избыточного модулярного кодирования, определять знаки целых чисел по упрощенной процедуре.

3. На базе созданного мультипликативно-субстрактивного метода вычисления денормирующего коэффициента для криптографических RSA-преобразований синтезирован эффективный алгоритм рекурсивного типа. На каждой итерации реализуемой вычислительной схемы в интервально-моду-

лярном коде выполняются операция умножения аккумулируемого произведения оснований МСС на очередное основание и серия вычитаний модуля криптосистемы из полученного произведения в целях приведения его к остатку по этому модулю. На множестве (1024—2048)-битовых модулей криптосистемы время работы предложенного алгоритма на ПЭВМ с четырехъядерным процессором Intel Core i5 (тактовая частота 2,27 ГГц) находится в секундном диапазоне.

Список литературы

1. **Харин Ю. С., Берник В. И., Матвеев Г. В.** и др. Математические и компьютерные основы криптологии. Минск: Новое знание, 2003. 382 с.
2. **Червяков Н. И., Евдокимов А.А., Галушкин А. И., Лавриненко И. Н., Лавриненко А. В.** Применение искусственных нейронных сетей и системы остаточных классов в криптографии. М.: Физматлит, 2012. 280 с.
3. **Каленик А. Н., Коляда А. А., Коляда Н. А., Протько Т. Г., Шабинская Е. В.** Компьютерно-арифметическая и реализационная база быстрых процедур умножения по большим модулям на основе модифицированной модулярной схемы Монтгомери // Электроника инфо. 2012. № 7. С. 114—118.
4. **Коляда А. А., Чернявский А. Ф., Шабинская Е. В.** Генерирование и функционально-структурная оптимизация базового комплекта таблиц для мультипликативной МИМА-схемы Монтгомери // Электроника инфо. 2013. № 4. С. 35—41.
5. **Kawamura S.** Cox-Rower architecture for fast parallel Montgomery multiplication // Eurocrypt 2000, LNCS. 2000. Vol. 1807. Berlin. P. 523—538.
6. **Alia G.** Fast modular exponentiation of large number with large exponents // J. Syst. Archit. 2002. Vol. 47, N. 14—15. P. 1079—1088.
7. **Bajard J.-C.** A Full RNS Implementation of RSA // IEEE Trans. Comp. 2004. Vol. 53, N. 6. P. 769—774.
8. **Hiasat A. A.** Semi-custo VLSI Design and Implimentation of are new efficient RNS division algorithm // Comput. I. 1999. Vol. 42, N. 3. P. 232—240.
9. **Talahmeh S.** Arithmetic Division in RNS Using Galois Field GF (p) // Comput. Artc. Math. Appl. 2000. Vol. 39, N. 5—6. P. 227—238.
10. **Нейронная** сеть основного деления модулярных чисел: А. с. № 2400813 РФ, МПК G06H3/02, G06F7/72 (2006.01). СВИС М-ва Обороны. Рос. Федерации / Червяков Н. И., Лавриненко И. Н., Лавриненко А. В., Головко А. Н. № 28150458/09, заявл. 22.12.2008., опубл. 27.09.2010.
11. **Коляда А. А.** Умножение по большому модулю в минимально избыточной модулярной системе счисления с применением операций масштабирования // Информатика. 2009. № 4. С. 49—65.
12. **Инютин С. А.** Основы модулярной алгоритмики. Ханты-Мансийск: Полиграфист, 2009. 347 с.
13. **Каленик А. Н., Коляда А. А., Коляда Н. А., Чернявский А. Ф., Шабинская Е. В.** Умножение и возведение в степень по большим модулям с использованием минимально избыточной модулярной арифметики // Информационные технологии. 2012. № 4. С. 37—44.
14. **Коляда А. А., Пак И. Т.** Модулярные структуры конвейерной обработки цифровой информации. Минск: Университетское, 1992. 256 с.
15. **Коляда А. А., Чернявский А. Ф.** Интегрально-характеристическая база модулярных систем счисления // Информатика. 2013. № 1. С. 106—119.
16. **Коляда А. А., Чернявский А. Ф.** Интервально-индексный метод четного модуля для расчета интегральных характеристик кода неизбыточной МСС с симметричным диапазоном // Доклады НАН Беларуси. 2013. Т. 57, № 1. С. 38—45.
17. **Чернявский А. Ф., Коляда А. А.** Вычисление интегральных характеристик минимально избыточного модулярного кода // Доклады НАН Беларуси. 2015. Т. 59, № 6. С. 40—46.

A. A. Kolyada, Associate Professor, Chief Researcher, e-mail: razan@tut.by,
N. A. Kolyada, Researcher, e-mail: razan@tut.by,
S. Yu. Protasenia, Junior Researcher, e-mail: Estellita@mail.ru,
E. V. Shabinskaya, Associate Professor, Senior Researcher, e-mail: shabinskaya@rambler.ru,
Research Establishment "Institute of Applied Physics Problems of A. N. Sevchenko"
Belarusian State University, Minsk, Belarus

Multiplicative-Subtractive Method of Calculating of Denormalization Factor for the Cryptographic RSA-transformation in the Modular Code

Article is devoted to a problem of a denormalization of basic transformations in the RSA cryptosystem with minimum excess modular code organization. To solve this problem we propose a new method, based on the multiplicative-subtractive calculation circuit recursive. The theoretical base of the applied approach is made by the device of interval and modular forms of numbers and interval and index characteristics. The used tools allow to reach essential simplification of the not modular operations which are a part of the synthesized procedure of calculation of denormalization coefficient for cryptographic RSA transformations. These operations include the operations of multiplication on the modular base of the number system and bringing the accumulated works of the bases to the residue modulo cryptosystem. On the set of cryptographic modules 1024—2048 digit bit time the proposed procedure on the PC with the Intel Core i5 processor (frequency of 2,27 GHz) is in the second range.

Keywords: *RSA cryptosystem, RSA cryptographic transform, residue number system, interval-index characteristics, the minimum redundant modular arithmetic, Montgomery multiplication*

References

1. **Harin Ju. S., Bernik V. I., Matveev G. V.** i dr. *Matematicheskie i kom-p'juternye osnovy kriptologii* (Mathematical foundations of cryptology and computer). Minsk: Novoe znanie, 2003. 382 p.
2. **Chervjakov N. I., Evdokimov A. A., Galushkin A. A., Lavrinenko I. N., Lavrinenko A. V.** *Primenenie iskusstvennyh nejronnyh setej i sistemy ostatochnyh klassov v kriptografii* (Application of artificial neural networks and residue number system in Cryptography). Moscow, Fizmatlit, 2012. 280 p.

3. **Kalenik A. N., Koljada A. A., Koljada N. A., Prot'ko T. G., Shabinskaja E. V.** Komp'juterno-arifmeticheskaja i realizacionnaja baza bystryh proce-dur umnozhenija po bol'shim moduljam na osnove modificirovannoj modu-ljarnoj shemy Montgomeri (The computer-arithmetic and realizable base fast multiplication of procedures for large modules on the basis of a modified modular Shem of Montgomery), *Jelektronika info*, 2012, no. 7, p. 114—118.

4. **Koljada A. A., Chernjavskij A. F., Shabinskaja E. V.** Generirovanie i funkcional'no-strukturnaja optimizacija bazovogo komplekta tablic dlja mul'tiplikativnoj MIMA-shemy Montgomeri (Generation and functional and structural optimization of the basic set of tables for the multiplicative IMSI shem of Montgomery), *Jelektronika info*, 2013, no. 4, pp. 35—41.

5. **Kawamura S.** Cox-Rower architecture for fast parallel Montgomery multiplication, *Eurocrypt* 2000, LNCS, 2000, vol. 1807. Berlin, pp. 523—538.

6. **Alia G.** Fast modular exponentiation of large number with large exponents, *J. Syst. Archit*, 2002, vol. 47, no. 14—15, pp. 1079—1088.

7. **Bajard J.-C.** A Full RNS Implementation of RSA, *IEEE Trans. Comp.*, 2004, vol. 53, no. 6, pp. 769—774.

8. **Hiasat A. A.** Semi-custo VLSI Design and Implimentation of are new efficient RNS division algorithm, *Comput. 1*, 1999, vol. 42, no. 3, pp. 232—240.

9. **Talahmeh S.** Arithmetic Division in RNS Using Galois Field GF (p), *Comput. Artc. Math. Appl.*, 2000, vol. 39, no. 5—6, pp. 227—238.

10. **Nejronnaya set'** *osnovnogo deleniya modulyarnyh chisel:* A. s. no. 2400813 Ros. Federaciya, MPK G06H3/02, G06F7/72 (2006.01). SVIS M-va Oborony. Ros. Federacii / Chervyakov N. I., Lavrinenko I. N., Lavrinenko A. V., Golovko A. N., № 28150458/09, zayavl. 22.12.2008., opubl. 27.09.2010.

11. **Kolyada A. A.** Umnozhenie po bol'shomu modulyu v minimal'no izby-tochnoj modulyarnoj sisteme schisleniya s primeneniem operacij masshtabi-rovaniya, *Informatika*, 2009, no. 4, pp. 49—65.

12. **Inyutin S. A.** *Osnovy modulyarnoj algoritmiki.* Hanty-Mansijsk, Poligrafist, 2009. 347 p.

13. **Kalenik A. N., Koljada A. A., Koljada N. A., Chernjavskij A. F., Shabinskaja E. V.** Umnozhenie i vozvedenie v stepen' po bol'shim moduljam s ispol'zovaniem minimal'no izbytochnoj moduljarnoj arifmetiki (The multiplication and exponentiation over large modules using the minimum excess modular arithmetic), *Informacionnye tehnologii*, 2012, no. 4, pp. 37—44.

14. **Koljada A. A., Pak I. T.** *Moduljarnye struktury konvejernoj obrabotki cifrovoj informacii* (Modular structures of conveyer processing of digital information). Minsk: Universitetskoe, 1992. 256 p.

15. **Koljada A. A., Chernjavskij A. F.** Integral'no-harakteristicheskaja baza moduljarnyh sistem schislenija (Integrated-characteristic modular base number systems), *Informatika*, 2013, no. 1, pp. 106—119.

16. **Koljada A. A., Chernjavskij A. F.** Interval'no-indeksnyj metod chetnogo modulja dlja rascheta integral'nyh harakteristik koda neizbytochnoj MSS s simmetrichnym diapazonom (Interval-index method is even a module for the calculation of the integral characteristics of non-redundant MSS with a symmetrical range of code), *Doklady NAN Belarusi*, 2013, vol. 57, no. 1, pp. 38—45.

17. **Chernjavskij A. F., Koljada A. A.** Vychislenie integral'nyh harakteristik minimal'no izbytochnogo moduljarnogo koda (Calculation of the integral characteristics of minimally redundant modular code), *Doklady NAN Belarusi*, 2015, vol. 59, no. 6, pp. 40—46.

УДК 004.056.53 DOI:10.17587/it.23.142-150

К. А. Щеглов, аспирант, **А. Ю. Щеглов,** д-р техн. наук, проф., e-mail: info@npp-itb.spb.ru,
Исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург

Сложность реализации угрозы безопасности и математическая модель потенциального нарушителя

Исследованы вопросы построения математической модели потенциального нарушителя в целях определения значений вероятности реализации им реальных угроз атак и определения актуальных угроз атак на защищаемую информационную систему с возможностью количественной оценки меры их актуальности. Предложенный подход к моделированию не требует использования каких-либо экспертных оценок. Введены характеристики безопасности: сложность реализации угрозы безопасности, вероятность реализации потенциальным нарушителем реальной угрозы безопасности, опасность потенциального нарушителя для информационной системы, при этом характеристика сложности реализации угрозы безопасности, позволяющая оценивать вероятность реализации потенциальным нарушителем реальной угрозы безопасности (угроз уязвимостей и угроз атак), определена как мера неопределенности для потенциального нарушителя присутствия в системе реальной угрозы безопасности и рассчитывается как частная энтропия — мера неожиданности возникновения в информационной системе реальной угрозы безопасности. Разработан подход к построению математической модели потенциального нарушителя.

Ключевые слова: *угроза атаки, угроза уязвимостей, сложность реализации угрозы, математическое моделирование, потенциальный нарушитель, мера неопределенности, частная энтропия*

Введение

Построение модели нарушителя в целях оценки готовности нарушителя реализовать ту или иную угрозу атаки (осуществить атаку на защищаемую информационную систему) является ключевой задачей при проектировании защищенной информационной системы, поскольку решение именно этой задачи моделирования позволяет определить актуальные угрозы атак на защищаемую информационную систему с учетом ценности для потенциального нарушителя обрабатываемой в ней информации,

в отношении которых следует реализовывать защиту. Сегодня практикуется построение неформальных моделей потенциального нарушителя, при этом в общем случае модель нарушителя определяет [1]:

- категории (типы) нарушителей, которые могут воздействовать на объект;
- цели, которые могут преследовать нарушители каждой категории, возможный количественный состав, используемые инструменты, принадлежности, оснащение и т.д.;
- типовые сценарии возможных действий нарушителей, описывающие последовательность (алгоритм) действий групп и отдельных нарушителей, способы их действий на каждом этапе.

Построение и последующее использование неформальной модели потенциального нарушителя не позволяет ввести какие-либо количественные метрики его опасности для информационной системы и, как следствие, — каким-либо образом учесть это при математическом моделировании. Это приводит к возможности определения актуальных для защищаемой информационной системы атак исключительно экспертным путем и к невозможности получения какой-либо объективной количественной оценки уровня актуальности той или иной угрозы атаки.

В работе [2] был предложен подход к построению математических моделей реализации угрозы атаки потенциальным нарушителем, где потенциальный нарушитель характеризовался вероятностью готовности к реализации угрозы атаки. Данные модели позволили ввести количественные характеристики (метрики) актуальности угроз атак. Однако в работе [2] не было определено то, как моделировать и рассчитывать вероятность готовности потенциального нарушителя к реализации угрозы атаки. Естественно, что задание ее значения при проектировании защищенной информационной системы экспертным путем сведет на нет все преимущества использования математического моделирования.

В данной работе рассмотрим подход к математическому моделированию потенциального нарушителя для расчета вероятности готовности потенциального нарушителя к реализации угрозы атаки. Требованием к его разработке является универсальность, т.е. возможность использования единой метрики (количественной оценки) для разнородных угроз атак и возможность определения значений требуемой характеристики без использования каких-либо экспертных оценок.

Характеристика безопасности "сложность реализации угрозы"

Говоря о моделировании потенциального нарушителя, имеет смысл обратить внимание на следующее. В современных условиях, когда существует нелегальный рынок инструментальных средств реализации атак на информационные системы и нелегальных предоставляемых услуг по реализации подобных атак, говорить о неформальной модели нарушителя в рассматриваемом ее виде, с учетом квалификации потенциального нарушителя, наличия у него соответствующих средств и т.д. особого смысла не имеет.

Сегодня, на наш взгляд, можно говорить о следующей логической модели нарушителя — существует только два параметра потенциального нарушителя: мера ценности для потенциального нарушителя обрабатываемой в информационной системе информации и мера сложности реализации той или иной атаки на эту информационную систему (в том числе выражаемая и в финансовом эквиваленте, при условии использования соответствующих нелегальных средств и/или услуг). Характеристикой же опасности для информационной системы потенциального нарушителя при этом будет его готовность к реализации атаки той или иной сложности на конкретную информационную систему, в которой обрабатывается информация, представляющая соответствующую ценность для нарушителя.

Заметим, что техническая реализуемость атак на угрозы уязвимостей учитывается при определении интенсивности их возникновения в информационной системе — статистика определяется только в отношении тех уязвимостей, для реализации которых были созданы (в случае необходимости) соответствующие эксплоиты, при этом эксплоиты, как правило, создаются нарушителями до "опубликования" сведений (до того, как они станут известны) о возникновении уязвимости [3].

Введем характеристику безопасности "сложность реализации угрозы", которую обозначим через S, соответственно "сложность реализации угрозы уязвимостей" через S_y, "сложность реализации угрозы атаки", создаваемой соответствующей совокупностью угроз уязвимостей, через S_a.

Сформулируем основные требования к моделированию характеристики "сложность реализации угрозы", которые достаточно очевидны [4]:

- должна определяться количественно, без использования каких-либо экспертных оценок, для возможности ее последующего использования при моделировании эксплуатационных параметров и характеристик безопасности;
- должна нелинейно зависеть от вероятности возникновения этой реальной угрозы при значительном опережении ее роста в области высокой готовности информационной системы к безопасной эксплуатации в отношении угрозы;
- должна быть универсальной, применимой к оценке разнородных угроз, что позволит использовать для них единую меру — единую шкалу сложности реализации;
- сложность реализации угрозы атаки должна определяться суммарной сложностью реализации создающих ее угроз уязвимостей.

В современных условиях моделирования потенциального нарушителя, с учетом сказанного ранее,

в предположении о технической реализуемости угрозы атаки существуют лишь два параметра безопасности, определяющие сложность реализации угрозы атаки (в определенном смысле — в равной мере) — интенсивность возникновения и продолжительность устранения реальной угрозы атаки (соответственно уязвимости).

В равной мере, так как высокая интенсивность возникновения реальных угроз атак одного и того же типа позволяет делать потенциальному нарушителю прогнозы о появлении в скором времени соответствующей реальной угрозы атаки в информационной системе. Как следствие, это дает ему возможность заранее подготовиться к реализации угрозы именно этой атаки, в том числе предварительно получить всю необходимую для этого информацию об информационной системе. Большая же продолжительность устранения позволяет реализовать реальную угрозу соответствующей атаки без сколько-нибудь значимой предварительной подготовки — пусть подобная реальная угроза атаки редка, но она позволяет реализовать все необходимые подготовительные этапы ее реализации уже при существовании этой реальной угрозы атаки.

По сути, оба этих параметра безопасности информации учитываются одной характеристикой — долей времени присутствия реальной угрозы атаки в информационной системе, определяемой с использованием марковских моделей вероятностью готовности системы к безопасной эксплуатации в отношении угрозы атаки [3].

Как следствие, можем говорить о том, что $S_{\text{у}} = f(P_{0\text{у}})$, $S_{\text{а}} = f(P_{0\text{а}})$, где $P_{0\text{у}}$ и $P_{0\text{а}}$ — это надежностные вероятностные характеристики безопасности соответственно угрозы уязвимостей и угрозы атаки [5].

Под характеристикой безопасности "сложность реализации угрозы атаки" будем понимать меру неопределенности (неожиданности или, соответственно, ожидаемости) для потенциального нарушителя присутствия в системе реальной угрозы этой атаки. Аналогичным образом определяется и "сложность реализации угрозы уязвимостей".

Мера неопределенности исхода некоторого события X (опыта) в теории информации характеризуется энтропией (информационной энтропией), которая определена К. Шенноном следующим образом. При n возможных исходах случайного события, определяемых вероятностями каждого исхода события P_i, $i = 1, ..., n$, энтропия, или неопределенность исхода события $H(X)$, вычисляется по формуле [7]:

$$H(X) = -\sum_{i=1}^{n} P_i \log P_i.$$

Шеннон предположил, что прирост информации равен утраченной неопределенности, и задал требования к ее измерению [8]:

- мера должна быть непрерывной, т.е. изменение значения вероятности на малую величину должно вызывать малое результирующее изменение функции;
- в случае, когда все варианты равновероятны, увеличение числа вариантов должно всегда увеличивать значение функции;
- должна быть возможность сделать выбор в два шага, в которых значение функции конечного результата должно являться суммой значений функций промежуточных результатов.

При этом Шеннон показал, что единственная функция, удовлетворяющая этим требованиям, имеет приведенный выше вид (вид логарифмической функции).

Поскольку в нашем случае два исхода опыта (угроза атаки реальна или нет), то используется логарифм по основанию 2:

$$H(X) = -(P_i \log_2 P_i + (1 - P_i)\log_2(1 - P_i)).$$

Энтропия (или средняя энтропия) $H(X)$ интерпретируется как математическое ожидание частных энтропий $I(X)$:

$$H(X) = M[I(X)],$$

где частная энтропия $I(X)$ при двух исходах случайного события определяется следующим образом:

$$I(X) = -\log_2 P_i.$$

Содержательный смысл частной энтропии — это мера неожиданности появления одного события. Чем меньше вероятность события, тем выше его частная энтропия, т.е. тем неожиданней появление события. Если средняя энтропия (или просто энтропия) отражает неопределенность всей системы и рассчитывается как среднее значение (математическое ожидание) частных энтропий всех ее состояний, то частная энтропия отражает неопределенность именно одного отдельного состояния системы.

В нашем случае неопределенностью или неожиданностью для потенциального нарушителя является событие присутствия в системе реальной угрозы безопасности — реальной угрозы уязвимостей, вероятность присутствия которой в системе определяется как $1 - P_{0\text{у}}$, либо соответственно реальной угрозы атаки, вероятность присутствия которой в системе определяется как $1 - P_{0\text{а}}$.

С учетом сказанного сложность реализации угрозы, как мера неопределенности или неожиданности для потенциального нарушителя присутствия в системе реальной угрозы безопасности (соответственно угрозы уязвимостей и угрозы атаки), может быть определена следующим образом:

$$S_{\text{у}} = f(P_{0\text{у}}) = I(P_{0\text{у}}) = -\log_2(1 - P_{0\text{у}});$$

$$S_{\text{а}} = f(P_{0\text{а}}) = I(P_{0\text{а}}) = -\log_2(1 - P_{0\text{а}}).$$

Рассмотрим, насколько, при определении подобным образом через частную энтропию характерис-

тики "сложность реализации угрозы", выполняются сформулированные к ней ранее требования.

Характеристика сложности реализации угрозы определяется количественно, что обусловливает возможность ее последующего использования при моделировании эксплуатационных характеристик безопасности информации.

Характеристика сложности реализации угрозы универсальна, так как применима к оценке разнородных угроз безопасности. Это позволяет использовать для них единую меру — шкалу сложности реализации. При оценке сложности реализации используются надежностные параметры безопасности угроз уязвимостей — интенсивности возникновения и устранения в информационной системе подобных угроз, т.е. не требуется получения какой-либо экспертной оценки.

Проанализируем вид функции $S_y = f(P_{0y}) = -\log_2(1 - P_{0y})$, проиллюстрированной на рис. 1. Аналогичный вид имеет и функция $S_y = f(P_{0a}) = -\log_2(1 - P_{0a})$.

Как видим, сложность реализации угрозы нелинейно зависит от вероятности возникновения угрозы при значительном опережении ее роста в области высокой готовности информационной системы к безопасной эксплуатации в отношении этой угрозы, причем при $P_{0y} = 0$ $S = 0$, а при $P_{0y} \to 1$ $S_y \to \infty$.

Единица сложности реализации угрозы $S_y = I(P_{0y}) = 1$, $S_a = I(P_{0a}) = 1$, задается условием $P_{0y} = 0{,}5$ (рис. 1) и $P_{0a} = 0{,}5$, определяющим равную вероятность событий (максимальное значение энтропии) — реальна или нет соответствующая угроза, что характеризует максимальную неопределенность соответствующего события для потенциального нарушителя.

В качестве примера оценки данной характеристики безопасности сравним, например, сложность реализации угрозы атаки на две различные угрозы уязвимостей. Пусть для угрозы одной из них значение характеристики P_{0y} составляет 0,7, для угрозы другой — 0,99. Видим, что в первом случае $S_{y1} = 1{,}74$, во втором случае $S_{y2} = 6{,}64$, т.е. реализация успешной атаки на вторую угрозу уязвимостей для потенциального нарушителя в 3,82 раза сложнее, чем на первую, в то время как отношение значений вероятностей для рассматриваемых угроз уязвимостей реализации P_{0y} составляет всего 1,41 (как видим, присутствует и достаточно заметен при больших значениях P_{0y} опережающий рост значения характеристики сложности реализации угрозы).

Теперь определимся с тем, как моделировать характеристику "сложность реализации угрозы безопасности". При этом необходимо обеспечить выполнимость сформулированного ранее требования в части того, что характеристика сложности реализации угрозы атаки должна определяться суммарной сложностью реализации создающих ее угроз уязвимостей реализации, что соответствует требованию Шеннона к энтропии — должна быть возможность сделать выбор в два шага, в которых значение функции конечного результата должно являться суммой значений функций промежуточных результатов.

Рассмотрим две марковские модели с дискретными состояниями и непрерывным временем для угрозы уязвимостей и для угрозы атаки, создаваемой двумя угрозами уязвимостей (именно подобные модели предложено строить для моделирования надежностных параметров и характеристик угроз уязвимостей [8] и угроз атак [3]). Для простоты и наглядности иллюстраций примем, что вероятностью одновременного возникновения в системе нескольких уязвимостей как одного, так и различных (для второго случая) типов можно пренебречь [5] (рис. 2).

На рис. 2, *а* через λ, μ соответственно обозначены интенсивности возникновения и устранения уязвимостей, на рис. 2, *б* индексы при этих параметрах указывают на тип уязвимости, через S_0 и S_1 на рис. 2, *а* обозначены состояния отсутствия и существования в системе реальной уязвимости, через S_{ij} на рис. 2, *б* — состояния отсутствия и существования в системе реальной уязвимости первого (i) и второго (j) типов.

Определим вероятность пребывания систем в состоянии отказа безопасности. Для первой модели — это состояние S_1, вероятность пребывания

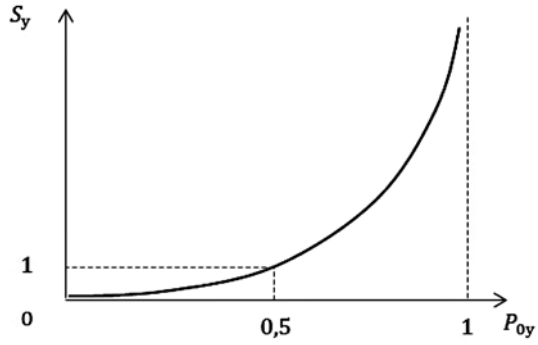


Рис. 1. Вид функции $S_y = f(P_{0y})$

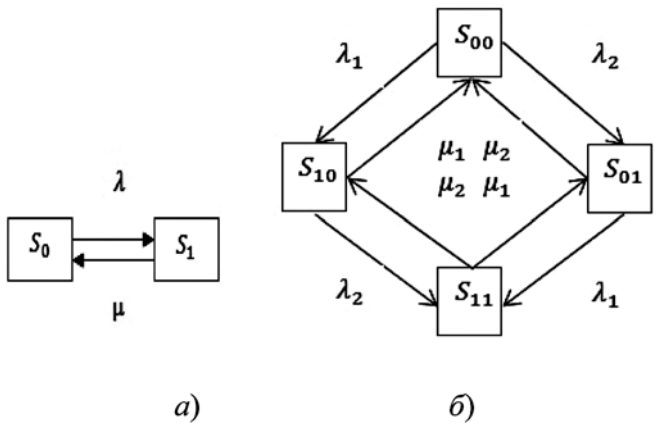


Рис. 2. Граф системы состояний случайного процесса для угроз безопасности:
а — модель угрозы уязвимостей; *б* — модель угрозы атаки

системы в котором P_1, для второй модели — это состояние S_{11} (выявлены все уязвимости, создающие угрозу атаки, — угроза атаки реальна), вероятность пребывания системы в котором P_{11}:

$$P_1 = \frac{\lambda}{\lambda + \mu};$$

$$P_{11} = \frac{\lambda_1 \lambda_2}{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2)}.$$

Замечание. Для определения P_{11} строится соответствующая система линейных алгебраических уравнений, описывающих стационарный (установившийся) режим:

$$\begin{cases} \mu_1 P_{10} + \mu_2 P_{01} = (\lambda_1 + \lambda_2) P_{00}; \\ \lambda_1 P_{00} + \mu_2 P_{11} = (\lambda_2 + \mu_1) P_{10}; \\ \lambda_2 P_{00} + \mu_1 P_{11} = (\lambda_1 + \mu_2) P_{01}; \\ \lambda_2 P_{10} + \lambda_1 P_{01} = (\mu_1 + \mu_2) P_{11}, \end{cases}$$

решаемая с учетом полной группы событий, т.е. с использованием условия

$$P_{00} + P_{01} + P_{10} + P_{11} = 1.$$

Докажем следующее ***утверждение***. Значение вероятности отказа безопасности информационной системы в отношении угрозы атаки P_a, создаваемой R угрозами уязвимостей, определяемой с использованием марковской модели угрозы атаки (см. рис. 2, *б*), равно произведению значений вероятностей отказов безопасности информационной системы в отношении угроз уязвимостей, определяемых с использованием марковской модели угроз уязвимостей (см. рис. 2, *а*), P_{yr}, т.е.

$$P_a = \prod_{r=1}^{R} P_{yr}. \qquad (1)$$

Замечание. Данное утверждение формулируется исходя из тех соображений, что события возникновения угроз уязвимостей различных типов, создающих угрозу атаки, взаимно не зависимы, что и имеет место в нашем случае.

Доказательство. Докажем это утверждение для частного случая — на примере моделей угроз безопасности, представленных на рис. 2. Для модели, представленной на рис. 2, *а*, соответствующие вероятности возникновения угроз уязвимостей каждого типа определяются следующим образом:

$$P_{y1} = \frac{\lambda_1}{\lambda_1 + \mu_1};$$

$$P_{y2} = \frac{\lambda_2}{\lambda_2 + \mu_2}.$$

Теперь определим P_a исходя из (1):

$$P_a = P_{y1} P_{y2} = \frac{\lambda_1}{\lambda_1 + \mu_1} \frac{\lambda_2}{\lambda_2 + \mu_2} = \frac{\lambda_1 \lambda_2}{(\lambda_1 + \mu_1)(\lambda_2 + \mu_2)}.$$

Как видим, получили тот же результат для P_a, как и при использовании соответствующей марковской модели угрозы атаки (см. рис. 2, *б*), создаваемой двумя угрозами уязвимостей реализации.

Докажем следующее ***утверждение***. Единственно корректным способом определения сложности реализации угрозы безопасности как меры неопределенности или неожиданности для потенциального нарушителя события присутствия в системе реальной угрозы является определение сложности реализации угрозы частной энтропией $I(X)$, которая при двух исходах случайного события рассчитывается следующим образом:

$$I(X) = -\log_2 P_i.$$

Доказывается данное утверждение с учетом предыдущего утверждения по аналогии с доказательством соответствующей теоремы Шеннона [7].

Для доказательства вернемся к требованию Шеннона к функции энтропии — должна быть возможность сделать выбор в два шага, в которых значение функции конечного результата должно являться суммой функций промежуточных результатов.

Вероятность P_{0a} того, что информационная система готова к безопасной эксплуатации в отношении угрозы атаки (как отмечали ранее, события выявления и устранения уязвимостей, создающих угрозу атаки, следует рассматривать как независимые события), создаваемой R угрозами уязвимостей реализации с соответствующими вероятностями P_{0yr} готовности системы к безопасной эксплуатации в их отношении определяется соотношением

$$P_{0a} = 1 - \prod_{r=1}^{R} (1 - P_{0yr}).$$

Исходя из того, что система готова к безопасной эксплуатации в отношении угрозы уязвимостей P_{0r}, создающих угрозу этой атаки, $r = 1, \ldots, R$, сложность реализации угрозы атаки

$$S_a = I(P_{0a}) = -\log_2(1 - P_{0a}) = -\log_2 \prod_{r=1}^{R} (1 - P_{0yr}),$$

что, используя соответствующее свойство логарифмов, можно записать следующим образом:

$$S_a = I(P_{0a}) = \sum_{r=1}^{R} S_{yr}.$$

Как видим, именно при использовании логарифмической функции для определения сложности реализации угрозы атаки существует возможность сделать выбор в два шага, при этом значение функции конечного результата будет являться суммой значений функций промежуточных результатов (основание логарифма, как ранее отмечали, определяется двумя исходами события).

Таким образом, моделирование характеристики сложности реализации угрозы атаки состоит в моделировании (с использованием соответствующих марковских моделей с дискретными состояниями и непрерывным временем) сложностей реализации угроз уязвимостей реализации, создающих эту угрозу атаки, с последующим суммированием полученных значений.

Характеристика безопасности "вероятность реализации потенциальным нарушителем реальной угрозы безопасности"

Предположим, что потенциальный нарушитель полностью готов к реализации атаки на реальную угрозу безопасности. В этих предположениях вероятность реализации атаки — реализации реальной угрозы безопасности, определяется исключительно надежностной характеристикой угрозы безопасности — вероятностью готовности информационной системы к безопасной эксплуатации в отношении этой угрозы (P_{0y}, P_{0a}).

Вероятность реализации нарушителем реальной угрозы атаки P_a и соответственно реальной угрозы уязвимостей реализации P_{ya} с учетом того, что:

$$S_y = -\log_2(1 - P_{0y});$$

$$S_a = -\log_2(1 - P_{0a}),$$

может быть определена через сложность реализации этих угроз безопасности (как количество информации в отношении присутствия в системе соответствующей угрозы безопасности, которое должен иметь (получить) потенциальный нарушитель для реализации успешной атаки) следующим образом:

$$P_a = 1/2^{S_a},\ P_{ay} = 1/2^{S_y}.$$

Вид функции $P_a = f(S_a)$ приведен на рис. 3. Аналогичный вид имеет и функция $P_{ay} = f(S_y)$.

Таким образом, с вероятностью P_a потенциальный нарушитель (при полной готовности к реализации угрозы атаки) реализует угрозу атаки сложности S_a, а с вероятностью P_{ay} — реализует атаку на реальную угрозу уязвимостей сложности S_y.

Теперь ответим на следующий вопрос. Если потенциальный нарушитель полностью готов к реализации угрозы атаки некой сложности S_{a1}, то насколько он при этом будет готов к реализации угрозы атаки большей сложности S_{a2} при условии, что $S_{a2} - S_{a1} = \Delta S_a > 0$, с какой вероятностью он реализует атаку сложности S_{a2}?

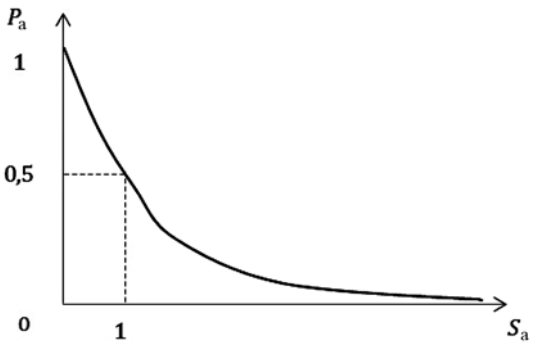


Рис. 3. Вид функции $P_a = f(S_a)$

Докажем следующее ***утверждение***. Увеличение сложности реализации угрозы атаки на $\Delta S_a = S_{a2} - S_{a1}$ приводит к снижению вероятности δP_a ее реализации потенциальным нарушителем, определяемому следующим образом:

$$\delta P_a = \log_2 \frac{P_{a1}}{P_{a2}}.$$

Доказательство. В теории информации существуют два взаимосвязанных понятия — количество информации и энтропия. Количество информации I и энтропия H характеризуют одну и ту же ситуацию, но с качественно противоположных сторон. I — это количество информации, которое требуется для снятия неопределенности H. По определению Бриллюэна информация есть отрицательная энтропия (негэнтропия) [6].

Когда неопределенность снята полностью, количество полученной информации равно изначально существовавшей неопределенности. При частичном снятии неопределенности полученное количество информации и оставшаяся неснятой неопределенность составляют в сумме исходную неопределенность $Ht + It = H$.

С учетом сказанного для рассматриваемых двух угроз атак можем записать:

$$I(P_{0a1}) = I(P_{0a2}) + \Delta I(P_{0a})$$

или то же, но с использованием характеристики сложности реализации угрозы атаки:

$$S_{a2} = S_{a1} + \Delta S_a,$$

где

$$S_{a1} = -\log_2 P_{a1},\ S_{a2} = -\log_2 P_{a2},$$

откуда:

$$\Delta S_a = S_{a2} - S_{a1} = \log_2 P_{a1} - \log_2 P_{a2} = \log_2 \frac{P_{a1}}{P_{a2}}.$$

Теперь докажем следующее ***утверждение***. При полной готовности потенциального нарушителя к реализации угрозы атаки некой сложности S_{a1} вероятность P_{a1} реализации им угрозы атаки большей сложности S_{a2} ($S_{a2} - S_{a1} = \Delta S_a > 0$) определяется следующим образом:

$$P_{a2} = 1/2^{\Delta S_a}.$$

Доказательство. Обратимся к полученному ранее равенству:

$$\Delta S_a = S_{a2} - S_{a1} = \log_2 \frac{P_{a1}}{P_{a2}},$$

из которого получаем

$$2^{\Delta S_a} = \frac{P_{a1}}{P_{a2}}.$$

Условие полной готовности потенциального нарушителя к реализации реальной угрозы атаки некой сложности S_{a1} предполагает выполнение равенства $P_{a1} = 1$, т.е. при возникновении реальной угрозы атаки подобной сложности потенциальный нарушитель ее неминуемо реализует. С учетом этого получаем, что вероятность P_{a2} реализации им реальной угрозы атаки большей сложности S_{a2} определяется выражением

$$P_{a2} = 1/2^{\Delta S_a}.$$

С учетом сказанного, в предположении, что потенциальный нарушитель полностью готов к реализации реальной угрозы уязвимостей некой сложности $S_{уг}$, вероятность $P_{ау}$ реализации им атаки на реальную угрозу уязвимостей в общем случае определяется следующим образом:

$$P_{ау} = \begin{cases} 1, \text{ если } S_у \leqslant S_{уг}; \\ 1/2^{\Delta S_у}, \text{ если } S_у > S_{уг}. \end{cases}$$

Аналогично можем записать и для вероятности реализации реальной угрозы атаки, где через $S_{аг}$ обозначена сложность реализации угрозы атаки, к реализации которой полностью готов потенциальный нарушитель:

$$P_а = \begin{cases} 1, \text{ если } S_а \leqslant S_{аг}; \\ 1/2^{\Delta S_а}, \text{ если } S_а > S_{аг}. \end{cases}$$

С учетом сказанного под эксплуатационной характеристикой безопасности "вероятность реализации потенциальным нарушителем реальной угрозы безопасности" будем понимать вероятность, определяемую сложностью реализации угрозы безопасности. Для реальной угрозы атаки эта характеристика определяется формулой

$$P_а = \begin{cases} 1, \text{ если } S_а \leqslant S_{аг}; \\ 1/2^{\Delta S_а}, \text{ если } S_а > S_{аг}, \end{cases}$$

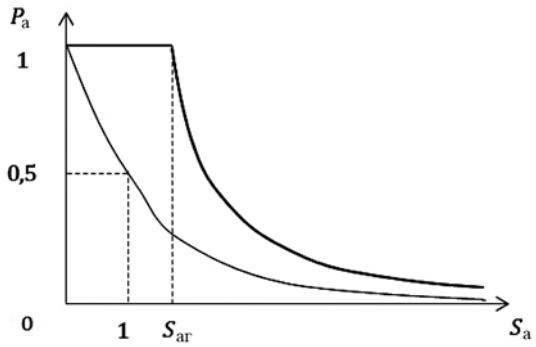


Рис. 4. Вид функции $P_а = f(S_а, S_{аг})$

а для реальной угрозы уязвимостей реализации

$$P_{ау} = \begin{cases} 1, \text{ если } S_у \leqslant S_{уг}; \\ 1/2^{\Delta S_у}, \text{ если } S_у > S_{уг}. \end{cases}$$

Вид функции $P_а = f(S_а, S_{аг})$ с учетом того, что потенциальный нарушитель полностью готов к реализации угрозы атаки сложности $S_{аг}$, приведен на рис. 4. Аналогичный вид имеет и функция $P_у = f(S_у, S_{уг})$.

Характеристика безопасности "опасность потенциального нарушителя для информационной системы". Модели потенциального нарушителя

Под опасностью потенциального нарушителя для информационной системы (для конкретной информационной системы, для которой проектируется система (как совокупность средств) защиты информации) будем понимать вероятность реализации им возникающих реальных угроз уязвимостей и реальных угроз атак определенной сложности в информационной системе, обрабатывающей информацию определенной ценности для потенциального нарушителя.

Характеристику ценности для потенциального нарушителя обрабатываемой в системе информации можно определить мерой сложности реализации угрозы атаки, которую он готов осуществить для получения несанкционированного доступа к этой информации.

Говоря же о конкретной информационной системе, для которой проектируется система защиты информации, подразумеваем, что речь идет об информационной системе, обрабатывающей определенную конкретную (содержание, объем) информацию, поскольку именно к обрабатываемой информации нарушителем и осуществляется несанкционированный доступ, т.е. информацию, обладающую определенной ценностью для потенциального нарушителя.

Развивая эту мысль, введем понятие подобной (или аналогичной) информационной системы.

Под *подобной (или аналогичной)* информационной системой, для которой проектируется система защиты информации будем понимать эксплуатируемую информационную систему, обрабатывающую подобную (содержание, объем), в идеале аналогичную, информацию, имеющую аналогичную ценность для потенциального нарушителя.

Отметим, что в той или иной мере подобная система при проектировании системы защиты конкретной информационной системы всегда может быть определена (всегда, с той или иной достоверностью, можно найти некий аналог защищаемой информационной системы).

Подобная информационная система на основании опыта ее практической эксплуатации (с использованием соответствующих средств аудита безопасности информации в процессе ее эксплуатации) может быть охарактеризована реализован-

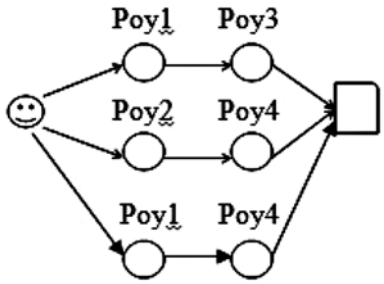


Рис. 5. Иллюстрация функциональной модели потенциального нарушителя

ными на нее атаками (успешными и нет) в соответствии с орграфом, который можно интерпретировать в качестве функциональной модели потенциального нарушителя. Иллюстрация подобного орграфа представлена на рис. 5.

Под функциональной моделью потенциального нарушителя безопасности информационной системы будем понимать орграф осуществленных атак (в том числе частично реализованных — отраженных на определенных этапах их реализации средствами защиты) на подобную информационную систему (аналог).

Орграф осуществленных атак потенциальным нарушителем на подобную информационную систему строится по полной аналогии с орграфом потенциально возможных угроз атак, представленным в [8].

Данный орграф внешне во многом будет совпадать с орграфом потенциально возможных угроз атак на конкретную информационную систему. Отличие его будет состоять в том, что последней вершиной в орграфе в функциональной модели потенциального нарушителя может быть не ресурс (объект), несанкционированный доступ к которому является целью осуществления атаки, а вершина, соответствующая последней угрозе уязвимостей, использованной нарушителем при реализации атаки.

Рассмотрим, как количественно (для последующего использования таких оценок при моделировании соответствующей характеристики безопасности и для проектирования системы защиты информации) можно описать потенциального нарушителя безопасности защищаемой информационной системы, какие с этой целью могут использоваться характеристики безопасности информации.

Пусть в процессе эксплуатации подобной (аналогичной) информационной системы зафиксировано множество осуществленных на нее атак (успешных и прерванных), каждая из которых может быть в общем случае охарактеризована соответствующей сложностью реализации угрозы атаки $\{S_{\text{ан}m},\ m = 1,\ ...,\ M\}$, а каждая m-я атака — соответствующими сложностями реализации, создающих эту атаку угроз уязвимостей реализации $\{S_{\text{ун}r_m},\ r_m = 1,\ ...,\ R_m\}$, при этом:

$$S_{\text{ан}m} = \sum_{r_m = 1}^{R_m} S_{\text{ун}r_m}.$$

Эти характеристики безопасности могут интерпретироваться в качестве математической модели потенциального нарушителя.

Под *математической моделью потенциального нарушителя безопасности* информационной системы будем понимать множество реализованных им атак на подобную информационную систему, характеризуемых сложностью реализации угрозы атаки $\{S_{\text{ан}m},\ m = 1,\ ...,\ M\}$, каждая из которых характеризуется множеством создающих угрозу атаки угроз уязвимостей, сложность реализации которых $\{S_{\text{ун}r_m},\ r_m = 1,\ ...,\ R_m\}$.

Определим на множествах значений данных характеристик безопасности, характеризующих потенциального нарушителя безопасности подобной (аналогичной) информационной системы, максимальные значения:

$$\max S_{\text{ан}} = \max\{S_{\text{ан}m},\ m = 1,\ ...,\ M\};$$

$$\max S_{\text{ун}} = \max\{S_{\text{ун}r_m},\ r_m = 1,\ ...,\ R_m,\ m = 1,\ ...,\ M\}.$$

Очевидно, что характеристики $\max S_{\text{ун}}$ и $\max S_{\text{ан}}$ соответственно определяют то, какой сложности атаку гарантированно готов реализовать потенциальный нарушитель на конкретную информационную систему — атаку на угрозу уязвимостей и атаку на угрозу атак. Значения данных характеристик безопасности можно интерпретировать как меру опасности потенциального нарушителя для конкретной информационной системы.

Под характеристикой безопасности "опасность потенциального нарушителя для информационной системы" будем понимать максимальные значения сложности реализованных им атак на угрозы уязвимостей и угрозы атак на подобную (аналогичную) информационную систему.

Рассмотрим, как можно использовать данные характеристики безопасности, полученные в результате построения математической модели потенциального нарушителя — значения $\max S_{\text{ун}}$ и $\max S_{\text{ан}}$.

Значение характеристики $\max S_{\text{ун}}$ можно принять для задания значения сложности реализации угрозы уязвимостей, к реализации которой полностью готов потенциальный нарушитель. В этом случае вероятность $P_{\text{ау}}$ реализации им атаки на реальную угрозу уязвимостей в общем случае будет определяться следующим образом:

$$P_{\text{ау}} = \begin{cases} 1, \text{ если } S_{\text{у}} \leqslant \max S_{\text{ун}}; \\ 1/2^{\Delta S_{\text{у}}}, \text{ если } S_{\text{у}} > \max S_{\text{ун}}, \end{cases}$$

где $\Delta S_{\text{у}} = S_{\text{у}} - \max S_{\text{ун}}$.

Характеристика же $\max S_{\text{ан}}$ может быть использована для определения совокупности актуальных угроз атак на защищаемую информационную систему из исходного орграфа потенциально возможных угроз атак на эту информационную систему.

Определив $\max S_{\text{ан}}$, можно говорить о том, что именно значение этой характеристики как раз и ха-

рактеризует меру ценности для потенциального нарушителя обрабатываемой в системе информации.

При этом, задав некое приращение $\Delta\max S_{\text{ан}}$, можно исключить при проектировании системы защиты информации для информационной системы угрозы атак, отнеся их к не актуальным для защищаемой информационной системы, для которых выполняется следующее условие:

$$S_{\text{а}} > \max S_{\text{ан}} + \Delta\max S_{\text{ан}}.$$

Таким образом, в результате решения рассмотренных задач моделирования может быть построен орграф актуальных угроз атак на защищаемую информационную систему, при этом может быть количественно оценена актуальность той или иной угрозы атаки, а также для каждого типа угрозы уязвимостей определено значение вероятности $P_{\text{ау}}$ реализации на нее атаки потенциальным нарушителем, что необходимо для построения математических моделей реализации реальных угроз атак потенциальным нарушителем [2]. При этом отметим, что принципиально важно — все задачи моделирования решаются без использования каких-либо экспертных оценок.

Заключение

Полученный в работе результат имеет ключевое значение для решения задачи проектирования систем защиты информации для конкретных информационных систем, характеризуемых соответствующей ценностью обрабатываемой в них информации для потенциального нарушителя. Без количественной оценки опасности потенциального нарушителя для информационной системы невозможно и количественно оценить актуальность для нее угроз атак, т.е. в принципе использовать методы математического моделирования при проектировании систем защиты информации и защищенных информационных систем, а также для оценки эффективности систем защиты информации.

Список литературы

1. **Гайкович В. Ю., Ершов Д. В.** Основы безопасности информационных технологий. М.: Изд-во МИФИ, 1995.
2. **Щеглов К. А., Щеглов А. Ю.** Эксплуатационная безопасность. Моделирование реализации угроз атак потенциальным нарушителем // Информационные технологии. 2017. № 1. С. 34—41.
3. **Щеглов К. А., Щеглов А. Ю.** Интерпретация и моделирование угрозы атаки на информационную систему. Часть 2. Моделирование угрозы атаки // Информационные технологии. 2016. Т. 22, № 1. С. 54—64.
4. **Щеглов К. А., Щеглов А. Ю.** Математические модели эксплуатационной информационной безопасности // Вопросы защиты информации. 2014. № 3. С. 52—65.
5. **Щеглов К. А., Щеглов А. Ю.** Вопросы корректности и универсальности подхода к моделированию угроз атак // Информационные технологии. 2016. Т. 22, № 11. С. 854—861.
6. **Габидулин Э. М., Пилипчук Н. И.** Лекции по теории информации. М.: Изд-во МФТИ, 2007.
7. **Шеннон К. Е.** Математическая теория связи // Работы по теории информации и кибернетике. М.: Изд-во иностр. лит. 1963. С. 243—332.
8. **Щеглов К. А., Щеглов А. Ю.** Интерпретация и моделирование угрозы атаки на информационную систему. Часть 1. Моделирование угрозы уязвимости и интерпретация угрозы атаки // Информационные технологии. 2015. Т. 21. № 12. С. 930—940.

K. A. Shcheglov, Graduate Student, **A. Yu. Shcheglov,** Professor, e-mail.ru: info@npp-itb.spb.ru,
University ITMO, St. Petersburg, Russia

Threat Implementation Complexity and Intruder Mathematical Model

We did research potential intruder mathematical model creation problems to determine the probability values of real attack threats implementation and defining actual attack threats on protected informational system, including the quantitative assessment of their actuality. The suggested modeling solution doesn't need any expert assessments usage. We introduce such security characteristics: security threat implementation complexity, rela security threat implementation possibility by potential intruder, intruder danger level towards the informational system. While this threat implementation complexity characteristic which allows to assess the real implementation threat possibility (vulnerability threats and attack threats) is defined as a uncertainty measure of real threat existence in the system for potential intruder and calculated like a private entropy — a measure of real threat sudden appearance inside the system. We suggest the approach to build a mathematical model of potential intruder.

Keywords: *attack threat, vulnerability threat, threat implementation complexity, math modeling, potential intruder, measure of uncertainty, private entropy*

References

1. **Gajkovich V. Yu., Ershov D. V.** *Osnovy bezopasnosti informatsionnykh tekhnologij*, Moscow: MIFI, 1995.
2. **Shcheglov K. A., Shcheglov A. Yu.** Ehkspluatatsionnaya bezopasnost'. Modelirovanie realizatsii ugroz atak potentsial'nym narushitelem, *Informacionnye tekhnologii*, 2017, vol. 23, no. 1, pp. 34—41.
3. **Shcheglov K. A., Shcheglov A. Yu.** Interpretatsiya i modelirovanie ugrozy ataki na informatsionnuyu sistemu. Chast' 2. Modelirovanie ugrozy ataki, *Informacionnye tekhnologii*, 2016, vol. 22, no. 1, pp. 54—64.
4. **Shcheglov K. A., Shcheglov A. Yu.** Matematicheskie modeli ehkspluatatsionnoj informatsionnoj bezopasnosti, *Voprosy zashhity informatsii*, 2014, no. 3, pp. 52—65.
5. **Shcheglov K. A., Shcheglov A. Yu.** Voprosy korrektnosti i universal'nosti podkhoda k modelirovaniyu ugroz atak, *Informacionnye tekhnologii*, 2016, vol. 22, no. 11, pp. 854—861.
6. **Gabidulin Eh. M., Pilipchuk N. I.** *Lektsii po teorii informatsii*, Moscow, MFTI, 2007.
7. **Shennon K. E.** Matematicheskaya teoriya svyazi, *Raboty po teorii informatsii i kibernetike*, Moscow, Izdatel'stvo inostrannoj literatury, 1963.
8. **Shcheglov K. A., Shcheglov A. Yu.** Interpretatsiya i modelirovanie ugrozy ataki na informatsionnuyu sistemu. Chast' 1. Modelirovanie ugrozy uyazvimosti i interpretatsiya ugrozy ataki, *Informacionnye tekhnologii*, 2015, vol. 21, no. 12, pp. 930—940.

ДИСКУССИОННЫЙ КЛУБ
DISCUSSION CLUB

УДК 004.912 DOI:10.17587/it.23.151-160

Г. В. Лосик, д-р псих. наук, гл. науч. сотр., e-mail: georgelosik@yahoo.com,
Объединенный институт проблем информатики Национальной академии наук Беларуси, г. Минск

Антропологическая информация о вариативности сообщения

Рассматривается частный случай в области теории кодирования и декодирования сообщения при его передаче от передатчика к приемнику. В этом случае в приемнике возникает дополнительная информация о сообщении, но декодируется она не из сообщения, а из физического строения самого приемника. Доказывается, что это возможно, когда строение передатчика и приемника как носителей информации — тождественно; в иных случаях возникновение указанной дополнительной информации априори невозможно.

Ключевые слова: *информация, объект вариативной формы, передатчик, приемник*

Введение

На стыке нескольких наук часто рождались ответы на вопросы, которые специалистам одной науки не удавалось решить, а поиск решения превращался по сути "толочь воду в ступе". В одной науке, информатике, статьями А. Я. Фридланда [11], К. К. Колина [7] показана критическая ситуация в дальнейшем изучении "голой" информации, канала ее передачи, исчисления объемов данных без исчисления "содержания" информационного сообщения, ставится под сомнение корректность понятий объективного и субъективного смысла в сообщении. В другой науке, психологии, статьями В. М. Аллахвердова [1] очерчена критическая ситуация в дальнейшем изучении психологии сознания, ставится вопрос причины возникновения сознания в филогенезе, роли его как когнитивного и адаптационного инструмента в жизни человека. Для выхода из критической ситуации в одной из указанных выше статей предложена гипотеза о двух компонентах информации и дается новое определение ее понятию [11, с. 77]. Суть гипотезы в предположении, что информационное сообщение содержит два компонента, в информации имеются две составляющие — *данные* и *смысл,* из которых вторая объявляется новой, не классической. Смысл сообщению придает сознание человека-индивида. Сознание является генератором смысла создаваемого индивидом сообщения и детектором смысла получаемого им сообщения. В контексте данной гипотезы становится актуальным вопрос, каким же образом в мозге кодируется смысл. Чтобы ответить на этот вопрос, в данной статье с позиции теории кодирования, теории распознавания образов рассматривается явление антропологического сходства в анатомическом строении сенсомоторной системы у человека как вида.

Понятие передатчика и приемника сообщения

В большинстве случаев сообщение передается от человека к человеку, от передатчика к приемнику в дискретном коде (например, словами), и приемнику достаточно знать правило декодирования сообщения, т.е. *алгоритм*. Человеку (компьютеру) достаточно знать информацию о языке сообщения, о смысле слов (об алгоритме декодирования) и необязательно быть физически тождественным передатчику [12]. Однако, как показано в данной работе, возможен частный случай, случай, например, двух человек, когда физическое строение передатчика и приемника как носителей информации — тождественно. Именно в этом случае возможно возникновение дополнительной информации о принимаемом сообщении.

Суть появления дополнительной информации в следующем. Если приемник сообщения имеет сходное строение с передатчиком сообщения, то приемник имеет возможность узнать дополнительную информацию о вариативности сообщения. Именно эту информацию о вариативности он может узнать не от источника сообщения, а без него, сам, из своего физического строения. Строго говоря, дополнительная информация возникает не о самом принимаемом сообщении, а лишь о законах его вариаций при многократной отправке его передатчиком. Поначалу в приемнике в его исходном состоянии этой информации о вариативности в готовом виде нет. Для возникновения дополнительной информации нужен специальный эксперимент, в самом уже приемнике, в виде специальной последовательности искажения того сообщения,

которое ранее принято как эталон. Почему такое возможно? Потому, что варианты искажения в приемнике в случае тождественности всецело повторяют варианты физической изменчивости передатчика. Поэтому приемник, приняв поначалу эталон, может указанным экспериментом не воспользоваться, а может и воспользоваться. В первом из таких случаев приемник набирает статистику об эталоне распознаваемого сообщения и о зоне его вариативности, но набирает ее, анализируя лишь поступающие от передатчика сообщения. Во втором случае от передатчика приемник получает информацию только об эталоне сообщения. Зоны его вариативности он отказывается узнавать. В приемнике формируется сенсорный эталон принимаемого сообщения. Это и есть начальная первая фаза эксперимента по вскрытию дополнительной информации. Далее во второй фазе приемник выучивается копировать сообщение, эталон которого он запомнил. Для этого у приемника должна существовать система подражания, копирования сообщений, сходных с теми, которые передавались передатчиком и которые приемник научился распознавать. После этого следует третья фаза. В ее период приемник симулирует всевозможные вариации исходного эталона. Он симулирует вариации им самим воспроизведенных сигналов путем синтеза таких вариантов эталона, которые потенциально возможны у приемника, а значит, и у передатчика. Эти вариации, заметим, совершаются уже в моторной, а не в сенсорной системе координат записи эталона. Такой эксперимент по симуляции неточностей, согласно нашей гипотезе, совершается у человека один раз. Приемник в это время не принимает внешние сигналы, а своей сенсорной системой принимает синтезированные в моторно-двигательной системе варианты эталона и обогащает сенсорные эталоны дополнительной информацией. В результате такого эксперимента приемник сам добавляет в сенсорный эталон в векторном коде информацию о его потенциальной вариативности. Дополнительная информация о передатчике появляется в приемнике не из самого передатчика, а передается приемнику как бы геномом, который сохраняет видовое постоянство организма.

Ниже мы подробно рассмотрим вопрос полезности данной дополнительной информации для повышения надежности распознавания вариативных по форме предметов. А здесь проиллюстрируем важность топологического сходства передатчика и приемника как источника дополнительной информации на примере антропологического (видового) подобия строения у всех людей слухоречевой и артикуляционной систем.

Усвоение ребенком устной речи и антропологическая информация

Рассмотрим полезность топологического сходства строения передатчика и приемника при кодировании информации на примере усвоения ребенком устной речи [2—4]. Исследованиями детской речи [8, с. 83—84] показано, что фонетические эталоны слогов устной речи (ба, ва, гу, бо) ребенок осваивает в четыре этапа, формируя эталон звучания слога, дублируя его запоминание в двух местах, двух признаковых пространствах. На первом этапе ребенок получает информацию о звучании слога в структуре слова, воспринимая его акустический сигнал, произнесенный взрослым. Ребенок узнает информацию о звучании слова, но информацию о допустимых неточностях звучания слога в слове узнать не может. Моторная программа артикуляции слова, хранящаяся в памяти взрослого, недоступна ребенку для воздействий на нее, чтобы узнать ее вариативность. Поэтому, воспринимая речь взрослых, ребенок лишен возможности услышать, например, два варианта произнесения одного и того же слова с некоторым отличием программы его артикуляции. Однако ребенок может устную речь не только воспринимать, но и воспроизводить. Для этого у него имеется артикуляторный аппарат, антропоморфный по строению у ребенка и взрослого.

На этом первом этапе в возрасте от 8 до 11 месяцев ребенок не владеет навыками артикуляции целого слова, не умеет осуществлять нужные артикуляторные движения для фонации слова, хотя лепетная речь у него уже развита. Тем не менее, в этом возрасте ребенок уже умеет узнавать простые слова на слух. На втором этапе в последующие сроки ребенок учится их произносить, повторять. У него постепенно формируется моторная программа произношения слова [3]. При этом структура моторной программы одной и той же речевой единицы у ребенка и взрослого оказывается одинаковой вследствие анатомо-физиологического изоморфизма строения их артикуляторного аппарата у первого и второго. И поэтому ребенок, чтобы узнать артикуляторные неточности произношения взрослого, может их проигрывать у себя, воспроизводя неточности в собственной программе той же речевой единицы. Это становится следующим, третьим этапом овладения словом [8, с. 89—90]. На этом этапе ребенок совершает плановые элементарные отклонения поочередно в разных звеньях указанной собственной программы. При этом он всякий раз воспринимает на слух одну за другой пару реализаций слова, произнесенных им самим. На четвертом этапе, "вычитая" первую реализацию из второй, ребенок получает и запечатлевает в сенсорной системе сведения о свойственных программе трансформациях, происходящих в его моторной системе.

Как показало проведенное нами исследование [8, с. 93—98], итерации лепетной речи служат для

ребенка тем специфическим источником, из которого он приобретает информацию о закономерностях модификации речевых сигналов взрослым. Лепетные итерации типа "ба-ба-ба", "ва-ва-ва" являются акустическими сигналами, которыми ребенок имитирует в своей речевоспроизводящей системе возможные неточности речепроизводства у взрослого. Другими словами, итерации лепета ребенка являются своеобразным видом перцептивных действий, которые слуховая система человека использует при изучении нестабильности работы речедвигательной системы.

Для этого уникального вида перцептивных действий субъект восприятия должен обязательно иметь в придачу к сенсорной системе систему синтеза тех сигналов, восприятие которых он выполняет. По данному принципу, описанному для слухоречевого восприятия, формируется образное восприятие у человека мимики и пантомимики окружающих людей, коммуникативных и выразительных жестов руки другого человека. Всякое подражание, которое совершает в действиях один человек (ученик) по отношению к другому (учителю), согласно предлагаемой модели обязательно сопровождается формированием у ученика умения не только подражать явлению, но и моделировать его неточности, чтобы константно в будущем его воспринимать.

Заметим, что указанные условия антропоморфности необходимы только для передачи в эталонный образ воспринимаемого предмета дополнительной информации о нестабильности его формы, но не самой формы. Минимальный же образ нового предмета может формироваться и без этой информации. В последнем случае восприятие предмета тоже становится возможным, но оно будет менее надежным, ибо будет нарушаться в случае, когда на вход анализатора поступает искаженный по форме предмет. Распознавание предмета в последнем случае будет осуществляться по его инвариантным признакам. Эти признаки не будут дополнены вариативными признаками и закономерности вариации окажутся не использованными.

Мы видим, что благодаря изоморфизму строения артикуляторной системы передатчика информации (взрослого) и приемника (ребенка), последний может не принимать от первого информацию о допустимых вариациях эталона. Ребенок узнает этот дополнительный компонент нужной информации об эталоне из своей собственной артикуляционной системы.

Рассмотрим в математическом плане четырехфазный процесс декодирования приемником информации о передатчике, которую последний завуалированно ему пересылает тем, что он с приемником тождественен по строению. Проиллюстрируем этот процесс на примере овладения ребенком устной речью в онтогенезе. Докажем, что информация у ребенка появляется не от взрослого, не извне, а рождается в его сенсомоторной системе в итоге срабатывания специального врожденного механизма.

Выражение математическим графом добавления антропологической информации

Проиллюстрируем этот процесс с помощью графов. Изобразим вначале набором точек некий набор сенсорных стимулов, которыми овладевает ребенок, слушая и наблюдая взрослого. Например, будем считать, что это набор слогов *ба*, *бу*, *бо* первых слов ребенка, которые от окружающих слышит младенец (рис. 1, *а*). Этими стимулами с таким же успехом может быть, например, набор указательных жестов руки взрослого, то ли набор эмоциональных экспрессий его лица, воспринимаемых ребенком зрительно.

Далее повторим изображение этих стимулов такими же точками, но в некотором пространстве теперь уже сенсорных шкал (рис. 1, *б*), в координатах которых в слуховой системе ребенка сформировались акустические среднестатистические эталоны, например, трех слогов: *ба*, *бу*, *бо*. Из фонетики известно, что гласные *а, у, о* отличаются первой F_1 и второй F_2 частотными формантами. Поэтому в слуховом пространстве ребенка, еще не умеющего их произносить, но умеющего их различать на

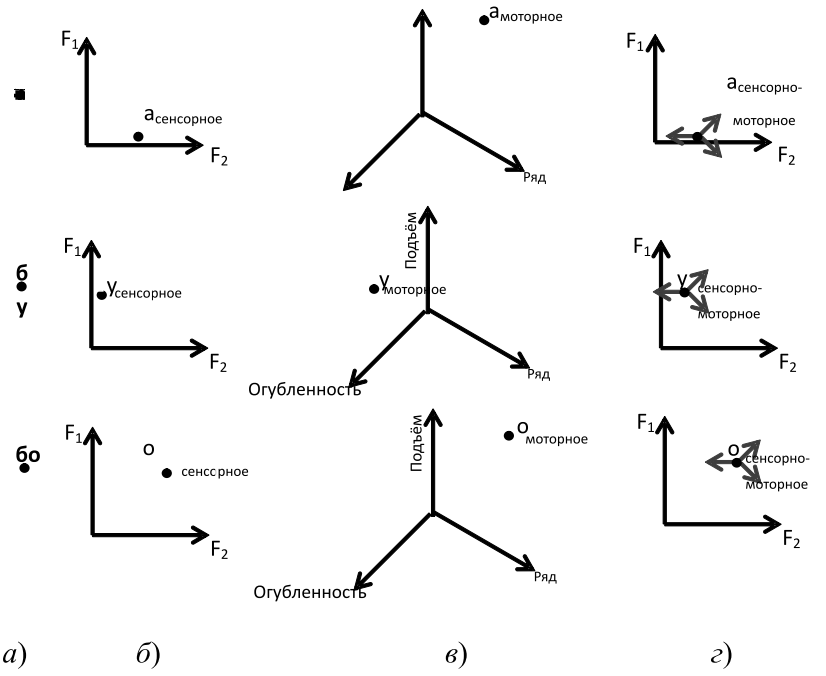


Рис. 1. Изображение точками (*а*) в координатном пространстве сенсорного анализатора (*б*) и моторного анализатора (*в*) трех фонем *а*, *у*, *о* по ходу формирования их образов в онтогенезе развития (*г*) речи ребенка

слух, данных три гласных звука, их эталоны отобразятся в системе значений сугубо двух признаков (рис. 1, *б*). Если ребенок на этом уровне развития слышит слоги с этими гласными, но с отклонениями их частотного спектра, то он их распознает сугубо в меру лишь акустической близости услышанного стимула и запомненного эталона. Это суть первый этап развития речи.

На следующем этапе развития речи ребенок овладевает произношением этих слогов. Из фонетики известно, что ребенок, уже имея сенсорные эталоны, подражает им и помещает, овладевая произношением данных звуков, их моторные эталоны в систему шкал артикуляторных признаков. Их, как известно, для гласных — три: ряд, подъем, огубленность гласной. Повторим далее изображение этих прежних стимулов точками в новом, трехмерном пространстве (рис. 1, *в*). Фонетистами доказано, что метрика акустического расстояния гласных (треугольник гласных) значимо отличается от метрики артикуляторного их расстояния [5]. После этого далее наступает третий, завуалированный информационный этап. Ребенок в возрасте 5—12 месяцев реализует в речи весьма специфическое явление — лепет. Это по сути многократные итерации одного и того же слога [8, с. 93—95]. И в итерации слогов *ба-ба-ба*, *бо-бо-бо*, *бу-бу-бу* в своей лепетной речи ребенок на своей собственной артикуляторной системе моделирует без информации от взрослого неточности артикуляции последнего. Ребенок при этом слушает внимательно свою собственную лепетную речь. Благодаря этому, у него в акустических эталонах двумерного сенсорного пространства добавляются три вектора трех возможных направлений артикуляторного искажения акустического эталона (рис. 1, *г*).

Изобразим далее системой графов возможные вариации произношения этих гласных в речи взрослого, возможные варианты путей их искажения. Ребенку в будущем, слушая взрослого, предстоит восстанавливать такие искаженные реализации в эталон. Изобразим этот переход акустической эталонной точки гласной фонемы в акустическую искаженную точку графом на рис. 2. Например, будем предполагать, что ребенок услышал цепь фонем, образующих слово "*БУЛОНКА*", которое произносит с типичной вариацией взрослый, повторяя это слово много раз.

Обозначим цепь фонем *Б*, *У*, *Л*, *О*, *Н*, *К*, *А* как цепь случайных событий, а одно из этих событий, событие реализации фонемы *О* обозначим как случайную величину *X*. Это событие повторяющееся, но не тождественно каждый раз. Ему свойственно подвергаться ряду артикуляционных трансформаций. Поэтому повторное произнесение фонемы *О* представим как ряд значений случайной величины *X*: X_1, X_2, ... X_i, ..., X_n. Каждое из событий *Б*, *У*, *Л*, *О*, *Н*, *К*, *А*, подобно событию *О*, может рассматриваться как некое случайное событие X_j, которое может из его эталонного $X_{j\ \text{эталон}}$ состояния подвергаться искажению и поэтому рождать ряд близких к $X_{j\ \text{эталон}}$ значений $X_{j1}X_{j2}X_{j3}$ ... . При этом потенциально *путей* искажений, например, у гласных *У*, *О*, *А* может быть несколько, но пока мы условно будем считать, что мы их не знаем. Это множество путей искажения может быть представлено как множество дуг графа (см. рис. 2, первый этап). Вероятность каждого из них правомерно представить как значение некоторой второй случайной величины *Y*.

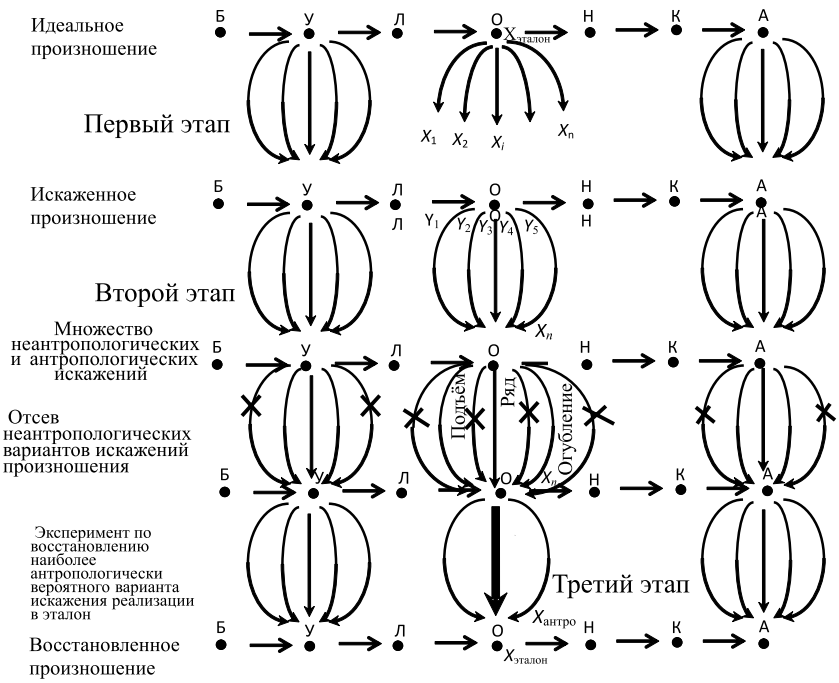


Рис. 2. Изображение графом перехода эталонных реализаций гласных фонем *У*, *О*, *А* в акустически искаженные реализации (первый этап), отсева антропологически нереальных акустических реализаций (второй этап), перехода искаженных реализаций опять в эталонные (третий этап)

В итоге мы представили точками ряд искажений гласной, которая поступает на вход приемника. Здесь ее искаженные варианты запечатляются сугубо точками в сенсорном пространстве приемника. Если приемник не повторяет антропологически строение передатчика, если не "знает" искажений, свойственных передатчику как материальному носителю, не знает его степеней свободы, то число дуг графа, принадлежащих множеству *Y*, число возможных путей, по которым $X_{\text{эталон}}$ как точка переходит в искаженное состояние $X_{\text{искажен}}$, как в иную точку, достаточно многочисленное. Это, строго говоря, множество значений второй случайной величины *Y*.

Теперь далее добавим информацию о множестве *Y* и объявим, что материальным источником, причиной искажений может быть не любой фактор, а

сугубо артикуляторный аппарат произношения слов и фраз. Антропоморфность, т.е. сходство артикуляторных систем ребенка (приемника) и говорящего взрослого (передатчика), резко уменьшает многочисленность вариантов путей возможного перехода $X_{\text{эталон}}$ в $X_{\text{искажен.}}$ Из дуг, принадлежащих множеству Y, как потенциально возможные могут быть оставлены только их часть, лишь "антропологически" возможные дуги (см. рис. 2, второй этап). И эта дополнительная информация резко облегчает ребенку при восприятии неизвестной фонемы подгонку искаженного ее стимула под эталон. Такое понижение неопределенности, что важно, ребенок в состоянии совершить сам (см. рис. 2, третий этап), лишь на основании сходства строения его речевого аппарата с речевым аппаратом взрослого.

Вероятно именно этим объясняется то обстоятельство, что неоднократные долгие эксперименты по обучению обезьян понимать на слух 300—400 слов человека не удались. Эксперименты однозначно показали, что распознавание обезьяной относительно большого числа слов человеческой речи невозможно. Это согласуется с тем, что у обезьяны голосовой аппарат принципиально иной, чем артикуляторный аппарат у человека. В то же время эксперименты показали, что распознавание жестов руки человека оказалось для обезьяны возможным. Соответственно, мы знаем, что кисть руки человека и обезьяны весьма схожи по анатомии и по нейронному механизму управления.

Вышеописанное явление, когда для восприятия речи необходимо ее проговаривание, известно как явление, лежащее в основе *моторной теории восприятия речи* [6]. Однако в ней не рассматривается столь подробно информационная сущность того, откуда рождается прибавка в надежности распознавания, когда к сенсорному описанию фонетической единицы добавляется ее моторное описание.

Представление в терминах теории вероятностей добавления антропологической информации

Мы рассмотрели в виде примера механизм добавления в приемник антропологической информации, механизм, описанный в терминах *теории графов*. Опишем этот же процесс и механизм в терминах *теории вероятностей*. Проиллюстрируем с иной математической выкладкой данный механизм завуалированного процесса добавления в приемник антропологической информации о передатчике [9, 10].

Означим сначала нулевой, классический случай, когда антропологическая информация о передатчике в расчет приемником не берется, когда заранее не совершается передача ее приемнику. В этом случае максимальное количество информации в цепи реализаций $X_{i1}, X_{i2}, ..., X_{in}$ сообщения α, которое получает сенсорная система приемника от передатчика, будет равно

$$I_{\max}(\alpha) = \sum_{X \in X_{\text{уч}}} P(\alpha) \cdot P(X/\alpha) \cdot \log[P(\alpha) \cdot P(X/\alpha)],$$

где $X_{\text{уч}}$ — учебная выборка.

Примем теперь далее в расчет случай изоморфизма строения сенсомоторной систем человека-передатчика и человека-приемника. В этом случае, как мы отмечали, правомерно ввести, кроме случайной величины X, вторую случайную величину Y. После этого выражение для вычисления максимального количества информации примет иной вид. Принятая сенсорной системой приемника в текущий момент порция информации от передатчика в виде цепи реализаций $X_{i1}, X_{i2}, ..., X_{in}$ сообщения α, будет дополнена еще одной порцией. После отсева неантропоморфных гипотез первая порция информации будет дополнена второй порцией информации от цепи реализаций переменной величины Y: $Y_1, Y_2, Y_3, ..., Y_i, ..., Y_{\overline{m}}$. Поэтому максимальное количество информации в цепи реализаций $X_{i1}, X_{i2}, ..., X_{in}$ сообщения α, которое получает сенсорная система приемника от передатчика, будет уже равно

$$I_{\max}(\alpha) = \sum_{X \in X_{\text{уч}}} P(\alpha) \cdot P(X/\alpha) \cdot \log[P(\alpha) \cdot P(X/\alpha)] + \sum_{Y \in Y_{\text{уч}}} P(\alpha) \cdot P(Y/\alpha) \cdot \log[P(\alpha) \cdot P(Y/\alpha)].$$

Итак, мы рассмотрели в виде примера добавление неучтенной антропологической информации в ходе овладения ребенком гласными фонемами в его устно-речевом общении со взрослым. Можно сделать вывод, что для возникновения в приемнике дополнительной информации о передатчике без его участия — нужны строгие условия. На приемном конце недостаточно простого физического подражания, недостаточно подражания на уровне акустического, оптического, баллистического сходства сигнала передатчика и сигнала приемника. При подражании приемника дискретным символам языка передатчика, например, фонеме, графеме, жесту, позе, недостаточно их физических сигналов. Необходимо подражание физического строения передатчика и приемника. Ибо только в этом случае информация о многообразии степеней свободы знака, о его способности реализовываться неточно может передаваться антропологическим сходством строения передатчика и приемника этой символьной информации.

Полезность информации о вариативности объекта для надежности распознавания

Воспринимающая система человека накапливает информацию об эталонах внешних явлений для последующего их распознавания. Поэтому следует рассмотреть вопрос, как затем используется дополнительная информация после появления ее в сен-

сорной системе. Ее применение, в частности, возможно для увеличения надежности распознавания новых реализаций ранее освоенных объектов, константности распознавания искаженных реализаций. Чтобы оттенить случай антропоморфности строения передатчика и приемника, сначала противопоставим его более общему случаю, когда приемник имеет дело с распознаванием вариативного по форме объекта, но не антропогенного явления, не речи, жеста, позы, которому передатчик может подражать. Такими вариативными по форме объектами восприятия являются деревья, цветы, туловища рыб, четвероногих животных, птиц, мяч, зонтик. Работа [9] специально посвящена доказательству того, что в общем плане для декодирования закономерностей вариации объекта необходимы активные перцептивные воздействия человека на этот объект. Только в условиях искусственной, инициированной человеком вариативности, воспринимающая система может совершить "декомпозицию" многообразия вариаций объекта на небольшой набор векторов элементарных вариаций. Изучение зон вариации менее эффективно, т.е. перцептивные действия нужны в любом случае, если объект значительно вариативен по форме. Но только в частном случае, если приемник может у себя создавать копию воспринятого объекта, он может совершать перцептивные действия не с самим объектом, а с копией его у себя.

Однако сначала рассмотрим более общий случай, когда объект вариативен, с ним совершаются перцептивные действия, а его распознавание ведется с учетом информации о его вариативности. Отметим, что, строго говоря, сеанс сенсорного контакта субъекта с внешним объектом передает субъекту лишь цепь фотографий. Но с точки зрения смысловой работы мозга это одновременно и цепь сообщений о *мотиве* воспринимающего субъекта относительно того, для какой цели ему полезно осваивать объект восприятия. В поступившей на вход приемника цепи "сырых" данных для предстоящего сенсорного их декодирования много неопределенности о том, какой путь (из множества сенсорно возможных путей трансформации) претерпел данный воспринимаемый объект, когда в реализации отклонился от эталона. Сохраняется неопределенность о возможных степенях его свободы отклонения, неопределенность выбора объектом конкретного пути перехода при трансформации его формы. Это связано с тем, что такие сырые данные не "антропомизированы" знанием материального строения передатчика, которое порождает неточности. Согласно нашей гипотезе, в сенсорную систему ребенка, в возникший в ней сенсорный эталон поступает информация об *антропологическом опыте* моторной системы взрослого. И эта информация уменьшает возможное число альтернатив.

Неклассический алгоритм распознавания объекта с вариативной формой

В работе [9] показано, что необходимо троекратное изучение объекта, которое позволяет определить вероятностные его характеристики $P(X/\alpha)$, $P(Y/\alpha)$, $P(\alpha)$, и в каждом из трех случаев объект изучается в разных условиях его существования. Очередность поиска $P(X/\alpha)$ и $P(Y/\alpha)$ может быть произвольной, в то время как поиск $P(\alpha)$ без учителя оказывается возможным лишь после нахождения $P(X/\alpha)$ и $P(Y/\alpha)$.

Будем считать, что система последовательно через три указанных этапа сначала обучалась распознавать ряд объектов $\alpha_1, \alpha_2, ..., \alpha_k, ..., \alpha_n$. Затем в режиме распознавания при поступлении на ее вход реализации X неизвестного объекта по этой реализации и имеющимся у системы сведениям о $P(X/\alpha_k)$ и $P(Y/\alpha_k)$ каждого из объектов α_k ($k = 1, 2, ..., n$) система вычисляет n гипотез о принадлежности X каждому из n объектов. По наиболее достоверной гипотезе принимается решение относительно объекта α_i, появившегося на входе.

Корректность этой задачи распознавания не уменьшится, если вместо группы объектов рассматривать один объект α, а его распознавание в оппозиции к объекту "не α". Если гипотеза об объекте α не подтверждается ($P(\alpha/X) < 0{,}5$), то из этого следует, что подтверждается с вероятностью $(1 - P)$ лишь обратная гипотеза в пользу объекта "не α", а не иные параллельные гипотезы о других объектах. Поэтому далее вместо вычисления n возможных гипотез будем рассматривать алгоритм вычисления гипотезы о принадлежности входной реализации X только одному объекту α (или объекту "не α").

Алгоритм принятия решения в случае вариативности объекта

Рассмотрим алгоритм принятия решения о распознавании объекта в случае, если он сильно вариативен по своей форме и если в связи с этим требуется получить дополнительную информацию о его вариативности и использовать ее. Обозначим вероятностное распределение формы изученного объекта α через $P(X/\alpha^0)$ и среднестатистическое значение его формы назовем эталоном объекта α и обозначим через α^0. Выше отмечалось, что если неизвестная новая реализация X распознаваемого объекта возникла вследствие незначительной трансформации его формы относительно α^0, то сходство X при сличении ее с α^0 обнаружится легко. В этом случае можно принимать решение о принадлежности X объекту α, не прибегая к информации о закономерностях его вариации. Для проверки гипотезы о принадлежности X объекту α достаточно использовать известную формулу Байеса следующего вида [19]:

$$P(\alpha_i/X) = P(\alpha_i) \cdot P(X/\alpha^i)/\Sigma P(\alpha_k) \cdot P(X/\alpha^k). \qquad (1)$$

Вместе с тем может оказаться, что у входной реализации не обнаруживается сходства с эталоном

объекта α и значение $P(\alpha/X)$ близко к 0,5. Поэтому ни гипотеза о принадлежности X к α, ни гипотеза о принадлежности X к "не α" не подтверждаются. Причиной тому может быть сильное несходство реализации X с эталоном α^0 вследствие значительной трансформации формы объекта α. В этом случае накопленная в анализаторе информация о "грамматике" и "лексике" антропологически возможных трансформаций объекта α становится полезной. Сведения о вариативности объекта α^0 и его частей α^ψ представлены в распределениях $P(Y/\alpha^0)$ и $P(Y/\alpha^\psi)$, где $\psi = 1, 2, \ldots$. В этом случае информация о законах вариативности может быть использована для увеличения сходства реализации X с эталоном того объекта, трансформацию которого она представляет.

Необходимость эксперимента по подгонке стимула под эталон

Возможны два пути использования информации о вариативности при сличении входной реализации и эталона. По *первому пути* система может пытаться их сблизить путем подгонки эталона под реализацию. В этом случае грамматика трансформаций с объектом, которая выявлена ранее и хранится в приемнике, совершается над объектом во внутреннем плане, т.е. интериоризованно: по этой грамматике осуществляются трансформации хранящегося в памяти эталона объекта. Воспроизведение в приемнике в интериоризованном плане не самих наблюдавшихся ранее трансформаций, а их грамматики делает приемник способным к экстраполяции всех теоретически возможных состояний объекта, которые, если не наблюдались при изучении объекта, могут встретиться после обучения.

Случай подгонки эталона под реализацию применяется, когда приемник не может создавать копию распознаваемого явления и он не антропоморфен по строению передатчику. Вместе с тем этот же эффект сближения реализации с эталоном может быть достигнут *вторым путем* — подгонкой не эталона под реализацию, а реализации под эталон. Для этого, зная грамматику трансформаций объекта вовне, анализатор совершает обратную операцию: делает попытки восстановить из трансформированной реализации X исходную эталонную реализацию α^0. С этой целью над поступившей на вход реализацией и ее частями по грамматическим правилам совершаются прежние трансформации, но не в прямом, а в обратном направлении. Интериоризованные трансформации по нормализации входного сигнала осуществляются по методу проб и ошибок поочередно в разных направлениях, пока какая-либо из них не приведет к увеличению сходства между нормализованной реализацией и эталоном одного из известных объектов.

Рассмотрим алгоритм принятия решения при распознавании входной реализации с использованием информации о вариативности воспринимаемого объекта. Будем исходить из предположения, что на этапе изучения объекта α приемник получил следующую антропологическую информацию о вариативности сообщения:

— статистическую информацию о нестабильности его формы (распределение вероятности $P(X/\alpha)$) и формы его частей до уровня ψ включительно (распределения вероятностей $P(X^\psi/\alpha)$), исходя из которых найдены среднестатистические эталоны α^0 формы объекта и его частей α^ψ;

— информацию о направлениях трансформации его формы (распределение $P(Y/\alpha)$) и ее частей (распределения $P(Y^\psi/\alpha)$), исходя из которых найдены среднестатистические значения этих направлений Y^0, Y^ψ;

— информацию о частоте $P(\alpha)$ появления объекта α.

Если при поступлении реализации X на вход не изменять эталон, т.е. не подгонять эталон под реализацию и рассчитывать степень ее сходства с эталоном по формуле (1), то информация о пути Y перехода объекта в состояние X останется не декодированной. Поэтому, чтобы по реализации X декодировать переменную Y, узнать ее значение, приемник должен устроить с поступившей на вход реализацией и эталоном как бы своеобразный эксперимент по подгонке их друг к другу.

Предположим, что выбран вариант, при котором совершается подгонка эталона под реализацию. Подгонку совершает та же переменная Y, отвечающая за преобразование эталона α^0 в реализацию X, но работающая в обратном направлении. Формула для проведения экспериментальных попыток найти случившееся преобразование α^0 в X будет следующей:

$$X^{\text{и}} = Y_X\{\Sigma\, Y_{X^1}[\Sigma\, Y_{X^2} \ldots (\Sigma\, Y_{X^\psi}(X^\psi))]\}, \qquad (2)$$

где X^ψ — части исходной реализации X неизвестного объекта, образующегося при ψ-кратном ее разбиении (при $\psi = 0$ величина X^ψ соответствует всей реализации X в целом); $X^{\text{и}}$ — искусственные реализации, образующиеся в результате экспериментальных попыток подогнать исходную реализацию X под эталон α^0 путем преобразования ее и ее частей X^ψ; Y_X — преобразования, по направлению обратные преобразованиям Y_{α^0}; Y_{X^ψ} — преобразования, по направлению обратные преобразованиям Y_{α^ψ}, распространяющиеся на части X^ψ реализации X ψ-кратного ее разбиения.

В формуле (2) известными величинами являются величина $X^{\text{и}}$, которая в итоге эксперимента должна принять значение, равное α^0, и величина X^ψ, в то время как величины $Y_X = -Y_{\alpha^0}$, $Y_{X^\psi} = -Y_{\alpha^\psi}$ оказываются неизвестными, так как могут принимать для одного и того же объекта одной и той же его

части несколько значений. Эксперимент в приемнике по подгонке α под X направлен именно на решение формулы (2). В математическом плане он выражается в том, что путем подстановки в формулу (2) разных значений Y_X и Y_{X^ψ} в разных их комбинациях находятся значения $Y_X^{\max}$ и $Y_{X^\psi}^{\max}$, т.е. те значения, которые в наилучшей степени удовлетворяют данному уравнению. Если входная реализация X является реализацией объекта α, то при последовательном переборе всех возможных трансформаций неизбежно будет обнаружена такая, которая из X образует искусственную реализацию $X^{\text{и max}}$, достаточно близкую к эталону α^0. Именно та трансформация $Y^{\max}$, при моделировании обеспечивающая превращение X в близкую к эталону реализацию $X^{\text{и max}}$, соответствует отыскиваемому значению случайной величины Y. Если входная реализация не является реализацией объекта α, то в эксперименте искусственной реализации, близкой к эталону α^0, обнаружено не будет. Приемник после такого эксперимента вправе принять решение о том, что входная реализация принадлежит объекту "не α" и приступить к новому эксперименту по подгонке реализации X к эталону иного известного ему объекта, например β.

Таким образом, лишь благодаря *эксперименту* с приходом реализации X на вход анализатора косвенно по текущему значению величины X и прошлым сведениям об антропологической вариативности объекта, хранящимся в памяти, находятся значения неявно представленной в реализации X второй величины — Y. В итоге приемник получает через реализацию X информацию не об одной, а о двух неизвестных величинах, характеризующих соответственно форму и вариативность воспринимаемого объекта. До эксперимента появление на входе описания реализации X снимает неопределенность только о форме объекта. После эксперимента и решения формулы (2) снимается еще неопределенность выбора объектом путей перехода его в состояние X.

Сведения о значениях величин $Y_X^{\max}$ и $Y_{X^\psi}^{\max}$ указывают путь обратного преобразования входной реализации X и ее частей в эталон α. Поэтому вместо вычисления гипотезы $P(\alpha/X)$ по формуле (1) появляется возможность применить приведенную ниже более точную формулу (3) для вычисления гипотезы о принадлежности реализации X объекту α. Условно можно считать, что сведения об опознаваемом объекте приемнику несут два поступающие на его вход события: X и Y, так как событие X всегда сопровождается событием Y. С учетом этого можно записать:

$$P(\alpha/X) \rightarrow \text{ЭКСПЕРИМЕНТ} \rightarrow P(\alpha/X^{\max},\ Y^{\max}) = \\ = P(\alpha)\cdot P(X^{\text{и max}},\ Y^{\text{и max}}/\alpha) = \\ = P(\alpha)\cdot P(X^{\text{и max}}/\alpha)\cdot P(Y^{\text{и max}}/\alpha). \qquad (3)$$

События X и Y по своей физической природе в отношении объекта α можно считать независимыми. В более строгом плане до появления события X у события Y имеется большой набор возможных значений, выявленных у всех изученных объектов. С приходом на вход конкретного события X после расчета по формуле (1) становится известным с определенной вероятностью объект α, соответствующий событию X. Так как событие X с некоторой вероятностью предопределяет своим значением объект α, то тем самым это значение детерминирует набор возможных значений события Y, сужает его до Y_α. Поэтому до принятия по формуле (1) промежуточного решения, событие Y можно считать зависимым от события X. Однако после промежуточного решения на этапе дальнейшего его уточнения с помощью формулы (3) события X и Y_α выступают как независимые, так как значение величины X не влияет на значение величины Y_α.

Таким образом, существенно то обстоятельство, что в формуле (3) для расчетов берется не исходное значение входной величины X, а искусственная реализация $X^{\text{и max}}$, являющаяся итогом наилучшей подгонки реализации X под эталон α^0. При этом рядом с величиной $X^{\text{и max}}$ в расчетах появляется величина $Y^{\max}$.

Таким образом, три рода информации об объекте α: $P(\alpha)$, $P(X/\alpha)$, $P(Y/\alpha)$, позволяют вычислять гипотезы о степени сходства с ним неизвестных реализаций X. Новым в таком расчете $P(Y/\alpha)$ является то, что критерий оценки сходства X и α представляется не метрикой измерения формы X и α, а метрикой измерения трансформируемости в реальной жизни X в α.

На основании рассмотрения данного феномена с позиции теории вероятности и в соответствии с работой [9] можно сделать следующие выводы.

1. Если объект восприятия отличается значительной вариативностью своей формы, по отношению к приемнику он выступает генератором не одной, а двух случайных величин. Второй из них является вид трансформации, которую избирает объект при переходе из одного состояния формы в другое. Чем больший у объекта выбор траекторий для трансформации своей формы, тем большую неопределенность для приемника представляет эта динамика и тем больше дополнительной информации может быть получено, если наряду с формой объекта будет изучена и динамика его формы.

2. Переменная величина, отражающая динамику объекта, может быть представлена векторной величиной Y. Для нахождения одного ее значения необходимо знать две скалярные величины — X_t и $X_{t+\Delta t}$, описывающие форму изучаемого объекта строго до и после искусственного воздействия на него.

3. Вычисление величины Y по обычной учебной выборке $X_{\text{уч}}$, состоящей из одинарных реализаций, принципиально невозможно. Реально возможен только искусственный способ, позволяющий по-

лучить учебную выборку $X_{уч}$, состоящую из пар реализаций для вычисления величины Y. Он выражается в переходе приемника в режим активного воздействия на объект и синхронного измерения его состояний в моменты начала и прекращения каждого нового воздействия.

4. Для вычисления степени сходства неизвестной реализации X с объектом α необходимы три вида информации: $P(\alpha)$, $P(X/\alpha)$ и $P(Y/\alpha)$. Новым видом информации в таком расчете является $P(Y/\alpha)$. При ее использовании оценка сходства X и α задается не только метрикой сходства их формы, но и метрикой трансформируемости в реальной жизни X в α.

5. Для формирования динамического образа (эталона) изучаемого объекта, если ему присуща значительная вариативность, необходимо наличие трех процессуальных условий (феноменов):

— приемник должен иметь механизм активного воздействия на объект;

— такие воздействия должны быть однокоординатными и быстротечными и порождать скачок в состоянии объекта;

— приемник должен иметь механизм приема описаний объекта в моменты до и после воздействия на него.

Обсуждение и выводы

Наряду с подражанием действиям введем понятие подражания строению. Передача геномом от человека к человеку изоморфного строения сенсомоторной системы у человека может рассматриваться как механизм "*подражания*" приемника информации повторять физическое строение, которое имеет передатчик. Благодаря сохранению антропоморфности строения сенсомоторной системы у людей как вида становится возможной передача между ними информации по аналоговому принципу кодирования, а не только по алгоритмическому, знаковому. Такая передача возможна применительно к информации о форме воспринимаемых человеком от другого человека жестов рук, артикуляций, мимики лица и пантомимики, поз.

В этом случае сохранение в филогенезе постоянства строения мозга, сенсорной и моторной систем объясняется не только с точки зрения *теории естественного отбора* Ч. Дарвина и выживания человека как телесного существа, представителя вида, но и "выживания" одного из способов передачи информации от индивида к индивиду. Изоморфность материального строения одного индивида как передатчика сообщения и второго как приемника в определенных случаях может выполнять функцию механизма шифровки и дешифровки информации о пространственно-топологических свойствах сообщения, передаваемого от передатчика к приемнику.

Знания, полученные и наделенные *смыслом* одной группой людей, закодированные устными, письменными, жестовыми символами, т.е. дискретным кодом, не должны прекратить существование. Они должны перейти из прошлого в будущее и должны быть декодированы другой группой людей. Мы приходим к выводу, что для передачи *смысла* знаний физическое строение носителя информации об этих знаниях у передатчика сообщения и приемника должно быть одинаковым. Это служит еще одним аргументом, почему в филогенезе человека как вида сохраняется постоянство строения мозга, строения тела. Одновременно появляется строгое объяснение, почему в роботостроении и "экзоскелетостроении" прослеживается тенденция "антропометрии" строения робота.

Можно сделать далеко идущий вывод о перспективах разработки систем автоматического распознавания образов и систем механического подражания человеку. При подражании фонеме, графеме, жесту, позе, т.е. дискретным символам языка передатчика, на приемном конце недостаточно подражания на уровне акустического, оптического, баллистического сходства физического сигнала фонемы, графемы, жеста, позы. Необходимо подражание материального строения передатчика и приемника. Ибо информация о многообразии степеней свободы знака реализовываться неточно не передаваема иначе, как антропологическим сходством строения передатчика и приемника этой символьной информации.

Список литературы

1. **Аллахвердов В. М.** Сознание в логике познания // Материалы пятой Международной конференции по когнитивной науке. 18—24 июня 2012, Калининград. С. 216.

2. **Бельтюков В. И.** Взаимодействие анализаторов в процессе восприятия и усвоения устной речи. М.: Педагогика, 1977. 176 с.

3. **Бельтюков В. И.** Программа овладения детьми произношением звуков речи. (К вопросу о соотношении социальных и биологических факторов) // Вопросы психологии. 1979. № 4. С. 66—78.

4. **Бельтюков В. И., Салахова А. Я.** Об усвоении ребенком звуковой (фонемной) системы языка // Вопросы психологии. 1975. № 4. С. 71—80.

5. **Бондарко Л. В.** Полезные признаки и иерархическая организация фонемной классификации. Звуковой строй языка. М., 1979. С. 20—26.

6. **Галунов В. И., Чистович Л. А.** О связи моторной теории с общей проблемой распознавания речи // Акустический журнал. 1965. Т. 2, вып. 4. С. 417—426.

7. **Колин К. К., Трошин Е. В.** Критика некоторых методологических подходов в информатике и информационное образование // Открытое образование. 2005. № 2. С. 81—89.

8. **Лосик Г. В.** Перцептивные действия в восприятии речи. Минск: Ин-т техн. кибернетики НАН Беларуси, 2000. 168 с.

9. **Лосик Г. В.** Перцептивные действия человека: кибернетический аспект. Минск: ОИПИ, 2008. 147 с.

10. **Лосик Г. В.** Кодирование информации в мозге. LapLambert Academic Publishing, 2015. 135 с.

11. **Фридланд А. Я.** О сущности информации: два подхода // Информационные технологии. 2008. № 5. С. 75—84.

12. **Шеннон К.** Работы по теории информации и кибернетике: пер. с англ. М.: Иностранная литература, 1963. 830 с.

G. V. Losik, D. Sc., General Researcher, e-mail: georgelosik@yahoo.com,
United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk

Anthropological Information about Variability Message

We consider the special case in the field of coding theory and decoding the message during its transfer from the transmitter to the receiver. It is a case when in the receiver there is additional information on the accepted message, but it is decoded not from the message, and from a physical structure of the receiver. This case is possible when the physical structure of the transmitter and receiver as the information carriers are identical. In other cases the appearance of this additional information a priori impossible.

Keywords: *information, the object of variable shape, transmitter, receiver*

References

1. **Allahverdov V. M.** Soznanie v logike poznanija, *Materialy 5-th Mezhdunarodnoj konferencii po kognitivnoj nauke,* 18—24 June 2012, Kaliningrad, pp. 216.

2. **Bel'tjukov V. I.** Vzaimodejstvie analizatorov v processe vosprijatija i usvoenija ustnoj rechi, Moscow, Pedagogika, 1977. 176 c.

3. **Bel'tjukov V. I.** Programma ovladenija det'mi proiznosheniem zvukov rechi. (K voprosu o sootnoshenii social'nyh i biologicheskih faktorov), *Voprosy psihologii,* 1979, no. 4, pp. 66—78.

4. **Bel'tjukov V. I., Salahova A. Ja.** Ob usvoenii rebenkom zvukovoj (fonemnoj) sistemy jazyka, *Voprosy psihologii,* 1975, no. 4, pp. 71—80.

5. **Bondarko L. V.** Poleznye priznaki i ierarhicheskaja organizacija fonemnoj klassifikacii, *Zvukovoj stroj jazyka,* Moscow, 1979, pp. 20—26.

6. **Galunov V. I., Chistovich L. A.** O svjazi motornoj teorii s obshhej problemoj raspoznavanija rechi, *Akusticheskij zhurnal*, 1965, vol. 2, is. 4, pp. 417—426.

7. **Kolin K. K. Troshin E. V.** Kritika nekotoryh metodologicheskih podhodov v informatike i informacionnoe obrazovanie, *Otkrytoe obrazovanie,* 2005, no. 2, pp. 81—89.

8. **Losik G. V.** *Perceptivnye dejstvija v vosprijatii rechi,* Minsk, In-t tehn. kibernetiki NAN Belarusi, 2000. 168 p.

9. **Losik G. V.** Osobennosti kodirovanija i obrabotki tekstovoj i analogovoj informacii v mozge cheloveka*, Materialy mezhdunar. konf. "RINTI-2012"*, Minsk, OIPI NAN Belarusi, 2012, pp. 211—220.

10. **Losik G. V.** Kodirovanie informacii v mozge, LapLambert Academic Publishing, 2015, 135 p.

11. **Fridland A. Ja.** O sushhnosti informacii: dva podhoda, *Informacionnye tehnologii,* 2008, no. 5, pp. 75—84.

12. **Shennon K.** *Raboty po teorii informacii i kibernetike,* per. s angl., Moscow, Inostrannaja literatura, 1963, 830 p.

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала **(499) 269-5510**

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова.*

Корректор *Е.В. Комиссарова.*

Сдано в набор 08.12.2016. Подписано в печать 25.01.2017. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT217. Цена договорная.

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

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Адвансед солюшнз". Отпечатано в ООО "Адвансед солюшнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1.