ISSN 1684–8853

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

2(63)/2013

НАУЧНЫЙ ЖУРНАЛ

2(63)/2013

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Адрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-02, e-mail: 80x@mail.ru, сайт: www.i-us.ru

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций. Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г. Перерегистрирован в Роскомнадзоре. Свидетельство о регистрации ПИ № ФС77-49181 от 30 марта 2012 г.

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

Журнал распространяется по подписке. Подписку можно оформить через редакцию, а также в любом отделении связи по каталогу «Роспечать»: № 48060 — годовой индекс, № 15385 — полугодовой индекс.

ЛР № 010292 от 18.08.98.
Сдано в набор 27.02.13. Подписано в печать 22.04.13. Формат 60×84 1/8.
Бумага офсетная. Гарнитура SchoolBookC. Печать офсетная.
Усл. печ. л. 12,1. Уч.-изд. л. 15,2. Тираж 1000 экз. Заказ 175.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.

Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.

УДК 621.396.96

ПОЛЯРИЗАЦИОННЫЕ ПРЕОБРАЗОВАНИЯ ЗОНДИРУЮЩИХ И ОТРАЖЕННЫХ СИГНАЛОВ РАДИОЧАСТОТНОЙ ИДЕНТИФИКАЦИИ

А. С. Вершинина,
магистрант
С. В. Кулаков,
доктор техн. наук, профессор
О. Д. Москалец,
канд. техн. наук, старший научный сотрудник
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Исследуются поляризационные преобразования сигналов систем радиочастотной идентификации в среде распространения и приемной антенне. Введены поляризационные спектры векторных сигналов. Метод исследования базируется на представлении поляризационных характеристик сигнала в форме вектора Джонса. Свойства среды распространения и приемной антенны, преобразующие состояние поляризации, описываются частотно-зависимой матрицей Джонса, при этом исходная матрица Джонса представлена в форме матричного ряда.

Ключевые слова *— радиочастотная идентификация, поляризационный спектр, вектор Джонса, матрица Джонса, ряд матрицы.*

Введение

Методы радиочастотной идентификации (РЧИД) находят все более широкое применение в различных сферах деятельности. За последние годы сегмент систем РЧИД оформился во вполне самостоятельную область, которую трудно отнести к какому-либо классическому разделу электроники. В качестве областей применения систем РЧИД можно отметить информационные системы, промышленное производство, автотранспорт и многое другое. Это выдвигает целый ряд задач, требующих неотложного решения [1–3], среди которых выделяются исследование поляризационных искажений принятых электромагнитных (ЭМ) сигналов и коррекция этих искажений в приемном устройстве.

Основным физическим носителем информации в современных радиоэлектронных системах (локационных, навигационных, РЧИД и др.) являются ЭМ-волны. В общем случае их электрическая $\mathbf{E}(\mathbf{r}, t)$ и магнитная $\mathbf{H}(\mathbf{r}, t)$ компоненты даются в декартовой системе координат как функции пространства $\mathbf{r} = (x, y, z)$ и времени t в форме разложения по ортам $\mathbf{i}, \mathbf{j}, \mathbf{k}$:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{i}E_x(\mathbf{r}, t) + \mathbf{j}E_y(\mathbf{r}, t) + \mathbf{k}E_z(\mathbf{r}, t);$$
$$\mathbf{H}(\mathbf{r}, t) = \mathbf{i}H_x(\mathbf{r}, t) + \mathbf{j}H_y(\mathbf{r}, t) + \mathbf{k}H_z(\mathbf{r}, t). \quad (1)$$

Векторная природа (1) ЭМ-поля требует учета не только всех временных и частотных характеристик ЭМ-сигналов — излучений, но и их поляризационных свойств. Существующие методы извлечения информации, переносимой ЭМ-волной, в большинстве основаны на анализе ее энергетических характеристик и в значительной мере исчерпали свои возможности. Дополнительную информацию, заключенную в поляризационных характеристиках ЭМ-волны, можно получить, применив векторную процедуру обработки принимаемых сигналов или коррекцию поляризационных искажений в приемном устройстве. Эти искажения носят частотно-зависимый характер, что делает необходимым ввести векторную модель сигнала в форме ЭМ-волны [4] и поляризационных спектров этих ЭМ-излучений [5].

Поляризационный спектр рассматривается как наиболее общая характеристика векторного ЭМ-сигнала, из которой путем соответствующих преобразований можно получить все остальные характеристики и параметры этого сигнала, в том

числе важнейшую в настоящем рассмотрении информацию о поляризационных искажениях сигналов систем РЧИД. Эта информация может быть использована для коррекции названных искажений в приемном устройстве.

Поляризационные измерения обязаны своему появлению решению задач оптического диапазона, где был разработан ряд методов описания и измерения состояния поляризации монохроматических и квазимонохроматических оптических излучений. В дальнейшем эти методы были успешно использованы в радиоастрономии [6, 7], а несколько позже нашли широчайшее применение в радиолокации [8–10]. Проводимые в настоящей работе исследования направлены на дальнейшее развитие методов и идей поляризационных измерений вообще и применительно к решению задач устройств РЧИД в частности.

Векторная модель динамического сигнала. Поляризационные спектры

Системы РЧИД работают с импульсными радиосигналами, и временной характер электрической $\mathbf{E}(\mathbf{r}, t)$ и магнитной $\mathbf{H}(\mathbf{r}, t)$ компонент ЭМ-поля описывается финитными функциями времени. Далее полагается, что ЭМ-волны — сигналы РЧИД — являются однородными и плоскими, причем в декартовой системе координат $\mathbf{r} = (x, y, z)$ в качестве направления распространения волны выбрана координата z.

Поведение векторов $\mathbf{E}(\mathbf{r}, t)$ и $\mathbf{H}(\mathbf{r}, t)$ в среде распространения описывается уравнениями Максвелла

$$\mathbf{rotH} = \varepsilon \frac{\partial \mathbf{E}}{\partial t};\ \ \mathbf{rotE} = -\mu \frac{\partial \mathbf{H}}{\partial t}, \qquad (2)$$

где ε, μ — диэлектрическая и магнитная проницаемости среды распространения.

Из уравнений (2) следует, что $\mathbf{E}(\mathbf{r}, t)$ и $\mathbf{H}(\mathbf{r}, t)$ являются гладкими функциями пространственных координат и времени. Гладкие финитные функции являются интегрируемыми, и во временном пространстве функции $\mathbf{E}(\mathbf{r}, t)$ и $\mathbf{H}(\mathbf{r}, t)$ удовлетворяют условиям Дини, и любая скалярная компонента в разложении (1) может быть представлена в форме двойного интеграла Фурье. Так, для скалярной электрической компоненты $E(t)$ имеет место двойной интеграл Фурье:

$$E(t) = \frac{1}{2\pi} vp \int_{-\infty}^{\infty} \mathrm{d}\omega \int_{-\infty}^{\infty} E(t') \exp[i\omega(t - t')]\mathrm{d}t', \qquad (3)$$

где vp означает главное значение интеграла при интегрировании по переменной ω (что далее опускается); ω — временная угловая частота.

Из формулы (3) следует пара преобразований Фурье:

$$S(\omega) = \int_{-\infty}^{\infty} E(t) \exp(i\omega t)\mathrm{d}t; \qquad (4)$$

$$E(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \exp(i\omega t)\mathrm{d}\omega. \qquad (5)$$

Спектральные компоненты $S(\omega)$ колебания $E(t)$ распространяются в линейной среде независимо друг от друга, и поведение скалярной волны, соответствующей колебанию $E(t)$, дается суперпозицией гармонических волн бесконечно малой амплитуды. Формула (5) позволяет представить такую скалярную волну, распространяющуюся вдоль оси z, в форме

$$E(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \exp[i(\omega t - kz)]\mathrm{d}\omega, \qquad (6)$$

где $k = \omega/c$ — волновое число (c — скорость света).

Соотношение (6) пригодно для описания вертикально или горизонтально поляризованной ЭМ-волны как частного случая. В общем случае состояния поляризации ЭМ-поля требуется ввести векторную модель динамического сигнала [4].

Векторная модель сигнала предполагает [4], что в форме (6) можно представить и горизонтальную, и вертикальную компоненту плоского ЭМ-поля. Тогда векторный сигнал запишется в виде

$$\mathbf{E}(t, z) = \mathbf{i}\frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) \exp[i(\omega t - kz)]\mathrm{d}\omega +$$

$$+ \mathbf{j}\frac{1}{2\pi} \int_{-\infty}^{\infty} S_y(\omega) \exp[i(\omega t - kz)]\mathrm{d}\omega =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left\{ \left[\mathbf{i}S_x(\omega) + \mathbf{j}S_y(\omega) \right] \exp[i(\omega t - kz)] \right\} \mathrm{d}\omega. \qquad (7)$$

Подобно тому, как скалярное соотношение (6) является суперпозицией бесконечно малых скалярных колебаний, выражение (7) представляет собой суперпозицию также бесконечно малых векторных колебаний и выступает как обобщение спектрального представления скалярной волны (6). В выражении (7) полагается, что каждая пара бесконечно малых спектральных волновых компонент с угловой частотой ω:

$$S_x(\omega) \exp[i(\omega t - kz)]\mathrm{d}\omega;$$

$$S_y(\omega) \exp[i(\omega t - kz)]\mathrm{d}\omega, \qquad (8)$$

определяет индивидуальное состояние поляризации (эллиптическое, круговое или линейное). В совокупности эти компоненты составляют векторный сигнал (7) с теми или иными поляризационными особенностями — от полной поляризации до полного ее отсутствия. Бесконечное континуальное множество совокупностей (8) составляет поляризационный спектр сигнала в форме ЭМ-волны [5].

Поляризационные преобразования монохроматических волн

Однородную плоскую монохроматическую ЭМ-волну, распространяющуюся вдоль оси z:

$$\dot{\mathbf{E}}(\mathbf{r}, t) = \mathbf{i}\dot{E}_x \exp[i(\omega t - kz)] + \mathbf{j}\dot{E}_x \exp[i(\omega t - kz)], \quad (9)$$

можно представить в форме

$$\dot{\mathbf{E}}(\mathbf{r}, t) = \begin{bmatrix} \dot{E}_x \\ \dot{E}_y \end{bmatrix} \exp[i(\omega t - kz)]. \quad (10)$$

Вектор-столбец в правой части выражения (10) называется вектором Джонса [8]:

$$\mathbf{J} = \begin{bmatrix} \dot{E}_x \\ \dot{E}_y \end{bmatrix} = \begin{bmatrix} H \exp(i\varphi_x) \\ V \exp(i\varphi_y) \end{bmatrix}, \quad (11)$$

где $\dot{E}_x = H\exp(i\varphi_x)$; $\dot{E}_y = V\exp(i\varphi_y)$; φ_x, φ_y — комплексные амплитуды и начальные фазы колебаний горизонтальной и вертикальной компоненты соответственно.

Вектор Джонса является комплексным и содержит полную информацию об амплитудах и фазах составляющих электрического вектора однородной плоской монохроматической ЭМ-волны.

Переходя к поляризационным преобразованиям, отметим, что системы координат волны, падающей на поляризующую систему, и волны выходящей совпадают.

Задача преобразования поляризационных характеристик ЭМ-поля ставится следующим образом [11, 12]. Рассматриваются плоские волны вида (9) в декартовой системе координат (x, y, z), причем в качестве направления распространения волны выбирается ось z. Несколько поляризующих приборов (систем) [12], соединенных последовательно, воздействуют на проходящую плоскую волну, создавая затем выходящую плоскую волну.

Преобразование состояния поляризации базируется на таком представлении плоской волны, которое однозначно связано с ней, и на описании поляризующего прибора неким математическим оператором $\mathbf{L}$. Этот оператор предполагается линейным, а векторная природа ЭМ-поля учитывается с помощью матриц.

Чтобы получить операторную матрицу $\mathbf{L}$ системы n приборов, расположенных последовательно, необходимо перемножить n операторов [12], т. е.

$$\mathbf{L} = \prod_{n}^{1} \mathbf{L}_n. \quad (12)$$

В зависимости от того или иного матричного представления однородной плоской ЭМ-волны рассматриваются два метода преобразования состояния ее поляризации плоской ЭМ-волны: метод Мюллера и метод Джонса [11, 12]. В настоящей работе применяется метод Джонса.

В результате взаимодействия падающей волны с поляризующей системой на выходе системы появляется одна или несколько модифицированных плоских волн. В данной работе рассматривается система, включающая передающую антенну, среду распространения ЭМ-волны и приемную антенну. Таким образом, на ЭМ-волну действуют две поляризующие системы: среда распространения и приемная антенна, — поляризующие свойства каждой из которых описываются соответствующими матрицами Джонса. Влияние среды распространения на поляризационные характеристики ЭМ-волны определяются следующей матрицей Джонса [11, 12]:

$$\mathbf{I}_M = \begin{bmatrix} M_{11} M_{12} \\ M_{21} M_{22} \end{bmatrix}. \quad (13)$$

Поляризующие свойства приемной антенны в общем случае выражаются матрицей Джонса [6]

$$\mathbf{I}_A = \begin{bmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{bmatrix}, \quad (14)$$

тогда совместное поляризующее действие среды распространения и приемной антенны в соответствии с формулой (12) дается матрицей Джонса в форме произведения матриц (13) и (14):

$$\mathbf{I}_0 = \mathbf{I}_A \times \mathbf{I}_M = \begin{bmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{bmatrix} \times \begin{bmatrix} M_{11} M_{12} \\ M_{21} M_{22} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \quad (15)$$

В общем случае передающая антенна излучает ЭМ-волну, которой соответствует вектор Джонса

$$\mathbf{J}_1 = \begin{bmatrix} H_1 \exp i\varphi_{x1} \\ V_1 \exp i\varphi_{y1} \end{bmatrix}. \quad (16)$$

Соотношения (13) — (15) позволяют записать результат поляризационных преобразований ЭМ-волны в форме

$$\mathbf{J}_2 = \begin{bmatrix} H_2 \exp i\varphi_{x2} \\ V_2 \exp i\varphi_{y2} \end{bmatrix} = \begin{bmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{bmatrix} \times \begin{bmatrix} M_{11} M_{12} \\ M_{21} M_{22} \end{bmatrix} \times \times \begin{bmatrix} H_1 \exp i\varphi_{x1} \\ V_1 \exp i\varphi_{y1} \end{bmatrix} = \begin{bmatrix} B_{11} B_{12} \\ B_{21} B_{22} \end{bmatrix} \times \begin{bmatrix} H_1 \exp i\varphi_{x1} \\ V_1 \exp i\varphi_{y1} \end{bmatrix}. \quad (17)$$

Вектор-столбец $\mathbf{J}_2$ в общем виде определяет поляризационные характеристики излученной однородной плоской монохроматической ЭМ-волны, искаженные средой распространения и приемной антенной.

Поляризационные преобразования электромагнитного сигнала

Подобно представлению (11) однородной плоской монохроматической ЭМ-волны, пару беско-

нечно малых монохроматических компонент (8) можно представить в виде

$$\mathrm{d}\mathbf{E}(z,t)=\begin{bmatrix}S_x(\omega)\\ S_y(\omega)\end{bmatrix}\exp[i(\omega t-kz)]\mathrm{d}\omega, \qquad (18)$$

и векторному сигналу (7) соответствует форма

$$\mathbf{E}(z,t)=\int_{-\infty}^{\infty}\begin{bmatrix}S_x(\omega)\\ S_y(\omega)\end{bmatrix}\exp[i(\omega t-kz)]\mathrm{d}\omega, \qquad (19)$$

где

$$\left|S_x(\omega)\right|^2+\left|S_y(\omega)\right|^2=\left|\Phi(\omega)\right|^2. \qquad (20)$$

Матрица-столбец в выражении (19) является вектором Джонса для поляризационных спектров [5]:

$$\mathbf{J}_s=\begin{bmatrix}S_x(\omega)\\ S_y(\omega)\end{bmatrix}=\begin{bmatrix}\left|S_x(\omega)\right|\exp i[\varphi_x(\omega)]\\ \left|S_y(\omega)\right|\exp i[\varphi_y(\omega)]\end{bmatrix}, \qquad (21)$$

она аналогична вектору Джонса для монохроматической волны. Вектор (21), по-видимому, введен в работе [5] и в дальнейшем получил подтверждение в публикациях [8, 13].

Вектор Джонса (21) описывает состояние поляризации исходного импульсного векторного ЭМ-сигнала, поляризационные характеристики которого преобразовываются прибором, изменяющим состояние поляризации.

Как отмечалось выше, матрица Джонса $\mathbf{I}_M$ в выражении (13) характеризует свойства среды распространения для плоской монохроматической волны определенной частоты. Передаваемый сигнал является суперпозицией (7) бесконечно малых монохроматических колебаний с разными частотами, которые входят в его состав. Следовательно, можно ввести матрицу Джонса $\mathbf{I}_M(\omega)$, которая описывает поляризационные свойства среды распространения, зависящие от частоты:

$$\mathbf{I}_M(\omega)=\begin{bmatrix}M_{11}(\omega) & M_{12}(\omega)\\ M_{21}(\omega) & M_{22}(\omega)\end{bmatrix}. \qquad (22)$$

Правило суммирования матриц позволяет представить матрицу Джонса (22) в виде ряда матриц Джонса. Для этого элементы матрицы $\mathbf{I}_M(\omega)$ следует разложить в ряды Тейлора в окрестности средней частоты ω_0 спектра сигнала:

$$M_{ij}(\omega)=M_{ij}(\omega_0)+\frac{1}{1!}\frac{\mathrm{d}M_{ij}}{\mathrm{d}\omega}\bigg|_{\omega=\omega_0}(\omega-\omega_0)+$$
$$+\frac{1}{2!}\frac{\mathrm{d}^2M_{ij}}{\mathrm{d}\omega^2}\bigg|_{\omega=\omega_0}(\omega-\omega_0)^2+\ldots. \qquad (23)$$

В итоге получается сумма матриц, элементами которых являются члены ряда (23):

$$\mathbf{I}_M(\omega)=\begin{bmatrix}M_{11}(\omega_0) & M_{12}(\omega_0)\\ M_{21}(\omega_0) & M_{22}(\omega_0)\end{bmatrix}+$$
$$+\begin{bmatrix}\frac{1}{1!}\frac{\mathrm{d}M_{11}}{d\omega}\Big|_{\omega=\omega_0}(\omega-\omega_0) & \frac{1}{1!}\frac{\mathrm{d}M_{12}}{d\omega}\Big|_{\omega=\omega_0}(\omega-\omega_0)\\ \frac{1}{1!}\frac{\mathrm{d}M_{21}}{\mathrm{d}\omega}\Big|_{\omega=\omega_0}(\omega-\omega_0) & \frac{1}{1!}\frac{\mathrm{d}M_{22}}{\mathrm{d}\omega}\Big|_{\omega=\omega_0}(\omega-\omega_0)\end{bmatrix}+$$
$$+\begin{bmatrix}\frac{1}{2!}\frac{\mathrm{d}M_{11}}{\mathrm{d}\omega^2}\Big|_{\omega=\omega_0}(\omega-\omega_0)^2 & \frac{1}{2!}\frac{\mathrm{d}^2M_{12}}{\mathrm{d}\omega^2}\Big|_{\omega=\omega_0}(\omega-\omega_0)^2\\ \frac{1}{2!}\frac{\mathrm{d}^2M_{21}}{\mathrm{d}\omega^2}\Big|_{\omega=\omega_0}(\omega-\omega_0)^2 & \frac{1}{2!}\frac{\mathrm{d}^2M_{22}}{\mathrm{d}\omega^2}\Big|_{\omega=\omega_0}(\omega-\omega_0)^2\end{bmatrix}+\ldots, \qquad (24)$$

где слагаемые после первого выражают частотную зависимость поляризационных свойств среды распространения.

Согласно правилу дифференцирования матриц, оператор дифференцирования можно вынести за знак матрицы, тогда с учетом этого имеем

$$\mathbf{I}_M(\omega)=\begin{bmatrix}M_{11}(\omega_0) & M_{12}(\omega_0)\\ M_{21}(\omega_0) & M_{22}(\omega_0)\end{bmatrix}+$$
$$+\sum_{n=1}^{\infty}\frac{1}{n!}(\omega-\omega_0)^n\frac{\mathrm{d}^n}{\mathrm{d}\omega^n}\bigg|_{\omega=\omega_0}\begin{bmatrix}M_{11}(\omega) & M_{12}(\omega)\\ M_{21}(\omega) & M_{22}(\omega)\end{bmatrix}. \qquad (25)$$

Поляризационные преобразования сигналов систем РЧИД предполагают, что передающая и приемная антенны являются поляризаторами (полуволновой вибратор или штыревая антенна), ориентированными вдоль вертикальной оси. При этом падающей на поляризующую систему ЭМ-волне ставится в соответствие вектор Джонса вертикально поляризованной ЭМ-волны. С учетом соотношений (20) и (21) этот вектор имеет вид

$$\mathbf{J}_1=\begin{bmatrix}0\\ S_y(\omega)\end{bmatrix},\ \left|S_y(\omega)\right|^2=\left|\Phi(\omega)\right|^2. \qquad (26)$$

Влияние среды распространения на поляризационные характеристики ЭМ-волны определяются матрицей (22) с частотно-зависимыми элементами. Далее предполагается, что приемная антенна является идеальным поляризатором, ориентированным вдоль вертикальной оси. Поляризационные свойства такой антенны описываются следующей матрицей Джонса:

$$\mathbf{I}_A=\begin{bmatrix}00\\ 01\end{bmatrix}. \qquad (27)$$

Тогда совместные поляризационные преобразования среды распространения и приемной антенны даются матрицей Джонса в форме произведения этих матриц в соответствии с выражением (15):

$$\mathbf{I}_0=\mathbf{I}_A\times\mathbf{I}_M=\begin{bmatrix}00\\ 01\end{bmatrix}\times\begin{bmatrix}M_{11}M_{12}\\ M_{21}M_{22}\end{bmatrix}=\begin{bmatrix}0 & 0\\ M_{21} & M_{22}\end{bmatrix}. \qquad (28)$$

Состояние поляризации выходящей из поляризующей системы волны с учетом выражения (22) запишется следующим вектором Джонса:

$$\mathbf{J}_2 = \begin{bmatrix} 0 & 0 \\ M_{21}(\omega) & M_{22}(\omega) \end{bmatrix} \times \begin{bmatrix} 0 \\ S_y(\omega) \end{bmatrix} = \begin{bmatrix} 0 \\ M_{22}(\omega) S_y(\omega) \end{bmatrix}. \quad (29)$$

Соотношение (29) определяет скалярный сигнал в частотном пространстве, его комплексный спектр

$$S_0(\omega) = M_{22}(\omega) S_y(\omega) = K_0(\omega) S_y(\omega). \quad (30)$$

Комплекснозначная функция $M_{22}(\omega)$ в выражении (30) играет роль коэффициента передачи $K_0(\omega)$ некоторого четырехполюсника, учитывающего итог поляризационных искажений, вносимых средой, в которой распространяется ЭМ-волна системы РЧИД. Эти искажения могут быть скомпенсированы введением в приемном устройстве корректирующего четырехполюсника с коэффициентом передачи

$$K_K(\omega) = \frac{1}{M_{22}(\omega)} = \frac{1}{K_0(\omega)}, \quad K_0(\omega) \neq 0. \quad (31)$$

В соотношении (31) величину $M_{22}(\omega) = K_0(\omega)$ можно назвать поляризационной передаточной функцией при сформулированных условиях относительно передающей и приемной антенн.

Заключение

Исследования поляризационных преобразований сигналов опирались на векторную модель излученного сигнала, поляризационные спектры последнего и матрицу Джонса, определяющую поляризующие свойства среды распространения ЭМ-волн. Элементы этой матрицы в общем случае являются комплексными функциями частоты, а сама матрица представлена в форме разложения в ряд по матрицам Джонса. С помощью ряда матриц можно описывать поляризационные искажения. Первое слагаемое матричного ряда не учитывает частотную зависимость поляризационных искажений, эту зависимость отражают остальные члены ряда.

Исследования установили искажения спектра ортогональных компонент ЭМ-поля. В случае систем РЧИД излученный сигнал представляет собой вертикально поляризованное ЭМ-поле; его прием осуществляется антенной, предназначенной также для приема вертикально поляризованного ЭМ-излучения. В этих условиях определены искажения спектра обрабатываемого сигнала и коэффициент передачи корректирующего четырехполюсника через соответствующий элемент матрицы Джонса, определяющей частотно-зависимые поляризационные преобразования среды распространения ЭМ-излучения.

Исследования выполнены в рамках государственного контракта № 14.527.12.0019; шифр лота 2011-2.7-527-025; шифр заявки 2011-2.7.-527-025-002.

Литература

1. **Койгеров А. С., Забузов С. А., Дмитриев В. Ф.** Исследование корреляционного метода для решения задач антиколлизии для систем радиочастотной идентификации на ПАВ // Информационно-управляющие системы. 2009. № 5. С. 48–55.
2. **Марковский С. Г., Марковская Н. В.** Разрешение конфликтов в системах радиочастотной идентификации с использованием идентификаторов меток и процедуры последовательной компенсации конфликтных сигналов // Информационно-управляющие системы. 2012. № 2. С. 48–55.
3. **Марковский С. Г., Марковская Н. В.** Расчет средней задержки алгоритма решения конфликтов в системах радиочастотной идентификации // Информационно-управляющие системы. 2012. № 4. С. 84–92.
4. **Москалец О. Д.** Модель сигнала при обработке векторных стохастических полей // Всесоюз. конф. по статистическим методам обработки данных дистанционного зондирования окружающей среды, Рига, сентябрь 1986 г. / АН СССР, 1986. С. 54.
5. **Москалец О. Д.** Учет поляризационных характеристик антенн при спектральных измерениях в радиоастрономии // Антенные измерения: IV Всесоюз. конф. «Метрологическое обеспечение антенных измерений» (ВКАИ-4). Ереван, 1987. С. 45–47.
6. **Дьяков Ю. П., Шишкин И. Ф.** К вопросу об описании свойств антенны поляриметра // Радиотехника. 1968. Т. 23. № 3. С. 98–99.
7. **Есепкина Н. А.** Поляризационные характеристики антенн радиотелескопов // Изв. вузов. Радиофизика. 1971. Т. XIV. № 5. С. 673–679.
8. **Козлов А. И., Логвин А. И., Сарычев В. А.** Поляризация радиоволн. Поляризационная структура радиолокационных сигналов. — М.: Радиотехника, 2005. — 704 с.
9. **Козлов А. И., Логвин А. И., Сарычев В. А.** Поляризация радиоволн. Радиолокационная поляриметрия. — М.: Радиотехника, 2007. — 640 с.
10. **Жу Мигун, Ян Рулян, Бай Ютян** и др. Особенности построения двухчастотной поляриметрической РСА с учетом разделения поляризационных сигналов // Радиотехнические тетради. 2000. № 22. С. 15–21.
11. **Джеррард А., Берч Дж. М.** Введение в матричную оптику: пер. с англ. — М.: Мир, 1978. — 341 с.
12. **О'Нейл Э.** Введение в статистическую оптику: пер. с англ. — М.: Мир, 1966. — 254 с.
13. **Слетков В. Л.** Аналитическое представление поляризованных сигналов // Изв. вузов. Радиоэлектроника. 2006. Т. 49. № 3. С. 17–23.

УДК 004.934

МЕТОД ФОНЕТИЧЕСКОГО ДЕКОДИРОВАНИЯ СЛОВ В ИНФОРМАЦИОННОЙ МЕТРИКЕ КУЛЬБАКА — ЛЕЙБЛЕРА ДЛЯ СИСТЕМ АВТОМАТИЧЕСКОГО АНАЛИЗА И РАСПОЗНАВАНИЯ РЕЧИ С ПОВЫШЕННЫМ БЫСТРОДЕЙСТВИЕМ

В. В. Савченко,
доктор техн. наук, профессор
Нижегородский государственный лингвистический университет
А. В. Савченко,
канд. техн. наук
Национальный исследовательский университет Высшая школа экономики, г. Нижний Новгород

Предложена новая разновидность метода фонетического декодирования слов в расчете на ограниченное множество минимальных звуковых единиц типа отдельных фонем как альтернатива большинству известных методов распознавания речи, основанных на скрытых марковских моделях речевых сигналов. В ее основе используется идея многократного (на порядок и более) сжатия данных за счет того, что слова и фразы из словаря отображаются на последовательность фонетических кодов. Достигаемый эффект, подтвержденный результатами экспериментальных исследований, состоит в увеличении скорости автоматической обработки речевого сигнала при сохранении достаточной точности и надежности распознавания речи.

Ключевые слова *— автоматическое распознавание речи, распознавание образов, распознавание с обучением, критерий минимума информационного рассогласования.*

Введение

Метод фонетического декодирования слов (МФДС) предложен в работах [1, 2] со ссылкой на новый математический аппарат информационной теории восприятия речи [3] как альтернатива большинству известных методов [4–6] автоматического распознавания речи (АРР) [7] с точки зрения вычислительных затрат на реализацию в режиме реального времени. Канонический подход к АРР основывается, как известно [7–13], на аппарате скрытых марковских моделей речевого сигнала [14] и поэтому неразрывно связан с многозатратной процедурой динамического выравнивания слов по темпу речи диктора. Неудивительно поэтому, что вопросу об увеличении скорости вычислений уделяется в настоящее время все большее внимание. Действительно, в тех случаях, когда объем рабочего словаря составляет несколько тысяч единиц, большинство известных алгоритмов, работающих на основе сегментирования слов на отдельные фонемы и их последующего выравнивания по динамике, для реализации в режиме реального времени требуют мощности, значительно превосходящей возможности современного персонального компьютера и тем более сотового телефона. В результате точная реализация классического подхода стала возможной лишь в проектах таких крупнейших корпораций, как Microsoft [4], Google [5], Apple [15] и Nuance Comminications [6]. При этом для распознавания в режиме реального времени и малопроизводительного оборудования используются облачные вычисления и технология клиент-сервер. К сожалению, клиент-серверный подход является недостаточно гибким: невозможна настройка системы на конкретную группу дикторов, рабочий словарь жестко фиксируется, работа системы требует подключения клиента к сети Internet, отсутствуют гарантии конфиденциальности.

Метод фонетического декодирования слов в своей первоначальной формулировке [1] также

использовал данную процедуру, хотя и в существенно более сжатом виде, рассчитанном на ограниченный объем R фонетической базы данных национального языка. Принцип действия предложенной ниже новой разновидности МФДС усилил эти различия: динамическое выравнивание слов в данном случае не предусматривается в принципе. В итоге вычислительные затраты на реализацию метода сократились на порядок и более, пропорционально повысилось его быстродействие в режиме реального времени. Исследованиям в этом актуальном направлении АРР и посвящена предлагаемая статья. Полученные результаты и сделанные по ним выводы рассчитаны на широкий круг специалистов в области современных речевых технологий, знакомых с основными положениями и терминологией информационной теории восприятия речи.

Фонетическая транскрипция речи

В большинстве известных методов АРР [7, 16] на первом этапе обычно выполняется автоматическое распознавание минимальных речевых единиц типа отдельных фонем. Пусть задан некоторый фонетический алфавит $\left\{\mathbf{x}_r^*\right\}, r=\overline{1,R},$ где R — количество фонем в алфавите. Задача состоит в том [16], чтобы поступившему на вход речевому сигналу X с частотой дискретизации F (в герцах) поставить в соответствие последовательность содержащихся в нем фонем $\left\{\mathbf{x}_r^*\right\}$.

Для решения задачи на первом этапе сигнал X разбивается на непересекающиеся сегменты $\{\mathbf{x}(t)\}$, $t=\overline{1,T},$ длиной $\tau = 0{,}01 - 0{,}015$ с, где T — общее число сегментов. Далее каждый парциальный сигнал $\mathbf{x}(t) = \|x_1(t) \ldots x_M(t)\|$ (здесь $M = \tau F$) рассматривается в пределах конечного списка фонем $\left\{\mathbf{x}_r^*\right\}$ и отождествляется с той из них, которая отвечает принципу минимума величины заданной исследователем меры близости между сигналом $\mathbf{x}(t)$ и эталоном $\mathbf{x}_r^*$. Для выбора меры близости воспользуемся широко использующейся в автоматической обработке речи [17] авторегрессионной (АР) моделью речевого сигнала на интервалах его квазистационарности $\tau = \text{const}$. Известно, что в этом случае при предположении о гауссовом распределении сигнала $\mathbf{x}(t)$ оптимальное в байесовском смысле решение дает принцип минимума информационного рассогласования (МИР) Кульбака — Лейблера [18]

$$\nu(t)=\underset{r\in\{1,\ldots,R\}}{\arg\min}\frac{1}{F}\sum_{f=1}^{F}\left(\frac{G_x(f)}{G_r(f)}-\ln\frac{G_x(f)}{G_r(f)}-1\right). \quad (1)$$

Здесь $G_x(f)$ — выборочная оценка спектральной плотности мощности (СПМ) сигнала $\mathbf{x}(t)$ в функции дискретной частоты f, а $G_r(f)$ — СПМ эталона r-й фонемы $\mathbf{x}_r^*$. Если воспользоваться АР-моделью речевого сигнала, то отношение СПМ в (1) приобретает вид [19]

$$\frac{G_x(f)}{G_r(f)}=\frac{\left|1+\sum_{m=1}^{p}a_{r,m}\exp(-j\pi mf/F)\right|^2}{\left|1+\sum_{m=1}^{p}a_{x,m}\exp(-j\pi mf/F)\right|^2}. \quad (2)$$

Здесь $j=\sqrt{-1}$ — мнимая единица, а $\left\{a_{x,m}\right\}$, $m=\overline{1,p}$ — оценка АР-коэффициентов сигнала $\mathbf{x}(t)$.

Важнейшее достоинство АР-модели в задачах АРР [17] — это возможность нормировки речевых сигналов по дисперсии порождающих процессов: $\sigma_0^2=\sigma_x^2$, где σ_x^2 — выборочная оценка дисперсии порождающего процесса $\mathbf{x}(t)$. В работе [20] показано, что при учете этого асимптотически оптимальное решение (1) эквивалентно адаптивному критерию

$$\nu(t)=\underset{r\in\{1,\ldots,R\}}{\arg\min}\frac{1}{2}\left[\frac{\sigma_r^2(\mathbf{x})}{\sigma_0^2}-1\right]. \quad (3)$$

Здесь $\sigma_r^2(\mathbf{x})$ — выборочная оценка дисперсии отклика r-го обеляющего фильтра

$$\mathbf{y}_r(t) = \|y_{r,1}(t) \ldots y_{r,M-p}(t)\|,$$

где p — порядок АР-модели, а

$$y_{r,j}(t)=x_{j+p}(t)-\sum_{m=1}^{p}a_{r,m}x_{j+p-m}(t),\ j=\overline{1,M-p}. \quad (4)$$

Каждый эталон из фонетической базы данных $\mathbf{x}_r^*$ задается своим вектором АР-коэффициентов $\mathbf{a}_r=\left\{a_{r,m}\right\}, m=\overline{1,p},$ полученным, например, с помощью алгоритма Берга и рекурсивной процедуры Левинсона — Дурбина [19].

Фонетическая транскрипция речевого сигнала сводится, как видим, к его АР-анализу. Подобные задачи обычно решаются [17, 19] с применением рекуррентных вычислительных процедур, обладающих высокой скоростью сходимости. При этом обработка речевого сигнала ведется здесь в R параллельных каналах с использованием набора обеляющих фильтров (3), (4), каждый из которых настроен на соответствующий эталон минимальной звуковой единицы $\mathbf{x}_r^*$. Решение принимается с периодом τ в пользу одной из возможных фонем по критерию МИР (1) или (3). В результате исходный речевой сигнал $\mathbf{x} = \mathbf{x}(l)$, $l = 1, 2, \ldots, L$, где $L = T_C/\tau$, на интервале его действия T_C преобразуется системой АРР в последовательность фонетических символов или букв национального языка $\mathbf{x}=\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_L\}, \mathbf{x}_i\in\left\{\mathbf{x}_r^*\right\}$. На этом завершается первый этап обработки речевого сигнала.

Метод фонетического декодирования слов

После выполнения фонетического транскрибирования задача АРР переходит далее в качественно иную плоскость, а именно фонетического декодирования слов или восстановления исходного речевого сообщения — в виде изолированного слова или целой фразы — по сформированной для него на первом этапе последовательности фонетических символов. При такой формулировке существует тривиальное решение при безошибочном фонетическом кодировании речевого сигнала. Такое решение сводится к многоканальному (по числу слов M из словаря (лексикона) системы АРР $\{\mathbf{y}_m\}$) поэлементному (на L смежных позициях) сравнению фонетического кода анализируемого слова $\mathbf{x}$ с аналогичными кодами слов-эталонов $\mathbf{y}_m = \{\mathbf{y}_{m,1}, \mathbf{y}_{m,2}, ..., \mathbf{y}_{m,L}\}$, $\mathbf{y}_{m,i} \in \left\{\mathbf{x}_r^*\right\}$, $m \le M$, из заданного словаря $\mathbf{Y} = \{\mathbf{y}_m\}$ объема M. Система отдает предпочтение тому из них, которое совпадает с фонетическим кодом $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_L\}$ слова на входе. Но, к сожалению, это практически недостижимый результат. Ввиду известных особенностей речевого механизма человека рассматриваемая задача принципиально не имеет безошибочного решения. Поэтому упростим ее, отбросив из рассмотрения все наиболее нестабильные (вариативные) фонемы национального языка. Останутся, главным образом, вокализованные фонемы [16], а в самом простом случае — гласные. Это видно, в частности, из следующей таблицы $\left\|\rho\left(\mathbf{x}_\nu^* / \mathbf{x}_r^*\right)\right\|$ значений величины информационного рассогласования (ВИР) по Кульбаку — Лейблеру (1) между отдельными фонемами русского языка, полученной по известной [20] методике экспериментальным путем с помощью встроенного аналого-цифрового преобразователя с частотой дискретизации речевого сигнала 8 кГц (табл. 1). При этом порядок АР-модели (3) был установлен $p = 20$. Серым фоном здесь отмечены строки с минимальными значениями ВИР. Все они относятся к случаю невокализованной фонемы на входе. Вокализованные и гласные фонемы характеризуются, напротив, существенно более высоким средним уровнем ВИР в пределах заданной фонетической базы данных $\left\{\mathbf{x}_r^*\right\}$, что говорит об их устойчивости в реализациях.

■ ***Таблица 1***. Матрица значений ВИР

ν	r									
	А	**З**	**М**	**О**	**Сь**	**У**	**Ф**	**Ц**	**Ш**	**Ы**
А	0,00	2933	1078	1224	3392	884	4242	1543	3154	905
З	17	66,3	9,04	2,35	2074	2,63	209,9	275	2857	8,06
М	4,86	2,92	0,00	18,61	28,3	10,9	15,95	23, 6	210	24,4
О	11,1	18,1	3,22	11,56	287,3	2,47	106	209	2615	2,61
Сь	2,87	4,27	4,02	12,82	16,7	3,25	2,54	5,65	109	1,70
У	284	1538	190	3,60	8700	0,00	1631	3384	6022	124
Ф	2,34	1,37	2,29	11,88	2,33	3,75	0,00	0,56	17,8	4,34
Ц	5,44	0,70	3,13	15,10	0,95	6,30	0,41	0,30	6,79	6,12
Ш	29,4	7,14	14,6	61,46	1,94	48,7	7,54	3,66	0,00	79,6
Ы	30,5	16,3	18,9	72,14	4,29	65,2	11	7,11	0,26	74,3

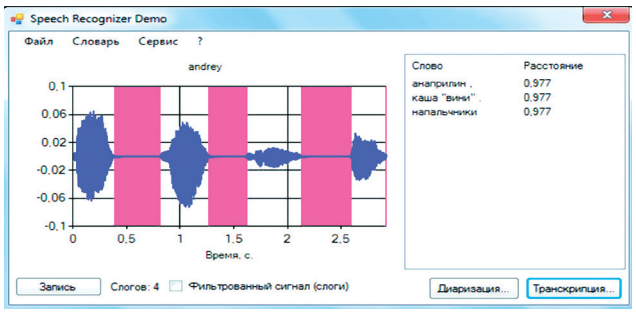
Отталкиваясь от множества гласных фонем русского языка (А, О, Е, И, У, Ы, Э) и следуя критерию (3), получим выражение для оптимальной решающей статистики [1, 2]

$$\rho(X / \mathbf{y}_m) = \sum_{l=1}^{L} \rho(\mathbf{x}_l / \mathbf{y}_{m,l}), \qquad (5)$$

определенное на множестве альтернатив $\mathbf{y}_m$, $m = \overline{1, M}$, из заданного рабочего словаря $\mathbf{Y}$. Отметим при этом важную деталь: длина каждого слова-эталона в данном случае равна длине L слова на входе и выражается в количестве выделенных из последнего на первом этапе гласных фонем. Иными словами, в вычислениях (5) и, значит, в дальнейшей проверке гипотез по критерию МИР участвует ограниченное количество слов-эталонов $K << M$ из словаря $\mathbf{Y}$ большого объема M — только с определенным количеством слогов L, меняющимся от одного слова на входе к другому. Причем наиболее точно распознаются слова большой длины: в словаре $\mathbf{Y}$ их меньшинство, а соответствующие им последовательности кодов существенно различаются между собой, поэтому вероятность их перепутывания зачастую оказывается незначительной. Более того, в ней не остается места для выравнивания слов по темпу речи: количество слогов от темпа не зависит, если не приводит к необратимым искажениям слов на входе. Но этот случай явно не актуален для задач АРР. Таким образом, выражения (3)–(5) в совокупности определяют модифицированный МФДС [1, 21] в расчете на сокращенное R-множество $\left\{\mathbf{x}_r^*\right\}$ фонетических единиц произвольного состава.

Программа и результаты экспериментальных исследований

Для оценивания эффективности МФДС согласно алгоритму (3)–(5) была разработана специальная компьютерная программа Speech Recognizer. Ее интерфейс представлен на рис. 1. Здесь на временной диаграмме слева показана запись речевого сигнала от диктора для слова «анаприлин». Хорошо видны границы его четырех слогов. Всего в рабочем словаре было задействовано более 2000 слов длиной от двух до десяти слогов, взятых из списка лекарств одной из аптек г. Нижний Новгород весной 2012 г. В окне справа перечислены в качестве примера три наилучшие по критерию МИР альтернативы из словаря этало-



■ ***Рис. 1.*** Главное окно программы Speech Recognizer в режиме распознавания

■ ***Таблица 2.*** Оценки вероятности ошибки распознавания слова

Режим распознавания	Диктор				
	1-й	2-й	3-й	4-й	5-й
Дикторозависимый	0,01	0,05	0,013	0,03	0,04
Дикторонезависимый	0,01	0,065	0,017	0,04	0,09
Режим распознавания	**Диктор**				
	6-й	7-й	8-й	9-й	10-й
Дикторозависимый	0,01	0,04	0,028	0,03	0,01
Дикторонезависимый	0,09	0,044	0,09	0,07	0,03

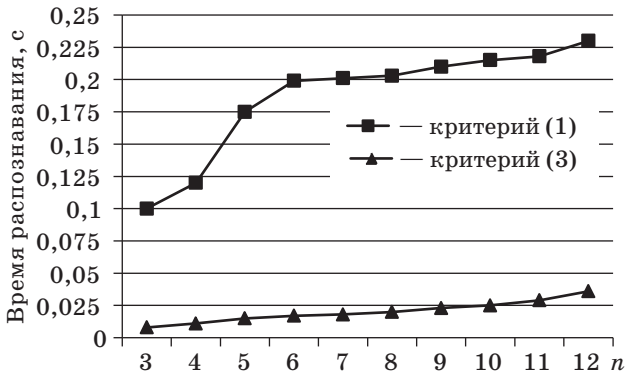
нов. Слово «анаприлин» в этом списке стоит на первом месте как наиболее близкое по критерию МИР (5) к произнесенному диктором слову.

Частота дискретизации речевого сигнала в АЦП была установлена по-прежнему равной 8 кГц, порядок АР-модели минимальной звуковой единицы $p = 20$, а длина одного сегмента сигнала для его обработки согласно (3), (4) составила $n = 80$ отсчетов или $\tau = 10$ мс по времени. Для ввода в программу речи каждого диктора применялся встроенный в ПК микрофон. При этом фонетическая база данных программы $\left\{\mathbf{x}_r^*\right\}$ варьировалась от диктора к диктору и составлялась в каждом случае из вышеупомянутых семи гласных фонем соответствующего диктора. На них и была настроена система обеляющего фильтра (3). В эксперименте приняла участие группа из десяти дикторов в возрасте от 25 до 57 лет, каждым из которых проговаривалось минимум по сто разных слов, причем в нескольких реализациях. По результатам их обработки согласно алгоритму (3)–(5) были получены оценки вероятности ошибочного распознавания отдельных слов (вероятности пропуска слова в итоговом списке решений (см. рис. 1, окно справа)). Основным требованием к речи дикторов при этом было разделение слов на открытые слоги с четкой паузой между ними. В процессе распознавания слоги выделялись простейшим амплитудным детектором паузы на интервале длительностью не менее 70 мс.

В указанных условиях в среднем по группе дикторов безошибочно было распознано около 97,37 % от суммы проговоренных ими слов (табл. 2, дикторозависимый режим). И это весьма высокий результат, особенно если учесть, что большинство (70 %) отмеченных ошибок в АРР приходится на короткие слова в 2—3 слога, которые могли нечетко проговариваться дикторами ввиду повышенной вариативности звукового строя их речи.

Для сравнения в табл. 2 (дикторонезависимый режим) представлены аналогичные оценки вероятности ошибки в той же системе АРР (см. рис. 1), но при ее настройке на фонемы одного и того же (в нашем случае — первого) диктора. Видно, что достоверность АРР если и ухудшилась, то все же осталась в приемлемых для практики пределах, при том, что обучение системы почти не потребовало в данном случае каких-либо существенных временных затрат на организацию и проведение. Продемонстрированные гибкость и малая критичность МФДС по отношению к используемому для настройки (обучения) системы АРР речевому материалу — это еще два ценных качества нового метода с точки зрения перспектив его применения.

В последнем эксперименте сопоставим вычислительную эффективность широко использующегося критерия сопоставления СПМ вида (1) с эквивалентной этому критерию адаптивной реализацией (3). Для ускорения процедуры распознавания в (4) сопоставлялись не все значения СПМ для $f = \overline{1, F}$, а только частоты с шагом $\Delta f = 10$ Гц. Для этого значения параметра скорость АРР в 10 раз превышает скорость распознавания согласно (1), при этом качество распознавания практически не отличается от точности критерия (1). Среднее время распознавания для критериев (1) и (3) в зависимости от числа слогов n во входном словосочетании показано на рис. 2.



■ ***Рис. 2.*** Зависимость времени распознавания для программы Speech Recognizer от рассогласования и числа слогов n во входном словосочетании

На этом рисунке хорошо видно, что вычислительная эффективность адаптивного критерия (3) на порядок превышает аналогичный показатель для традиционного сопоставления сигналов по их СПМ (1).

Заключение

Решению проблемы вычислительной сложности алгоритмов АРР для больших словарей в последние годы исследователями уделяется повышенное внимание. В представленной работе для этого предложен новый подход на основе метода фонетического декодирования слов — в терминах слоговой фонетики [22]. Его основное преимущество перед известными методами и подходами состоит в существенном (на порядок и более) сокращении вычислительных затрат на реализацию за счет отказа от трудоемкой процедуры динамического выравнивания слов по темпу речи. При этом, как убедительно показал пример из актуальной области голосовых заказов лекарств по телефону, и точность, и надежность АРР обеспечиваются на высоком уровне. В задачах подобного рода требование к слоговому произношению слов диктором является более чем приемлемой платой за достигаемые преимущества в скорости, точности и надежности их распознавания.

В области будущего исследования МФДС можно определить следующие задачи.

1. Повышение точности АРР за счет использования дополнительной информации о типе согласных звуков (шумовые, взрывные, модулированные) в слогах. Такая информация, как известно [7, 17], может быть выделена с достаточной степенью надежности.

2. Предварительная сегментация слогов на последовательности стационарных фонем для выделения в речевом сигнале стационарных фрагментов.

3. Автоматическое выделение речевых команд, произнесенных по слогам, в потоке непрерывной речи.

Работа выполнена при финансовой поддержке Минобрнауки РФ по государственному контракту № 07.514.11.4137 ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы» и в рамках программы «Научный фонд НИУ ВШЭ» в 2013–2014 гг., проект № 12-01-0003.

Литература

1. **Савченко В. В.** Метод фонетического декодирования слов в задаче автоматического распознавания речи // Изв. вузов России. Радиоэлектроника. 2009. Вып. 5. С. 41–49.
2. **Савченко В. В.** Автоматическое распознавание речи на основе кластерной модели минимальных речевых единиц в информационной метрике Кульбака — Лейблера // Изв. вузов России. Радиоэлектроника. 2011. Вып. 3. С. 9–19.
3. **Савченко В. В.** Информационная теория восприятия речи // Изв. вузов России. Радиоэлектроника. 2007. Вып. 6. С. 3–9.
4. **Pat.** US 6301560 B1, Int.CI G10L15/22 Discrete speech recognition system with ballooning active grammar / Inventor: Masters S. P. Assignee: Microsoft Corporation. Pub. Date 09.10.2001.
5. **Schuster M.** Speech Recognition for Mobile Devices at Google: Proc. of the 11th Pacific Rim Intern. Conf. on Trends in Artificial Intelligence // LNCS. 2010. Vol. 6230. P. 8–10.
6. **Pat.** US 8175883 B2, Int.CI G10L21/00 (2006.01) Speech recognition system and method / Inventor: Grant. R., Gregor. P. Assignee: Nuance Communications Inc., Pub. Date 08.05.2012.
7. **Benesty J., Sondh M., Huang Y.** (eds.). Springer Handbook of Speech Recognition. — N. Y.: Springer, 2008. — 1159 p.
8. **Савченко В. В., Акатьев Д. Ю.** Результаты экспериментальных исследований методики формирования фонетической базы данных диктора из непрерывного потока его разговорной речи // Информационно-управляющие системы. 2012. № 6. С. 38–42.
9. **Ронжин А. Л., Глазков С. В.** Метод автоматического распознавания голосовых команд и неречевых акустических событий // Информационно-управляющие системы. 2012. № 4. С. 74–77.
10. **Кипяткова И. С.** Комплекс программных средств обработки и распознавания разговорной русской речи // Информационно-управляющие системы. 2011. № 4. С. 53–59.
11. **Кипяткова И. С., Карпов А. А.** Автоматическая обработка и статистический анализ новостного текстового корпуса для модели языка системы распознавания русской речи // Информационно-управляющие системы. 2010. № 4. С. 2–8.
12. **Ронжин А. Л., Карпов А. А., Кагиров И. А.** Особенности дистанционной записи и обработки речи в автоматах самообслуживания // Информационно-управляющие системы. 2009. № 5. С. 32–38.
13. **Ронжин А. Л.** и др. Фонетико-морфологическая разметка речевых корпусов для распознавания и синтеза русской речи // Информационно-управляющие системы. 2006. № 6. С. 24–34.

14. **Rabiner L.** A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition // Proc. of the IEEE. 1989. Vol. 77. N 2. P. 257–285.
15. **Pat.** US 0016678 A1, Int.CI G10L21/00 (2006.01) Intelligent automated assistant / Inventor: Gruber T., Cheyer A., Kittlaus D., Guzzoni D., Brigham C., Giuli R., Bastea-Forte M., Saddler H., Assignee: Apple Inc., Pub. Date 19.01.2012.
16. **Савченко А. В.** Автоматическое построение фонетической транскрипции речи на основе принципа минимума информационного рассогласования // Вестник компьютерных и информационных технологий. 2012. № 8. С. 14–19.
17. **Levinson S. C.** Mathematical models for speech technology. — Chichester: John Wiley&Sons Ltd, 2005. — 261 p.
18. **Kullback S.** Information Theory and statistics. — N. Y.: Dover Publications, 1997. — 399 p.
19. **Марпл С. Л.-мл.** Цифровой спектральный анализ и его приложения. — М.: Мир, 1990. — 584 с.
20. **Савченко В. В.** Автоматическая обработка речи по критерию минимума информационного рассогласования на основе метода обеляющего фильтра // Радиотехника и электроника. 2005. Т. 50. № 3. С. 309–314.
21. **Патент** на полезную модель № 111944. Устройство для фонетического анализа и распознавания речи / В. В. Савченко, А. В. Савченко, Д. Ю. Акатьев (Роспатент). — № 2011125526/08; заявл. 21.06.2011; опубл. 27.12.2011; Бюл. № 6.
22. **Белявский В. М., Светозарова Н. Д.** Слоговая фонетика и три фонетики Л. В. Щербы: Теория языка, методы его исследования и преподавания. — Л.: Наука, 1981. С. 36–40.

УДК 004.8

СУБОПТИМАЛЬНАЯ ЗВЕЗДЧАТАЯ СТРУКТУРА АЛГЕБРАИЧЕСКОЙ БАЙЕСОВСКОЙ СЕТИ

***А. А. Фильченков*[1],**
аспирант
Санкт-Петербургский государственный университет

Выделен подкласс минимальных графов смежности — звездчатые графы смежности. Доказано, что множество таких графов содержит оптимальные вторичные структуры, т. е. графы смежности с минимальным диаметром. Представлен алгоритм синтеза такого множества. На основе этого предложен способ ускорения синтеза оптимальной вторичной структуры алгебраической байесовской сети.

Ключевые слова *— алгебраическая байесовская сеть, звездчатый граф, обучение глобальной структуры, граф смежности, машинное обучение, субоптимальная вторичная структура.*

Введение

Алгебраическая байесовская сеть (АБС) — активно развивающаяся теоретико-компьютерная модель, относящаяся к классу логико-вероятностных графических моделей, которая позволяет использовать как скалярные, так и интервальные оценки вероятности для представления неопределенности [1–6]. АБС может быть применима как одна из математических моделей анализа информационной безопасности [7], в частности, для оценки защищенности системы от социо-инженерных атак [8]. Кроме того, АБС могут применяться в задаче распознавания изображений [9] — традиционной области, в которой используются вероятностные графические модели [10]. Обработка интервальных оценок также оказывается востребованной в анализе гранулярных данных о рекордных интервалах между последними эпизодами [11]. Известные алгоритмы глобального логико-вероятностного вывода для АБС опираются на ее вторичную структуру, которая представляет собой граф, построенный над ее первичной структурой [1, 2, 6, 12]. Первичная структура АБС в свою очередь представляет собой набор математических моделей фрагментов знаний (далее — фрагментов знаний) — сложным образом устроенном представлении случайных элементов.

Вторичной структурой могут выступать только особые графы — минимальные (по числу ребер) графы смежности [6, 12–14], формальное определение которых будет дано в следующем разделе. В силу специфики алгоритмов логико-вероятностного вывода, предложенных в работах [1, 2, 6, 12], одной из важных характеристик минимальных графов смежности является их диаметр. Глобальное обучение АБС [15] — это один из видов машинного (автоматического) обучения, развитие методов, моделей и алгоритмов которого — одна из самых актуальных задач в искусственном интеллекте [6, 16, 17]. В рамках задачи обучения глобальной структуры было бы желательно строить графы смежности с минимальным диаметром [18]. Следует отметить, что кроме минимальности диаметра существуют и другие требования [18], поэтому минимальные графы смежности с минимальным диаметром являются оптимальной вторичной структурой.

Построение минимального графа смежности, одновременно минимального по диаметру, возможно за счет построения всего множества минимальных графов смежности и выбора из него графа с минимальным диаметром. В рамках данной работы мы рассмотрим звездчатые графы — особый класс минимальных графов смежности, который, как будет показано, характеризуется тем, что множество звездчатых графов смежности содержит в себе некоторые минимальные графы смежности с минимальным диаметром. Мощность этого множества меньше, чем мощность множества всех минимальных графов смежности, поэто-

[1] Научный руководитель — доктор физико-математических наук, доцент, заведующий лабораторией теоретических и междисциплинарных проблем информатики СПИИРАН *А. Л. Тулупьев.*

му поиск минимального графа смежности с минимальным диаметром в таком множестве будет быстрее. Будет также предложен алгоритм синтеза указанного множества, который благодаря свойствам звездчатых графов отличается от алгоритма синтеза множества всех минимальных графов смежности лишь небольшими изменениями.

Графы смежности

В качестве вторичной структуры АБС выступает минимальный (по числу ребер) граф смежности. Дадим необходимые понятия, следуя введенной в работах [13, 14, 19] терминологии.

Согласованный нагруженный граф G — граф $G = (V, E)$ без петель, на вершинах и ребрах которого задана функция нагрузки W, множество значений которой состоит из подмножеств некоторого алфавита A:

$$W : V \cup E \to 2^A,$$

причем

$$\forall e = (u, v) \in E \quad W(e) = W(u) \cap W(v).$$

Сепаратором двух вершин называется пересечение их нагрузок. В согласованном нагруженном графе, как видно из определения, вес ребра равен сепаратору его концов. Множество сепараторов для данной первичной структуры обозначим как Separator. Две вершины называются *сочлененными*, если их сепаратор непуст.

Магистральный путь между двумя вершинами согласованного нагруженного графа — это такой ненаправленный путь между ними, на котором нагрузка любой его вершины содержит сепаратор исходных вершин. Согласованный нагруженный граф магистрально связен, если между каждой парой сочлененных вершин существует магистральный путь, но не обязательно ребро.

Граф смежности — это магистрально связный согласованный нагруженный граф, ребра в котором возможны только между сочлененными вершинами (т. е. нагрузка ребер непуста). *Минимальный граф смежности* — минимальный по числу ребер граф смежности.

В рамках исследования графов смежности вместо математических моделей фрагментов знаний (достаточно сложных конструкций, описание которых можно найти в работах [5, 6]) обычно рассматривается база таких моделей — подалфавиты некоторого алфавита A. Набор таких подалфавитов PS образует разбиение A, причем подалфавиты не поглощают друг друга:

$$\mathrm{PS} = \{A_i\}_{i=1,\ldots,n} : \bigcup_{i=1..n} A_i = A \text{ и}$$

$$\forall A_i, A_j \in \mathrm{PS}, i \neq j \Rightarrow A_i \subsetneq A_j.$$

Введенный таким образом PS задает значения функций нагрузки (либо просто нагрузку) для вершин графа. Задачей глобального обучения вторичной структуры АБС является построение минимального графа смежности для заданного множества вершин — именно на них осуществляются алгоритмы логико-вероятностного вывода.

Синтез множества минимальных графов смежности

Далее необходимо ввести систему терминов, связанную с синтезом множества минимальных графов смежности, поскольку при помощи нее будут определены звездчатые графы смежности. Будем следовать работам [13, 19, 20].

Для каждого сепаратора определим множество его *сыновей*: U' является сыном сепаратора U тогда и только тогда, когда $U' \subset U$ и $\nexists U'' : U \subset U'' \subset U'$. Другими словами, если на множестве сепараторов задать частичный порядок, индуцированный отношением включения, то сыновья для сепаратора будут множеством его последователей (в терминах теории частично-упорядоченных множеств). Множество сыновей сепаратора U будем обозначать как $\mathrm{Sons}(U)$.

Сужение $G \downarrow A'$ согласованного нагруженного графа G на подалфавит A' — это ненаправленный граф, в который входят только те вершины и ребра исходного графа G, нагрузки которых содержат или равны A':

$$G \downarrow A' = (\{V_i \mid V_i \in V(G), A' \subseteq W(V_i)\};$$
$$\{E_i \mid E_i \in E(G), A' \subseteq W(E_i)\}).$$

Сильное сужение $G \Downarrow U$ — сужение $G \downarrow U$ на сепаратор U, из которого удалили все ребра с нагрузкой U:

$$G \Downarrow U = (\{V_i \mid V_i \in V(G), U \subseteq W(V_i)\};$$
$$\{E_i \mid E_i \in E(G), U \subset W(E_i)\}).$$

Рассмотрим *максимальный по числу ребер* граф смежности $G_{\max.}$ Сильное сужение $G_{\max} \Downarrow U$ разбивается на компоненты связности, которые называются *владениями* для нагрузки U, их будем обозначать как $\mathrm{Holdings}(U)$. Вершина, которая образует одноэлементное владение для нагрузки U, называется *доменной вершиной* для этой нагрузки. Известно, что множество вершин во владении для нагрузки U является либо множеством вершин, входящим в сужение $G_{\max}$ на какого-либо ее сына, либо объединением таких множеств, либо доменной вершиной.

Жила S_U с нагрузкой U — набор ребер с нагрузкой U, число которых равно числу владений для нагрузки U, уменьшенному на единицу,

а граф, полученный добавлением к $G_{\max} \downarrow U$ ребер из S_U, является связным: $S_U \subseteq E(G_{\max})$:

1) $\forall e \in S_U \; W(e) = U$;

2) $|S_U| = |\text{Holdings}(U)| - 1$;

3) граф $((V(G_{\max} \downarrow U), E(G_{\max} \downarrow U) \cup S_U))$ связен.

В определенном смысле жила является деревом на владениях (строгая формализация этого рассуждения, не требующаяся в данной работе, приведена в работах [13, 14]).

Теорема 1 (о минимальных графах смежности) [13, 14]. Граф смежности является минимальным тогда и только тогда, когда множество ребер каждого значимого веса является в нем жилой.

Теорема о минимальных графах смежности создает основу для конструктивного представления множества минимальных графов смежности как декартова произведения множеств жил, построенных для каждого сепаратора. Последнее лежит в основе алгоритмов синтеза такого множества, схема которого будет приведена ниже.

Звездчатый граф смежности

Введем теперь определение звездчатого графа смежности. *Звездой* называется дерево, одна из вершин которого является концом каждого ребра. Такая вершина называется *центром звезды*, а диаметр такого графа равен двум. Верно также и то, что любое дерево с диаметром, равным двум, является звездой.

Звездчатым графом смежности будем называть граф, в котором каждая жила является звездой. Поскольку звезда является деревом, то, согласно теореме о минимальных графах смежности, звездчатый граф смежности является минимальным. *Звездчатым графом смежности* будем называть такой минимальный граф смежности, что:

1) любая его жила является звездой;

2) концы любого ребра в любой жиле с нагрузкой U являются либо доменными вершинами для этой нагрузки, либо центрами жилы с нагрузкой U', $U' \in \text{Sons}(U)$.

Теорема 2 (о структуре жил графа смежности). В жиле графа смежности не существует трех вершин, попарно соединенных ребрами одного веса.

Доказательство: Пускай существуют три такие вершины, попарно соединенные ребрами одной жилы. Они входят не более чем в три разных владения. По определению, в жиле число ребер на единицу меньше, чем число владений. Владения, в которые входят эти вершины, соединенные тремя ребрами, представляют компонент связности. Рассмотрим другое владение. Укажем ребро, которое соединяет его с указанным компонентом связности. Добавим это владение и это ребро к уже рассмотренным компонентам связности. Будем повторять эту операцию со всеми владениями, добавляя каждое из них вместе с ребром к компоненту связности. Когда мы станем рассматривать последний компонент связности, то окажется, что мы рассмотрели уже достаточное для жилы число ребер, ни одно из которых не соединяет это владение с оставшимися. Следовательно, предположение неверно.

Теорема 3 (о звездчатом графе смежности с минимальным диаметром). Существует минимальный граф смежности с минимальным диаметром, который является звездчатым.

Доказательство: Рассмотрим произвольный минимальный граф смежности G с минимальным диаметром и будем последовательно преобразовывать его к звездчатому графу. Для этого рассмотрим произвольную жилу U графа G. В ней рассмотрим все кратчайшие пути между двумя вершинами в G, длина которых максимальна среди таких путей (и по определению равна диаметру), которые содержат ребра веса U. Множество таких путей обозначим как Shortest_U.

Выберем произвольный путь P из Shortest_U, возьмем какой-либо из двух его концов (назовем его s) и будем последовательно рассматривать вершины P, начиная с выбранного конца. Первую вершину, из которой выходит ребро нагрузки U, обозначим g. Последнюю вершину, в которую входит ребро нагрузки U, обозначим g'. Другой конец пути обозначим s'. Вершины g и g' будем называть *воротами*. Если бы мы обходили путь в обратном направлении, то ворота остались бы прежними, поменялся бы лишь порядок их прохода. По построению все ребра нагрузки U, принадлежащие пути P, принадлежат и пути между g и g'. Ненаправленные пути между s и g и между g' и s' будем называть *тупиками*. Тупики не содержат ребер с нагрузкой U. Тупику мы можем однозначно сопоставить ворота, являющиеся его концом.

Теперь будем для каждого сепаратора менять жилу соответствующей нагрузки на жилу, являющуюся звездой (на самом деле, семейство таких жил), таким образом, чтобы не увеличить диаметр.

Пускай мы на данном шаге выбрали сепаратор U. Рассмотрим множество тупиков для данного сепаратора, среди них выберем один наибольшей длины. Выбранный тупик обозначим b_1, а его длину — n_1. Ворота тупика b_1 сделаем центром звезды, соединяя ее ребрами нагрузки U с произвольными вершинами из других владений. Рассмотрим, как изменилась длина кратчайших путей, содержащих ребра веса U. Каждый такой путь состоял из двух тупиков (например, b_2 длины n_2 и b_3 длины n_3), и длина такого пути была не менее $n_2 + n_3 + 1$, поскольку этот путь должен содержать не менее одного ребра веса U, которые не входят в эти тупики. Соответственно, любой путь, содержащий тупик b_1, по построению со-

держит ровно одно ребро веса U, поэтому его длина не увеличилась.

Предположим, что увеличилась длина кратчайшего пути, не содержащего b_1 (т. е. ни один из тупиков b_2 и b_3 не совпадает с b_1). Обозначим исходный путь $p_{2,3}$, а получившийся — $p'_{2,3}$. Ворота, соответствующие b_2 и b_3, обозначим g_2 и g_3. По построению длина $p'_{2,3}$ равна $n_2 + n_3 + 2$, следовательно, длина $p_{2,3}$ была равна $n_2 + n_3 + 1$, следовательно, в G между g_2 и g_3 было ребро. Согласно лемме, в G не могло быть ребер одновременно между g_1 и g_2, между g_1 и g_3. Пускай ребра не было между g_1 и g_2, следовательно, путь между ними состоял не менее чем из двух ребер. Следовательно, путь, включающий тупики b_1 и b_2, в G имел длину не менее $n_1 + n_2 + 2$, что в силу максимальности n_1 не меньше, чем $n_2 + n_3 + 2$, откуда следует, что $p_{2,3}$ не был максимальным по длине. Следовательно, замена жилы на звездчатую не изменит длины максимальных путей.

Проведя замену жил для каждого сепаратора, мы придем к звездчатому графу смежности, диаметр которого не больше, чем диаметр рассмотренного минимального графа смежности с минимальным диаметром, а, следовательно, минимален.

Схемы синтеза множеств минимальных графов смежности и звездчатых графов смежности

Звездчатые графы смежности представляют интерес, в первую очередь, потому, что синтез множества таких графов похож на синтез множества минимальных графов смежности, поскольку основан на конструктивном представлении таких графов.

Сначала приведем схему алгоритма синтеза множества минимальных графов смежности. Следует указать, что подобных алгоритмов существует значительное число, но приведем лишь наиболее общий, а затем покажем, как за счет внесения небольших изменений можно привести его к алгоритму синтеза множества звездчатых графов смежности. Подобные изменения будут применимы и для любого другого алгоритма синтеза множества минимальных графов смежности.

На вход такого алгоритма подается множество взаимно непоглощающих подалфавитов некоторого алфавита A, соответствующих первичной структуре алгебраической байесовской сети.

1. Построить множество Separator, содержащее все сепараторы.

2. Для каждого сепаратора U построить соответствующее ему сильное сужение, найти компоненты связности такого сильного сужения H_U.

3. Для каждого набора таких владений H_U перебрать все возможные деревья $T_{U,\,i}$, построенные над владениями H_U как над вершинами.

4. Для каждого такого дерева $T_{U,\,i}$ построить все возможные жилы $S_{U,\,I,\,j}$, ребра в которых соединяют вершины из двух владений в том и только том случае, если подобные владения соединены ребром в $T_{U,\,i}$.

5. Объединить все жилы, построенные для одного сепаратора, в множество S_U.

6. Перебрать все возможные кортежи жил, выбранных по одной для каждого сепаратора.

Согласно теореме о минимальных графах смежности, объединение жил в каждом кортеже является множеством ребер какого-то минимального графа смежности, причем каждому минимальному графу смежности соответствует один из кортежей подобного вида. Следовательно, будет построено множество минимальных графов смежности.

Схема алгоритма синтеза множества звездчатых графов совпадает с приведенной выше схемой во всех шагах, кроме четвертого и пятого, поэтому приведем только их.

4. Для каждого набора построенных на шаге 3 владений H_U перебрать все возможные звезды $Z_{U,\,i}$, построенные над владениями H_U как над вершинами.

5. Для каждой такой звезды $Z_{U,\,i}$ построить все возможные жилы $Z_{U,\,I,\,j}$, ребра в которых соединяют вершины из двух владений в том и только том случае, если подобные владения соединены ребром в $Z_{U,\,i}$, причем один из концов каждого ребра является одной и той же вершиной (лежащей во владении, соответствующем центру $Z_{U,\,i}$).

Результатом работы алгоритма станет множество звездчатых графов.

Поскольку все известные алгоритмы синтеза множества минимальных графов смежности различаются в первых трех шагах, то каждый из них можно использовать в качестве основы для синтеза множества звездчатых графов смежности.

Заключение

В работе рассмотрен особый класс минимальных графов смежности — звездчатых графов смежности, содержащий минимальные по диаметру графы смежности. При отсутствии точных алгоритмов построения минимального по диаметру графа смежности поиск такого графа эффективнее осуществлять не среди вообще всех минимальных графов смежности, а среди именно звездчатых графов смежности, которые выступают субоптимальными структурами, т. е. структурами, множество которых содержит оптимальные структуры. Таким образом, представленный в работе результат развивает теорию алгебраических байесовских сетей в срезе более общей проблемы синтеза интеллектуальных систем, основанных на вероятностных графических моде-

лях, — автоматического обучения глобальной структуры таковых.

Для решения задачи синтеза графа смежности с минимальным диаметром необходимо предложить конструктивное описание таких графов. Исследования в этом направлении показывают, что, вероятно, минимальность диаметра не является локальным свойством, т. е. графы с минимальным диаметром не могут быть описаны через особого вида жилы минимальных графов смежности. Поэтому в рамках решения той же задачи актуален вопрос поиска подкласса минимальных графов смежности, который меньше, чем звездчатые графы смежности, но обладает описанным выше свойством локальности и содержит хотя бы один граф смежности с минимальным диаметром.

Работа выполнена при поддержке РФФИ, гранты № 12-01-00945-а и 12-01-31202-мол_а.

Литература

1. **Тулупьев А. Л.** Алгебраические байесовские сети: глобальный логико-вероятностный вывод в деревьях смежности: учеб. пособие. — СПб.: СПбГУ; Анатолия, 2007. — 40 с. (Сер. Элементы мягких вычислений.)
2. **Тулупьев А. Л.** Алгебраические байесовские сети: система операций глобального логико-вероятностного вывода // Информационно-измерительные и управляющие системы. 2010. № 11. С. 65–72.
3. **Тулупьев А. Л.** Апостериорные оценки вероятностей в алгебраических байесовских сетях // Вестн. С.-Петербург. ун-та. Сер. 10. 2012. Вып. 2. С. 51–59.
4. **Тулупьев А. Л.** Байесовские сети: логико-вероятностный вывод в циклах. — СПб.: Изд-во СПбГУ, 2008. — 140 с. (Сер. Элементы мягких вычислений.)
5. **Тулупьев А. Л., Николенко С. И., Сироткин А. В.** Байесовские сети: логико-вероятностный подход. — СПб.: Наука, 2006. — 607 с.
6. **Тулупьев А. Л., Сироткин А. В., Николенко С. И.** Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. — СПб.: Изд-во СПбГУ, 2009. — 400 с.
7. **Котенко И. В., Степашкин М. В., Юсупов Р. М.** Математические модели, методы и архитектуры для защиты компьютерных сетей: аналитический обзор перспективных направлений исследований по результатам международного семинара MMM-ACNS-2005 // Тр. СПИИРАН. 2006. Т. 2. Вып. 3. С. 11–29.
8. **Азаров А. А., Тулупьев А. Л., Тулупьева Т. В.** SQL-представление реляционно-вероятностных моделей социо-инженерных атак в задачах расчета агрегированных оценок защищенности персонала информационной системы // Тр. СПИИРАН. 2012. Вып. 3(22). С. 31–44.
9. **Сергеев М. Б., Соловьев Н. В., Стадник А. И.** Методы повышения контрастности растровых изображений для систем цифровой обработки видеоинформации // Информационно-управляющие системы. 2007. № 1. С. 2–7.
10. **Bishop C.** Pattern Recognition and Machine Learning. — Springer, 2006. — 740 p.
11. **Суворова А. В.** и др. Вероятностные графические модели социально-значимого поведения индивида, учитывающие неполноту информации // Тр. СПИИРАН. 2012. Вып. 3(22). С. 101–112.
12. **Тулупьев А. Л.** Дерево смежности с идеалами конъюнктов как ациклическая алгебраическая байесовская сеть // Тр. СПИИРАН. 2006. Т. 1. Вып. 3. С. 198–227.
13. **Фильченков А. А., Тулупьев А. Л.** Структурный анализ систем минимальных графов смежности // Тр. СПИИРАН. Вып. 11. С. 104–127.
14. **Фильченков А. А., Тулупьев А. Л., Сироткин А. В.** Структурный анализ клик максимальных графов смежности алгебраических байесовских сетей // Вестн. Тверск. гос. ун-та. Сер. Прикладная математика. 2011. № 20. С. 139–151.
15. **Тулупьев А. Л., Фильченков А. А., Вальтман Н. А.** Алгебраические байесовские сети: задачи автоматического обучения // Информационно-измерительные и управляющие системы. 2011. Т. 9. № 11. С. 57–61.
16. **Alpaydin E.** Introduction to Machine Learning. 2nd ed. — Cambridge: Mass. MIT Press, 2010. — 581 p.
17. **Тихомиров А. В., Шалыто А. А.** Применение генетического подхода для генерации клеточных автоматов // Науч.-техн. вестн. С.-Петербург. гос. ун-та информационных технологий, механики и оптики. 2011. Вып. 2(72). С. 62–66.
18. **Фильченков А. А., Тулупьев А. Л., Сироткин А. В.** Минимальные графы смежности алгебраической байесовской сети: формализация основ синтеза и автоматического обучения // Нечеткие системы и мягкие вычисления: Научный журнал Российской ассоциации нечетких систем и мягких вычислений. 2011. Т. 6. № 2. С. 145–163.
19. **Фильченков А. А., Тулупьев А. Л.** Совпадение множеств минимальных и нередуцируемых графов смежности над первичной структурой алгебраической байесовской сети // Вестн. С.-Петербург. ун-та. Сер. 1. Математика. Механика. Астрономия. 2012. Вып. 2. С. 65–74.
20. **Фильченков А. А.** Алгоритмы построения элементов третичной полиструктуры алгебраической байесовской сети // Тр. СПИИРАН. 2011. Вып. 3(18). С. 237–266.

УДК 681.3

АНАЛИЗ НАДЕЖНОСТИ ФУНКЦИОНАЛЬНЫХ УЗЛОВ ЦИФРОВЫХ СБИС СО СТРУКТУРНЫМ РЕЗЕРВИРОВАНИЕМ И ПЕРИОДИЧЕСКИМ ВОССТАНОВЛЕНИЕМ РАБОТОСПОСОБНОГО СОСТОЯНИЯ

С. Л. Максименко,
старший преподаватель
В. Ф. Мелехин,
доктор техн. наук, профессор
Санкт-Петербургский государственный политехнический университет

Проводится анализ влияния радиационных воздействий на цифровые устройства со структурным резервированием на уровне функциональных узлов интегральных схем в составе информационно-управляющих систем. Предлагается математическая модель, позволяющая оценить надежность узла, представленного на уровне регистровых передач, с учетом цикличности вычислительных процессов и периодического восстановления информации при сбоях в элементах. Показано, что при организации цикличности работы узлов с периодическим восстановлением информации достигается существенное улучшение показателей надежности.

Ключевые слова *— информационно-управляющие системы, радиационные эффекты, интегральные схемы, сбои, отказы, восстановление, надежность, модель, структура, троирование, мажоритар.*

Введение

В работе [1] проведен анализ радиационных эффектов в полупроводниковых структурах и их влияния на информационные процессы в цифровых устройствах. Обоснована актуальность повышения радиационной стойкости за счет соответствующей функциональной организации устройств и организации периодических процессов восстановления информации в случаях сбоев.

Сбои (восстанавливаемые отказы) вызываются попаданием одиночных частиц высоких энергий (Single Event Effects — SEE). Попадание частицы в триггер может вызвать информационный отказ триггера (Single Event Upset — SEU), т. е. искажение хранящейся информации. Информационный отказ является устранимым (soft error), поскольку может быть исправлен записью правильной информации в пораженный элемент. Попадание отдельной частицы в логический вентиль комбинационной схемы может привести к появлению сбоя в виде «иголки» на его выходе (этот эффект носит название Single Event Transient — SET). Распространяясь по цепочке элементов, такой импульс может привести к нежелательной смене состояния триггеров.

Основными методами борьбы с информационными отказами являются структурное резервирование (троирование и мажорирование), введение информационной избыточности (помехоустойчивое кодирование), использование временной избыточности (повторное выполнение операций). Информационная и временная избыточность преимущественно используются в запоминающих устройствах и системах передачи данных. Троирование и мажорирование применяют для устройств преобразования информации, которые можно представить композицией операционных и управляющих автоматов.

В данной работе рассматриваются задачи повышения надежности устройств преобразования информации.

Важным вопросом при проектировании системы является выбор уровня резервирования. Резервирование на уровне отдельных комбинаторных и запоминающих элементов [1] имеет ряд недостатков: ограничение по энергии частиц, невозможность фиксировать факт отказа (например, для выявления неустранимых отказов или получения информации о приближении к пороговой дозе), наибольшее влияние на быстродействие си-

стемы (необходимость использовать элементы с внутренней избыточностью на критических цепях). Резервирование на уровне готовых сложных функциональных узлов (IP-ядер) имеет другой недостаток: восстановление информации в таком узле обычно требует остановки вычислений. Возможность временной остановки работы отдельных узлов должна быть поддержана архитектурой устройства. Для реализации обоих вариантов, как правило, необходима еще и глубокая переработка схемы узла, сводящаяся к внедрению в нее цепей, реализующих восстановление.

Чем ниже уровень резервирования узлов, тем проще становится восстановление узла при информационном отказе. С другой стороны, при понижении уровня резервирования увеличивается количество контролируемых узлов, суммарная сложность средств контроля и восстановления возрастает. Повышается также сложность системы учета частоты отказов.

Процессы в узлах информационно-управляющей системы носят циклический характер. Каждый цикл работы узла связан с получением новой информации для обработки и выдачей результатов. Результаты работы на предыдущем цикле, как правило, не важны. В противном случае устройство, хранящее предыдущие результаты, можно рассматривать как отдельный узел со своей цикличностью. Таким образом, на каждом цикле происходит восстановление информации в узле без необходимости его останова. Обычно чем ниже уровень иерархии узла, тем короче цикл его работы (поскольку проще выполняемая операция) и тем чаще происходит восстановление информации.

Отказ одного из запоминающих элементов узла не обязательно сразу приводит к выдаче узлом неправильного результата. Отказ может проявиться через некоторое время. Чем выше длительность цикла работы узла, тем дольше узел может проработать до обнаружения отказа. За это время может возникнуть отказ в другом экземпляре узла, что приведет к аварии системы в целом.

Правильный выбор уровня резервирования требует учета цикличности процессов в системе. Для этого необходимо построить математическую модель, позволяющую связать эффекты радиационного воздействия на отдельные вентили с параметрами надежности системы при заданном уровне резервирования и известных параметрах цикличности.

Для СБИС со сложным поведением (например, для микропроцессоров) характерны латентность проявления отказов, невозможность по их проявлению определить первоисточник отказа (например, оценить влияния SET на общее количество отказов), что делает крайне сложным получение оценки устойчивости СБИС в физическом эксперименте.

В целях получения экспериментальных данных по устойчивости отдельных элементов разрабатывают специальные интегральные схемы, обладающие высокой наблюдаемостью, такие как сдвигающий регистр с дополнительными цепочками инверторов [2].

Перенос результатов физических экспериментов над отдельными элементами на СБИС со структурным резервированием требует применения соответствующей математической модели СБИС. Современная СБИС обработки информации — это сложная система, которую можно представить в виде сети функциональных узлов. В данной работе рассматривается модель функционального узла. Расширение модели узла до уровня системы требует дополнительного рассмотрения.

Построение параметризируемой модели узла

Обоснование способа абстрактного представления узла, не связанного с его функциональностью

Основной целью построения модели является анализ надежности узла при выбранном варианте резервирования и восстановления для заданной библиотеки элементов. Поэтому можно исключить рассмотрение процессов внутри элемента, считая, что интенсивность отказов элементов известна (например, из физических экспериментов над элементами).

В соответствии с разделением эффектов SET и SEU [1] будем рассматривать узел как сеть регистров, связанных комбинаторными преобразователями. Триггеры, составляющие регистры, подвержены отказам типа SEU. Комбинаторные элементы подвержены сбоям типа SET, которые, распространяясь, могут привести к формированию неверных сигналов на «информационных» входах триггеров в момент записи или к появлению «ложных» импульсов в цепях управления (асинхронных установок и тактирования).

Критерий исправности узла определяем следующим образом: узел исправен, если формирует выходные сигналы, соответствующие спецификации.

Из-за своей малой длительности импульсы, порожденные SET, не могут непосредственно привести к сбоям на внешних выводах устройства (высокоскоростные интерфейсы типа LVDS, SSTL и т. п. рассматривать не будем). Поэтому к отказу может привести только ситуация, когда некоторые запоминающие элементы хранят неверные значения.

Сам по себе информационный отказ элемента не является отказом устройства. Отказ произой-

дет, только если неверное значение распространится до выхода устройства.

Информационный отказ не приведет к отказу устройства, если значение пораженного элемента памяти не влияет на формирование выходов и на вычисление новых значений других элементов памяти. Это возможно, например, когда к некоторому адресуемому регистру не обращаются либо когда информационный отказ не влияет на функцию перехода автомата. Понятно, что данная ситуация не может продолжаться бесконечно долго, иначе данный элемент памяти просто не используется в устройстве.

В общем случае чтобы определить, влияет ли значение элемента памяти на другие элементы, необходимо проанализировать все логические функции, их связывающие. Будем рассматривать худший случай, считая, что любой узел — это автомат с памятью, и сбой в элементе памяти — это изменение состояния автомата, что неизбежно приводит к отказу всего узла.

Приняв это предположение, мы избавляемся от необходимости рассматривать отдельные комбинаторные преобразователи внутри узла. Можно рассматривать все комбинаторные преобразователи как один блок, являющийся источником потока сбоев, некоторая известная доля которых будет запомнена в триггерах и приведет к отказу узла.

Выходные комбинаторные преобразователи не могут повлиять на отказ своего узла. Однако они могут привести к отказу других узлов, подключенных к нему по входу. Если учесть, что в системе со структурным резервированием различные узлы связаны между собой только через мажоритары (узлы, блокирующие распространение отказа), то для запоминания сбоя, порожденного в выходных преобразователях, должно наступить одно из двух событий:

— одновременно происходят два сбоя на одном выходе в разных экземплярах узла;

— сбой происходит одновременно с отказом другого экземпляра узла.

Первое событие является крайне маловероятным, второе — более вероятно. Кроме того, если сбой будет пропущен мажоритаром, он приведет к отказу всех экземпляров узла-приемника, т. е. к отказу системы. (Если мажоритар троирован, его экземпляры, скорее всего, среагируют одинаково.) Таким образом, сбой проявится так же, как если бы отказал второй узел в тройке.

Из изложенного следует, что выходные формирователи влияют на исправность системы так же, как внутренние, но только когда один из экземпляров узла отказал. Если считать, что доля времени, когда один из узлов отказал, мала, то учитывать сбои в выходных преобразователях не требуется. Это справедливо для узлов с периодическим восстановлением их исправного состояния. Если же рассматривать поведение системы при невосстановимых отказах, такой учет требуется.

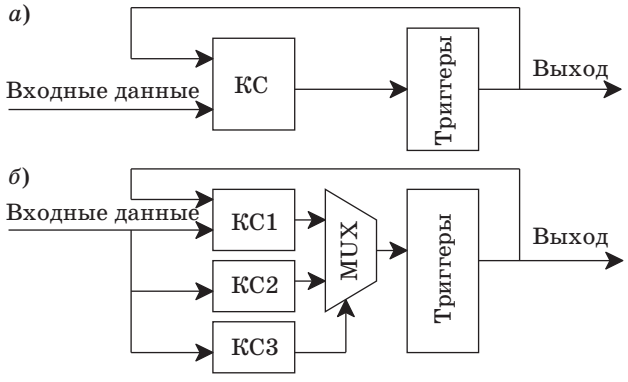
При разбиении схемы на узлы для последующего резервирования лучше размещать комбинаторные преобразователи на выходе узла.

Структурная модель узла

Будем рассматривать узел как автомат с памятью, состоящий из комбинационной схемы *КС* и триггеров (рис. 1, *а*). На входы КС поступают входные данные и выходные сигналы триггеров. Автомат работает в двух режимах: в рабочем режиме, когда автомат перерабатывает входные последовательности (данные) в выходные; в режиме сброса или предустановки, когда КС воспринимает только внешние входные сигналы и ее выходы не зависят от выходов триггеров. Не обязательно среди входов узла есть сигнал, напрямую определяющий загрузку входных данных в регистры. Режим загрузки может активироваться при некоторых комбинациях входных данных, при этом в регистры могут заноситься не сами входные данные, а результаты вычисления некоторой функции над ними.

Для анализа работы узла во всех режимах рассмотрим рис. 1, *б*. На нем комбинационные преобразователи в составе узла представлены в виде трех комбинационных схем: *КС*1 реализует преобразование информации в рабочем режиме, *КС*2 — преобразование входной информации при загрузке, а *КС*3 — активацию режима загрузки. Выходной комбинационный преобразователь, реализующий функцию выхода автомата, на рис. 1 не приведен по причинам, изложенным выше.

Узел, реализующий конвейерную обработку данных, не может быть непосредственно описан схемой, представленной на рис. 1, поскольку загрузка данных выполняется только на первой



■ ***Рис. 1.*** Структурная модель узла: *а* — как автомата с памятью; *б* — с периодической загрузкой информации

ступени. Если данные поступают на вход первой ступени, каждый такт и есть k ступеней, вся информация, относящаяся к одному входному набору данных, остается в конвейере k тактов, но в каждой ступени она хранится только один такт. Поэтому для оценки надежности при информационном отказе конвейеризованный узел может быть представлен узлом со схемой, рассмотренной на рис. 1, *а*, у которого загрузка выполняется каждый такт.

В каждом триггере поток информационных отказов — пуассоновский [1] с интенсивностью λ_T [3]. Будем считать, что искаженное значение состояния хотя бы одного триггера приводит к неверному состоянию триггеров на следующем такте, и этот отказ не устраняется до момента записи нового значения с входов узла (а при записи он устраняется, если в момент записи не возникнет сбой в КС).

Комбинационная схема порождает сбои, которые могут быть запомнены триггерами. Запоминание конкретного сбоя имеет случайный характер, причем вероятность запоминания зависит как от параметров импульса сбоя, так и от свойств и условий работы триггера. Принимая, что ложные импульсы от эффекта SET могут распространяться по цепочкам логических элементов на существенное расстояние, можно считать, что для фиксированных условий количество сбоев на выходе КС пропорционально занимаемой ею площади на кристалле, т. е. количеству логических элементов.

Для конкретной элементной базы и условий работы узла путем моделирования [4] либо экспериментально [2, 5] можно определить, какая доля сбоев запоминается. В дальнейшем можно учитывать только такие «запоминаемые» сбои, считая, что интенсивность сбоев на выходе КС связана с количеством ее элементов некоторым коэффициентом, фиксированным для конкретной задачи анализа: $\lambda_{КС} = kN_{л.э}$.

Поскольку отказ в любом триггере или «запоминаемый» сбой в любом логическом элементе в нашей модели приводят к информационному отказу узла, вероятность его безотказной работы на одном цикле выполнения операции можно оценить как $P(t) = e^{-(N_T\lambda_T + kN_{л.э})t}$.

Причиной отказа узла могут быть неверные данные, поступившие со входа. Мажоритар, подверженный импульсам SET, может стать причиной информационных отказов в подключенных к нему узлах, если сбой в мажоритаре будет далее запомнен. Как и с КС внутри узла, вероятность запоминания сбоя зависит от многих факторов. Будем считать, что мы можем оценить эту вероятность, и будем учитывать только те сбои на выходе мажоритаров, которые будут запомнены узлами, т. е. введем поправочный коэффициент. Тогда интенсивность потока «запоминаемых» сбоев на выходе мажоритара можно оценить величиной, пропорциональной количеству логических элементов в мажоритаре, как $\lambda_{маж} = kN_{л.э\_маж}$.

Количество элементов в мажоритаре пропорционально его разрядности и зависит от типа мажоритара (побитный или пословный, с формированием признака ошибки или без).

Как сказано выше, восстановление информации в узле происходит за счет загрузки в него правильной информации. Для того чтобы на вход узла поступала правильная информация, отказы в узлах-источниках должны отсутствовать или быть блокированными. Если блокирующий узел (мажоритар) отсутствует, то узел-источник и узел-приемник можно рассматривать как один узел с конвейерной структурой (с соответствующим периодом обновления информации). Поэтому будем считать, что узлы в устройстве соединены только через мажоритары.

Далее будем рассматривать системы, для которых интенсивности возникновения отказов в узлах и мажоритарах известны (т. е. они войдут в формулы для расчета надежности как параметры).

Оценка надежности отдельного узла с троированным мажоритаром

Рассмотрим систему (рис. 2), которая состоит из троированного узла M (module) и троированного мажоритара V (voter). Пусть интенсивность потока отказов узла составляет λ_m, а интенсивность потока отказа мажоритара — λ_v.

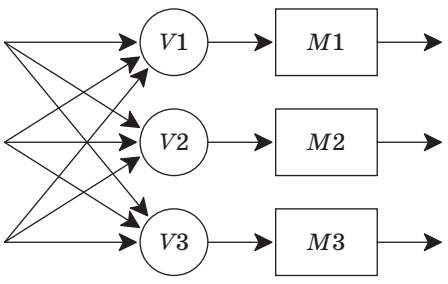
Зададим схему событий, позволяющую оценивать исправность работы системы. Будем считать, что:

— система исправна, если хотя бы два из трех экземпляров узла M выдают правильное значение;

— узел выдает правильное значение тогда и только тогда, когда он сам исправен и на его вход поступают правильные данные;

— мажоритар выдает правильное значение тогда и только тогда, когда он сам исправен и хотя бы на два его входа поступают правильные данные.

Вероятность исправности узла в некоторый момент времени t составляет $P_m(t) = e^{-\lambda_m t}$.



■ ***Рис. 2.*** Троированный узел с троированным мажоритаром

Вероятность исправности мажоритара

$$P_v(t) = e^{-\lambda_v t}.$$

Нетрудно показать для приведенной схемы событий, что вероятность исправности системы

$$P(t) = 3P_s(t)^2 - 2P_s(t)^3,$$

где $P_s(t) = P_m(t) \cdot P_v(t)$. Выражая вероятность исправного состояния системы через $\lambda_s = \lambda_m + \lambda_v$, получим $P_s(t) = 3e^{-2\lambda_s t} - 2e^{-3\lambda_s t}$.

Рассмотрим теперь систему, у которой информация в узлах восстанавливается в начале каждого интервала длительностью t_R. Восстановление происходит за счет цикличности работы узла.

Очевидно, что восстановление произойдет только в том случае, если в момент восстановления на узел поступает правильная информация с соответствующего мажоритара. Физическая природа отказов мажоритара (проявляющихся как «иголки» на его выходе) позволяет следующим образом учитывать их при анализе: отказ мажоритара, возникший во время восстановления, мы заменяем отказом, возникшим «сразу после» восстановления, что по влиянию на работу системы эквивалентно.

Оценим вероятность исправности системы в момент $T = Nt_R$. Поскольку возникновение отказа на некотором цикле восстановления не зависит от отказов на предыдущем цикле, $P(T) = P(t_R)^N$.

Для рассматриваемых класса систем и природы отказов $\lambda_S t_R << 1$, при этом N очень велико, и вычислить $P(T)$ напрямую невозможно из-за ограничения точности представления чисел в ЭВМ.

Воспользовавшись тем, что $\lambda_S t_R << 1$, построим приближенную оценку $P(T)$:

$$P(T) = P(t_R)^N = e^{\ln(P(t_R))N};$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

При $|x| < H$ остаточный член $\gamma_n(x) < \frac{e^H}{(n+1)!} x^{n+1}$. Тогда

$$P_s(t_R) = 3e^{-2\lambda_s t_R} - 6e^{-3\lambda_s t_R} = 3 - 6\lambda_s t_R + 6(\lambda_s t_R)^2 -$$
$$- 2 + 6\lambda_s t_R - 9(\lambda_s t_R)^2 + O((\lambda_s t_R)^3) \approx 1 - 3(\lambda_s t_R)^2.$$

Если ограничиться степенью 2 в разложении, то погрешность оценки не превысит

$$3\frac{e^H}{6} \cdot 8(\lambda_s t_R)^3 + 2\frac{e^H}{6} \cdot 27(\lambda_s t_R)^3 = 13e^H(\lambda_s t_R)^3.$$

При $\lambda_s t_R < 0{,}001$ погрешность меньше $14(\lambda_s t_R)^3$.

Воспользуемся разложением функции ln в степенной ряд: $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n}$. Обозначив $\lambda_s t_R$ через x и учитывая, что q мало, оставим только первый член ряда: $\ln(1 + x) \approx x$. При этом погрешность составит менее $x^2/2$. Тогда

$$P(T) = e^{\ln(1-3(\lambda_s t_R)^2)N} \approx e^{-3\lambda_s^2 t_R^2 \cdot T/t_R} = e^{-3\lambda_s^2 t_R T} \quad (1)$$

и систему можно рассматривать как элемент с простейшим потоком отказов с интенсивностью $3\lambda_s^2 t_R$.

Выражение (1) показывает, что интенсивность потока отказов прямо пропорциональна периоду восстановления информации.

На основании этого можно сделать вывод, что при определении уровня резервирования и декомпозиции системы на резервируемые узлы необходимо учитывать период восстановления информации в каждом узле и стремиться минимизировать этот период.

При $a, b > 0$ $(a + b)^2 > a^2 + b^2$. Если считать, что интенсивность потока отказов узла пропорциональна его размеру, и не учитывать влияние отказов в мажоритаре на надежность системы, то из выражения $P(T) = e^{-\left(3\lambda_e^2 t_R\right)T}$ следует, что максимальная надежность достигается при минимальном уровне резервирования.

Минимальный уровень — это отдельные триггеры: каждый триггер троируется, и на вход триггеров поступают данные, вычисленные по результатам мажорирования. Запись в триггеры должна производиться на каждом такте: если алгоритмом работы устройства запись новых данных на некотором такте не предусмотрена, в триггер записывается результат мажорирования старых значений в тройке триггеров (это дополнительный мультиплексор).

Описанный вариант резервирования требует максимального количества мажоритаров, максимально ухудшает быстродействие устройства за счет наличия мажоритаров и мультиплексоров на всех критических путях. Кроме того, из-за максимального количества мажоритаров наиболее сложной становится схема сбора данных об отказах.

Кроме того, известно [6, 7], что при снижении проектной нормы увеличивается вероятность одновременного поражения нескольких близлежащих триггеров, что снижает эффективность резервирования на уровне отдельных триггеров.

Необходимо выбирать оптимальный уровень резервирования с учетом периода восстановления и суммарной сложности мажоритаров.

Для принятия решений в процессе проектирования при рассмотрении различных вариантов резервирования узлов необходимо построить оценку надежности системы как сети из резервированных узлов с различными периодами восстановления.

Заключение

Предложен подход, позволяющий оценить надежность цифровых СБИС к информационным отказам, вызванным радиационными воздействиями, основанный на анализе цикличности функционирования узлов. В рамках данного подхода предлагается рассматривать устройство как сеть из групп резервированных узлов, соединенных через мажоритары. В каждом узле выделяются регистровая и комбинаторная составляющие. Обосновывается выбор параметров для оценки надежности этих составляющих. Для предложенной структурной модели получена оценка зависимости вероятности безотказной работы резервированного узла от периода восстановления информации. Полученная оценка показывает, что организация циклической работы резервированного узла в составе системы существенно повышает показатели надежности. Результат при учете параметров мажоритаров, связывающих узлы, может быть обобщен на устройство (систему) в целом. Многократное восстановление после информационных отказов за время наработки на невосстанавливаемый отказ устройства позволяет организовать ведение статистики информационных отказов и определить приближение к неустранимому отказу из-за накопления изменений в полупроводниковой структуре.

Литература

1. **Максименко С. Л., Мелехин В. Ф., Филиппов А. С.** Анализ проблемы построения радиационно-стойких информационно-управляющих систем // Информационно-управляющие системы. 2012. № 2. С. 18–25.
2. **Hass K. J., Ambles J. W.** Single Event Transients in Deep Submicron CMOS // 42nd Midwest Symp. on Circuits and Systems. 2000. Vol. 1. P. 122–125.
3. **Глухих М. И., Максименко С. Л., Мелехин В. Ф., Филиппов А. С.** Организация и проектирование высоконадежных вычислительных систем // Научно-технические ведомости СПбГПУ. 2011. № 6.1(138). С. 54–61.
4. **Rao R. R., Chopra K., Blaauw D. T., Sylvester D. M.** Computing the Soft Error Rate of a Combinational Logic Circuit Using Parameterized Descriptors // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2006. Vol. 26. N. 3. P. 468–479.
5. **Benedetto J. M.** et al. Variation of Digital SET Pulse Widths and the Implications for Single Event Hardening of Advanced CMOS Processes // IEEE Transactions on Nuclear Science. 2005. Vol. 52. P. 2114–2119.
6. **Amusan O. A.** et al. Single event upsets in deep-submicrometer technologies due to charge sharing // IEEE Transactions on Device and Materials Reliability. 2008. Vol. 8. N 3. P. 582–589.
7. **Haghi M., Draper J.** The 90 nm Double-DICE storage element to reduce Single-Event upsets // IEEE Intern. Midwest Symp. on Circuits and Systems. 2009. P. 463–466.

УДК 519.71

ПОИСК НЕИСПРАВНОСТЕЙ В БОРТОВЫХ СИСТЕМАХ УПРАВЛЕНИЯ В ПРОЦЕССЕ ПРИЕМОЧНОГО КОНТРОЛЯ

***В. А. Смирнов*[1],**
аспирант
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассмотрена методика выбора оптимальной последовательности процедур тестирования электронных блоков, входящих в бортовую систему управления. Определены случаи целесообразного использования предлагаемой методики. На простых примерах показана работа алгоритма поиска неисправного электронного блока.

Ключевые слова *— сложная техническая система, диагностирование, байесовская сеть доверия, тестирование, методика выбора, апостериорный вывод, пример решения задачи поиска.*

Введение

Современный этап развития бортовых автоматизированных систем управления (БАСУ) летательных аппаратов характеризуется повышением сложности как самих систем, так и их составных частей. Это усложняет процесс их контроля, эксплуатации и обслуживания. Значительная часть расходов при проверке работоспособности БАСУ определяется временем, необходимым для установления причин отказа и замены вышедших из строя ее составных элементов.

Вместе с тем большая продолжительность диагностических работ, возникающие в процессе электрических испытаний отказы элементов системы, а также необходимость повторять отдельные этапы, а иногда и циклы испытаний вынуждают искать пути оптимизации процессов контроля и диагностирования таких систем. Целью оптимизации является уменьшение времени проведения испытаний при сохранении неизменного качества контроля. Разработка новых и совершенствование известных эффективных методов, методик и алгоритмов диагностирования является актуальной и практически востребованной задачей. Ее актуальность возрастает по мере усложнения объектов контроля и условий их функционирования, а также с ростом несоответствия между уровнем сложности технических систем и диагностическими возможностями методов и средств, применяемых для поиска дефектов.

Проблемы технической диагностики электронных приборов решали как российские, так и зарубежные ученые. Известны работы Пархоменко П. П., Карибского В. В., Согомоняна Е. С., Каравая М. Ф., Лобанова А. В., Кузнецова П. И., Schlichting R., Rennels D. A., Dolev D. и многих других [1, 2]. В последние годы количество публикаций, посвященных вопросам оценки работоспособности и поиска дефектов, заметно увеличилось. Дальнейшее развитие методы и алгоритмы диагностирования сложных систем получили в работах Тулупьева А. Л., Нечаева Ю. И., Дегтярева А. Б., Портнягина Н. Н., Пюкке Г. А., Бидюка П. И., Фефелова А. О., Абдуллаева П. Ш., Жернакова С. В. и др. [3–6]. Анализ опубликованных работ показывает, что в настоящее время одним из основных направлений развития систем контроля и диагностики является совершенствование процессов обработки информации с привлечением новых методов анализа данных, поддержки принятия решений, формализации и решения задач диагностирования в условиях неопределенности, которые дополняют и развивают классические статистические методы исследований. Ключевым направлением совершенствования систем диагностирования становится интеграция результатов различных подходов и направлений, построение моделей, которые более гибко и адекватно описывают интеллектуальное поведение в условиях неопределенности и неполноты информации.

Целью работы является рассмотрение путей и методик повышения качества диагностирования

[1]Научный руководитель — доктор технических наук, профессор, заведующий кафедрой микро- и нанотехнологий аэрокосмического приборостроения Санкт-Петербургского государственного университета аэрокосмического приборостроения *В. П. Ларин*.

БАСУ в реальном времени, расширение функциональных возможностей средств технической диагностики за счет использования современного метода анализа данных. Поставленная цель достигается путем решения следующих задач:

— определение основных особенностей, присущих исследуемой технической системе, с точки зрения ее контроля и диагностирования;

— построение модели объекта диагностирования, разработка методики выбора последовательности проведения проверок, которая обеспечивает минимальную затрату ресурсов на поиск неисправного электронного блока технической системы;

— выбор оптимальной последовательности проведения процесса контроля и диагностирования при неполном обнаружении неисправности;

— демонстрация на практических примерах возможности использовать данный подход для решения диагностических задач.

Основные особенности БАСУ с точки зрения ее контроля и диагностирования

В основу методики диагностирования рассматриваемой БАСУ положен иерархический принцип постепенного увеличения глубины поиска неисправностей в соответствии с необходимостью реализации заданного уровня глубины поиска и возможностью деления объекта диагностирования на уровни составных частей по степени конструктивной сложности. Данная техническая система имеет встроенную систему самодиагностики, которая осуществляет проверку после каждого включения БАСУ и основным назначением которой является самодиагностика во время наземной подготовки под управлением корабельной автоматизированной системы управления. Комплекс проверок самодиагностики является достаточным, но не таким полным, как в контрольно-проверочной аппаратуре БАСУ, которая используется при изготовлении, эксплуатации и обслуживании. Контрольно-проверочная аппаратура представляет собой автоматизированную многопроцессорную контрольно-измерительную систему, построенную на базе специализированных промышленных компьютеров, осуществляющих управление процессом контроля — выдачей в объект контроля стимулирующих воздействий и анализом принимаемой информации, содержащей ответную реакцию.

Алгоритм диагностирования состоит из определенной совокупности элементарных проверок, а также правил, устанавливающих последовательность реализации элементарных проверок, и правил анализа результатов последних. Каждая проверка состоит из определенного количества проверяемых параметров. Численные значения измеренных параметров в зависимости от попадания в интервал допустимых значений получают лингвистическую оценку «годен» или «не годен». Если хотя бы один из параметров элементарной проверки получил оценку «не годен», то результату проверки присваивается лингвистическая оценка «не в норме». Дальнейшая локализация неисправностей производится с использованием ремонтно-эксплуатационной документации по таблице, где приведены перечень возможных неисправностей в процессе эксплуатации и рекомендации по действиям при их возникновении. Выявленный неисправный блок необходимо демонтировать из БАСУ. Дальнейшее диагностирование целесообразно выполнять на отдельном, специализированном для этого блока, стенде диагностики в целях локализации неисправностей.

Одним из недостатков существующей контрольно-проверочной аппаратуры является невозможность локализовать неисправность до конкретного блока при некоторых сочетаниях итогов проверок.

Выбор, обоснование и построение модели объекта диагностирования

На сегодняшний день одной из наиболее подходящих моделей, предназначенных для работы с неполной и неточной информацией при диагностировании сложных систем, являются байесовские сети доверия (БСД).

Несмотря на то, что байесовским сетям уделяется много внимания в зарубежной литературе, принципы их построения, обучения и использования применительно к решению задач диагностирования еще недостаточно освещены в отечественных публикациях.

В технической диагностике вероятностный подход, основанный на использовании математического аппарата байесовских сетей, может быть вполне эффективен для решения задачи поиска места и типа отказа технической системы. При этом модели событий и процессов графически представляются в виде БСД на основе объединения некоторых положений теории вероятностей и теории графов. Граф в БСД ациклический, т. е. в нем отсутствуют направленные циклы. Граф состоит из узлов и дуг, которые соединяют эти узлы. Узлы представляют собой случайные переменные, которые могут быть дискретными или непрерывными. Дуги отображают причинно-следственные связи между переменными, благодаря чему БСД еще иногда называют причинно-следственными сетями. В причинно-следственных сетях родительские вершины представляют собой причины (гипотезы), а дочерние — следствия (сви-

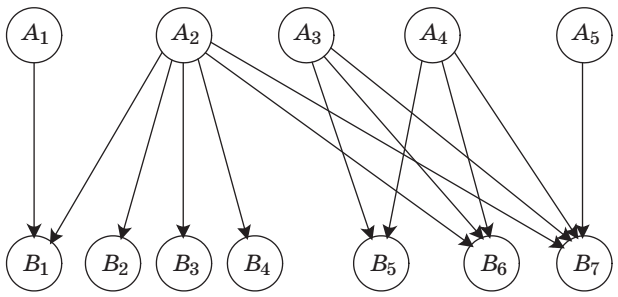
детельства). Каждой переменной сети ставится в соответствие таблица условных вероятностей, в которой перечислены вероятности всех возможных значений этой переменной при условии принятия тех или иных значений ее родителями.

При построении БСД была выполнена следующая последовательность действий: анализ процесса диагностирования системы, генерация топологии сети, определение условных вероятностей для сетевых связей. В результате анализа ситуаций, когда возникают неопределенности при локализации неисправностей с точностью до конкретного блока, были определены конкретные электронные блоки и тестовые проверки, связанные с таким положением, и установленные причинно-следственные связи между ними.

Байесовская сеть доверия была построена таким образом, что в корневых узлах находятся ненаблюдаемые переменные (электронные блоки), а наблюдаемые (тестовые проверки) располагаются в нижнем уровне сети. Между узлами сети установлены причинно-следственные связи.

Оценка безусловных вероятностей для ненаблюдаемых и условных вероятностей для наблюдаемых переменных сети была получена на основе анализа имеющейся базы данных результатов предыдущих испытаний и опроса опытных специалистов (экспертов). Здесь следует сказать, что данной оценке подвергаются лишь те переменные, которые имеют непосредственное влияние (прямую связь) на своего потомка. Результаты оценки выражаются в виде таблиц условных вероятностей, в которых перечислены вероятности всех возможных значений текущей переменной при условии принятия всех возможных значений ее родительскими переменными.

В приведенной на рис. 1 БСД диагностирования фрагмента БАСУ A_1 — радиовысотомер, A_2 — прибор преобразования информации БЦВМ и смежных систем, A_3 — усилитель рулевого агрегата, A_4 — прибор коммутации, A_5 — датчик угловых скоростей, $B_1 ... B_7$ — тестовые проверки № 1...7. Целевым состоянием узлов сети является неработоспособное состояние и отрицательный результат тестовой проверки.



■ ***Рис. 1.*** БСД диагностирования фрагмента БАСУ

■ ***Таблица 1.*** Значения условных вероятностей $P(B_1|A_1, A_2)$

A_1	A_2	$P(B_1=1\|A_1, A_2)$	$P(B_1=0\|A_1, A_2)$
1	1	0,95	0,05
1	0	0,95	0,05
0	1	0,15	0,85
0	0	0,05	0,95

Переменные в узлах БСД являются булевыми. Значения переменных A_1, ..., A_5, соответствующие неработоспособному состоянию, равняются 1, а работоспособному, соответственно, 0. Значения переменных B_1, ..., B_7, соответствующие отрицательному результату тестовой проверки, равняются 1, а положительному результату, соответственно, 0. На рис. 1 дано графическое представление БСД, однако это только качественное представление. Количественным представлением БСД является множество таблиц условных вероятностей.

Например, для переменной B_1 (тестовая проверка № 1) таблица условных вероятностей выглядит следующим образом (табл. 1).

Методика выбора оптимальной последовательности проведения процедур тестирования электронных блоков

Предлагаемая методика конкретизирует отдельные положения известного метода [6] и распространяет его на новый класс технических систем. Диагностическая модель, построенная на основе БСД, используется для вывода суждений, основанных на поступившей информации о результатах прохождения тестовых проверок. По результатам тестирования в диагностической модели происходит установка значений переменных B_1, ..., B_7, соответствующих результатам тестовой проверки. Вывод суждений делается на основе изменения степеней доверия к другим случайным переменным A_1, ..., A_5, соответствующим состоянию неработоспособности каждого электронного блока. Следовательно, можно сказать, что информация, приходящая в наблюдаемые переменные, распространяется внутри байесовской сети и изменяет вероятностные распределения ненаблюдаемых переменных. Вычисление вероятностей неработоспособности электронных блоков дает возможность их ранжировать и сравнивать. Диагностирование начинают с того узла, у которого вероятность отказа является максимальной. Выбор такого блока означает принятие решения о его демонтаже из БАСУ и проведении дальнейшего диагностирования на отдельном специализированном стенде.

Алгоритм поиска отказавшего электронного блока БАСУ конкретизирует известный обобщенный алгоритм поиска отказа технической системы [6]. Он представляет собой следующую последовательность шагов.

Шаг 1. Получение информации от специализированного исполняющего процессора контрольно-проверочной аппаратуры о результатах тестовых проверок. Установка соответствующих значений переменных B_1, ..., B_7 БСД.

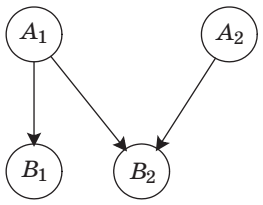
Шаг 2. Вычисление апостериорных вероятностей отказа электронных блоков БАСУ A_1, ..., A_5 в соответствии с установленными значениями переменных B_1, ..., B_7.

Шаг 3. Сортировка вычисленных вероятностей отказа электронных блоков БАСУ A_1, ..., A_5 в направлении убывания их значений.

Шаг 4. Выполнение тестовых проверок на специализированном стенде, начиная с того электронного блока БАСУ, чья вероятность отказа характеризуется наибольшим значением.

Пример решения задачи поиска отказа фрагмента БАСУ

Рассмотрим пример поиска отказа фрагмента БАСУ, модель которого в виде БСД представлена на рис. 2, где A_1 — радиовысотомер, A_2 — прибор преобразования информации БЦВМ и смежных систем, B_1, B_2 — тестовая проверка № 1 и 2 соответственно. Значения вероятностей, указанные в табл. 2–5, получены на основе анализа имеющейся базы данных результатов предыдущих испытаний и опроса опытных специалистов (экспертов).



■ ***Рис. 2.*** Пример байесовской сети диагностирования технической системы

■ ***Таблица 2.*** Значения априорных вероятностей $P(A_1)$

$P(A_1 = 1)$	$P(A_1 = 0)$
0,11	0,89

■ ***Таблица 3.*** Значения априорных вероятностей $P(A_2)$

$P(A_2 = 1)$	$P(A_2 = 0)$
0,1	0,9

■ ***Таблица 4.*** Значения условных вероятностей $P(B_1|A_1)$

B_1	$P(B_1\|A_1 = 1)$	$P(B_1\|A_1 = 0)$
1	0,96	0,04
0	0,04	0,96

■ ***Таблица 5.*** Значения условных вероятностей $P(B_2|A_1, A_2)$

B_2	$P(B_2\|A_1, A_2 = 1)$		$P(B_2\|A_1, A_2 = 0)$	
	$A_1 = 1$	$A_1 = 0$	$A_1 = 1$	$A_1 = 0$
1	0,94	0,6	0,9	0,03
0	0,06	0,4	0,1	0,97

Пример 1. По результатам тестовых проверок переменные B_1 и B_2 принимают следующие значения: $B_1 = 1$, $B_2 = 1$, т. е. оба теста показали наличие неисправностей. Необходимо вычислить вероятности $A_1 = 1$ и $A_2 = 1$, т. е. вероятности неработоспособности радиовысотомера и прибора преобразования информации БЦВМ и смежных систем.

Согласно формуле, предложенной С. А. Тереховым [7, с. 160], вероятность совместного появления событий A_1, A_2, B_1, B_2 равна

$$P(A_1, A_2, B_1, B_2) = P(A_1)\cdot P(A_2|A_1)\times$$
$$\times P(B_2|A_1, A_2)\cdot P(B_1|A_1, A_2, B_2).$$

В БСД (см. рис. 2) ориентированные ребра графа отражают те вероятности, которые реально имеют место в данном примере. Так как A_2 не зависит от A_1, а B_1 не зависит от A_2, B_2, то это позволяет нам представить совместное распределение вероятностей более компактно:

$$P(A_1, A_2, B_1, B_2) =$$
$$= P(A_1)\cdot P(A_2)\cdot P(B_2|A_1, A_2)\cdot P(B_1|A_1).$$

Воспользуемся другой формулой [7, с. 162] и вычислим вероятности неработоспособности приборов A_1 и A_2, просуммировав совместное распределение по означиванию всех остальных переменных:

$$P(A_1 = 1|B_1 = 1, B_2 = 1) = \frac{P(A_1 = 1, B_1 = 1, B_2 = 1)}{P(B_1 = 1, B_2 = 1)} =$$
$$= \frac{\sum\limits_{A_2=\{1,0\}} P(A_1 = 1, A_2, B_1 = 1, B_2 = 1)}{\sum\limits_{\substack{A_1=\{1,0\}\\A_2=\{1,0\}}} P(A_1, A_2, B_1 = 1, B_2 = 1)};$$

$$\sum_{A_2=\{1,0\}} P(A_1 = 1, A_2, B_1 = 1, B_2 = 1) =$$
$$= P(A_1)\cdot P(A_2 = 1)\cdot P(B_2|A_1, A_2 = 1)\times$$
$$\times P(B_1|A_1) + P(A_1)\cdot P(A_2 = 0)\times$$
$$\times P(B_2|A_1, A_2 = 0)\cdot P(B_1|A_1) = 0{,}11\cdot 0{,}1\cdot 0{,}94\times$$
$$\times 0{,}96 + 0{,}11\cdot 0{,}9\cdot 0{,}9\cdot 0{,}96 = 0{,}0954624;$$

$$\sum_{\substack{A_1=\{1,0\}\\A_2=\{1,0\}}} P(A_1, A_2, B_1=1, B_2=1) = P(A_1=1)\times$$
$$\times P(A_2=1)\cdot P(B_2|A_1=1, A_2=1)\cdot P(B_1|A_1=1)+$$
$$+ P(A_1=0)\cdot P(A_2=1)\cdot P(B_2|A_1=0, A_2=1)\times$$
$$\times P(B_1|A_1=0)+P(A_1=1)\cdot P(A_2=0)\times$$
$$\times P(B_2|A_1=1, A_2=0)\cdot P(B_1|A_1=1)+$$
$$+ P(A_1=0)\cdot P(A_2=0)\cdot P(B_2|A_1=0, A_2=0)\times$$
$$\times P(B_1|A_1=0)=0{,}11\cdot 0{,}1\cdot 0{,}94\cdot 0{,}96+0{,}89\cdot 0{,}1\times$$
$$\times 0{,}6\cdot 0{,}04+0{,}11\cdot 0{,}9\cdot 0{,}9\cdot 0{,}96+0{,}89\cdot 0{,}9\times$$
$$\times 0{,}03\cdot 0{,}04=0{,}0985596;$$

$$P(A_1=1|B_1=1, B_2=1)=\frac{0{,}0954624}{0{,}0985596}=$$
$$=0{,}968575359\approx 0{,}97;$$

$$P(A_2=1|B_1=1, B_2=1)=\frac{P(A_2=1, B_1=1, B_2=1)}{P(B_1=1, B_2=1)};$$

$$P(A_2=1, B_1=1, B_2=1)=$$
$$=\sum_{A_1=\{1,0\}} P(A_1, A_2=1, B_1=1, B_2=1)=P(A_1=1)\times$$
$$\times P(A_2)\cdot P(B_2|A_1=1, A_2)\cdot P(B_1|A_1=1)+P(A_1=0)\times$$
$$\times P(A_2)\cdot P(B_2|A_1=0, A_2)\cdot P(B_1|A_1=0)=$$
$$=0{,}11\cdot 0{,}1\cdot 0{,}94\cdot 0{,}96+0{,}89\cdot 0{,}1\cdot 0{,}6\cdot 0{,}04=0{,}0120624;$$

$$P(A_2=1|B_1=1, B_2=1)=\frac{0{,}0120624}{0{,}0985596}=$$
$$=0{,}12238686\approx 0{,}12.$$

Таким образом, при отрицательных результатах прохождения тестовых проверок B_1 и B_2 значение вероятности неработоспособности радиовысотомера больше, чем прибора преобразования информации БЦВМ и смежных систем. Следовательно, радиовысотомер должен быть демонтирован из БАСУ для проведения дальнейших тестовых проверок на специализированном стенде.

Пример 2. По результатам тестовых проверок переменные B_1 и B_2 принимают следующие значения: $B_1 = 0$, $B_2 = 1$. Необходимо вычислить вероятности $A_1 = 1$ и $A_2 = 1$:

$$P(A_1=1|B_1=0, B_2=1)=\frac{P(A_1=1, B_1=0, B_2=1)}{P(B_1=0, B_2=1)}=$$
$$=\frac{0{,}0039776}{0{,}0783104}=0{,}050792742\approx 0{,}05;$$

$$P(A_2=1|B_1=0, B_2=1)=\frac{P(A_2=1, B_1=0, B_2=1)}{P(B_1=0, B_2=1)}=$$
$$=\frac{0{,}0516776}{0{,}0783104}=0{,}65990724\approx 0{,}66.$$

При заданных результатах тестовых проверок должен быть демонтирован из БАСУ прибор преобразования информации БЦВМ и смежных систем для проведения дальнейших тестовых проверок на специализированном стенде.

Заключение

Определены основные особенности БАСУ с точки зрения ее контроля и диагностирования. Предложена методика оценки состояния диагностируемой аппаратуры и алгоритм оптимального обнаружения отказавшего электронного блока, входящего в БАСУ, с помощью БСД. Методика дает возможность уменьшить затраты времени и средств на проведение дальнейших тестовых проверок. Показана на практических примерах возможность использования данного подхода для решения диагностических задач. Экспериментальные исследования подтверждают эффективность разработанной методики, которая может быть реализована в качестве отдельного компонента системы встроенного диагностирования БАСУ, контрольно-проверочной аппаратуры БАСУ, а также специализированных стендов диагностирования отдельных блоков, входящих в БАСУ. Данная методика может быть применена в других отраслях промышленности.

Литература

1. **Пархоменко П. П., Согомонян Е. С.** Основы технической диагностики. — М.: Энергия, 1981. — 320 с.
2. **Кузнецов П. И., Пчелинцев Л. А., Гайденко В. С.** Контроль и поиск неисправностей в сложных системах. — М.: Сов. радио, 1969. — 240 с.
3. **Тулупьев А. Л., Николенко С. И., Сироткин А. В.** Байесовские сети: Логико-вероятностный подход. — СПб.: Наука, 2006. — 607 с.
4. **Портнягин Н. Н., Пюкке Г. А.** Теория и методы диагностики судовых электрических средств автоматизации / КамчатГТУ. — Петропавловск-Камчатский, 2003. — 112 с.
5. **Нечаев Ю. И., Дегтярев А. Б., Сиек Ю. Л.** Принятие решений в интеллектуальных системах реального времени с использованием концепции мягких вычислений // Искусственный интеллект. 2000. № 3. С. 525–533.
6. **Фефелов А. А.** Использование байесовских сетей для решения задачи поиска места и типа отказа сложной технической системы // Автоматика. Автоматизация. Электротехнические комплексы и системы. 2007. № 2(20). С. 87–93.
7. **Терехов С. А.** Введение в байесовы сети: лекции по нейроинформатике / МИФИ. — М., 2003. Ч. 1. — 188 с.

УДК 005.8:615.478

МЕТОД ОЦЕНКИ РИСКОВ В МУЛЬТИАГЕНТНОЙ СИСТЕМЕ УПРАВЛЕНИЯ ПРОЕКТАМИ НИР И ОКР В РЕАЛЬНОМ ВРЕМЕНИ

Е. М. Клейменова,
руководитель
А. Л. Феоктистов,
заместитель Генерального конструктора по информационным технологиям
НТЦ «Корпоративные информационные технологии», ОАО «РКК «Энергия», г. Королев
П. О. Скобелев,
доктор техн. наук, ведущий научный сотрудник
Институт проблем управления сложными системами РАН, г. Самара
В. Б. Ларюхин,
директор по разработкам
И. В. Майоров,
ведущий специалист научно-исследовательского отдела
Е. В. Симонова,
канд. техн. наук, доцент, ведущий аналитик
ООО «НПК «Разумные решения», г. Самара
Е. В. Полончук,
начальник отдела
НТЦ «Корпоративные информационные технологии», ОАО «РКК «Энергия», г. Королев

Предлагаются удобные для практики количественная модель и метод оценки рисков проектной деятельности по срокам выполнения, позволяющие интерактивно учитывать ход выполнения проектов НИР и ОКР в мультиагентной системе управления проектами в реальном времени. Агенты представляют подразделения, проекты, задачи и сотрудников, причем задачи каждого проекта являются связанными и распределяются на общем поле ресурсов подразделений. Модель основана на линейной аппроксимации вероятностных распределений и вычислении рисков по задачам, связанным отношениями следования. Метод представляет собой порядок расчетов, позволяющий оценивать риск и перепланировать цепочки связанных задач непосредственно в реальном времени, когда распределение задач по сотрудникам постоянно меняется в связи с непредвиденными событиями. Разработанный метод предназначен для снижения рисков при планировании проектов НИР и ОКР в аэрокосмических приложениях.

Ключевые слова *— управление проектами, оценка рисков, вероятностный подход, мультиагентные системы, адаптивное планирование, реальное время.*

Введение

Современные проекты научно-исследовательских и опытно-конструкторских работ (НИР и ОКР) в современной технике имеют все возрастающую организационную, техническую и управленческую сложность и, по определению, содержат высокую степень неопределенности и постоянной динамики распределения ресурсов, вызванную новизной создаваемых продуктов и возникающими непредвиденными событиями разного рода.

При этом задачи создания образцов новой техники, особенно в уникальных аэрокосмических приложениях, могут состоять из десятков и сотен этапов, разбиваться на тысячи подзадач и требовать сотен и тысяч исполнителей различной квалификации, которые должны работать согласованно. Чтобы координировать деятельность таких научных коллективов, необходимы

значительные управленческие усилия для достижения поставленных целей и получения практических результатов с заданным качеством, в установленные сроки, в ограниченный бюджет, с минимальными рисками. В условиях возрастающей сложности, уникальности и разнородности таких проектов адекватные программные средства должны учитывать влияние внешних условий, конъюнктуру рынка труда, изменения в календарных планах и составах исполнителей, возникновение новых задач в ходе выполнения проекта и т. д.

Однако гибко и эффективно реагировать на все эти требования позволяют лишь системы, функционирующие в реальном времени, поскольку в противном случае возникают длительные задержки, простои или, наоборот, дефицит ресурсов. Одной из важных задач управления проектной деятельностью при этом становится определение реалистичных сроков выполнения задач и постоянный учет рисков в целях минимизации влияния внешних неблагоприятных факторов, так как НИР и ОКР относятся как раз к той сфере, где риски особенно велики и адекватный их учет особенно востребован.

Управление проектной деятельностью осложняется тем, что существуют значительные неопределенности, связанные с внешними и внутренними факторами: неполнотой и возможной недостоверностью информации обо всех параметрах и обстоятельствах, затрудняющих выбор оптимального решения; принципиальной невозможностью адекватно и точно учитывать все доступные сведения; вероятностной сущностью поведения исследуемой среды и создаваемого продукта и т. п. Случайные внешние факторы обычно заранее просто невозможно корректно прогнозировать и предусмотреть даже в вероятностной интерпретации из-за большой степени неопределенности их влияния на результат.

С другой стороны, присутствие субъективных факторов, когда взаимодействие руководителей и исполнителей проектных работ напоминает ситуацию конкуренции и кооперации партнеров с несовпадающими интересами и целями на внутреннем рынке предприятия, еще более затрудняет процесс принятия управленческих решений, способствует росту неопределенности и риска.

В этой связи в настоящей работе предлагаются удобные для практики количественная модель и метод оценки рисков проектной деятельности по срокам выполнения, позволяющие интерактивно учитывать ход выполнения проектов НИР и ОКР в интеллектуальной системе управления проектами в реальном времени [1].

Разработанная система построена на основе мультиагентной технологии адаптивного планирования ресурсов [2–4], находящей все большее применение при управлении ресурсами в реальном времени.

В разработанной системе агенты представляют подразделения, проекты, задачи и сотрудников, причем задачи каждого проекта являются связанными и распределяются на общем поле ресурсов подразделений. При этом исполнитель не рассматривается как «винтик» в проекте, а участвует в определении и согласовании сроков работы, может откладывать или изменять длительность работы, вводить новые работы и т. д.

Соответствующая модель оценки рисков основана на линейной аппроксимации вероятностных распределений и вычислении рисков по задачам, связанным отношениями следования. Метод представляет собой порядок расчетов, позволяющий оценивать риск и перепланировать цепочки связанных задач непосредственно в реальном времени, когда распределение задач по сотрудникам постоянно меняется в связи с непредвиденными событиями (появлением нового проекта, задержками по этапам и т. д.).

Общая постановка и основные подходы к решению задачи

Будем далее полагать под неопределенностью в проектной деятельности неполноту и неточность исходной и текущей информации о стадиях в реализации проекта, в том числе о результатах и затратах.

Под риском понимается потенциальная, численно измеримая возможность неблагоприятных ситуаций и связанных с ними последствий в виде потерь, ущерба, убытков денежных средств в связи с неопределенностью, т. е. со случайным изменением условий выполнения проекта, а также возможность получения непредсказуемого результата в зависимости от принятых управленческих решений и действий.

Согласно теории и практике современного управления проектами (Project Management Body of Knowledge — PMBOK) [5], управление рисками включает в себя процессы, касающиеся идентификации, анализа и оперативного реагирования на риски, возникающие в проекте. Управление рисками должно приводить к максимизации положительных влияний и минимизировать отрицательные последствия. Основными процессами, согласно PMBOK, являются:

1) идентификация рисков — определение, какие риски воздействуют на проект, и документирование характеристик каждого из них;

2) качественная оценка — оценка условий возникновения рисков и их последствий; количественная оценка — оценка вероятностей рисков и влияния их на успех проекта;

3) разработка мер реагирования на риск — определение способов и методов ослабления отрицательных последствий;

4) мониторинг рисков — постоянное слежение за рисками, определение рисков, оставшихся к данному моменту, и оценка эффективности принятых мер.

Подробно рассмотренные в работе [6] методы оценки рисков следующие:

— количественная оценка рисков с помощью методов математической статистики;

— методы экспертной оценки рисков;

— методы имитационного моделирования рисков;

— комбинированные методы, представляющие собой объединение нескольких методов или их отдельных элементов.

В настоящей работе мы рассмотрим подход, позволяющий в реальном времени оценивать изменения рисков выполнения задач по мере ввода факта выполнения задач и сразу же осуществлять перепланирование, направленное на уменьшение рисков.

Этот подход важен для построения интеллектуальной системы, поддерживающей интерсубъективную теорию управления, учитывающую множество точек зрения участников процесса, действенную роль акторов и социальную самоорганизацию в коллективе по исполнению проектов [7].

Обзор методов учета рисков

В современных стандартах PMBOK признаны обязательными методы экспертной оценки и SWOT-метод, когда все факторы, влияющие на проект, делятся на категории: сильные стороны, слабые стороны, возможности и угрозы (strengths, weakness, opportunities, threats). Затем эксперты оценивают качественно или полуколичественно влияние этих факторов на вероятность успешного завершения проектов [8, 9].

Качественно-количественные шкалы для оценки вероятностей рисков и их последствий часто используются в методах построения матрицы рисков, когда в зависимости от степени опасности факторов и влияния их последствий составляют соответствующие таблицы. Матрицы рисков используются для анализа и прогнозов проектного планирования [10].

Для оценки последствий различных сценариев развития проектов применяется метод оценки, основанный на построении дерева решений. Такие деревья используются обычно в двух вариантах — деревья событий и деревья отказов. Метод дерева решений может быть применен в случае обозримого количества вариантов. Строятся узлы, представляющие собой основные события проекта, и ветви начала и окончания работ. Затем могут быть рассчитаны вероятности развития событий по каждому сценарию [11]. Такой подход часто применяется в оценке надежности промышленных объектов [12].

Оценка рисков в проектах может быть произведена на основе построения PERT-диаграмм (Project Evaluation Review Technique) и расчета наиболее вероятной продолжительности стадий работ на основе бета-распределения вероятностей и нахождения критического пути. Несмотря на длительный (с середины 1950-х гг.) период использования PERT-методов, в настоящее время они признаны рискованными сами по себе, поскольку дают слишком оптимистические оценки [13, 14].

Один из мощных методов количественного анализа и оценки рисков базируется на применении метода имитационного моделирования Монте-Карло. При моделировании этапов выполнения проекта оцениваются диапазоны результатов влияния различных случайных факторов на исследуемый проект при розыгрыше различных сценариев. Такая технология помогает предсказывать наиболее вероятные результаты внешних воздействий. Большое количество вариантов и факторов, потенциально влияющих на проект, трудно оценить без моделирования. Для решения этой проблемы анализируются наилучший и наихудший сценарии развития, после чего рассчитываются промежуточные сценарии. В большинстве случаев получаемые оценки выходных параметров соответствуют нормальному закону распределения вероятностей. Однако методы Монте-Карло довольно трудно использовать в реальных проектах из-за отсутствия достаточно удобных программных средств, сложности выбора факторов и их диапазонов [15, 16].

Значительный вклад в современные математические методы управления в организационных структурах и построения оценки рисков составляют работы [17, 18], где рассмотрены действенные механизмы выявления и реакции на риски, например самоопережающий контроль (стимулирование к раннему обнаружению рисков) или компенсационные реакции (направленные на повышение оплаты исполнителей или вовлечение новых исполнителей). Эти методы хорошо согласуются с мультиагентным подходом, поскольку могут применяться индивидуально к участникам в ходе выполнения проектов в реальном времени, но не рассчитаны на постоянное перевычисление рисков в ходе поступления непредвиденных событий.

В целом можно утверждать, что многие разработанные модели исходят из сложившейся классической централизованной и статической природы проектов, и, как следствие, адаптивный пе-

ресмотр плана проекта не происходит при возникновении непредвиденных событий. Кроме того, рассматриваются такие отношения центра и исполнителей, в которых лишь центр является активным, а исполнитель не способен к активному добавлению или изменению задач и не может динамически менять планы.

Модели и методы, рассматриваемые в статье, напротив, используются при построении системы, которая призвана позволять оценивать риски в ходе выполнения проекта и на основе этих оценок перепланировать задачи в реальном времени.

Предлагаемый подход к оценке рисков

В настоящем подходе в качестве основы мы используем разработанную интеллектуальную систему управления проектами НИР и ОКР в реальном времени [1], реализованную на основе концепции сетей потребностей и возможностей [2–4].

Мультиагентные технологии в последнее время находят все больше применений и позволяют решать сложные и динамичные задачи адаптивного планирования в различных предметных областях — от грузовых перевозок и производственного планирования до управления грузопотоком Международной космической станции и роем спутников [19].

Разработанная система представляет собой новое поколение интеллектуальных систем управления ресурсами на основе адаптивного планирования, в котором реализуется полный цикл управления (планирование — исполнение — контроль — анализ) в реальном времени. При этом система ориентирована не на планирование «сверху-вниз», когда начальник все решает за подчиненных, а наоборот, стимулирует планирование «снизу-вверх», когда исполнители договариваются между собой. Механизмы адаптивного планирования позволяют более гибко реагировать на все поступающие события в реальном времени и предотвращать критические ситуации.

В разработанной системе на основе подхода, изложенного в работах [2–4], задачи каждого проекта представляются множеством агентов потребностей, которые непрерывно ищут наилучшее размещение на ресурсах, представленных агентами возможностей. В результате строится сеть связанных операций (задач), формирующая сеть потребностей и возможностей, открытую к изменениям под действием событий любого рода: приход нового проекта, приостановка или завершение существующего, отказ или задержка исполнителя, перенос задачи на более поздний срок и т. д.

Более подробно разработанный подход адаптивного планирования, реализованный в системе, рассматривается в работе [1].

Очевидно, что при таком подходе в реальном времени меняются и риски проекта, вследствие чего та часть проекта, где ранее все было в «зеленом цвете» (без риска), становится «в красном цвете» (получает риск) и наоборот, т. е. постоянно меняются оценки, которые должны вызывать адекватные изменения в планах, например, сложная задача должна быть переброшена на менее загруженный ресурс и т. д.

Нами предлагается метод аппроксимации вероятностей завершения проектов в срок, который позволяет приближенно количественно оценивать в реальном времени риски завершения НИР и ОКР для сокращения объема вычислений.

Математическая модель оценки риска для одного проекта

Рассмотрим множество проектов $Projects = \{Project_j\}$, $j = 1..n$, n — количество проектов. Проект состоит из множества подзадач $Subproblems = \{Subproblem_i\}$, $i = 1..m$, m — количество подзадач в проекте. Обозначим C_j — запланированный срок выполнения проекта j, d_j — предельный срок выполнения проекта, τ_j — реальный срок выполнения проекта, p_{ij} — длительность выполнения подзадачи i в проекте j.

Под риском r_j в задаче проектного планирования будем понимать вероятность выхода за предельный срок d_j реального времени выполнения τ_j данного проекта j:

$$r_j = P(\tau_j > d_j). \quad (1)$$

Длительность выполнения проекта определяется как время окончания последней подзадачи в проекте j, которое вычисляется по цепочке от первой подзадачи и зависит от всех предыдущих запланированных подзадач перед каждой подзадачей данного проекта на каждом ресурсе (исполнителе). Поэтому точное вычисление вероятности выхода за предельный срок потребует большого количества вычислений.

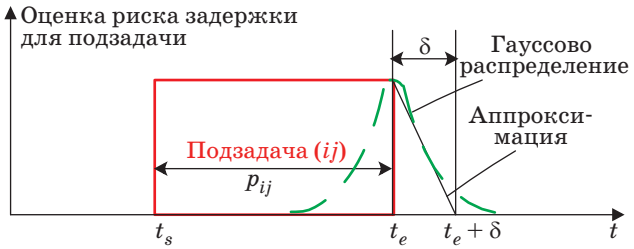
Рассмотрим некоторые упрощения, которые позволят приближенно вычислить значение вероятности.

Будем рассматривать один проект, $j = 1$. Предположим, что каждая подзадача проекта выполняется на отдельном ресурсе (исполнителе) i, множество исполнителей $Performers = (Performer_i)$, $i = 1..m$. Предположим, что разного рода неопределенные факторы, которые в принципе очень трудно или даже невозможно учесть, случайным образом влияют на среднюю продолжительность выполнения конкретной подзадачи и на время ее завершения, поэтому τ_j является случайной величиной. Согласно теории вероятностей, распределение времени окончания проекта при учете

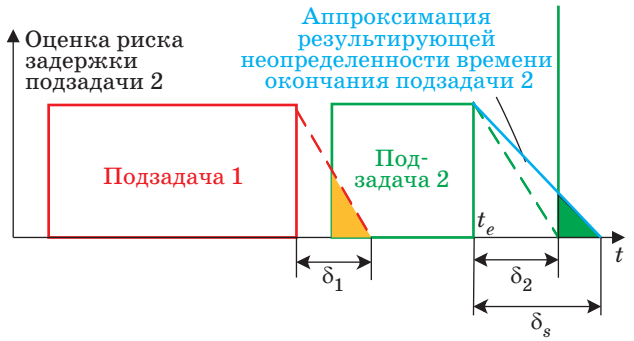
несистематических факторов подчиняется нормальному закону, который описывается гауссовой функцией распределения. Предположим также, что в целом запас времени исполнителей достаточен для выполнения каждой подзадачи в срок в идеальном случае, без учета несистематических факторов.

Оценим неопределенность времени выполнения подзадачи (т. е. той части распределения времени выполнения предыдущей подзадачи, с которой перекрывается следующая подзадача) за счет хвостов гауссовой функции распределения вправо от каждой подзадачи величиной $\delta(p_{ij}) \sim \sigma_{ij}$, σ_{ij} — стандартное отклонение, i — номер исполнителя, j — номер проекта. Таким образом, время выполнения каждой подзадачи описывается, кроме длительности p_{ij}, величиной неопределенности $\delta(p_{ij})$. Аппроксимируем хвост гауссовой функции с помощью линейной функции (рис. 1), при этом получим трапециевидную форму функции распределения. Считаем, что момент времени начала каждой подзадачи t_s достоверно известен.

Поскольку величина $\delta(p_{ij})$ оценивает интервал неопределенности времени окончания t_e подзадачи, начало следующей подзадачи может попасть в этот интервал. Поэтому при выстраивании цепочки подзадач вправо по оси времени суммарная неопределенность будет накапливаться (по теореме сложения дисперсий). Это означает, что время окончания последней подзадачи может выйти за предельный срок (рис. 2).



■ ***Рис. 1.*** Аппроксимация части гауссова распределения линейной функцией



■ ***Рис. 2.*** Аппроксимация распределения времени выполнения одного проекта

Отметим, что суммарное распределение времени выполнения подзадач может иметь очень сложную форму. Например, сумма двух равномерно распределенных величин имеет треугольное распределение, распределение суммы бесконечного количества произвольных случайных величин стремится к нормальному распределению согласно центральной предельной теореме теории вероятностей. Поэтому упрощаем рассуждение и считаем, что итоговое распределение также можно аппроксимировать трапецией, правый треугольник которой имеет основание, равное результирующей неопределенности цепочки подзадач δ_s. Например, для проекта, состоящего из двух подзадач, δ_s вычисляется по формуле

$$\delta_s = t_{e_1} + \delta_1 - t_{s_2}. \qquad (2)$$

Согласно рис. 2, где рассматривается единственный проект ($j = 1$), вероятность $P(\tau > d)$ выхода за предельный срок окончания проекта, состоящего из двух подзадач, равна площади закрашенного прямоугольного треугольника, одну из сторон которого образует плотность вероятности $f(t) = At + B$. Здесь A и B — коэффициенты, определяемые из условия прохождения графика $f(t)$ через точку $(0, t_e + \delta_s)$ и условия нормировки:

$$A(t_e + \delta_s) + B = 0; \qquad (3)$$

$$P = 1 = \int_{t_e}^{t_e+\delta_s} f(t)\mathrm{d}t = \int_{t_e}^{t_e+\delta_s} (At + B)\mathrm{d}t =$$

$$= \frac{A}{2}\delta_s(2t_e + \delta_s) + B\delta_s. \qquad (4)$$

Из (3) и (4) получаем значения коэффициентов $A = -\frac{2}{\delta_s^2}$ и $B = \frac{2}{\delta_s^2}(t_e + \delta_s)$. Тогда плотность вероятности

$$f(t) = \frac{2}{\delta_s^2}(t_e + \delta_s - t). \qquad (5)$$

Вероятность $P(\tau > d)$ выхода срока выполнения проекта за предельный срок вычисляется как площадь прямоугольного треугольника с катетами $f(d) = 2\frac{(t_e + \delta_s - d)}{\delta_s^2}$ и $(t_e + \delta_s - d)$:

$$P(\tau > d) = \int_{d}^{t_e+\delta_s} f(t)\mathrm{d}t = \left(1 - \frac{d - t_e}{\delta_s}\right)^2. \qquad (6)$$

Соответственно, риск r равен $\left(1 - \frac{d - t_e}{\delta_s}\right)^2$ при условии $t_e + \delta_s - d \geq 0$, т. е. предельный срок вы-

полнения проекта находится внутри интервала результирующей неопределенности.

Таким образом, риск выхода времени выполнения проекта за предельный срок определяется выражением

$$r=\begin{cases}1, & d\le t_e\\ \left(1-\dfrac{d-t_e}{\delta_s}\right)^2, & t_e\le d\le t_e+\delta_s.\\ 0, & t_e+\delta_s\le d\end{cases} \qquad (7)$$

С учетом (5) среднее время окончания проекта

$$\overline{t_e}=\int_{t_e}^{t_e+\delta_s} tf(t)\mathrm{d}t=\int_{t_e}^{t_e+\delta_s}\left(At^2+Bt\right)\mathrm{d}t=t_e+\frac{\delta_s}{3}. \qquad (8)$$

Следует отметить, что возрастающие значения неопределенности объясняют существование естественного горизонта планирования. Если результирующая неопределенность будет сравнима со средним временем выполнения подзадачи, дальнейшее планирование теряет смысл, так как накопившаяся неопределенность однозначно делает невыполнимым план последней подзадачи.

Математическая модель оценки риска для нескольких проектов

При рассмотрении нескольких проектов следует учитывать, что исполнитель может выполнять подзадачи различных проектов. На рис. 3 приведена диаграмма Гантта для двух исполнителей, представляющая график выполнения двух проектов, каждый из которых включает по две подзадачи.

Оценим нижнюю границу риска каждого проекта *r*(*Project*) упрощенным способом, учитывая максимум из двух значений: результирующей неопределенности по ресурсу (исполнителю) — δ(*Performer*) и результирующей неопределенности по подзадачам проекта — δ(*Subproblems*). На том ресурсе, где находится финальная подзадача данного проекта, последовательно суммируются перекрытия распределений времени выполнения всех подзадач, которые стоят до финальной подзадачи:

$$\delta(Performer)=\sum_{i\in Performer}^{all\,intersected}\delta_i'+\delta_{last}, \qquad (9)$$

где δ_i' — область перекрытия неопределенности предыдущей подзадачи со временем выполнения следующей подзадачи по исполнителю; δ_{last} — неопределенность последней подзадачи.

Аналогично подсчитывается неопределенность по последовательности подзадач в данном проекте:

$$\delta(Subproblems)=\sum_{j\in Project}^{all\,intersected}\delta_j''+\delta_{last}, \qquad (10)$$

где δ_j'' — область перекрытия неопределенности предыдущей подзадачи со временем выполнения следующей подзадачи по проекту.

Далее на основании δ(*Performer*) и δ(*Subproblems*) с использованием формулы (7) вычисляем риск по исполнителю *r*(*Performer*) и риск по цепочке подзадач проекта *r*(*Subpoblems*), после чего определяем риск проекта по формуле

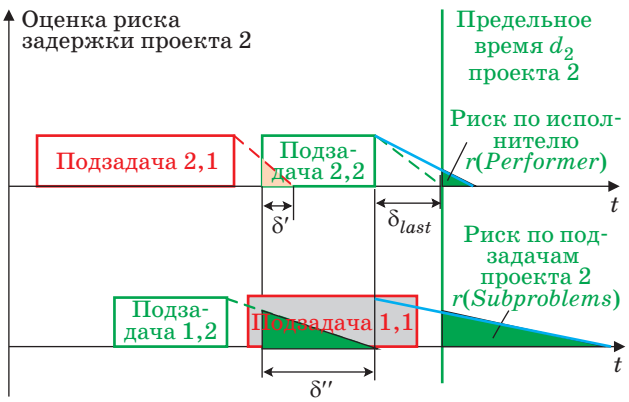
$$r(Project)=\max(r(Performer),r(Subproblems)). \qquad (11)$$

Следует отметить, что даже треугольная аппроксимация правой части функции вероятности потребовала бы пересчета цепочек распределений, поскольку распределения меняли бы свою дисперсию в сторону увеличения, при этом возникали бы дополнительные вычислительные сложности. Поэтому предлагается считать дисперсии распределений постоянными или оценивать их по статистике истории выполнения подобных задач и выбирать максимум из оценки по ресурсу и по последовательности подзадач.

Грубые оценки можно сделать, не учитывая перекрытия распределений, а вычисляя полную сумму дисперсий. Эту величину можно считать близкой к верхней оценке риска. Реальный риск будет больше, чем вычисленный, поэтому естественно предположить, что найдена нижняя оценка. Однако вполне возможно, что в реальности время выполнения подзадач может быть не только больше, но и меньше, чем среднее.

Можно предположить, что неопределенность во времени для последовательности подзадач различна для разных исполнителей, и чем меньше неопределенность, тем больше стоимость выполнения подзадачи конкретным исполнителем.

Таким образом, можно найти вероятности выхода проектов за предельный срок, используя расписание, существующее в текущий момент



■ ***Рис. 3.*** Аппроксимация распределения времени выполнения нескольких проектов

времени, например сразу после внесения в него отметок об очередном событии.

Далее должно быть принято решение о снижении рисков при выходе их значений за допустимые границы и выбрана стратегия парирования рисков. Решение зависит от величины штрафных санкций по каждому проекту, от стоимости снижения разброса выполнения подзадач (обычно плохо определенная и быстро растущая величина) и от стоимости привлечения дополнительных ресурсов. Поэтому потенциально возможны следующие способы снижения рисков, которые требуют определения критического пути для каждого проекта:

1) перепланирование наиболее рискованных подзадач;

2) уменьшение неопределенности проекта путем дополнительных затрат на работу исполнителей, уже участвующих в проекте (выплата премий);

3) привлечение дополнительных исполнителей;

4) разрешение сверхурочной работы исполнителей.

Выбор наилучшего сценария зависит от конкретного плана на данный момент, степени допустимых рисков, наличия и стоимости работ дополнительных исполнителей.

Пример применения предлагаемого метода расчета рисков

Рассчитаем риск выполнения некоторого проекта на гипотетических модельных данных. Пусть имеются три проекта, состоящие из подзадач, для выполнения которых могут быть привлечены три исполнителя. Неопределенности δ одинаковы для всех подзадач и равны четырем единицам времени. Характеристики проектов приведены в таблице, план выполнения проектов представлен на рис. 4.

Найдем риски с учетом перекрытий неопределенностей:

— *по исполнителю 1*: неопределенность по ресурсу (*Performer* 1) равна 0 + 0 + 4 = 4 (время начала последней подзадачи (1,2) превышает неопределенности предыдущих подзадач), момент окончания проекта равен 34, предельное время окончания проекта 36. По формуле (7) r(*Performer* 1) = = 1/4 = 0,25;

— *по последовательности подзадач*: учитываются только перекрытие подзадач (2,2) и (1,2) и распределение последней подзадачи (1,2) проекта 2, при этом неопределенность равна 0 + 4 + 4 = 8, момент окончания равен 34, предельный срок равен 36. По формуле (7) риск по подзадачам проекта 2 равен r(*Subproblems* 2) = (1 − 2/8) · 2 = 9/16 = 0,56.

Таким образом, риск проекта 2 составляет r(*Project* 2) = 0,56.

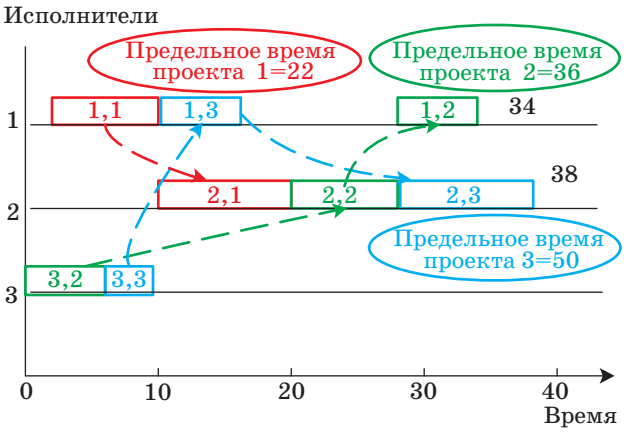
Обнаружение риска, его увеличение или сокращение могут идентифицироваться в разрабатываемой системе на «бизнес-радаре» руководителя проекта и вызывать автоматическое или ручное перестроение плана для минимизации риска соответствующего проекта.

Предположим, что штрафные санкции за опоздание проекта 2 велики и равны 10 единицам за каждую единицу времени отставания от графика (например, от срока окончания проекта 2 зависят другие дорогостоящие проекты). Проект 1 предусматривает простые штрафы, пропорциональные времени опоздания. Имеется также более дешевый на 30 % ресурс (исполнитель 2′), который, однако, дает и вдвое большую неопределенность по срокам — 8 единиц. Допустим, что снижение неопределенностей выполнения подзадач исполнителя 2 требует больших затрат и невозможно. Проанализируем два варианта снижения рисков:

1) перепланирование подзадач проектов 1 и 2 на 2-м исполнителе;

2) привлечение исполнителя 2′.

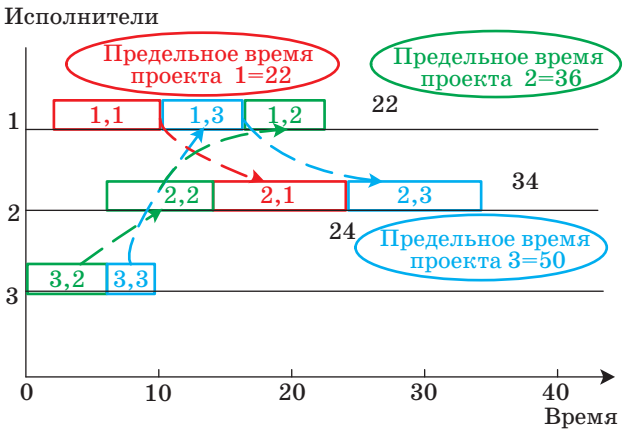
Средние потери при существующем плане складываются из средних потерь за счет штрафов по проектам 1 и 2. Среднее время завершения проекта 1 с учетом неопределенности по подзадачам, равной 8, по формуле (8) будет 20 + 8/3 = 22,7. По-



■ ***Рис. 4.*** План выполнения трех проектов с подзадачами для трех исполнителей

■ Характеристики проектов

Проект j	Вес штрафа w_j	Начало t_{sj}	Предельное время d_j	Последовательность подзадач у исполнителей i	Длительность подзадач p_{ij}
1	1	2	22	1, 2	8, 10
2	2	0	36	3, 2, 1	6, 8, 6
3	2	0	50	3, 1, 2	4, 6, 10



■ ***Рис. 5.*** План выполнения проекта при перепланировании подзадач (2,1) и (2,2)

этому средний штраф по проекту 1 составит 0,7. Среднее время завершения проекта 2 с учетом неопределенности по подзадачам, равной 8, будет 34 + 8/3 = 36,7. Тогда средние потери для данного плана составят 0,7 + 10 · (36,7 – 36) = 7,7.

В результате перепланирования подзадач (2,1) и (2,2) проектов 1 и 2 получается расписание, представленное на рис. 5.

В данном варианте плана средние потери по проекту 2 равны 0, потому что среднее время окончания проекта 2 не выходит за предельный срок. Среднее время окончания проекта 1 зависит от неопределенности, связанной с исполнителем 2, и равно 24 + 8/3 = 26,6. Средний штраф за отставание по проекту 1 составит 26,6 – 22 = 4,6. Поэтому план, полученный в результате перепланирования, выгоднее первоначального плана на (7,7 – 4,6) = 3,1 единицы.

Рассмотрим теперь вариант привлечения дополнительного исполнителя 2′. Стоимость выполнения минимальной по длительности подзадачи (2,2) равна 8 единицам и 5,6 единицам в первоначальном и новом планах соответственно. Следовательно, привлечение дополнительного исполнителя невыгодно.

Разработанный метод в настоящее время реализуется в рамках указанной выше интеллектуальной системы, находящейся в опытной эксплуатации в подразделениях ОАО «РКК «Энергия».

Заключение

Автоматизация планирования, мониторинга и контроля выполнения НИР и ОКР при создании сложных технических объектов требует учета специфики, связанной с высокой степенью неопределенности и динамики в этапах проектов.

В свою очередь планирование в условиях высокой неопределенности связано со значительными рисками. Рекомендуемые в настоящее время стандартами PMBOK методы носят скорее качественно-количественный характер и на практике не позволяют оценивать риски непосредственно в ходе проекта в реальном времени, особенно при интерактивном изменении планов исполнителями.

Предлагаемый в создаваемой интеллектуальной системе управления проектами НИР и ОКР метод приближенного расчета рисков выполнения проектов дает возможность в реальном времени на практике учитывать и контролировать возникающие риски для оперативного принятия решения по перераспределению ресурсов или привлечению новых ресурсов и минимизации рисков.

Предложенный в статье метод предполагается в дальнейшем развивать в направлении многокритериального подхода к планированию, что позволит комплексно учитывать многие важные критерии, например качество и стоимость работ, равномерность загрузки персонала, индивидуальные предпочтения и ограничения отдельных исполнителей и многие другие.

Литература

1. **Разработка** принципов построения многоуровневой мультиагентной системы для управления проектами НИР и ОКР «РКК «Энергия» / А. Л. Феоктистов, Е. М. Клейменова, П. О. Скобелев, И. А. Сюсин, В. Б. Ларюхин, А. В. Царев, Е. В. Симонова // Проблемы управления и моделирования в сложных системах (ПУМСС'2012): тр. XIV Междунар. конф. / СНЦ РАН. Самара, 2012. С. 718–723.
2. **Виттих В. А., Скобелев П. О.** Мультиагентные модели взаимодействия для построения сетей потребностей и возможностей в открытых системах // Автоматика и телемеханика. 2003. № 1. С. 162–169.
3. **Виттих В. А., Скобелев П. О.** Метод сопряженных взаимодействий для управления распределением ресурсов в реальном масштабе времени // Автометрия. 2009. № 2. С. 78–87.
4. **Скобелев П. О.** Мультиагентные технологии в промышленных применениях: к 20-летию основания Самарской научной школы мультиагентных систем // Мехатроника. Автоматизация. Управление. 2010. № 12. С. 33–46.
5. **A Guide** to the Project Management Body of Knowledge: PMBOK Guide. 4 Ed. / Project Management Institute. — 2008.

6. **Мазур И. И., Шапиро В. Д., Ольдерогге Н. Г.** Управление проектами: учеб. пособие / под общ. ред. И. И. Мазура. — М.: Омега-Л, 2004. — 664 с.
7. **Виттих В. А.** Введение в теорию интерсубъективного управления / СНЦ РАН. — Самара, 2013. — 64 с.
8. **Nedeljakova I.** Review of risk assessment methods // J. of Information, Control and Management Systems. 2007. Vol. 5. N 2/1. P. 277–284.
9. **Berg Heinz-Peter.** Risk management: procedures, methods and experiences, reliability // Theory & Applications. 2010. Vol. 1. N 2(17). P. 79–95.
10. **Cox L. A. Jr.** What's Wrong with Risk Matrices? // Risk Analysis. 2008. Vol. 28. N 2. P. 497–512.
11. **Rausand M., Hoyland A.** System Reliability Theory: Models, Statistical Methods and Applications. — N. Y.: Wiley, 2004. — 636 p.
12. **Srivastava Anurag, Bowles David S., Chauhan Sanjay S.** Generalized event tree algorithm and software for dam safety risk assessment // Proc. of the Intern. Conf. Dam Safety'12. Denver: ASDSO, 2012. P. 295–324.
13. **Bowen Ronda, Gundlach Marlene.** Project Management Methods & Ideologies, Disadvantages of the PERT Formula. http://www.brighthubpm.com/methods-strategies/15188-disadvantages-of-the-pert-formula/ (дата обращения: 14.12.12).
14. **Иванов В.** PMBOK 4-й редакции. Революция или Эволюция? http://www.microsoftproject.ru/articles.phtml?aid=158#risk (дата обращения: 14.12.12).
15. **Young Hoon Kwak, Ingall Lisa.** Exploring Monte Carlo simulation applications for project management // Risk management. 2007. Is. 9. P. 44–57.
16. **Tilo Nemuth.** Practical Use of Monte Carlo Simulation for Risk Management within the International Construction Industry // Proc. of the 6th Intern. Probabilistic Workshop. Darmstadt: Grauber, Schmidt & Proske, 2008. P. 471–481.
17. **Бурков В. Н., Новиков Д. А.** Как управлять проектами. — М.: Синтег, 1997. — 188 с.
18. **Бурков В. Н., Коргин Н. А., Новиков Д. А.** Введение в теорию управления организационными структурами / под ред. чл.-корр. РАН Д. А. Новикова. — М.: Либроком, 2009. — 264 с.
19. **Скобелев П. О.** Интеллектуальные системы управления ресурсами в реальном времени: принципы разработки, опыт промышленных внедрений и перспективы развития // Приложение к теоретическому и прикладному научно-техническому журналу «Информационные технологии». 2013. № 1. С. 1–32.

УДК 681.5

ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ РАЗВИТИЯ АВАРИЙНЫХ СИТУАЦИЙ В ЭНЕРГЕТИЧЕСКИХ УСТАНОВКАХ

А. Е. Городецкий,
доктор техн. наук, профессор
В. Г. Курбанов,
канд. физ.-мат. наук, старший научный сотрудник
И. Л. Тарасова,
канд. техн. наук, старший научный сотрудник
Институт проблем машиноведения РАН, г. Санкт-Петербург

Предложена имитационная модель, которая на основе комбинации логико-вероятностного и логико-лингвистического моделирования позволяет прогнозировать аварийные ситуации в энергетических установках большой единичной мощности.

Ключевые слова *— имитационное моделирование, логико-вероятностные переменные, логико-лингвистические переменные, функция принадлежности, вероятность безотказной работы, база данных.*

Введение

При оценке функционирования оборудования ГЭС и возможных неисправностей принято руководствоваться СТО 17330282.27.140.001-2006 «Методика оценки технического состояния основного оборудования гидроэлектростанций» и СТО 17330282.27.140.0019-2008 «Генераторы. Условия поставки. Нормы и требования». Каких-либо автоматизированных систем оценки возможных аварийных ситуаций путем анализа текущего состояния гидроагрегатов и показаний приборов не предусмотрено. Однако создание систем, способных подсказывать операторам возможные развития аварийных ситуаций и рекомендовать возможные действия для сохранения живучести, весьма актуально [1–4]. При этом необходимо решить проблемы моделирования аварийных ситуаций и быстрого анализа большого объема количественной и качественной информации в условиях неполной определенности, связанные с тем, что чем сложнее система, тем труднее дать точные и в то же время имеющие практическое значение суждения о ее поведении [5]. Такая ситуация определяется термином «принцип несовместимости» [6]. Следствие из этого принципа кратко можно выразить так: «Чем глубже мы анализируем реальную задачу, тем неопределеннее становится ее решение». Именно в этом смысле точного количественного анализа поведения сложных систем для практического исследования реальных задач, по-видимому, недостаточно. Поэтому при отсутствии принципиальной возможности получить четкую модель системы в целом или каких-либо ее частей целесообразно строить нечеткие модели [7–9].

Необходимость использовать такой подход может быть оправдана следующими обстоятельствами:

— при решении некоторых проблем не нужна точная оценка параметров объектов и явлений;

— по утверждению Л. Заде, с ростом сложности системы постепенно падает способность человека делать точные и в то же время значащие утверждения относительно ее поведения, так как существует порог, за которым точность и значимость становятся взаимоисключающими характеристиками.

В нечетких задачах моделирования логические переменные, как аргументы логических функций, обычно характеризуются набором атрибутных данных, среди которых наиболее используемые: вероятность логической переменной, являющейся в данном случае случайным событием; интервал значений переменной, которому присваивается имя данной логической переменной; функция принадлежности, характеризующая степень принадлежности текущей логической переменной к заданному интервалу [7].

Правила вычисления вероятностей описываются в разделах теории вероятности [10], вычисления

интервалов изучаются в интервальной математике [11], а вычисление функций принадлежности — в теории лингвистических переменных [12]. Однако при вычислении атрибутов логических функций по известным атрибутам аргументов, за исключением простейших функций (И, ИЛИ, НЕ), возникают определенные сложности и неоднозначности. В данной статье рассматриваются принципы моделирования развития аварийных ситуаций при функционировании гидроагрегатов и возможные пути решения проблемы вычисления вероятностей и функций принадлежности логических переменных, соответствующих наступлению предаварийных и аварийных ситуаций.

Принципы моделирования

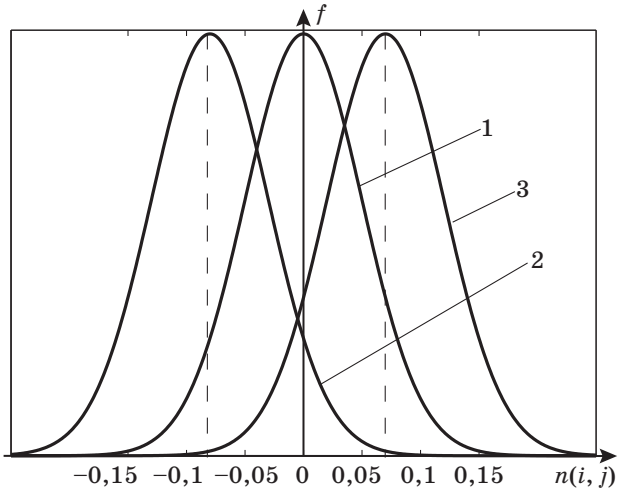
Моделируя развитие аварийных ситуаций, можно применять методы, основанные на представлении логических функций, описывающих те или иные аварийные и предаварийные ситуации, как упорядоченные множества. При этом можно использовать комбинаторные (не символьные) приемы их преобразования или такие методы, когда для упорядочивания множеств строится декартово произведение, элементы которого лексикографически упорядочены. Тогда нет необходимости записывать явно все его члены, а достаточно знать, как вычислить любой из них. Поэтому благодаря арифметическим свойствам получаемых систем логических уравнений, которые они проявляют при их представлении в виде алгебраических структур по модулю 2, т. е. в алгебре Жегалкина, оказывается возможным сведение логических задач к «арифметическим» или подобным арифметическим. Это в общем случае позволяет представлять логические системы как линейные структуры, уравнения которых не содержат конъюнктивных элементов, а для анализа и синтеза их структурных свойств использовать математический аппарат векторно-матричной алгебры [7].

В рассматриваемой модели на первом шаге имитируются отклонения $n(i, j)$ j-х параметров i-х блоков оборудования с помощью генератора случайных чисел с нормальным законом распределения с математическим ожиданием $m(i) = m_0 = 0$ (нулевое отклонение) и среднеквадратическим отклонением $\sigma_i = \sigma_0 = 0{,}05$ (5 % отклонения), запускаемого для каждого $n(i, j)$ (рис. 1).

По полученным значениям $n(i, j)$ можно определить для следующего шага эволюции новые значения математического ожидания

$$m(i) = m(i) + \sum_{i}^{A(i,j)} \frac{n(i, j)}{A(i, j)}, \qquad (1)$$

где $A(i, j)$ — количество j-х контролируемых параметров для каждого i-го оборудования.



■ ***Рис. 1.*** Нормальный закон распределения: *1* — исходная кривая распределения ($m(i) = 0$); *2*, *3* — кривые распределения после моделирования ($m(i)$ — расчетное)

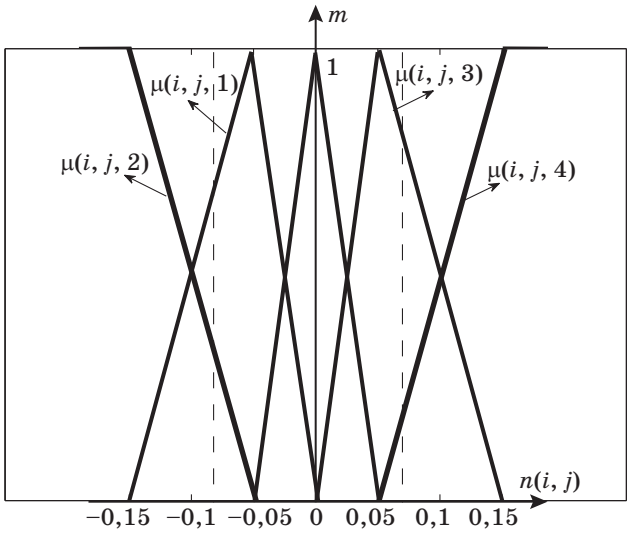
Для фаззификации имитируемых параметров, т. е. для получения логических величин $x(i, j, k)$ и соответствующих им функций принадлежности $\mu(i, j, k)$, где k — индикатор отклонения ($k = 1$ — «параметр ниже нормы»; $k = 2$ — «параметр значительно ниже нормы»; $k = 3$ — «параметр выше нормы»; $k = 4$ — «параметр значительно выше нормы») вначале необходимо установить опасную границу $b(i)$. Далее можно воспользоваться следующими правилами получения $\mu(i, j, k)$ (рис. 2).

1. Если $n(i, j) \le -b\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 1$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 0$.

2. Если $-b\sigma_0 < n(i, j) \le -0{,}5b\sigma_0$, то

$\mu(i, j, 1) = (n(i, j) + b\sigma_0)/0{,}5b\sigma_0$, $\mu(i, j, 2) = -(n(i, j) + 0{,}5b\sigma_0)/0{,}5b\sigma_0$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 0$.



■ ***Рис. 2.*** Фаззификация

3. Если $-0{,}5b\sigma_0 < n(i, j) \le 0$, то

$\mu(i, j, 1) = -n(i, j)/0{,}5b\sigma_0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 0$.

4. Если $0 < n(i, j) \le 0{,}5b\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = n(i, j)/0{,}5b\sigma_0$, $\mu(i, j, 4) = 0$.

5. Если $0{,}5\ b\sigma_0 < n(i, j) \le b\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = -(n(i, j) - b\sigma_0)/0{,}5b\sigma_0$, $\mu(i, j, 4) = (n(i, j) - 0{,}5b\sigma_0)/0{,}5b\sigma_0$.

6. Если $n(i, j) > b\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 1$.

Кроме того, при моделировании нужно вычислять вероятности отказа $P_o(i)$ i-х блоков, задавшись предельно допустимым значением контролируемого параметра $n_d(i, j)$. При этом:

1) если $m(i) \le 0$, то $y(i) = (-n_d(i, j) - m(i))/\sigma(i)$ и $P_o(i) = \Phi(y(i))$;

2) если $m(i) \ge 0$, то $y(i) = (n_d(i, j) - m(i))/\sigma(i)$ и $P_o(i) = 1 - \Phi(y(i))$, где $\Phi(y(i))$ выбирается по табл. 1 приложения из работы [13] по значению $y(i)$.

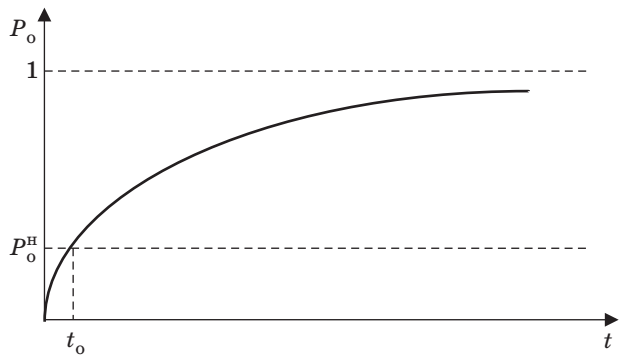
В процессе моделирования этапов эволюции (деградации) оборудования происходит сдвиг математического ожидания $m(i)$ в соответствии с уравнением (1). Поэтому может наступить такой момент, когда $-n_d(i, j) < m(i) \le -0{,}5n_d(i, j)$ либо $0{,}5n_d(i, j) \le m(i) < n_d(i, j)$. Это означает, что i-й блок находится в предаварийном, опасном состоянии, время наступления которого можно вычислить:

$$t_{па}(i) = (-1/\alpha_0(i))\ln(1 - P_o(i)),$$

где $\alpha_0(i)$ — показатель надежности i-го блока, вычисляемый по экспоненциальному закону убывания вероятности безотказной работы $P_{бо}(t) = 1 - P_o(t)$ (рис. 3, где $P^н_o$ — заданная вероятность отказа при заданном времени t_o наработки на отказ) с течением времени t и заданных значениях $b(i)$ и t_o по формуле

$$\alpha_0(i) = (-1/t_o(i))\ln(\Phi(y(i)), \qquad (2)$$

где $\Phi(y(i))$ определяется по таблице [13] при $|y(i)| = n_d(i, j) = b(i)$.



■ ***Рис. 3.*** Вероятность отказа

При дальнейшем моделировании этапов эволюции (деградации) оборудования может наступить такой момент, что $m(i) < -n_d(i, j)$ либо $n_d(i, j) < m(i)$. Это означает, что i-й блок находится в аварийном состоянии, и для него можно аналогично вычислить время наступления данного состояния:

$$t_a(i) = (-1/a_o(i))\ln(1 - P_o(i)). \qquad (3)$$

Прогнозирование времени $t_a(i)$ наступления аварийной ситуации очень важно для своевременного ремонта или замены оборудования в процессе его эксплуатации.

Описание алгоритма компьютерного моделирования

Введем следующие обозначения:

$x(i, j, k)$ — аварийное событие i-го объекта по j-му контролируемому параметру с индикатором отклонения k;

$\alpha_0(i)$ — исходный показатель надежности i-го оборудования;

$t_o(i)$ — время наработки на отказ i-го оборудования;

$t(i)$ — время работы i-го оборудования;

$P^н_{бо}(i)$ — вероятность безотказной работы i-го оборудования;

$P^н_o(i)$ — вероятность отказа нового i-го оборудования;

$P_o(i)$ — вероятность отказа i-го оборудования в процессе моделирования;

N — количество анализируемых объектов (оборудования);

$A(i, j)$ — количество j-х контролируемых параметров для каждого i-го оборудования;

V — конечное число изменений состояния системы (эволюций);

$m(i)$ — математическое ожидание контролируемых параметров i-го оборудования;

$\sigma(i)$ — среднеквадратическое отклонение контролируемых параметров i-го оборудования;

$b(i)$ — опасная граница выхода оборудования из строя;

$n_d(i, j)$ — предельно допустимое отклонение контролируемого j-го параметра i-го оборудования.

Шаги алгоритма.

1. Задание начальных условий.

Для всех i, j, k $t_o(i) = t_o = 27\ 000$, $P^н_o(i) = P^н_o = 0{,}004$, $t = 0$, $m(i, j) = m_0$, $m_0 = 0$, $P^н_{бо}(i) = P^н_{бо} = 0{,}996$, $n_d(i, j) = b = 0{,}15$, где $i \in [1, N]$, $j \in [1, A(i, j)]$, $k \in [1, 4]$; начальные значения счетчиков: $v = 0$, $q = 0$.

Значения m_0, V, σ_0 задаются экспертами — специалистами по оборудованию.

Значения N, $A(i, j)$ задаются оператором ГЭС либо берутся из базы данных (БД) (таблицы) или задаются экспертами — специалистами по оборудованию.

Кроме того, оператор ГЭС может изменять значения t_o, P_o для всех i, j, k.

Величины t_o = 27 000 ч и $P_{бо}^{н}$ = 0,996 взяты из СТО 17330282.27.140.0019-2008 «Генераторы. Условия поставки. Нормы и требования».

2. Вычисление α_0 и b.

По таблице [13] определяем $\Phi(y(i))$ при $|y(i)| = n_d(i, j) = b$;

α_0 вычисляем по формуле (2);

$\sigma_0 = \sigma(i, j) = n_d(i, j)/3$.

Записываем в БД α_0 и σ_0.

3. Вычисление математического ожидания $m(i)$ для i-го объекта.

Запускаем $A(i, j)$ раз генератор случайных чисел $x(i)$, распределенных по нормальному закону с математическим ожиданием $M(x) = m(i)$ и среднеквадратическим отклонением $\sigma_x = \sigma_0$ в соответствии с формулой

$$x(i) = M(x) + \sigma_x \left(\sum_{i=1}^{m} \xi_i - 6 \right),$$

где ξ_i — случайное число, генерируемое генератором случайных чисел с равномерным распределением в интервале от 0 до 1, $m = 12$.

Получаем случайные числа $x(i) = n(i, j)$ с нормальным распределением для каждого j-го контролируемого параметра i-го объекта и вычисляем математическое ожидание по формуле (1).

4. Определяем величину $\mu(i, j, k)$ и заполняем ее в БД для каждого i-го объекта по j-му контролируемому параметру с индикатором отклонения k по следующим правилам.

Если $n(i, j) \le -b(i)\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 1$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 0$.

Если $-b(i)\sigma_0 < n(i, j) \le -0{,}5b(i)\sigma_0$, то

$\mu(i, j, 1) = (n(i, j) + b(i)\sigma_0)/0{,}5b(i)\sigma_0$, $\mu(i, j, 2) = -(n(i, j) + 0{,}5b(i)\sigma_0)/0{,}5b(i)\sigma_0$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 0$.

Если $-0{,}5b(i)\sigma_0 < n(i, j) \le 0$, то

$\mu(i, j, 1) = -n(i, j)/0{,}5b(i)\sigma_0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 0$.

Если $0 < n(i, j) \le 0{,}5b(i)\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = n(i, j)/0{,}5b(i)\sigma_0$, $\mu(i, j, 4) = 0$.

Если $0{,}5\, b(i)\sigma_0 < n(i, j) \le b(i)\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = -(n(i, j) - b(i)\sigma_0)/0{,}5b(i)\sigma_0$, $\mu(i, j, 4) = (n(i, j) - 0{,}5b(i)\sigma_0)/0{,}5b(i)\sigma_0$.

Если $n(i, j) > b(i)\sigma_0$, то

$\mu(i, j, 1) = 0$, $\mu(i, j, 2) = 0$, $\mu(i, j, 3) = 0$, $\mu(i, j, 4) = 1$.

5. Вычисляем вероятность отказа P_o для i-го объекта:

— если $m(i) \le 0$, то $y(i) = (-n_d(i, j) - m(i))/\sigma(i)$ и $P_o(i) = \Phi(y(i))$, где $\Phi(y(i))$ выбирается по таблице [13] по значению $y(i)$;

— если $m(i) \ge 0$, то $y(i) = (n_d(i, j) - m(i))/\sigma(i)$ и $P_o(i) = 1 - \Phi(y(i))$, где $\Phi(y(i))$ выбирается по таблице [13] по значению $y(i)$.

6. Вычисляем и записываем в БД время наступления аварийной ситуации t_a i-го объекта, если $m(i) < -0{,}15$ или $m(i) > 0{,}15$, по формуле (3) и выдаем сообщение «аварийное состояние по i-му оборудованию».

7. Вычисляем и записываем в БД время наступления опасной ситуации t_υ i-го объекта, если $-0{,}15 < m(i) \le -0{,}05$ или $0{,}05 \le m(i) < 0{,}15$, по формуле $t_\upsilon(i) = (-1/\alpha_0)\ln(1 - P_o(i))$ и выдаем сообщение «опасное состояние по i-му оборудованию».

8. Увеличиваем значение счетчика q: $q = q + 1$:

— если $q \le N$, переходим к п. 3;

— если $q > N$, переходим к п. 9 и печатаем БД.

9. Увеличиваем счетчик υ: $\upsilon = \upsilon + 1$:

— если $\upsilon \le V$, то $q = 1$ и возвращение к шагу 3;

— если $\upsilon > V$, то остановка процесса эволюции.

Пример моделирования

По результатам моделирования эволюции (деградации) исследуемого оборудования ГЭС заполняется БД (таблица). Надо отметить, что $\mu(i, j, k)$ может использоваться по запросу оператора для вычисления значения j-го параметра i-го оборудования одним из методов дефаззификации [7].

Блоки оборудования, отказы которых моделировались:

— рабочие колеса поворотно-лопастных гидротурбин ($i = 1$);

— маслоприемник рабочего колеса поворотно-лопастной гидротурбины ($i = 2$);

— направляющий аппарат гидротурбины ($i = 3$);

— крышка гидротурбины ($i = 4$);

— металлические элементы проточной части гидротурбины ($i = 5$);

— аварийные, аварийно-ремонтные затворы, сороудерживающие решетки гидротурбинного блока ($i = 6$);

■ Фрагмент заполненной БД

i	$\alpha(i)$	$P_o(i)$	$m(i)$	k	$\Phi(y(i))$	$y(i)$	$t_\upsilon(i)$	j	Сообщение
11	$2{,}15 \cdot 10^{-7}$		0,03	3				1	
11	$2{,}15 \cdot 10^{-7}$	0,0228	0,05	3	0,9772	2	1072	2	Опасное состояние щеточно-контактного аппарата
11	$2{,}15 \cdot 10^{-7}$	0,0548	0,07	3	0,9452	1,6	2621	3	Опасное состояние щеточно-контактного аппарата
11	$2{,}15 \cdot 10^{-7}$	0,1587	0,1	3	0,8413	1	8037	4	Опасное состояние щеточно-контактного аппарата

— обмотка статора ($i = 7$);
— стальные конструкции статора ($i = 8$);
— стальные конструкции ротора ($i = 9$);
— обмотка возбуждения и демпферная система ($i = 10$);
— щеточно-контактный аппарат ($i = 11$);
— подпятники гидрогенераторов ($i = 12$);
— направляющие подшипники ($i = 13$);
— валы гидроагрегата ($i = 14$);
— система автоматического регулирования гидротурбин ($i = 15$);
— система технического водоснабжения ($i = 16$);
— система охлаждения и вентиляции ($i = 17$);
— система смазки ($i = 18$);
— система перевода гидроагрегатов в режим синхронного компенсатора ($i = 19$);
— система торможения гидроагрегата ($i = 20$).

Заключение

Предложено имитационное моделирование развития аварийных ситуаций в энергетических установках, основанное на использовании комбинации логико-вероятностного и логико-лингвистического описания развития и анализа аварийных ситуаций. Такое моделирование позволяет анализировать и прогнозировать аварийные ситуации для большинства гидроэнергетических агрегатов большой единичной мощности с учетом влияния основных технических и эксплуатационных показателей, вводимых операторами в БД перед началом сеанса работы. При этом учитываются «Методика оценки технического состояния основного оборудования гидроэлектростанций и влияние основных технических и эксплуатационных показателей» в соответствии с требованиями СТО 17330282.27.140.001-2006 и СТО 17330282.27.140.0019-2008 «Генераторы. Условия поставки. Нормы и требования».

Модель развития аварийных ситуаций в энергетических установках реализована в виде компьютерной программы. Достоверность прогноза и адекватность модели составляют от 70 до 85 %, зависят от точности задаваемых исходных параметров и могут быть повышены при коррекции модели по результатам апробации на характерных примерах.

Работа выполнена при поддержке государственного контракта № 16.515.12.5002.

Литература

1. **Кавалеров Б. В., Казанцев В. П., Шмидт И. А.** Компьютерные и полунатурные испытания средств управления энергетических газотурбинных установок // Информационно-управляющие системы. 2011. № 4. С. 34–41.
2. **Поршнев С. В., Соломаха И. В.** О возможности повышения качества многомерных математических моделей технологической информации, собираемой на тепловых электрических станциях // Информационно-управляющие системы. 2011. № 2. С. 29–36.
3. **Миленин А. А., Шишлаков В. Ф.** Система автоматического управления ГЭС малой мощности методом частотного регулирования // Информационно-управляющие системы. 2009. № 6. С. 25–29.
4. **Шмидт И. А., Кавалеров Б. В., Один К. А., Шигапов А. А.** Сопряжение программных сред в задачах моделирования и тестирования систем управления энергетическими газотурбинными установками // Информационно-управляющие системы. 2009. № 5. С. 25–31.
5. **Городецкий А. Е., Курбанов В. Г., Тарасова И. Л.** Экспертная система анализа и прогнозирования аварийных ситуаций в энергетических установках // Информационно-управляющие системы. 2012. № 4. С. 59–63.
6. **Zadeh L. A.** Fuzzy sets // Inform. Contr. 1965. Vol. 8. P. 338–353.
7. **Городецкий А. Е., Тарасова И. Л.** Нечеткое математическое моделирование плохо формализуемых процессов и систем. — СПб.: Изд-во Политехн. ун-та, 2010. — 336 с.
8. **Чернов В. Г.** Нечеткие деревья решений (нечеткие позиционные игры) // Информационно-управляющие системы. 2010. № 5. С. 8–14.
9. **Суконщиков А. А., Яковлев С. А.** Обобщенная модель системы ситуационного интеллектуально-агентного моделирования // Информационно-управляющие системы. 2010. № 2. С. 9–14.
10. **Феллер В.** Введение в теорию вероятностей и ее приложения. — М.: Мир. Т. 1. 1964. 500 с.; Т. 2. 1967. 752 с.
11. **Алефельд Г., Херцбергер Ю.** Введение в интервальные вычисления. — М.: Мир, 1987. — 360 с.
12. **Заде Л. А.** Понятие лингвистической переменной и его применение к принятию приближенных решений. — М.: Мир, 1976. — 168 с.
13. **Венцель Е. С.** Теория вероятностей. — М.: Наука, 1969. — 576 с.

УДК 681.52

МОДЕЛЬ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ДИАГНОСТИКЕ ВОСПАЛИТЕЛЬНЫХ ПРОЦЕССОВ ОРГАНИЗМА ПО ВИДУ ИНТОКСИКАЦИИ ИОНАМИ HS^- и Fe^{2+}

Г. А. Машевский,
ассистент
З. М. Юлдашев,
доктор техн. наук, профессор
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Рассматривается возможность использования Pt- и Ag_2S-электродов для контроля развития воспалительных процессов в организме человека, связанных с интоксикацией организма ионами HS^- и Fe^{2+}, а также электрохимическая модель работы данных электродов в присутствии сульфгидрильных соединений. Предложена модель принятия решения по виду интоксикаций, а также алгоритм распознавания данных патологий.

Ключевые слова *— система мониторинга, ионометрия, диагностика воспалительного процесса, математическая модель.*

Введение

В практической медицине сегодня прилагаются значительные усилия для решения проблемы опережающего распознавания различных патологий, а также проведения контроля лечения пациентов.

Интоксикации организма ионами HS^- или Fe^{2+} являются распространенными формами послеоперационных осложнений, способными существенно затруднить процесс послеоперационной реабилитации пациента, а в тяжелых случаях — привести к летальному исходу. Механизм токсического действия гидросульфид-аниона (HS^-) подобен цианиду (CN^-) и угарному газу (CO) и заключается в комплексовании атома меди в цитохроме А митохондрий, приводящему к его ингибированию. Результатом этого становится невозможность генерировать АТФ и накопление восстановителей в цепи переноса электронов в митохондриях. Избыток Fe^{2+} при воспалении приводит в действие белковый механизм острой фазы, который ограничивает поступление Fe в ткани и снижает его доступность для микроорганизмов, улавливает и транспортирует в макрофаги этот элемент. Данный механизм также связан с разрушением цитохромов, содержащих Fe. Патологические отклонения организма, вызванные повышением концентрации Fe^{2+}, при лечении онкологических больных составляют одну из наиболее часто встречающихся форм осложнений (28,4 % из 1364 контрольных измерений). Указанные интоксикации часто сопровождают развитие в организме воспалительного процесса, следовательно, появившиеся в моче ионы HS^- или Fe^{2+} становятся его маркерами. Таким образом, решение задачи разработки методики диагностики и лечения интоксикации организма ионами гидросульфида и двухвалентного железа позволит повысить эффективность клинического сопровождения пациентов. Данная проблема была рассмотрена нами в рамках исследования по созданию метода и системы мониторинга состояния водно-солевого обмена пациента в постоперационный период [1, 2].

Теоретические исследования

В основе возможности потенциометрического контроля лежат реакции твердокристаллического Ag_2S- и Pt-электродов в присутствии ионов HS^- или Fe^{2+}. Контроль мочи с помощью Ag_2S-электрода в присутствии сульфгидрильных компонентов основан на электрохимических реакциях

$$2Ag^0 + S^{2-} \leftrightarrow Ag_2S + 2e^- \quad (1)$$

$$2Ag^0 + HS^- + OH^- \leftrightarrow Ag_2S + H_2O + 2e^- \quad (2)$$

Электродные функции для этих реакций соответственно описываются уравнениями

$$\varphi_1 = -0{,}688 - 0{,}029\lg[S^{2-}]\ [В]; \quad (3)$$

$$\varphi_2 = -0{,}282 - 0{,}029\lg[HS^-] - 0{,}029pH\ [В]. \quad (4)$$

Ag_2S-электрод является классическим электродом для определения концентрации сульфидных и гидросульфидных ионов, что определяется произведением растворимости Ag_2S (10^{-51}).

Контроль с помощью платинового электрода возможен благодаря окислительно-восстановительным системам, присутствующим в компартментах человеческого организма. В частности, для описанного выше случая redox-система определяется электрохимическими уравнениями

$$S^{2-} \leftrightarrow S^0 + 2e^- \quad (5)$$

$$HS^- + OH^- \leftrightarrow S^0 + H_2O + 2e^- \quad (6)$$

которым соответствуют электродные функции

$$\varphi_3 = -0{,}480 - 0{,}029\lg[S^{2-}]\ [В]; \quad (7)$$

$$\varphi_4 = -0{,}074 - 0{,}029\lg[HS^-] - 0{,}029pH\ [В]. \quad (8)$$

Результатом решения полученной системы уравнений является выражение, отражающее связь между потенциалами электродов:

$$E(Ag_2S) = -208 + E(Pt). \quad (9)$$

Отклонения от данной зависимости свидетельствуют о присутствии в биосубстрате других сильных восстановителей помимо ионов гидросульфида, прежде всего ионов Fe^{2+}. Таким образом, существует возможность построить, опираясь на зависимость (9), методику распознавания видов интоксикации.

В теории оксредметрии [3, 4] известно, что обратимые органические окислительно-восстановительные системы проявляют одно общее свойство: их окислительный потенциал определенным образом зависит от pH. В результате исследования на большом статистическом массиве выявлена линейная зависимость для здорового организма:

$$E_{Pt} = 202{,}56 - 33{,}48pH. \quad (10)$$

Поэтому физиологическое значение имеет не абсолютное значение электродного потенциала, а его отклонение от величины (10):

$$\Delta Pt = Pt_{факт} - (202{,}56 - 33{,}48pH)\ [мВ]. \quad (11)$$

Сопоставляя значения данной величины с показаниями Ag_2S-электрода, можно провести распознавание таких опасных патологий, как HS^--интоксикация и интоксикация катионами Fe^{2+}.

Таким образом, при анализе результатов измерений при мониторинге больных с подозрением на воспалительный процесс следует оперировать не абсолютными значениями потенциалов, а значениями отклонений от зависимости (9):

$$\Delta pS = EAg_2S_{факт} - (-208 + Pt_{факт})\ [мВ]. \quad (12)$$

Если эта величина меньше нуля, то имеет место интоксикация организма ионами HS^-. При значениях потенциала Ag_2S больше –300 мВ патологии не наблюдается (концентрация HS^- в пределах нормы):

$$\Delta pS = E_{Ag_2S,\ измер} - (-208 + E_{Pt,\ измер}) > 0 \rightarrow Fe^{2+}\text{-интоксикация}; \quad (13)$$

$$\Delta pS = E_{Ag_2S,\ измер} - (-208 + E_{Pt,\ измер}) \le 0 \rightarrow HS^-\text{-интоксикация при } E_{Ag_2S,\ измер} < -300\ мВ. \quad (14)$$

Присутствие в моче Fe^{2+} подтверждено качественной реакцией по методике Лурье [5]. Розовое окрашивание образца мочи наблюдается только после его подкисления HNO_3, добавки H_2O_2 ($Fe^{2+} \rightarrow Fe^{3+}$) и NH_4NCS. Здесь уместно отметить, что содержание железа в моче очень мало — 0,7–5,7 нмоль/сут (0,04–0,3 мкг) [6], и это количество не может быть уловлено обычными методами. Такая возможность контроля появляется только после приема железосодержащих препаратов или комплексообразователя, который способствует выведению железа с мочой.

Известно, что при изменении pH субстрата изменяется диссоциация FeS-протеиновых комплексов. При высоких значениях водородного показателя в контролируемой среде появляются OH^- и наблюдается денатурация белков, которую мы можем зафиксировать по появлению сульфидных ионов в моче с помощью Ag_2S-электрода. Поэтому основу эксперимента составляло титрование проб мочи с помощью раствора NaOH.

Анализ результатов экспериментальных исследований

Экспериментальные исследования проводились на базе городской больницы № 26 Санкт-Петербурга. Всего было обследовано 1884 чел., проанализировано 7785 проб мочи, из них 1242 измерения относятся к обследуемым, считающим себя здоровыми, 2432 измерения выполнены для пациентов с распространенными форма-

ми ракового заболевания, 120 зафиксировано для пациентов с летальным исходом. Для измерений были использованы следующие электроды:

электрод pH-селективный: $E = 380 - 56{,}5pH$;

Ag_2S-электрод: $E = -688 - 29lg[S^{-2}]$;

Pt-электрод: $E_{Pt} = 201{,}8 - 33{,}4pH$;

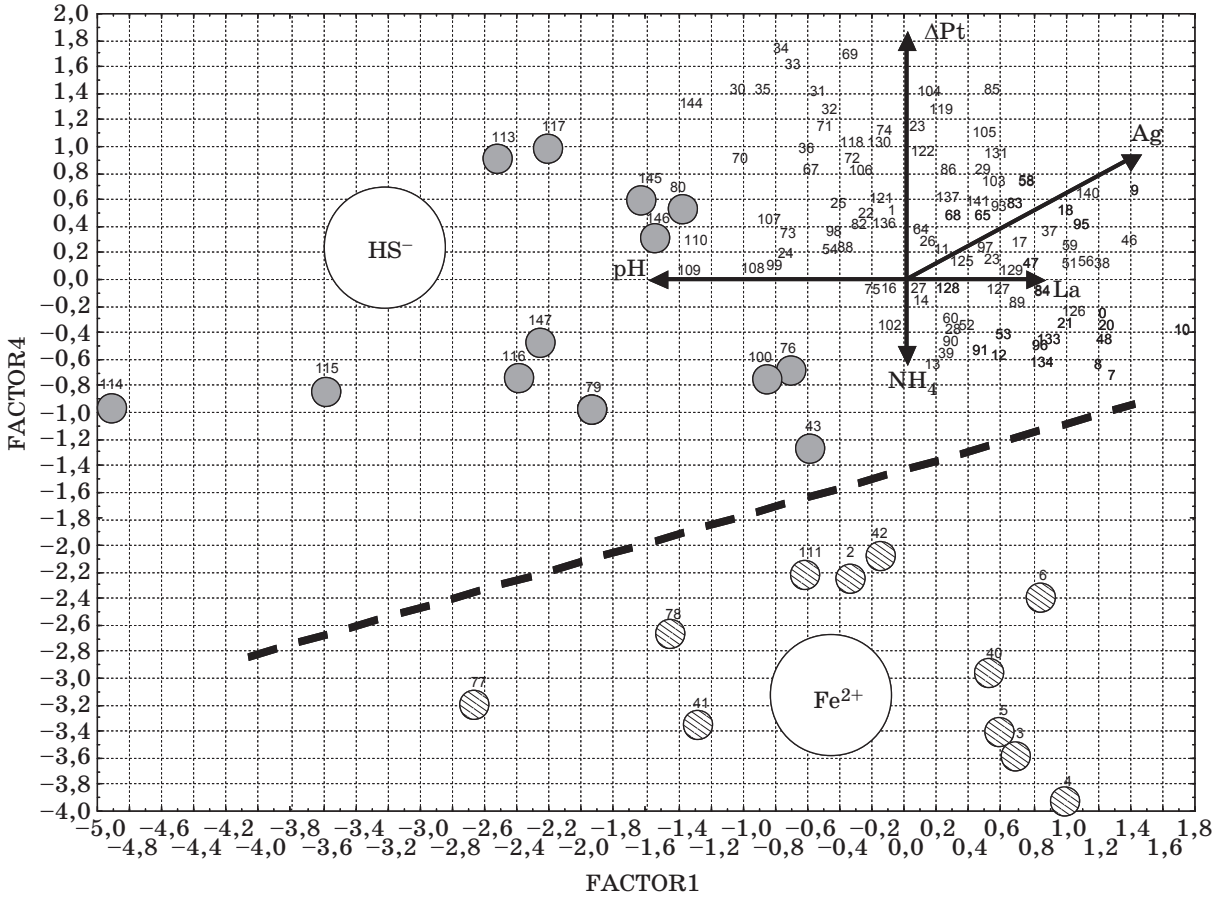
электрод сравнения ЭВЛ–1М1 (ЗИП, Гомель).

Для измерений использовались пробы мочи, параллельно велся учет данных анамнеза и характера проводимой терапии.

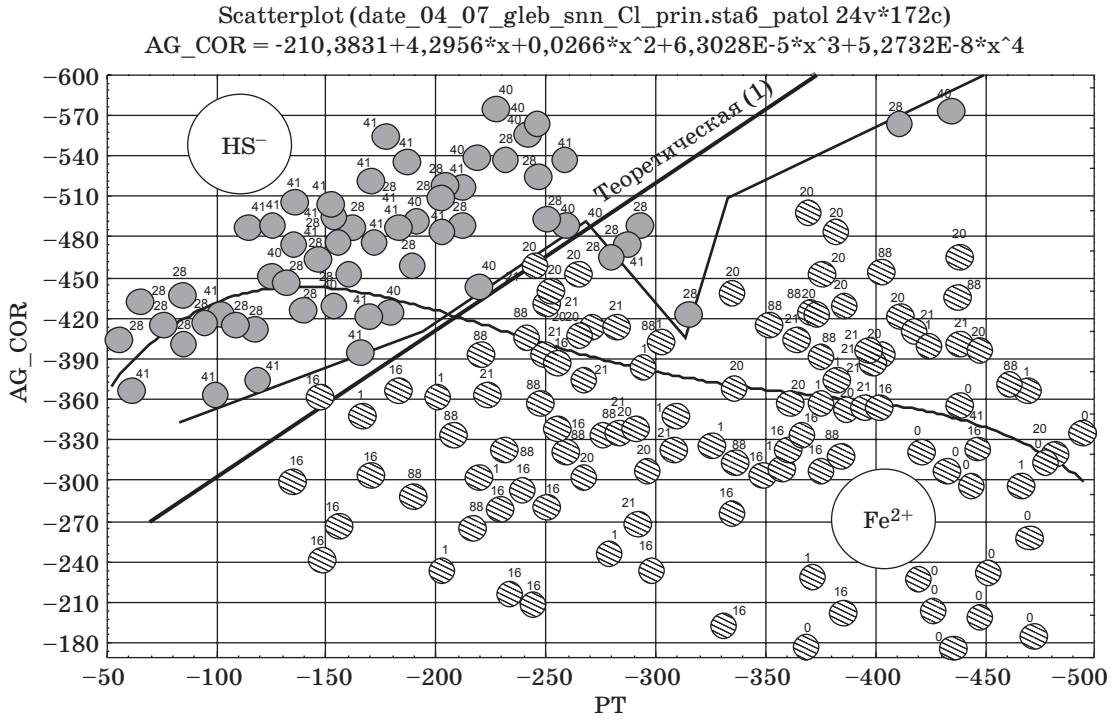
Классификация измерений была выполнена при помощи нейронной сети Кохонена формата 37 × 4. Для интерпретации результатов полученная топологическая карта была подвергнута факторному анализу. На рис. 1 представлена проекция полученного четырехмерного пространства на плоскость *F*1–*F*4, на которую также нанесены проекции векторов исходных параметров. Можно видеть, что нейроны, характеризующиеся наличием HS^-- или Fe^{2+}-интоксикации, группируются в две непересекающиеся области. Из рисунка также видно, что все нейроны с данными патологическими отклонениями лежат в области низких значений как для Ag_2S-электрода, так и для Pt-электрода. Разделительная линия, проведенная на рисунке, является весьма условной, поскольку, как показали наши дальнейшие исследования, имеется единство этиологии воспалительных процессов, и такой границы теоретически не существует.

Сходная картина была получена по результатам анализа экспериментального массива с помощью сети Кохонена формата 20 × 5 с архитектурой СОКК 8:8–100:1. Нейроны 0, 1, 16, 20, 21, 88 отражают интоксикацию, связанную с катионами Fe^{2+}, а нейроны 17, 22, 28, 40, 41 – HS^--интоксикацию. Совместное расположение этих нейронов на плоскости «Pt – Ag» представлено на рис. 2. На рисунке хорошо прослеживается разделение наблюдений теоретической линией по модели (9), хотя и отмечается некоторое отклонение в области высоких значений потенциалов электродов.

Исследование функциональных связей между выходной функцией ΔpS и входными параметрами Na, pH, Ag_2S выполнено с помощью ОРНС 3:3-495-2-1:1. Исследованный массив составил 987 наблюдений, непосредственно связанных с воспалительными процессами, диагностированны-



■ ***Рис. 1.*** Проекция многофакторного пространства на плоскость *F*1–*F*4, отражающая расположение нейронов воспалительного процесса



■ ***Рис. 2.*** Характеристика Fe^{2-}- и HS^{-}-интоксикации по расположению нейронов на плоскости «Pt – Ag»

ми с помощью Pt-, Ag_2S-электродами и pH. Критерием сортировки исходного массива принято значение Pt-электрода < −50 мВ. Результаты модели представлены в табл. 1.

Результаты анализа чувствительности по входным параметрам даны в табл. 2, оценки регрессии — в табл. 3.

Представленная на рис. 3 поверхность отклика проявляет главные особенности развития интоксикации в ходе воспалительного процесса. В области ацидоза развивается Fe-интоксикация, в области алкалоза — HS-интоксикация. Снижение натриевого потенциала увеличивает вероятность развития воспалительного процесса типа Fe-интоксикации.

Результаты выполненных исследований позволили предложить алгоритм распознавания и лечения данной патологии (рис. 4), основанный на только что приведенных теоретических предпосылках.

■ ***Таблица 2.*** Анализ чувствительности — 2 (date_HS_Fe_patologij_cor-50_St6.sta)

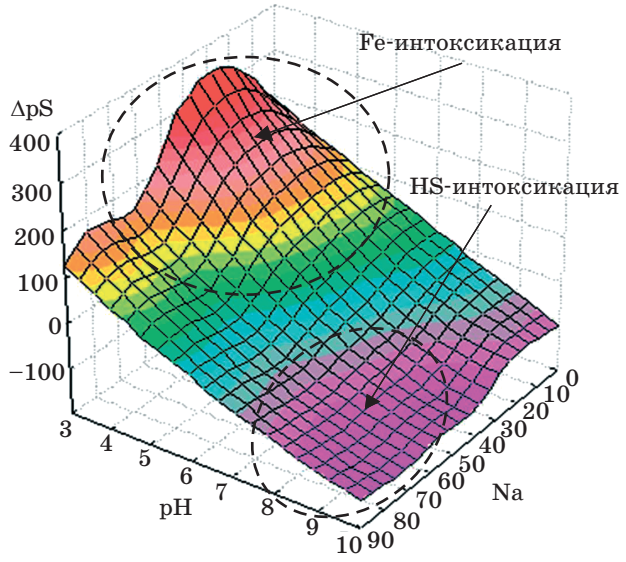
Характеристика	Na	pH	Ag_2S
Отношение	1,15	1,30	1,45
Ранг	3	2	1

■ ***Таблица 3.*** Регрессия (2) (date_HS_Fe_patologij_cor-50_St6.sta)

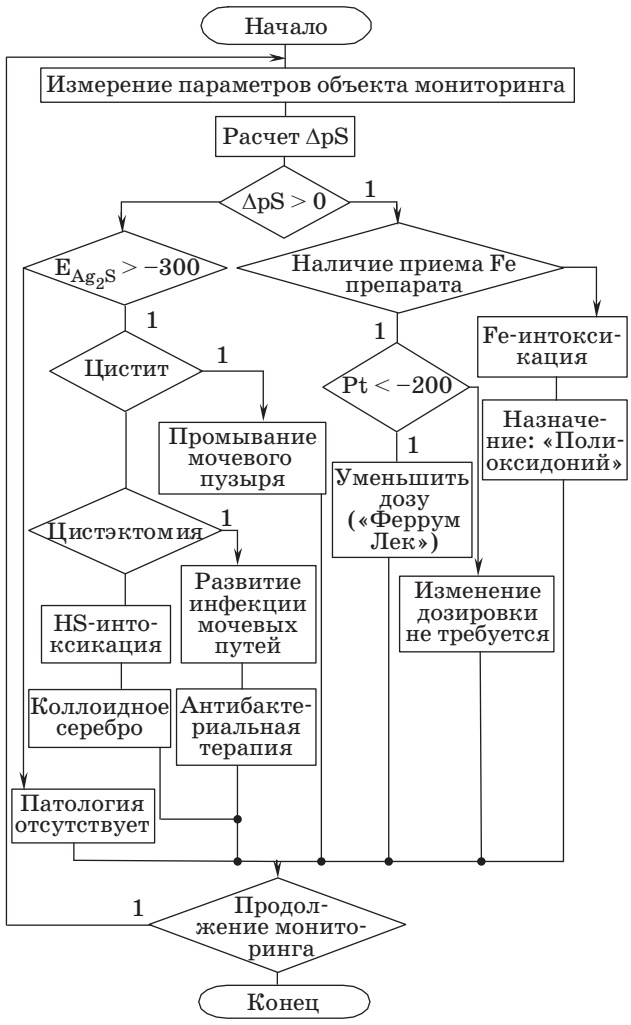
Характеристика	Значение
Среднее данных	60,60
Статистическое отклонение данных	150,11
Среднее ошибки	−0,70
Статистическое отклонение ошибки	104,75
Среднее абсолютной ошибки	71,71
Отношение статистического отклонения	0,70
Корреляция	0,72

■ ***Таблица 1.*** Подробные результаты моделей (date_HS_Fe_patologij_cor-50_St6.sta)

Архитектура	Производительность обучения	Контрольная производительность	Тестовая производительность	Ошибка обучения	Контрольная ошибка	Тестовая ошибка	Примечания	Входы	Скрытые (1)	Скрытые (2)
ОРНС 3:3-495-2-1:1	0,39	0,90	0,91	0,003	0,006	0,006		3	495	2



■ ***Рис. 3.*** Поверхность отклика модели интоксикации ОРНС 3:3-495-2-1:1



■ ***Рис. 4.*** Алгоритм диагностики и лечения интоксикаций организма ионами HS^- и Fe^{2+}

В алгоритме учтена возможность протекания у пациента воспалительного процесса в форме HS-интоксикации либо интоксикации катионами Fe^{2+}. Соответственно, проводится дифференциация этих двух состояний на основе значений параметра ΔpS с последующим уточнением окончательного диагноза. Если $\Delta pS < 0$ и $E_{Ag_2S} > -300$, делается вывод об отсутствии у пациента патологии. Дополнительно учитывается присутствие факторов, способных оказать влияние на показания электродов: прием пациентом железосодержащих препаратов (в случае если $\Delta pS > 0$) либо наличие у пациента цистита или цистэктомии (в случае если $\Delta pS < 0$ и $E_{Ag_2S} < -300$).

Заключение

В результате выполненных исследований была предложена электрохимическая модель принятия решений по виду интоксикации организма ионами HS^- и Fe^{2+}, основанная на реакции Pt- и Ag_2S-электродов в присутствии сульфгидрильных соединений. Достоверность предложенной модели была подтверждена экспериментально. На основе модели был разработан алгоритм распознавания интоксикаций, учитывающий возможность наличия у пациента сопутствующих патологий и содержащий рекомендации по возможным способам их лечения. Предложенные модель принятия решений и алгоритм распознавания интоксикаций позволяют повысить эффективность послеоперационного сопровождения пациентов.

Литература

1. **Машевский Г. А.** Исследование влияния ионов фторида и фосфата на состояние организма человека с помощью LaF_3-электрода // Биомедицинская радиоэлектроника. 2010. № 11. С. 69–73.
2. **Машевский Г. А., Юлдашев З. М.** Оценка энергетического потенциала организма человека по данным ионометрии мочи // Биомедицинская радиоэлектроника. 2009. № 11. С. 40–44.
3. **Михаэлис Л.** Окислительно-восстановительные потенциалы и их физиологическое значение. — М.: ОНТИ. Гл. ред. хим. лит., 1936. — 284 с.
4. **Никольский Б. П., Пальчевский В. В., Пенжин А. А.** и др. Оксредметрия. — Л.: Химия, 1975. — 304 с.
5. **Лурье Ю. Ю.** Унифицированные методы анализа вод. — М.: Химия, 1973. — 376 с.
6. **Анализы.** Полный справочник. — М.: Эксмо, 2008. — 767 с.

УДК 681.3.067

ОБОСНОВАНИЕ МЕРОПРИЯТИЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В. Ю. Осипов,
доктор техн. наук, профессор
И. А. Носаль,
аспирант
Санкт-Петербургский институт информатики и автоматизации РАН

Предложен подход к обоснованию мероприятий информационной безопасности с учетом ценности защищаемых информационных ресурсов. Рассмотрена модель ценности этих ресурсов. Приведены математическая формулировка и алгоритм решения задачи поиска целесообразных мероприятий информационной безопасности, предусматривающие синтез и анализ возможных программ деструктивных воздействий на защищаемые информационные ресурсы. Отражены результаты моделирования.

Ключевые слова *— информационная безопасность, оптимизация, методы, программы.*

Введение

Одной из актуальных научно-технических задач в области информационной безопасности (ИБ) выступает обоснование мероприятий ее обеспечения. От успешности решения этой задачи во многом зависят затраты на разработку (модернизацию) систем ИБ, а также потенциальные потери из-за нарушений безопасности. Из известных подходов к решению этой задачи [1–8] некоторые сводятся к минимизации стоимости мероприятий ИБ при выполнении требований к показателям безопасности. Когда затраты на реализацию мероприятий ИБ ограничены, при поиске целесообразного варианта минимизируют потери от нарушения ИБ или максимизируют эффект защиты, получаемый от проведения этих мероприятий [3–5]. Есть попытки обоснования мероприятий ИБ при минимизации суммы потерь на их реализацию и потерь в виде возможных информационных ущербов из-за деструктивных действий злоумышленников [1, 8]. В качестве показателей информационного ущерба используют абсолютные и относительные положительные или отрицательные приращения характеристик (свойств) защищаемых информационных ресурсов (ЗИР), систем, в которых они хранятся и обрабатываются, а также их потребителей. При этом в интересах оценки мероприятий ИБ во многих случаях строят графы (схемы программ) атак на ЗИР со стороны злоумышленников и оценивают их с помощью математических методов [3, 6–8].

Несмотря на значительное число работ, посвященных ИБ, многие аспекты оставлены без должного внимания, например вопросы в части обоснования мероприятий ее обеспечения по новым показателям (целевым функциям) с условиями, отражающими объективные закономерности неисследованных процессов. В частности, не уделяется должного внимания ценности ЗИР, потенциальной осведомленности и мотивациям злоумышленников, изменению во времени свойств каналов и программ деструктивного воздействия, характеристик применяемых средств защиты. Несовершенны также методы генерации и анализа таких программ с циклами.

Предлагается подход к обоснованию мероприятий ИБ по новым показателям с применением методов автоматического синтеза и анализа возможных программ деструктивного воздействия на ЗИР со стороны злоумышленников.

Задача обоснования мероприятий ИБ

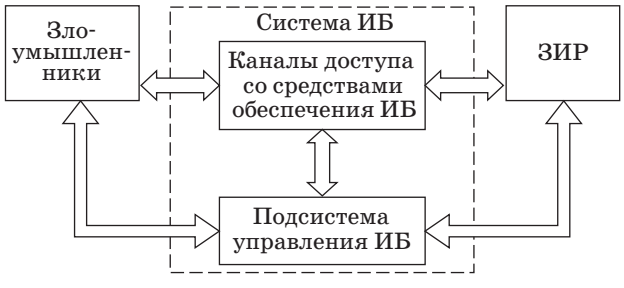
Рассмотрим постановку задачи на примере обеспечения ИБ в учреждениях высшей школы и науки. Определены виды защищаемых информационных ресурсов, форма их хранения и обработки. Такими ресурсами могут выступать персональные данные, отдельные сведения об учебной, научной и воспитательной работе, о финан-

сово-экономической деятельности, результаты перспективных научных исследований и разработок, учебно-методические материалы, имеющие существенную ценность, различные компьютерные программы, базы данных и др. Храниться информационные ресурсы могут в электронной форме (магнитные, лазерные диски, флеш-память), на бумажных носителях, в памяти сотрудников и обучаемых. Каждому такому информационному ресурсу свойственна своя ценность и возможные последствия от нарушений ИБ (незаконного копирования и распространения, искажения, уничтожения). Известно, что злоумышленники будут стараться деструктивно воздействовать на ЗИР по различным программам (схемам атак), используя уязвимости в системе ИБ и в системе обработки информации. При этом ценность ЗИР, как и потенциальные угрозы, со временем могут существенно изменяться.

Уже существует некоторая организационно-техническая система ИБ в учреждении, но она несовершенна.

Согласно схеме (рис. 1), комплекс мероприятий ИБ может предусматривать как исключение или затруднение деструктивных воздействий на ЗИР по многим каналам, так и непосредственное воздействие на злоумышленников. Среди этих мероприятий выступают мониторинг, применение различных средств защиты, изменение их параметров, ограничения физического доступа лиц к ЗИР и др. Кроме этого, на систему ИБ возлагаются функции по устранению последствий деструктивных воздействий и восстановлению ЗИР. Обоснование и реализация мероприятий ИБ осуществляются подсистемой управления ИБ, в качестве которой в частном случае могут выступать лица, ответственные за безопасность, со средствами управления.

Для совершенствования обеспечения ИБ необходимо разработать метод обоснования мероприятий ИБ, учитывающий особенности возникающих ситуаций и их динамику, позволяющий минимизировать возможные потери ценности ЗИР и затраты на реализацию этих мероприятий.



■ ***Рис. 1.*** Обобщенная структура системы информационной безопасности

Модель ценности защищаемых информационных ресурсов

В интересах разработки метода обоснования мероприятий ИБ определимся сначала с моделью ценности ЗИР. Опираясь на работу [9], определим ценность ЗИР как $V(t) = V_2(t) - V_1(t)$. Здесь $V_2(t)$, $V_1(t)$ — конечные эффекты на момент времени t, пересчитанные к входу системы, — санкционированного потребителя информации при наличии и отсутствии ЗИР соответственно. Пересчет конечных эффектов к входу предусматривает в нашем случае вычитание из них затраченных материальных и других ресурсов на получение интересующей информации. При этом предполагается, что потребитель использует полученную информацию оптимальным способом. Потребителем в частном случае может выступать сама система — обладатель ЗИР.

В соответствии с этими исходными посылками конечные эффекты $V_1(t)$, $V_2(t)$ могут быть представлены следующей аналитической зависимостью:

$$V_{1(2)}(t) = \max_{i(j)\in I(J)} \left\{ W_{1(2)i(j)}(t) - \sum_{r=1}^{N} a_r \cdot C_{1(2)r_{i(j)}}(t) \right\}, \quad (1)$$

где $W_{1i}(t)$, $W_{2j}(t)$ — эффекты, получаемые потребителем на момент времени t при достижении одних и тех же целей, соответственно, без интересующих ЗИР и при их наличии; I, J — множества всех возможных способов получения конечных эффектов в первом и втором случаях; $C_{1ri}(t)$, $C_{2rj}(t)$ — расходы r-го ресурса потребителя информации на достижение результатов $W_{1i}(t)$, $W_{2j}(t)$ соответственно; a_r — коэффициент приведения расхода r-го ресурса потребителя к единицам измерения конечных эффектов; N — число видов ресурсов потребителя, которые он может расходовать на получение и использование ЗИР.

Если выделить среди всех затраченных ресурсов на достижение эффекта $V_2(t)$ ресурсы, которые израсходованы на получение использованных ЗИР (на их разработку, покупку, восстановление, добывание), тогда

$$\sum_{r=1}^{N} a_r \cdot C_{2rj}(t) = \sum_{r=1}^{N} a_r \cdot C_{2.1rj}(t) + \sum_{r=1}^{N} a_r \cdot C_{2.2rj}(t). \quad (2)$$

Первое слагаемое в правой части выражения (2) соответствует затратам ресурсов на достижение эффекта $W_{2j}(t)$ при условии, что необходимые ЗИР в наличии, а второе — на получение ЗИР. Заметим, что наличие ЗИР означает и присутствие когда-то сделанных затрат на их получение.

С учетом (1), (2) ценность $V(t)$ ЗИР можно определить как

$$V(t)=\max_{j\in J}\min_{i\in I}\left\{W_{2j}(t)-W_{1i}(t)-\sum_{r=1}^{N}a_r\times\right.$$
$$\left.\times(C_{2.1rj}(t)-C_{1ri}(t))-\sum_{r=1}^{N}a_r\cdot C_{2.2rj}(t)\right\}. \quad (3)$$

Проанализируем это выражение. В условиях, когда разница между затратами ресурсов на достижение $W_{1i}(t)$, $W_{2j}(t)$ сводится к затратам на приобретение ЗИР:

$$\sum_{r=1}^{N}a_r(C_{2.1rj}(t)-C_{1ri}(t))=0, \quad (4)$$

их ценность относительно потребителя равна

$$V(t)=\max_{j\in J}\min_{i\in I}\left\{W_{2j}(t)-W_{1i}(t)-\sum_{r=1}^{N}a_r\cdot C_{2.2rj}(t)\right\}. \quad (5)$$

Снижение ценности $V(t)$ ЗИР для потребителя возможно, например, за счет искажения или внедрения в них ложных данных или раскрытия конфиденциальности.

В ситуации, когда можно успешно восстановить утраченные ЗИР до момента их использования, ценность ЗИР определяется как минимум затрат на их восстановление:

$$V(t)=\min_{j\in J}\sum_{r=1}^{N}a_r\cdot C_{2.2rj}(t). \quad (6)$$

Таким образом, в самом простом случае ценность ЗИР можно оценивать согласно (6), а при учете отдаленных последствий — по формуле (3) или (5).

Метод обоснования мероприятий ИБ

Учитывая особенности исследуемого процесса, задачу обоснования мероприятий ИБ математически можно сформулировать в следующем виде. Требуется найти комплекс M_o целесообразных мероприятий ИБ, при котором на момент времени t достигается минимум суммарных потерь $L_o(M_o, t)$:

$$L_o(M_o,t)=\min_{k\in Q}\left\{B_k(M_k,t)+\sum_{z=1}^{Z}V_z(t)\times\right.$$
$$\left.\times\left(1-\prod_{s=1}^{S_{kz}}(1-P_{kzs}(PRG_{kzs}(M_k),t))\right)\right\} \quad (7)$$

и выполняются условия

$$P_{kzs}(PRG_{kzs}(M_k),\ t)\ge P_E; \quad (8)$$
$$PRG_{kzs}(M_k)\in R, \quad (9)$$
$$k=1,\ 2,\ \ldots,\ K;\ z=1,\ 2,\ \ldots,\ Z;\ s=1,\ 2,\ \ldots,\ S_z.$$

В формулах (7)–(9) приняты обозначения: Q — область допустимых мероприятий ИБ; $B_k(M_k, t)$ — суммарные затраты на реализацию комплекса M_k мероприятий ИБ; $V_z(t)$ — текущая ценность z-го ЗИР; K — число мероприятий ИБ; Z — число ЗИР; $P_{kzs}(PRG_{kzs}(M_k),\ t)$ — вероятность деструктивного воздействия на z-й ЗИР по возможной s-й программе $PRG_{kzs}(M_k)$ при реализации комплекса M_k мероприятий ИБ; S_{kz} — число возможных альтернативных программ деструктивных воздействий на z-й ЗИР при комплексе M_k мероприятий ИБ; P_E — вероятность, при превышении которой угроза принимается во внимание; R — область допустимых результативных программ деструктивного воздействия на ЗИР.

В правой части выражения (7) второе слагаемое — это ожидаемые потери ценности ЗИР (риски). Потеря ценности z-го ЗИР имеет место, если он подвергся деструктивному воздействию хотя бы по одной из s-х программ. Согласно (8), принимаются во внимание только деструктивные программы с эффектом не ниже заданного. В соответствии с (9) анализируются только результативные программы, приводящие к нарушениям ИБ за конечное число шагов.

Решение этой задачи предусматривает генерацию потенциально возможных программ деструктивного воздействия на ЗИР, которые злоумышленники могут разработать с учетом их осведомленности, технической оснащенности и мотивации.

Результаты потенциальной осведомленности возможных злоумышленников о системе ИБ относительно текущего момента времени с формальной точки зрения для каждого анализируемого комплекса мероприятий ИБ предлагается задавать вектором исходных данных $\mathbf{d}_b = (d_{b1}, d_{b2}, d_{b3}, \ldots, d_{bN})$ и вектором интересующих результатов (целей) $\mathbf{d}_w = (d_{w1}, d_{w2}, d_{w3}, \ldots, d_{wM})$, а также условиями

$$F_{zv}(d_{zv\_e},\ e=1,\ 2,\ \ldots,\ E_z)\to d_{zv\_a},$$
$$z=1,\ 2,\ \ldots,\ Z;\ v=1,\ 2,\ \ldots,\ V_z, \quad (10)$$

связывающими исходное состояние с конечным. Для упрощения изложения материала индекс k здесь опущен. В (10) могут входить, например, функции получения доступа к ресурсам сервера с преодолением средств защиты, открытия защищаемых файлов, копирования, изменения параметров средств защиты для облегчения последующего доступа и др. Каждой из них ставятся в соответствие временные затраты на их реализацию. В выполнении этих функций могут участвовать как аппаратно-программные средства, так и злоумышленники. Условия (10) могут быть заданы в следующем виде (таблица).

■ Формализованные условия задачи

№ пп.	Функциональное выражение	Логическая запись	Статус	Условия истинности
1	$d_5 = F_{11}(d_1, d_2, d_3)$	$F_{11}(.), d_1, d_2, d_3 \to d_5$	F	
2	$d_5 >= F_{21}(d_2, d_4)$	$F_{21}(.), d_2, d_4 \to d_5$	R	7,12
3	$d_8 = F_{41}(d_7)$	$F_{41}(.), d_7 \to d_8$	S	N
.	.	.	.	
n	$d_{37} = F_{zv}(d_8, d_5)$	$F_{zv}(.), d_8, d_5 \to d_{37}$	F	
.	.	.	.	
N	$d_{504} = F_{ZVz}(d_{92})$	$F_{ZVz}(.), d_{92} \to d_{504}$	F	

В таблице F — основное условие; R — предусловие; S — постусловие. В графе «Условия истинности» для предусловий определены номера основных условий, при которых они должны быть истинными, а для постусловий — номера основных условий, при которых они должны быть ложными.

Ищутся результативные программы, с применением которых за конечное число шагов, исходя из $\mathbf{d}_b$, может быть достигнут конечный результат $\mathbf{d}_w$.

Синтез программ на заданном множестве условий предлагается осуществлять, исходя из стремления найти программу с наибольшим числом интерпретаций исходных данных, при которых достигается положительный результат [10].

Для доказательства существования результативной программы необходимо установить выполнимость основных и разрешимость вспомогательных условий задачи. Основное условие (F) считается выполнимым, если все входящие в него переменные являются свободными, а аргументы определены. Переменные могут считаться свободными в двух случаях. Во-первых, если они не связаны вспомогательными условиями. Во-вторых, если они связаны, но используются совместно с вспомогательными условиями, причем эти условия являются разрешимыми, и переменные, входящие в них, свободны. Вспомогательные условия (предусловия и постусловия) разрешимы, если все входящие в них переменные определены.

Если поставить всем элементам в условиях (10) соответствующие им предикаты, $P_{zv0}(F_{zv}(.))$, $P_{zv1}(d_{zv1})$, ..., $P_{zva}(d_{zv\_a})$, принимающие значение 1, когда переменные определены (истинны), и 0 в противном случае, то условия задачи могут быть записаны в виде

$$\{P_{zv0} \wedge P_{zv1} \wedge, ..., P_{ZVEz} \to P_{zv\_a} \,|\, S_{zv}, M_{zv}, z = 1, 2, ..., Z; v = 1, 2, ..., V_z\},$$

где S_{zv} — статус zv-го условия; M_{zv} — множество номеров основных условий, при которых предусловия и постусловия, соответственно, истинны и ложны. Для основных условий $M_{zv} = \varnothing$.

С учетом этого основное условие со свободными переменными выполнимо, если

$$P_{zv0} \wedge P_{zv1} \wedge, ..., P_{ZVEz} = 1.$$

Вспомогательное условие разрешимо, если

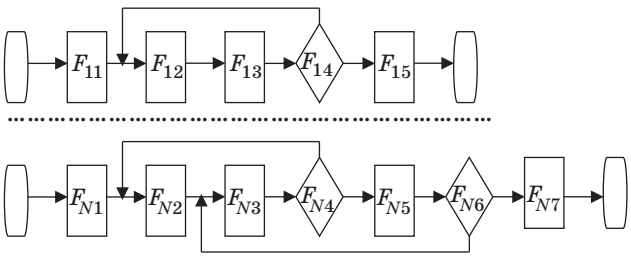
$$P_{zv0} \wedge P_{zv1} \wedge, ..., P_{ZVEz} \wedge P_{zv\_a} = 1.$$

Принимая это во внимание, синтез потенциально возможных деструктивных программ в интересах обоснования комплексов мероприятий ИБ предлагается осуществлять по следующему обобщенному алгоритму.

1. Ввод исходных данных и дополнительных условий.
2. Доказательство существования результативной программы.
3. Если доказательство не существует, то завершение синтеза.
4. Извлечение опорной программы из доказательства (вывода).
5. Если в опорной программе отсутствуют логические условия, то переход к п. 9.
6. «Сжатие» полученного вывода.
7. Обработка логических условий и получение подпрограмм.
8. Сведение подпрограмм в общую программу.
9. Выдача результативной программы.

Доказательство существования результативной программы выполняется от исходных данных к результату с использованием принципа резолюций [11]. Для выделения результативной программы из этого доказательства, если она существует, рекомендуется использовать идею обратного вывода. Этот вывод следует осуществлять по схеме зеркальной прямому выводу. При «сжатии» полученного вывода условия, взаимный порядок расположения которых в линейной программе не влияет на результат, прижимаются к условиям, использующим их. В интересах синтеза результативных программ с циклами на заданном множестве условий ищутся программы с наибольшим числом интерпретаций исходных данных, при которых достигается положительный результат.

При синтезе мы получаем схемы программ, подлежащих оцениванию. Простые примеры таких схем программ показаны на рис. 2, где F_{ij} —



■ ***Рис. 2.*** Примеры синтезированных схем программ

функции, которые потенциально могут реализовываться злоумышленниками при доступе и деструктивном воздействии на ЗИР. Если условиям (10) однозначно ставятся элементы программного кода на одном из известных языков программирования, то в результате такого синтеза получаем машинные программы.

Зная структуры таких программ и принимая во внимание случайный характер подлежащих анализу процессов, вероятности деструктивных воздействий можно рассчитать с применением математического аппарата полумарковских процессов [12]. Частным случаем его выступает аппарат марковских процессов. Предлагается по структурам синтезированных программ автоматически составлять системы интегральных (для полумарковских процессов) или дифференциальных (для марковских процессов) уравнений и разрешать их, получать искомые вероятности $P_{kzs}(PRG_{kzs}(M_k), t)$ деструктивных воздействий на z-е ЗИР по возможным s-м программам $PRG_{kzs}(M_k)$ при реализации комплекса M_k мероприятий ИБ. Технология автоматического составления таких систем уравнений известна. Методы разрешения их реализованы в ряде пакетов прикладных программ (MatLab, MathCad). В некоторых случаях могут быть использованы и другие методы оценки [13]. Особенностью такого анализа выступает необходимость учета неопределенности по разрешению логических условий в синтезированных программах. Эта неопределенность численно характеризуется числом и длительностью циклов в соответствующих программах, которые необходимо реализовать, чтобы ее преодолеть. Ее также можно определять через относительные частоты анализируемых переходов.

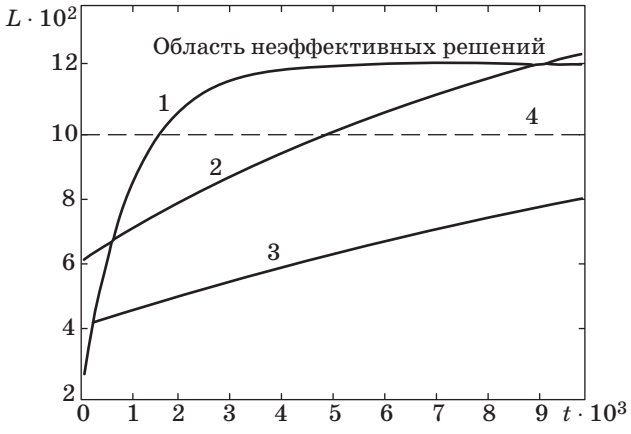
Суммарные затраты $B_k(M_k, t)$ на реализацию комплекса M_k мероприятий ИБ могут складываться из стоимости приобретаемых и устанавливаемых средств защиты, затрат на управление ими и восстановление ИР. Определение их может осуществляться простым суммированием затрат на реализацию отдельных мероприятий.

Заметим, что с течением времени исходные данные, защищаемые ресурсы, возможные мероприятия ИБ и условия (10) могут изменяться. С учетом этого обоснование мероприятий ИБ должно осуществляться для предварительно заданного интервала времени T. Период повторения $\Delta t << T$ такого обоснования и реализации найденных решений должен выбираться в зависимости от интенсивности угроз.

Результаты моделирования

Проверка работоспособности предлагаемого метода осуществлялась на простых примерах. Информация о защищаемых ресурсах и самой системе ИБ задавалась в виде правил (10). Защищались три вида информационных ресурсов и оценивались три различных комплекса мероприятий ИБ. Каждому комплексу мероприятий с формальной точки зрения ставилась своя система правил (10), с соответствующими им временными характеристиками и затратами со стороны системы ИБ. В качестве таких систем выступали совокупности из 18, 24 и 40 правил (10). Достаточно просто и легко синтезировались возможные линейные программы доступа к информационным ресурсам. Время t_c синтеза таких программ на скалярных процессорах прямо пропорционально квадрату числа n условий (10) задачи: $t_c = cn^2$. Синтез программ с логическими условиями (программ с циклами) требовал больше времени: $t_c \approx bn_1^2(1 + n_2)^2$, где b — постоянный коэффициент; n_1 — число основных условий задачи; n_2 — число логических условий. Для расчета вероятностей деструктивного воздействия на ЗИР по синтезированным программам использовался аппарат марковских процессов. При оценке ценности ЗИР применяли выражение (6). Установлено, что если с течением времени не тратить средства на поддержание требуемого уровня ИБ, существенных потерь ценности ЗИР не избежать. Затраты на обеспечение ИБ должны быть ниже ценности защищаемых информационных ресурсов, иначе мы всегда будем в минусе. На рис. 3 показаны результаты математического моделирования процессов обеспечения ИБ.

Из трех проанализированных вариантов предпочтение следует отдать третьему комплексу мероприятий ИБ. Затраты на его реализацию $B_3(M_3, t = 0) = 400$ усл. ед. больше, чем для первого, но меньше, чем для второго комплекса.



■ ***Рис. 3.*** Результаты обоснования комплексов мероприятий ИБ: *1*, *2*, *3* — зависимости основного показателя L суммарных потерь от времени t при реализации, соответственно, первого, второго и третьего комплексов мероприятий ИБ; *4* — линия, разделяющая области эффективных и неэффективных решений

При применении третьего комплекса мероприятий ИБ на всем анализируемом временном интервале суммарные потери не превышают ценности ЗИР. Она в данном случае составляет 1000 усл. ед. Однако заметим, что первый комплекс в связи с малыми затратами на его реализацию обладает временными преимуществами над третьим вариантом на незначительном интервале времени t = = 0...200. Следовательно, при принятии решений по обеспечению ИБ необходимо учитывать временной интервал безопасности. Принимая это во внимание, в качестве целевой функции задачи обоснования мероприятий ИБ вместо (7) в ряде случаев предлагается использовать минимальную площадь $S_o(M_o, T)$ под кривой суммарных потерь на заданном интервале времени T:

$$S_o(M_o, T) = \min_{k \in Q} \int_0^T L_k(M_k, t)\mathrm{d}t; \qquad (11)$$

$$L_k(M_k, t) = B_k(M_k, t) + \\ + \sum_{z=1}^{Z} V_z(t)\left(1 - \prod_{s=1}^{S_{kz}} (1 - P_{kzs}(PRG_{kzs}(M_k), t))\right). \qquad (12)$$

Параметры, входящие в (11), (12), раскрыты при пояснении (7)–(9).

Заключение

В результате выполненного исследования разработан новый метод обоснования мероприятий ИБ.

Обоснование мероприятий ИБ предлагается осуществлять, исходя из минимума общих потерь, среди которых ключевое место занимают потери из-за снижения ценности ЗИР. В интересах этого уточнена модель ценности ЗИР. В ряде случаев ценность ЗИР предлагается оценивать как минимум затрат на их восстановление. В общем случае рекомендуется учитывать отдаленные последствия из-за возможных деструктивных воздействий на ЗИР. Для этого целесообразно разработать математические модели систем — потребителей конкретной информации.

При обосновании мероприятий ИБ предлагается автоматически синтезировать на знаниях потенциально возможные для злоумышленников программы деструктивных воздействий на ЗИР и оценивать их. При этом возможен автоматический синтез не только линейных программ, но и программ с циклами.

Особенностью предлагаемого метода выступает его ориентированность на широкий круг возможных ситуаций обеспечения ИБ, учет ценности ЗИР и потенциальной информированности злоумышленников на текущий момент времени.

Отдельные положения метода могут быть применимы также при решении частных задач ИБ.

В целом предлагаемый метод расширяет взгляды и возможности по обоснованию мероприятий ИБ в различных условиях.

Литература

1. **Астахов М. А., Ростовцев Ю. Г., Яфраков М. Ф.** Информационная борьба. — М.: ТОМ, 2007. — 334 с.
2. **Осипов В. Ю., Юсупов Р. М.** Информационный вандализм, криминал и терроризм как современные угрозы обществу // Тр. СПИИРАН. 2009. Вып. 8. С. 34–45.
3. **Мальцев Г. Н., Теличко В. В.** Оптимизация состава средств защиты информации в информационно-управляющей системе с каналами беспроводного доступа на основе графа реализации угроз // Информационно-управляющие системы. 2008. № 4. С. 29–33.
4. **Миронов В. В., Носаль И.А.** Моделирование и оценка системы обеспечения информационной безопасности на примере ГОУ ВПО «СыктГУ» // Информация и безопасность. 2011. № 2. С. 209–211.
5. **Молдованин Т. В.** Решение задачи выбора оптимального варианта комплексной защиты информации с помощью метода экспертного оценивания // Информационно-управляющие системы. 2007. № 3. С. 39–44.
6. **Wang L., Yao C., Singhal A., Jajodia S.** Implementing interactive analysis of attack graphs using relational databases // J. of Computer Security. 2008. N 16. P. 419–437.
7. **Burgess M., Canright G., Engo-Monsen K.** A graph-theoretical model of computer security // Intern. J. of Information Security. 2004. N 3. P. 70–85.
8. **Dewri R., Ray I., Poolsappasit N., Whitley D.** Optimal security hardening on attack tree models of networks: a cost-benefit analysis // Intern. J. of Information Security. 2012. N 11. P. 167–188.
9. **Осипов В. Ю., Кондратюк А. П.** Оценка информации в интересах рефлексивного управления конкурентами // Программные продукты и системы. 2010. № 2. С. 64–68.
10. **Осипов В. Ю.** Синтез результативных программ управления информационно-вычислительными ресурсами // Приборы и системы управления. 1998. № 12. С. 24–27.
11. **Искусственный** интеллект. В 3 кн. Кн. 2. Модели и методы: справочник / под ред. Д. А. Поспелова — М.: Радио и связь, 1990. — 304 с.
12. **Новиков И. С.** Методы расчета количественных показателей надежности сложных программных комплексов на стадии проектирования и разработки // Тр. СПИИРАН. 2008. Вып. 6. С. 86–111.
13. **Осипов В. Ю.** Оценка защищенности информационно-вычислительных ресурсов от несанкционированного доступа // Приборы и системы управления. 1996. № 7. С. 16–19.

УДК 519.727, 621.391

ОСОБЕННОСТИ ПЕРЕДАЧИ И ОБРАБОТКИ ИНФОРМАЦИИ В СВЕРХСКОРОСТНЫХ ВОЛОКОННО-ОПТИЧЕСКИХ ЛИНИЯХ СВЯЗИ

Ю. И. Шокин,
доктор физ.-мат. наук, академик РАН
А. С. Скидин,
канд. физ.-мат. наук
Институт вычислительных технологий Сибирского отделения РАН, г. Новосибирск
М. П. Федорук,
доктор физ.-мат. наук, профессор
Новосибирский государственный университет, г. Новосибирск

Проведен анализ особенностей искажения сигнала в высокоскоростных волоконно-оптических линиях связи. На основе анализа предложены методы кодирования и обработки оптического сигнала, учитывающие специфику воздействия на сигнал в волоконном световоде при передаче данных на высокой скорости.

Ключевые слова *— волоконная оптика, математическое моделирование, теория кодирования, нелинейные эффекты.*

Введение

В настоящий момент самой быстродействующей и надежной средой передачи информации является волоконный световод. Около 70 % всего мирового информационного трафика передается через оптоволоконные линии связи. За десятилетия, прошедшие с момента появления оптоволокна, суммарная длина всех проложенных кабелей составила 1 млрд км, а скорость передачи информации выросла до 100 Тбит/с [1]. Чтобы достичь таких скоростей передачи, разработчики используют технологию WDM (Wavelength Division Multiplexing — мультиплексирование с разделением по длинам волн), позволяющую передавать информацию по нескольким частотным каналам одновременно. Применяются также иные методы, в частности, все активнее прибегают к использованию форматов модуляции с высокой спектральной эффективностью (например, к квадратурно-амплитудной модуляции [1, 2]).

Однако при уплотнении оптического сигнала в спектральной области на него начинают оказывать существенное влияние особенности среды передачи информации, главной из которых является наличие нелинейных воздействий на сигнал. Влияние нелинейных воздействий на передачу информации прямо пропорционально мощности сигнала, поэтому его имеет смысл учитывать только при плотном использовании спектральной полосы. В данной работе влияние нелинейного взаимодействия описывается как теоретически, так и экспериментально, после чего предлагаются методы преобразования информации, использующие особенности искажений сигнала в волоконном световоде, причиной которых является нелинейное воздействие среды на сигнал. Акцент делается на фазовые форматы модуляции ввиду их важности в современных волоконно-оптических системах связи.

Математическое описание передачи информации по оптоволоконной линии связи

С теоретической точки зрения распространение сигнала по волоконному световоду можно описать уравнением Шредингера [3]

$$i\frac{\partial A}{\partial z} = -\frac{i}{2}\alpha A + \frac{1}{2}\beta\frac{\partial^2 A}{\partial T^2} - \gamma\left|A\right|^2 A. \qquad (1)$$

Здесь $A(z, T)$ — амплитуда огибающей импульса; $|A(z, T)|^2$ — мощность импульса; T — время, измеряемое в системе отсчета, движущейся с импульсом и его групповой скоростью v_g ($T = t - z/v_g$).

Правая часть уравнения содержит три слагаемых, которые описывают влияние рассеяния импульса $\left(-\frac{i}{2}\alpha A\right)$, влияние дисперсии $\left(\frac{1}{2}\beta\frac{\partial^2 A}{\partial T^2}\right)$ и влияние нелинейности $(-\gamma|A|^2A)$ на распространение импульсов по волоконному световоду, с соответствующими коэффициентами. Данное уравнение используется для расчета при длительности импульсов более 0,1 пс. При анализе распространения более коротких импульсов в уравнение (1) включается член $\frac{\partial^3 A}{\partial T^3}$, однако, исходя из практических приложений, далее в работе будут рассматриваться импульсы длительностью, большей 0,1 пс, поэтому третью частную производную в рассмотрение можно не включать.

Компоненты рассеяния и дисперсии не зависят в явном виде от величины самого сигнала (модуля амплитуды огибающей), в отличие от нелинейного элемента $(-\gamma|A|^2A)$, величина которого прямо пропорциональна квадрату модуля амплитуды огибающей сигнала или, что то же самое, пропорциональна мощности сигнала. Нелинейным данный компонент уравнения (1) называется потому, что при его наличии с физической точки зрения реакция среды зависит от мощности возмущения, а в волоконной оптике такие явления называются нелинейными.

Теоретико-информационный аспект нелинейных явлений в волоконной оптике

Нелинейные явления существенно меняют представление о применимости классических результатов теории передачи информации к оптоволоконным линиям связи. Как известно, если в канале связи искажения являются гауссовыми, то по теореме Шеннона — Хартли между пропускной способностью канала (C), отношением сигнал/шум (SNR) и шириной полосы пропускания (B) имеется зависимость: $C = B \cdot \log_2(1 + \mathrm{SNR})$. Данное соотношение выполняется для абсолютного большинства каналов связи, так как обычно сигнал в канале претерпевает либо гауссовы воздействия, либо воздействия, которые могут быть хорошо аппроксимированы гауссовыми.

Применительно к оптическому сигналу теорема Шеннона — Хартли не выполняется по двум причинам. Прежде всего, искажения сигнала могут быть аппроксимированы гауссовым шумом лишь при небольшой мощности сигнала. Однако помимо этой имеется другая причина: в оптоволокне сигнал ведет себя таким образом, что увеличение SNR не приводит к увеличению пропускной способности канала, а, напротив, вызывает усиление нелинейных искажений, которые, в свою очередь, становятся причиной ухудшения характеристик линии. То есть теорема Шеннона — Хартли в данном случае не выполняется не только количественно, но и качественно.

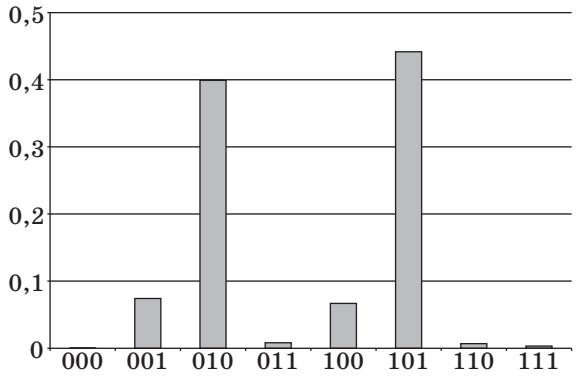
В силу данного обстоятельства в современной теории информации активно ведутся исследования с целью найти аналог теоремы Шеннона — Хартли для каналов с воздействием нелинейностей [4–6]. В целом результаты говорят о том, что на практике возможно применить схему кодирования сигнала, которая может дать емкость канала, превышающую эквивалентную гауссову емкость канала из теоремы Шеннона — Хартли, но при этом, разумеется, не превышает емкость канала, вычисленную с учетом нелинейных воздействий [7].

Особенности искажений сигнала в волоконно-оптических линиях связи

Последние исследования искажений оптического сигнала позволяют выявить как количественную, так и качественную составляющую их влияния. В количественном отношении исследования велись с целью выяснить статистику ошибок при передаче информации с помощью амплитудных [8, 9] и фазовых [10, 11] форматов модуляции. С качественной точки зрения проводился анализ положения принятых точек на фазовой плоскости и распределение энергий в принятых символах [12].

В экспериментах, проведенных на достаточном уровне мощности исходного сигнала (порядка 10 мВт), наблюдалась сильная зависимость количества ошибок от вида передаваемых данных (паттерн-эффект). Так, для пятиканальной системы с фазовым двоичным форматом по разности фаз [10] (прямое детектирование) наблюдалась статистика ошибок по двоичным триплетам (рис. 1).

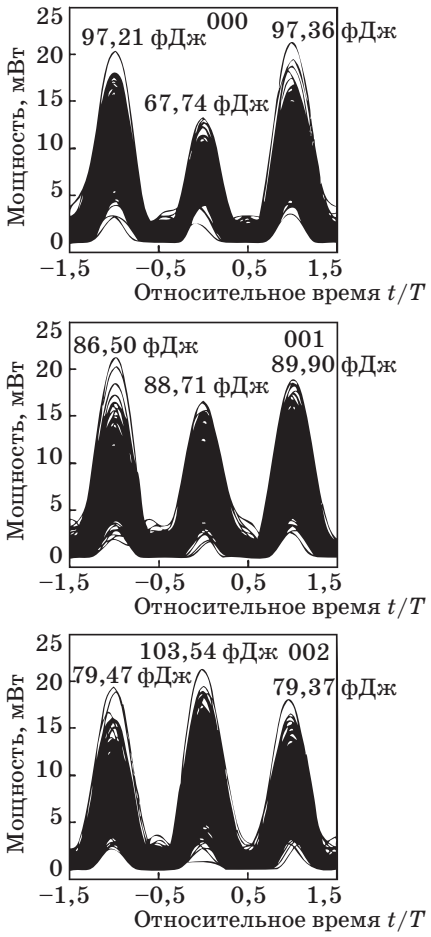
В данном численном эксперименте было передано всего 4 млн бит и получено 97 189 ошибок, т. е. частота ошибок составляет $2{,}4 \cdot 10^{-2}$. Можно видеть, что в сумме 85 % этих ошибок приходят-



■ ***Рис. 1.*** Гистограмма веса статистики ошибок по двоичным триплетам

ся на центральные биты триплетов 010 и 101. При этом величина ошибок на выходе системы такова, что средства коррекции (FEC — Forward Error Correction) не могут дать приемлемое качество декодирования сигнала, а те, которые могут (например, турбокоды), не могут быть использованы в оптоволоконных линиях связи ввиду большой сложности их декодирования (быстрые алгоритмы декодирования турбокодов на данный момент неизвестны). В связи с чем возникает проблема обеспечения качественной передачи информации в таких условиях. В данной работе проблема решается путем использования специальных кодов, снижающих количество ошибок в принятой последовательности данных.

В другом эксперименте [11] анализировалась передача 4-ичного QPSK-сигнала (фазово-модулированного сигнала) с когерентным приемом. Показано, что статистика ошибок по-прежнему остается неравномерной, при этом была проанализирована не только сама статистика ошибок, но и причины, которые ее обусловливают. Так, на рис. 2 [11] представлены амплитудные диаграммы по различным триплетам QPSK-сигнала. Видно, что при передаче на расстояние 1400 км из центрального нулевого символа триплета 000 «уходит» энергия, в то время как у других триплетов это не наблюдается, а энергия центрального «нуля» триплета 002, напротив, увеличивается, что можно объяснить нелинейными взаимодействиями, так как при передаче сигнала на низкой мощности подобных эффектов не наблюдается. Вместе с тем максимальная частота ошибок наблюдается именно в триплете 000. Это позволяет предположить, что главным источником ошибок является в данном случае падение амплитуды символа, и что в таком случае возможно учесть данное обстоятельство при детектировании сигнала.



■ ***Рис. 2.*** Диаграммы мощности QPSK-сигнала при передаче на 1400 км для триплетов 000, 001 и 002

Использование модулирующего кодирования для улучшения качества передачи оптического сигнала

Для того чтобы улучшить характеристики передачи информации со статистикой ошибок, показанной на рис. 1, рассмотрим блочный код, который отбирает из всех передаваемых кодовых слов заданной длины m (длина блока) только те q из них, которые имеют вероятность быть переданными с ошибкой, не превосходящей заданного значения P. Такого рода коды называют модулирующими (modulation codes) в том смысле, что они придают определенный вид передаваемому сообщению (модулируют его); похожие коды были применены для других целей [13–15]. Кодовая скорость (code rate) кода, описываемого ниже, будет равна $R = \dfrac{\lfloor \log_2 q \rfloor}{m}$.

Для построения кода рассмотрим вектор $\mathbf{W} = (w_1, w_2, \ldots, w_q)$, который получается путем упорядочивания всех возможных кодовых слов (число которых равно 2^m) по убыванию вероятности их безошибочной передачи по каналу $P(w_i)$. Таким образом, для любых i справедливо $P(w_{i+1}) \le P(w_i)$. Таблица $\mathbf{W}$ будет далее использоваться для кодирования и декодирования.

Для нахождения вероятностей представим в двоичном виде кодовое слово $\mathbf{a} = a_0 a_1 \ldots a_{m-1}$. Обозначим $q_i(\mathbf{a})$ i-й триплет слова $\mathbf{a}$. Так, $q_0(\mathbf{a}) = a_0 a_1 a_2$, $q_1(\mathbf{a}) = a_1 a_2 a_3$, ..., $q_{m-3}(\mathbf{a}) = a_{m-3} a_{m-2} a_{m-1}$, всего $m - 2$ триплетов. Таким образом, вероятность ошибки в кодовом слове $\mathbf{a}$ может быть определена по формуле

$$P(w_i) = 1 - \sum_{j=0}^{m-3} \left(1 - R(q_j(w_i))\right),$$

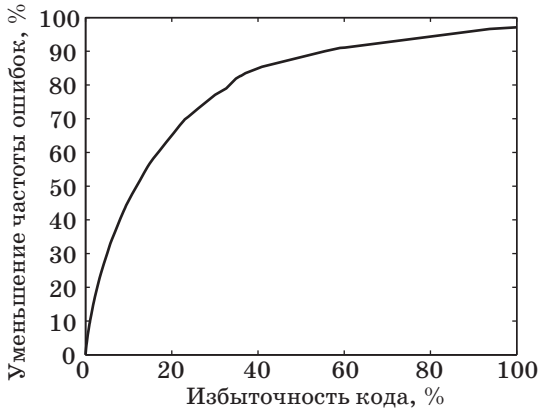
где $R(q_j(w_i))$ — вероятность появления ошибки в триплете $q_j(w_i)$.

Будем представлять входной битовый поток как последовательность небольших блоков дли-

ны n бит ($n \le m$): $S = (s_1, s_2, ..., s_p)$. Обозначим через D закодированное сообщение: $D = (d_1, d_2, ..., d_p)$, где каждый блок d_i имеет размер m бит. Теперь рассмотрим адаптивный блочный код $AC(m, n)$, который отображает блоки s_i на блоки d_i по следующему правилу: $d_i = W(s_i)$, $i = 1..p$. Для декодирования требуется построить «обратную» таблицу W^{-1}, $|W^{-1}| = 2^m$: $W^{-1}(r) = j$, если $W(j) = r$. Таким образом, декодирование осуществляется по формуле $s_i = W^{-1}(d_i)$, $i = 1..p$. Относительная избыточность кода будет равна $\frac{m-n}{n}$. Очевидно, чем меньше значение n, тем сильнее код может улучшить качество передачи данных, поскольку в этом случае для передачи будут использоваться наиболее «надежные» битовые последовательности, вероятность ошибки в которых очень мала; но при этом можно существенно проиграть в реальной пропускной способности линии ввиду того, что вместе с качеством растет доля избыточных данных в передаваемом потоке, которые не несут реальной информации и необходимы лишь для декодирования сигнала.

В целях наглядного представления возможностей кода был построен код с длиной блока 16 бит на основе статистики ошибок (см. рис. 1). Зависимость снижения количества ошибок относительно первоначального их количества показана на рис. 3. В 2 раза ошибки уменьшаются при применении кода с избыточностью примерно 12 %, в то время как возможно снизить частоту ошибок на порядок, но ценой большой потери в полезном объеме передаваемых данных.

При применении кода преимуществом также является то, что исходя из потребностей конкретной системы можно определить параметры таким образом, чтобы уменьшить ошибки до заданного уровня за счет добавления в сообщение небольшой избыточности. Снижение может быть критически важно для схемы коррекции ошибок, так как при определенных уровнях ошибок без дополнительного снижения ошибок она корректировать качественно не может, а с ним — может. С технической точки зрения вопросы сочетаемости различных типов кодов не представляют сложности, разные варианты сочетания рассматривались достаточно давно и хорошо известны [13].



■ ***Рис. 3.*** Результаты применения адаптивного кода

Усовершенствование детектирования сигнала

Стоит отметить, что коды с ограничениями, хотя и являются эффективным, как показано выше, средством улучшить качество принимаемого сигнала, имеют один недостаток — они обладают избыточностью. Это снижает реальную скорость передачи информации на 10–15 %, что во многих случаях является существенным обстоятельством.

Ниже будет предложен метод детектирования символа, который улучшает качество приема и при этом не добавляет избыточности в передаваемые данные, в отличие от методов кодирования. Предлагаемый метод основан на результатах наблюдений (см. рис. 2), из которых видно, что основной причиной возникновения ошибки в триплетах является падение энергии в центральном символе триплета. С математической точки зрения результат детектирования определяется по максимуму функции правдоподобия, которая имеет вид весовой функции, связывающей фазовые различия между принятым символом QPSK и возможными его значениями, а также амплитудные различия в символах относительно его «соседей».

Описание метода

Обозначим x_i символ, который обрабатывается в настоящий момент в детекторе. Его соседей слева и справа обозначим соответственно x_{i-1} и x_{i+1}. Будем считать, что символы x_{i-1} и x_{i+1} детектируются стандартным способом, т. е. по значению фазы в средней точке импульса. Обозначим $E(x)$ энергию символа x, а $P(x)$ — фазу символа x. Цель детектора, таким образом, заключается в том, чтобы получить значение x_i.

Определим для каждого триплета данных вектор $\mathbf{D}_T = \left(\overline{D}_1^T, \overline{D}_2^T, \overline{D}_3^T\right)$, где $\overline{D}_k^T$ — усредненная амплитуда среднего символа для k-го символа в триплете T. Множество $D = (D_{000}, D_{001}, D_{002}, ..., D_{333})$ задает, таким образом, распределение энергии в триплетах. Данное распределение может быть получено из экспериментальных данных. На амплитудных диаграммах (см. рис. 2) можно видеть, что в общем случае два различных триплета имеют различное распределение по символам.

Предлагаемый метод детектирования отличается от стандартного тем, что разбивает процесс детектирования на две составляющие — анализ

амплитуды принятого сигнала и анализ его фазы. В результате метод дает на выходе показатели правдоподобия по фазе для каждого из возможных значений обрабатываемого символа x_i: $P(x_i = 0)$, $P(x_i = 1)$, $P(x_i = 2)$, $P(x_i = 3)$ — и показатели правдоподобия по амплитуде $E(x_i = 0)$, $E(x_i = 1)$, $E(x_i = 2)$, $E(x_i = 3)$; каждое из значений заключено между 0 и 1. Общая функция правдоподобия дается следующей формулой:

$$L(x_i = k) = p \cdot P(x_i = k) + (1 - p) \cdot E(x_i = k),$$

где p — вес фазового показателя правдоподобия ($0 \le p \le 1$), который является свободным параметром и может быть полезен в качестве оптимизирующего параметра при настройке детектора на конкретную систему. Символ x_i определяется таким образом, чтобы показатель правдоподобия был максимальным при данном x_i.

Расчет показателя правдоподобия

Представим, что символ x_i детектирован по средней точке и эта точка равна $A(a, b)$. Преобразуем точку A в точку $A'\left(\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}\right)$, т. е. нормируем амплитуду точки A на единицу. Это необходимо для того, чтобы детектор мог работать по одному и тому же алгоритму вне зависимости от мощности принятого сигнала. Таким образом, «эталонными» точками формата QPSK будут $S_0 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $S_1 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $S_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ и $S_3 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$; с этими точками будет сравниваться по положению точка A' при формировании функции правдоподобия.

С формальной точки зрения стандартный детектор вычисляет расстояния $\rho(A', S_0)$, $\rho(A', S_1)$, $\rho(A', S_2)$, $\rho(A', S_3)$, а затем выбирает из полученных значений минимум. То значение символа, которое соответствует минимуму, и будет ответом.

В отличие от стандартного детектора предлагаемый способ не требует нахождения минимума указанных расстояний. Вместо этого ему необходимо найти показатели правдоподобия $P(x_i = k)$ и $E(x_i = k)$. Значение $P(x_i = k)$ получается по следующей формуле, нормирующей расстояния $\rho(A', S_k)$ на единицу:

$$P(x_i = k) = 1 - \frac{\rho(A', S_k)}{\sum\limits_{j=0,1,2,3} \rho(A', S_j)}.$$

Амплитудное правдоподобие $E(x_i = k)$ показывает, насколько амплитудные характеристики триплета, центральным символом которого является обрабатываемый символ x_i, являются «похожими» на характеристики триплетов $x_{i-1}0x_{i+1}$, $x_{i-1}1x_{i+1}$, $x_{i-1}2x_{i+1}$ и $x_{i-1}3x_{i+1}$:

$$E(x_i = k) = 1 - \frac{\xi(k)}{\sum\limits_{j=0,1,2,3} \xi(j)},$$

где

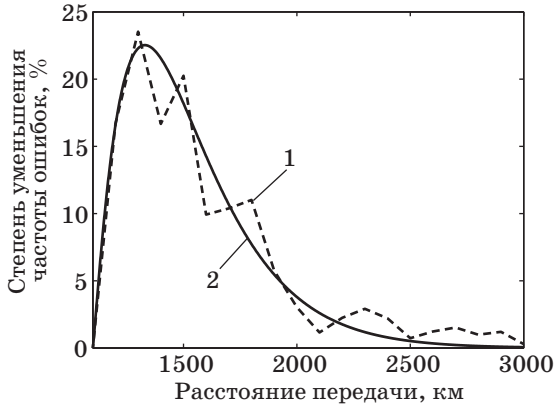
$$\xi(K) = \left|\overline{D}_1^{x_{i-1}\,K\,x_{i+1}} - W(x_{i-1})\right| + \left|\overline{D}_2^{x_{i-1}\,K\,x_{i+1}} - W(x_i)\right| + \left|\overline{D}_3^{x_{i-1}\,K\,x_{i+1}} - W(x_{i+1})\right|,$$

а $W(x_i)$ — принятая амплитуда символа x_i.

Результаты применения метода

Результаты применения метода для различных расстояний передачи QPSK-сигнала на конфигурации линии из работы [11] показаны на рис. 4. Как можно видеть, максимальный эффект в виде 20 %-го снижения количества ошибок достигается на дистанции передачи 1200–1600 км, т. е. на среднемагистральных дистанциях. Этот эффект важен с точки зрения приложений, так как на данных дистанциях полученная [11] частота битовых ошибок равняется величине $10^{-3} - 10^{-2}$, т. е. находится на границе возможностей применения классических схем коррекции ошибок. Таким образом, применение описанной выше схемы способно увеличить максимальную дистанцию передачи, на которой возможно полностью декодировать принятый сигнал с помощью FEC-кодов.

Отдельно стоит отметить, что оптимальное значение коэффициента p в эксперименте варьировалось в очень узких пределах около среднего 0,75. Фактически можно считать, что оптимальный коэффициент фазового «веса» в функции правдоподобия составляет 3/4, а вклад амплитудных характеристик равен 1/4. Это показывает, что каче-



■ ***Рис. 4.*** Результаты улучшенного детектирования оптического QPSK-сигнала: *1* — достигнутые значения эффекта для различных дистанций; *2* — аппроксимация полученных экспериментальных результатов

ственно характер искажений не зависит от текущего состояния сигнала, т. е. не зависит от того, насколько сигнал искажен в настоящий момент.

Таким образом, преимуществом данного подхода является его безызбыточность, а также наличие свободного параметра p, который в случае каждой системы может быть настроен таким образом, чтобы обеспечивать минимум ошибок детектирования. Кроме того, величина правдоподобия L_i после преобразования из символьного вида в битовый вид и соответствующего отображения $[0; 1] \times [0; 1] \rightarrow \mathbf{R}^2$ может быть использована в «мягких» декодерах (soft decoder) типа декодеров с низкой плотностью проверок на четность (LDPC-декодерах), что способно дать при декодировании дополнительную информацию и улучшить характеристики декодера.

Заключение

Предложенные выше методы обработки информации, передаваемой по волоконно-оптическим линиям связи, существенно улучшают характеристики приема сигнала за счет применения информации о характере искажений в линии, обусловленных действием внутри линии нелинейных взаимодействий сигнала со средой передачи.

Улучшение, достигаемое предложенными методами, наблюдается даже в области большого количества ошибок, при котором стандартные схемы помехоустойчивого кодирования не могут снизить на выходе частоту битовых ошибок до стандартной величины $10^{-9} - 10^{-12}$. Однако при совместном применении предложенных методов со стандартными корректирующими кодами возможно получить на выходе частоту ошибок, соответствующую стандартам, что подтверждается полученным при реалистичных условиях передачи данных эффектом снижения ошибок вдвое при добавлении в код 12 % избыточности.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации, проект № 11.519.11.4018.

Литература

1. **Qian D.** et al. 101.7-Tb/s (370x294-Gb/s) PDM-128QAM-OFDM Transmission over 3x-55-km SSMF using Pilot-based Phase Noise Mitigation // OFC/NFOEC. PDPB5. Optical Society of America, USA, 2011. P. 1–3.
2. **Yu J., Zhou X.** 16x107-Gb/s 12.5-GHz-Spaced PDM-36QAM Transmission Over 400 km of Standard Single-Mode Fiber // IEEE Photonics Technology Letters. 2010. Vol. 22(17). P. 1312–1314.
3. **Agrawal G. P.** Nonlinear Fiber Optics. — N. Y.: Academic Press, 2001. — 467 p.
4. **Essiambre R.-J.** et al. Capacity limits of fiber-optic communication systems // OFC/NFOEC. OThL1. Optical Society of America, USA, 2009. P. 1–37.
5. **Narimanov E., Mitra P.** The channel capacity of a fiber optics communication system // OFC/NFOEC. ThQ1. Optical Society of America, USA, 2002. P. 504–505.
6. **Tang J.** The Shannon channel capacity of dispersion-free nonlinear optical fiber transmission // J. of Lightwave Technology. 2001. Vol. 19(8). P. 1104–1109.
7. **Turitsyn K. S., Turitsyn S. K.** Nonlinear communication channels with capacity above the linear Shannon limit // Optics Letters. 2012. Vol. 37(17). P. 3600–3602.
8. **Shapiro E. G., Fedoruk M. P., Turitsyn S. K.** Direct modeling of error statistics at 40 Gbit/s rate in SMF/DCF link with strong bit overlapping // Electronics Letters. 2004. Vol. 40(22). P. 1436–1437.
9. **Skidin A., Redyuk A., Shtyrina O., Fedoruk M., Shafarenko A.** The analysis of the error statistics in a 5x40 Gbit/s fibre link with hybrid amplification // Optics Communications. 2011. Vol. 284(19). P. 4695–4698.
10. **Turitsyn S. K., Fedoruk M. P., Shtyrina O. V.** et al. Patterning effects in a WDM RZ-DBPSK SMF/DCF optical transmission at 40 Gbit/s channel rate // Optics Communications. 2007. Vol. 277(2). P. 264–268.
11. **Редюк А. А.** и др. Математическое моделирование экспериментального прототипа высокоскоростной линии связи на основе дифференциального фазового формата модуляции без возвращения к нулю // Квантовая электроника. 2011. Т. 41(10). С. 929–934.
12. **Redyuk A. A., Skidin A. S., Shafarenko A. V., Fedoruk M. P.** Direct modelling of error statistics for data transmission through a high data rate communication line using a four-level phase modulation format // Quantum Electronics. 2012. Vol. 42(7). P. 645–649.
13. **Shafarenko A., Skidin A., Turitsyn S.** Weakly-constrained codes for suppression of patterning effects in digital communications // IEEE Transactions on Communications. 2010. Vol. 58(10). P. 2845–2854.
14. **Immink K. A. S.** A Practical Method for Approaching the Channel Capacity of Constrained Channels // IEEE Transactions on Information Theory. 1997. Vol. 43(5). P. 1389–1399.
15. **Медведева Ю. С., Рябко Б. Я.** Быстрый алгоритм нумерации слов с заданными ограничениями на длины серий единиц // Проблемы передачи информации. 2010. Т. 46(4). С. 130–139.

УДК 681.3

ПОДХОД К ПОСТРОЕНИЮ КРИПТОСХЕМ НА ОСНОВЕ НЕСКОЛЬКИХ ВЫЧИСЛИТЕЛЬНО ТРУДНЫХ ЗАДАЧ

А. А. Демьянчук,
младший научный сотрудник
Д. Н. Молдовян,
младший научный сотрудник
Санкт-Петербургский институт информатики и автоматизации РАН
Е. С. Новикова,
канд. техн. наук, ассистент
Д. Ю. Гурьянов,
канд. техн. наук, ассистент
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Предлагается подход к построению криптосхем, основанных на двух вычислительно трудных задачах, который обеспечивает формирование подписи небольшой длины. Определены требования к выбору системных параметров криптосхем и личных ключей пользователя. Разработанный способ построения криптосхем обладает свойством универсальности и может быть применим для построения протоколов различного типа, таких как протокол открытого распределения ключей, протокол аутентификации с нулевым разглашением секрета.

Ключевые слова *— электронная цифровая подпись, протокол открытого шифрования, протокол обмена ключами, задача дискретного логарифмирования, задача факторизации.*

Введение

В основе криптосистем с открытым ключом (ОК), например протоколов электронной цифровой подписи (ЭЦП), лежит некоторая трудная математическая задача, которая определяет верхнюю границу безопасности соответствующей схемы. Криптосистемы такого типа используются для защиты и аутентификации информации в информационно-телекоммуникационных системах при условии, что неизвестны алгоритмы взлома криптосхемы и вероятность появления в обозримом будущем практически реализуемых прорывных решений используемой трудной задачи является достаточно малой. В настоящее время на практике наиболее широко используются две трудные задачи: 1) задача факторизации (ЗФ) целых чисел специального вида и 2) задача дискретного логарифмирования (ЗДЛ) по простому модулю (т. е. в простом конечном поле). Данные задачи независимы, и вероятность появления прорывного решения каждой из них в обозримом будущем имеет достаточно низкое значение. Для повышения безопасности алгоритмов ЭЦП, достигаемого за счет снижения вероятности взлома путем применения качественно новых прорывных решений используемых трудных задач, в работах [1–3] предложены схемы ЭЦП, взлом которых требует одновременного решения ЗФ и ЗДЛ. В этих криптосхемах используется ЗДЛ по простому модулю p, имеющему специальную структуру: $p - 1 = erq$, где r и q — 512-битовые простые числа и e — четное число небольшого размера, а в качестве основания дискретных логарифмов выбирается число, имеющее порядок $n = rq$. Параметры r и q являются элементами секретного ключа, а значения p, α и y, где $y = \alpha^x \bmod p$ (x — элемент секретного ключа), составляют ОК. Один из элементов подписи вычисляется по модулю n, поэтому суммарный размер ЭЦП превышает 1024 бит. Данный подход применяется для построения протоколов слепой ЭЦП [3–5], открытого шифрования и открытого согласования ключей [6]. Однако предложенные в работах схемы характеризуются сложностью построения и большой длиной вырабатываемой подписи.

В настоящей работе предлагается подход к построению криптосхем различного типа, основанных на трудности одновременного решения ЗФ и ЗДЛ по простому модулю. Использование

данного подхода обеспечивает снижение размера подписи в схемах ЭЦП и устраняет громоздкость построения криптографических протоколов других видов.

Подход к построению криптосхем

Для построения криптографических протоколов предлагается использовать ЗДЛ по трудно разложимому модулю n, для решения которой необходимо выполнить факторизацию составного модуля и решить ЗДЛ по простым модулям, являющимся делителями числа n, или применить один из общих методов дискретного логарифмирования (метод больших и малых шагов, переборный метод, метод Полларда [7]), используемых для решения ЗДЛ в любых конечных группах. Следует отметить, что общие методы становятся вычислительно нереализуемыми при сравнительно малых порядках конечной группы, равных примерно значению 2^{256}. Однако существование специальных методов решения ЗДЛ по простому модулю p, таких как метод вычисления индексов [7], обладающих субэкспоненциальной сложностью, требует использования в криптосхемах чисел p, имеющих достаточно большой размер, не менее 1024 бит. Появление прорывных специализированных методов решения ЗДЛ и ЗФ в обозримом будущем оценивается достаточно малыми значениями вероятности, тем не менее снижение вероятности взлома криптосхем в результате применения прорывных решений является важным моментом для криптосхем, применяемых на практике. Если криптосхема устроена таким образом, что для ее взлома требуется решить обе указанные задачи, то вероятность ее взлома существенно снижается, так как в этом случае необходима одновременная реализация двух маловероятных событий. Получение точных оценок рассматриваемых вероятностей проблематично, однако существенность указанного снижения вероятности взлома, основанного на прорывных решениях трудных задач, достаточно очевидна.

Следует отметить, что необходимость решать две трудные задачи практически не приводит к повышению стойкости криптосхем, поскольку при взломе криптосхемы задачи решаются независимо друг от друга. Однако если сложности решения ЗФ и ЗДЛ по простому модулю примерно равны, то появление прорывного решения одной из этих задач не приводит к снижению стойкости заданной криптосхемы. Известно, что ЗФ составного модуля n и ЗДЛ по простому модулю p имеют субэкспоненциальную сложность, причем сложности решения этих задач примерно одинаковы, если размеры чисел n и p равны и делители числа n имеют примерно одинаковый размер. Если делители числа n имеют различный размер, то сложность ЗФ определяется делителем меньшего размера [8]. Идея предлагаемого подхода состоит в построении криптосхем с использованием трудности ЗДЛ по трудно разложимому модулю n, для которого выполняется следующее условие: размер минимального делителя r модуля в 2 раза меньше разрядности второго делителя q. В этом случае сложность решения ЗФ примерно равна сложности ЗДЛ по простому модулю q. Построение криптосхем выполняется по аналогии с известными криптосхемами, основанными на трудности ЗДЛ по простому модулю, с учетом того, что значения оснований дискретных логарифмов следует выбирать таким образом, чтобы их нельзя было бы использовать для выполнения вычислительно осуществимых алгоритмов факторизации модуля n.

Выбор параметров криптосхем

В криптосхемах, создаваемых в рамках предложенного подхода, используется ОК, представляемый тройкой чисел $\{n, \alpha, y\}$, где y вычисляется по формуле

$$y = \alpha^x \bmod n.$$

Личным секретным ключом (ЛСК) пользователя является тройка чисел (r, q, x), где $n = rq$, q — простое 1024-битовое число, r — простое 512-битовое число; x — случайное число, меньшее, чем порядок числа α по модулю n, который обозначим как число γ. Требования к генерации элементов секретного ключа рассмотрены в работе [9], где показано, что число α с достаточно малым значением порядка (требование малого порядка генератора группы α необходимо для построения схем ЭЦП с малым размером подписи) может быть использовано для факторизации числа n, если числа r, q и α не удовлетворяют одному из следующих двух требований.

1. Простые числа r и q имеют следующую структуру: $r = N_r\gamma + 1$ и $q = N_q\gamma + 1$, где N_r и N_q — два больших четных числа, содержащих большой простой делитель. Параметр γ имеет размер не менее 160 бит и не является секретным.

2. Простые числа r и q представляются в виде $r = N_r\gamma' + 1$ и $q = N_q\gamma'' + 1$, где N_r и N_q — два больших четных числа, содержащих большой простой делитель. Значение порядка числа α равно $\gamma = \gamma'\gamma''$. Каждое из чисел γ' и γ'' имеет размер не менее 80 бит, а параметр γ является дополнительным элементом секретного ключа.

Генерация ОК и ЛСК в соответствии с одним из этих требований может быть легко выполнена [9], поэтому указанные требования не препятствуют практическому применению криптосхем на основе ЗДЛ по модулю n специальной структуры.

Криптографические протоколы, основанные на сложности ЗДЛ по трудно разложимому модулю

В этом разделе описаны протоколы ЭЦП, слепой подписи, коллективной подписи, а также протоколы открытого шифрования, открытого распределения ключей. Системные параметры и ключи пользователя, открытый и закрытый, формируются согласно требованиям, определенным в предыдущем разделе.

Протокол ЭЦП

Генерация подписи к сообщению M выполняется следующим образом.

1. Сформировать случайное число $k < \gamma$ и вычислить параметр $R = \alpha^k \bmod n$.

2. Используя некоторую специфицированную хэш-функцию F_H, вычислить ее значение от сообщения с присоединенным к нему числом R: $E = F_H(M, R)$. Параметр E является первым элементом подписи.

3. Вычислить второй элемент подписи: $S = k + xE \bmod \gamma$.

Проверка подлинности подписи (E, S) к сообщению M выполняется по ОК владельца подписи следующим образом.

1. Вычислить значение $\tilde{R} = y^{-E}\alpha^S \bmod n$.

2. Вычислить значение $\tilde{E} = F_H\left(M, \tilde{R}\right)$. Если $\tilde{E} = E$, то подпись признается подлинной.

Протокол слепой подписи

Протокол слепой подписи используется в тех случаях, когда пользователь желает получить подпись к сообщению M таким образом, чтобы подписывающий не мог впоследствии при получении M и соответствующей подписи идентифицировать этого пользователя. Протокол слепой подписи построен на основе алгоритма ЭЦП, описанного в предыдущем подразделе, и реализуется по аналогии со способом, впервые предложенным в работе [10], следующим образом.

1. Пользователь инициирует взаимодействие с подписывающим лицом.

2. Подписывающий генерирует случайное число k и вычисляет значение параметра $\bar{R} = \alpha^k \bmod n$, которое затем отправляет пользователю.

3. Пользователь генерирует случайные числа τ и ε (ослепляющие параметры, не превосходящие γ) и вычисляет значения $R = \bar{R}y^{\tau}\alpha^{\varepsilon} \bmod n$, $E = F_H(M, R)$ и $\bar{E} = E + \tau \bmod \gamma$, после чего отправляет подписывающему значение $\bar{E}$.

4. Подписывающий вычисляет значение $\bar{S}$ такое, что $\bar{R} = y^{-\bar{E}}\alpha^{\bar{S}} \bmod n$ (т. е. $\bar{S} = k + x\bar{E} \bmod \gamma$), и направляет $\bar{S}$ пользователю.

5. Пользователь вычисляет второй элемент подписи (E, S) к сообщению M по формуле $S = \bar{S} + \varepsilon \bmod \gamma$.

Процедура проверки подписи выполняется так же, как и в схеме ЭЦП. Подлинность подписи доказывается путем подстановки значения (E, S) на вход процедуры проверки подлинности:

$$\tilde{R} \equiv y^{-E}\alpha^S \equiv y^{-\bar{E}+\tau}\alpha^{\bar{S}+\varepsilon} \equiv$$

$$\equiv y^{-\bar{E}}y^{\tau}\alpha^{\bar{S}}\alpha^{\varepsilon} \equiv \left(y^{-\bar{E}}\alpha^{\bar{S}}\right)y^{\tau}\alpha^{\varepsilon} \equiv \bar{R}y^{\tau}\alpha^{\varepsilon} \bmod n \Rightarrow$$

$$\Rightarrow \tilde{R} = R \Rightarrow \tilde{E} = E.$$

При этом проблема анонимности решена, поскольку произвольная подпись (E, S), сформированная подписывающим, может быть сопоставлена с любой слепой подписью $\left(\bar{E}, \bar{S}\right)$. Действительно, если выполняются равенства $R = y^{-E}\alpha^S \bmod n$ и $\bar{R} = y^{-\bar{E}}\alpha^{\bar{S}} \bmod n$, то верны и следующие сравнения: $R/\bar{R} \equiv y^{\bar{E}-E}\alpha^{S-\bar{S}} \equiv y^{\tau}\alpha^{\varepsilon} \bmod n$, т. е. при равновероятном случайном выборе «ослепляющих» параметров τ и ε подпись (E, S) с равной вероятностью могла быть порождена из любой слепой подписи, формировавшейся когда-либо подписывающим.

Следует отметить, что в данном протоколе приемлемы два варианта построения модуля n. В первом варианте параметр γ не является составным и входит в состав ОК пользователя, поскольку его раскрытие не может быть использовано для разложения модуля. Во втором варианте параметр γ является составным и является частью ЛСК пользователя. В этом случае вместо формул $\bar{E} = E + \tau \bmod \gamma$ и $S = \bar{S} + \varepsilon \bmod \gamma$ можно использовать формулы $\bar{E} = E + \tau \bmod 2^g$ и $S = \bar{S} + \varepsilon \bmod 2^g$ соответственно. В остальном описание протокола не меняется. В последних двух формулах значение g является специфицируемым параметром протокола и превосходит на единицу значение разрядности параметра γ.

Протокол коллективной ЭЦП

Протокол коллективной ЭЦП необходим в случаях, когда несколько пользователей должны одновременно сформировать подпись к документу. В настоящей работе предлагается протокол коллективной ЭЦП, аналогичный протоколам с формированием общего параметра рандомизации [11, 12], однако системные параметры (n, α, γ) протокола, генерируемые доверительным центром, и ОК y_i и ЛСК x_i m пользователей, где $i = 1, 2, ..., m$, вычисляются с учетом требований, определенных в предыдущем разделе. Протокол коллективной ЭЦП включает следующие шаги.

1. Участники протокола вычисляют коллективный ОК $Y = y_1 y_2 ... y_m \bmod n$.

2. Каждый i-й пользователь формирует случайное число $k_i < \gamma$ и вычисляет значение $R_i = \alpha^{k_i} \bmod n$ и рассылает его остальным пользователям.

3. Пользователи вычисляют общий рандомизирующий параметр $R = R_1 R_2 ... R_m \bmod n$ и первый элемент коллективной подписи $E = F_H(M, R, Y)$.

4. Каждый i-й пользователь вычисляет свою долю во втором элементе коллективной подписи: $S_i = k_i + x_i E \bmod \gamma$, и рассылает ее остальным пользователям.

5. После этого пользователи вычисляют второй элемент коллективной подписи (E, S) по формуле $S = S_1 + S_2 + \ldots + S_m \bmod \gamma$.

Проверка подлинности коллективной подписи (E, S) к сообщению M выполняется следующим образом.

1. Вычислить коллективный ОК $Y = y_1 y_2 \ldots y_m \bmod n$ и значение $\tilde{R} = Y^{-E} \alpha^S \bmod n$.

2. Вычислить значение $\tilde{E} = F_H\left(M, \tilde{R}, Y\right)$. Если $\tilde{E} = E$, то подпись признается подлинной.

Используя в качестве прототипа протоколы, описанные в работах [13–15], легко разработать протокол коллективной слепой подписи, в котором общая тройка значений n, α и γ генерируется доверительным центром.

Протокол открытого шифрования

Используя ОК некоторого пользователя, можно послать секретное сообщение этому пользователю по открытым каналам связи. Для этого сообщение следует зашифровать по ОК, применяя следующий алгоритм, построенный по аналогии с алгоритмом открытого шифрования Эль-Гамаля [16].

1. Сгенерировать случайное число k.

2. Вычислить число $R = \alpha^k \bmod n$.

3. Используя ОК получателя y, вычислить значение $Q = y^k \bmod n$.

4. Зашифровать сообщение M путем умножения сообщения на значение Q, играющее роль разового ключа шифрования: $C = QM \bmod n$.

5. Отправить получателю криптограмму в виде пары чисел (R, C).

Получатель криптограммы (R, C), используя свой ЛСК x, выполняет процедуру дешифрования сообщения M, которая описывается следующими шагами.

1. Вычислить разовый общий секретный ключ $Q' = R^x \bmod n$.

2. Используя расширенный алгоритм Евклида, вычислить значение Q'^{-1}, обратное значению Q' по модулю n. (Легко показать, что число Q' является взаимно простым с модулем n, поэтому обратное значение Q'^{-1} существует и легко вычисляется с помощью расширенного алгоритма Евклида.)

3. Расшифровать сообщение M путем умножения значения C на целое число Q'^{-1}: $M = CQ'^{-1} \bmod n$.

Корректность описанной схемы шифрования легко доказать самостоятельно.

Протокол открытого распределения ключей

Для реализации протокола открытого распределения необходимо участие доверительного центра, который генерирует системные параметры протокола, отвечающие требованиям, определенным в предыдущем разделе. Пользователи генерируют случайный ЛСК в виде числа x и вычисляют свой ОК по формуле $y = \alpha^x \bmod n$. Протокол включает стандартные шаги схемы Диффи — Хеллмана.

1. Пользователь А вычисляет общий секретный ключ с удаленным пользователем В по ОК последнего y_B и своему ЛСК x_A по формуле $Z_{AB} = y_B^{x_A} \bmod n$.

2. Пользователь В вычисляет общий секретный ключ с пользователем А по ОК последнего y_A и своему ЛСК x_B по формуле $Z_{AB} = y_A^{x_B} \bmod n$.

В результате этих шагов оба пользователя получают одно и то же значение, которое известно только им, и для этого не потребовалось использовать защищенный канал связи.

Если атаку на протокол проводить с участием доверительного центра, то она окажется эффективной в случае появления прорывного решения ЗДЛ по простому модулю. Таким образом, стойкость протокола к таким атакам примерно равна его стойкости к атакам без участия доверительного центра, однако его безопасность существенно выше, если учесть вероятность взлома в результате появления прорывных решений трудных задач.

Протокол аутентификации с нулевым разглашением секрета

Протоколы с нулевым разглашением используются в процедурах строгой аутентификации удаленных абонентов телекоммуникационных систем. Пользователь, подлинность которого устанавливается, называется *доказывающим*. Пользователь, который проверяет подлинность доказывающего, называется *проверяющим*. Термин «*нулевое разглашение секрета*» подчеркивает, что при обмене информацией между доказывающим и проверяющим не происходит какой-либо утечки информации о ЛСК доказывающего.

Рассмотрим протокол с нулевым разглашением на основе сложности ЗДЛ по составному модулю n, в котором ОК вычисляется по формуле $y = \alpha^x \bmod n$. Доказывающий в ходе протокола показывает, что он знает ЛСК x, соответствующий его ОК. Протокол состоит из многократного выполнения следующего раунда.

1. Доказывающий выбирает текущий разовый секрет k, вычисляет значение $R = \alpha^k \bmod n$, которое играет роль разового ОК, и передает его проверяющему.

2. Проверяющий случайным образом выбирает значение бита e и направляет его доказывающему (случайный запрос проверяющего).

3. Доказывающий направляет проверяющему ответ w в виде числа, вычисляемого по формуле $w = ex + k \bmod \gamma$, и направляет его проверяющему.

Проверяющий проверяет выполнимость соотношения $\alpha^w \equiv y^e R \bmod n$. При положительной проверке делается заключение, что доказывающий знает значение x. Нарушитель может дать правильный ответ с вероятностью 0,5, поэтому протокол включает в себя многократное выполнение описанного раунда, при котором достигается приемлемо малая вероятность обмана 2^{-h}, где h — число повторенных раундов.

С целью уменьшить большое число интерактивных шагов можно использовать трехшаговый протокол с нулевым разглашением, в котором ОК доказывающего является набор из h значений y_i, $i = 0, 1, 2, ..., h - 1$, вычисленных по формуле $y_i = \alpha^{x_i} \bmod n$. Набор чисел x_i, $i = 0, 1, 2, ..., h - 1$, составляет ЛСК доказывающего. Протокол включает три следующих шага.

1. Доказывающий выбирает случайное число k такое, что $1 < k < \gamma$, вычисляет значение $R = \alpha^k \bmod n$ и посылает его проверяющему (значение R называется *фиксатором*).

2. Проверяющий отправляет доказывающему запрос в виде случайной равновероятной h-битовой строки $E = (e_0, e_1, ..., e_{h-1})$, в которой каждый бит e_i с вероятностью 0,5 равен 1.

3. Доказывающий вычисляет ответ W на запрос E по формуле $W = k + \sum_{i=0}^{h-1} x_i e_i \bmod \gamma$ и направляет его проверяющему.

Проверяющий считает ответ положительным, если выполняется соотношение

$$\alpha^W = R \prod_{i=0}^{h-1} y_i^{e_i} \bmod n.$$

Легко показать, что вероятность обмана проверяющего в этом протоколе составляет 2^{-h}.

Описанный трехпроходный протокол с нулевым разглашением может быть преобразован в схему ЭЦП, пригодную для практического использования, по аналогии с построениями, выполненными в работе [17]. В таком преобразовании рассматривается следующий сценарий формирования цифровой подписи. Подписывающее лицо генерирует конкретное значение фиксатора R. Затем в зависимости от фиксатора и подписываемого документа M он вычисляет значение запроса E, после чего формирует ответ S на запрос. Пара чисел (E, S), включающая запрос и ответ, является цифровой подписью к документу. Для того чтобы подделка подписи была практически невозможной, схема ЭЦП строится таким образом, чтобы после вычисления значения запроса изменить значение фиксатора без изменения запроса было вычислительно трудно. Это требование может быть достигнуто путем задания запроса как значения стойкой хэш-функции, вычисляемой от значения фиксатора с присоединенным к нему сообщением. В этом значение запроса зависит от каждого бита фиксатора и каждого бита подписываемого документа, и без знания ЛСК формирование подписи становится вычислительно невыполнимым. Согласно описанному способу преобразования трехпроходного протокола с нулевым разглашением в протокол ЭЦП, алгоритм генерации ЭЦП включает следующие шаги.

1. Подписывающий выбирает случайное число k $(1 < k < \gamma)$, вычисляет значение фикатора $R = \alpha^k \bmod n$.

2. Затем, используя специфицированную хэш-функцию F_H, он вычисляет значение $E = F_H(M, R) = (e_0, e_1, ..., e_{h-1})$, которое может быть рассмотрено как случайный запрос со стороны документа.

3. Подписывающий вычисляет ответ S на запрос E по формуле $S = k + \sum_{i=0}^{h-1} x_i e_i \bmod \gamma$, который является вторым элементом цифровой подписи к документу M.

Проверка подлинности ЭЦП (E, S) состоит в проверке выполнимости соотношения $\alpha^S = R \prod_{i=0}^{h-1} y_i^{e_i} \bmod n$, которое является проверочным соотношением в исходном протоколе с нулевым разглашением. Процедура проверки подписи включает следующие шаги.

1. Вычисляется значение фиксатора

$$\tilde{R} = \alpha^S \prod_{i=0}^{h-1} y_i^{-e_i} \bmod n.$$

2. Вычисляется значение хэш-функции

$$\tilde{E} = F_H\left(M, \tilde{R}\right).$$

3. Сравниваются значения $\tilde{E}$ и E. Если $\tilde{E} = E$, то подпись (E, S) считается подлинной. В противном случае подпись отклоняется как ложная.

Нарушитель, который пытается подделать подпись, не может вычислить правильный ответ на запрос, формируемый по значению документа, поскольку ему неизвестен ЛСК. Однако он может попытаться сгенерировать случайные значения запроса $E' = \left(e'_0, e'_1, \ldots, e'_{h-1}\right)$ и ответа S' и вычислить по формуле $R' = \alpha^{S'} \prod_{i=0}^{h-1} y_i^{-e'_i} \bmod n$ значение фиксатора R', которое вместе со значениями E' и S' будет удовлетворять проверочному соотношению. С вероятностью 2^{-h} будет выполняться соотношение $E' = F_H(M, R')$, и подпись (E', S') пройдет процедуру проверки как подлинная подпись. Однако чтобы такая атака с вероятностью 50 % привела к удачной подделке подписи, потребуется сформировать примерно 2^{h-1} вариантов подписи (E', S'), поэтому при $h \geq 80$ атака вычислительно невыполнима в настоящее время.

Недостатком данной схемы ЭЦП является большой размер ОК, который составляет не менее $1536h = 122\,880$ бит. Сокращение размера ОК

можно достигнуть приемом, который состоит в том, что генерируется ЛСК в виде случайного числа x, по которому вычисляется значение $y_0 = \alpha^x \bmod n$, а значения y_i, $i = 1,..., h - 1$, определяются формулой $y_i = y_0^{2^i} = \alpha^{2^i x} = \alpha^{x_i} \bmod n$. Тогда в описанной схеме ЭЦП $x_i = 2^i x \bmod \gamma$ и выражение $S = k + \sum_{i=0}^{h-1} x_i e_i \bmod \gamma$ принимает вид $S = k + xE \bmod \gamma$, а выражение $\alpha^S = R\prod_{i=0}^{h-1} y_i^{e_i} \bmod n$ приводится к виду

$$\alpha^S = R\prod_{i=0}^{h-1} y_i^{e_i} = R\prod_{i=0}^{h-1} y_0^{2^i e_i} = Ry_0^{\sum_{i=0}^{h-1} 2^i e_i} = Ry_0^E \bmod n,$$

где битовая строка запроса E рассматривается как двоичное число. Заменяя обозначение y_0 на y, можно перейти к схеме ЭЦП, описываемой следующими шагами.

1. Подписывающий выбирает случайное число k ($1 < k < \gamma$), вычисляет значение фиксатора $R = \alpha^k \bmod n$.

2. Затем он, используя специфицированную хэш-функцию F_H, вычисляет первый элемент подписи в виде значения $E = F_H(M, R) = (e_0, e_1, ..., e_{h-1})$, которое рассматривается как двоичное число $E = \sum_{i=0}^{h-1} 2^i e_i$.

3. Подписывающий вычисляет ответ S на запрос E по формуле $S = k + xE \bmod \gamma$, который является вторым элементом цифровой подписи к документу M.

Процедура проверки подписи (E, S) включает следующие шаги.

1. Вычисляется значение фиксатора

$$\tilde{R} = y^{-E}\alpha^S \bmod n.$$

2. Вычисляется значение хэш-функции

$$\tilde{E} = F_H\left(M, \tilde{R}\right).$$

3. Сравниваются значения $\tilde{E}$ и E. Если $\tilde{E} = E$, то подпись (E, S) считается подлинной. В противном случае подпись отклоняется как ложная.

То есть получен протокол ЭЦП, описанный в первом подразделе. Его «вывод» из протокола с нулевым разглашением может быть использован как формальное доказательство его стойкости. Примеры такого доказательства приведены в работе [17].

Общее обсуждение предложенных криптосхем

Во всех криптографических протоколах и алгоритмах, описанных в предыдущем разделе, составной модуль n может быть сформирован доверительным центром. Для некоторых криптосхем это условие является обязательным (протоколы коллективной подписи, открытого распределения ключей, коллективной слепой ЭЦП), для других — альтернативным вариантом реализации (протоколы обычной и слепой подписи, алгоритм открытого шифрования, протокол с нулевым разглашением). Следует отметить, что в криптосхемах последнего типа генерация трудно разложимого модуля самими пользователями является предпочтительным вариантом использования, поскольку в этом случае устраняются атаки с участием недобросовестного доверительного центра. При этом параметры n и α являются уникальными для каждого пользователя и должны быть включены в состав ОК, чтобы предоставить к ним доступ другим пользователям. В результате размер ОК увеличивается.

При индивидуальной генерации модуля n и числа α значение γ может оставаться секретным. Для последнего случая можно использовать как простое 160-битовое значение γ, так и составное значение γ, равное произведению двух 80-битовых простых чисел γ' и γ''. В случае генерации модуля n и числа α доверительным центром последний должен также вычислить и сделать общедоступным значение порядка числа γ, поэтому вариант составного числа γ является неприемлемым. (Это связано с тем, что значение γ выбирается сравнительно малого размера, поэтому его можно разложить на множители, использование которых дает возможность достаточно просто факторизовать модуль n.)

В рассмотренных криптосхемах размеры их параметров выбирались с учетом обеспечения минимально приемлемой стойкости, оцениваемой как 2^{80} модульных умножений. Для обеспечения более высокого уровня стойкости размер параметров должен быть соответствующим образом увеличен.

В предложенных схемах ЭЦП обеспечивается достаточно малый размер подписи (≈ 240 бит). Производительность криптосхем примерно в 2,25 раза меньше производительности аналогичных известных криптосхем, использующих вычисления по простому 1024-битовому модулю. Однако, поскольку последние обладают высоким быстродействием, то это снижение производительности не является существенным для практического применения.

Заключение

В данной работе описан и обоснован подход к построению криптосхем на основе трудности ЗДЛ по трудно разложимому модулю. Показано, что в рамках предложенного подхода повышается безопасность криптосхем по сравнению со схемами, использующими ЗФ или ЗДЛ по простому модулю. Производительность разработанных схем снижается незначительно. По сравнению с ранее известными подходами к синтезу криптосхем, основанными на трудности одновременного решения ЗФ и ЗДЛ по простому модулю, предложенный подход обеспечивает существенное со-

кращение размера подписи в протоколах обычной, коллективной и слепой ЭЦП. Кроме того, он может быть использован для разработки других типов криптографических протоколов, взлом которых требует одновременного решения ЗФ и ЗДЛ по простому модулю.

В предложенных криптосхемах применяется циклическая подгруппа мультипликативной группы кольца вычетов по составному модулю. Значительный интерес представляет синтез криптосхем, построенных на нециклических подгруппах, а именно мультипликативных подгруппах с двухмерной цикличностью, т. е. подгруппах, базис которых содержит два элемента α и β, имеющих один и тот же простой порядок γ. В последнем случае элемент ОК y вычисляется по формуле $y = \alpha^x \beta^w \bmod n$, где значения x и w являются элементами ЛСК. Применение нециклических подгрупп такого типа описано в работах [18–20] для построения протокола слепой ЭЦП со сравнительно малым размером подписи, основанной на трудности задачи факторизации. Однако использование таких подгрупп для построения алгоритмов и протоколов, основанных на сложности одновременного решения ЗФ и ЗДЛ, не рассматривалось и представляет собой предмет отдельного обсуждения.

Работа выполняется при финансовой поддержке РФФИ (проект № 12-07-31164 мол_а).

Литература

1. **Дернова Е. С., Молдовян Н. А.** Синтез алгоритмов цифровой подписи на основе нескольких вычислительно трудных задач // Вопросы защиты информации. 2008. № 1. С. 22–26.
2. **Дернова Е. С., Молдовян Н. А.** Протоколы коллективной цифровой подписи, основанные на сложности решения двух трудных задач // Безопасность информационных технологий. 2008. № 2. С. 79–85.
3. **Tahat N. M. F., Shatnawi S. M. A., Ismail E. S.** New Partially Blind Signature Based on Factoring and Discrete Logarithms // J. of Mathematics and Statistics. 2008. Vol. 4(2). P. 124–129.
4. **Tahat N. M. F., Ismail E. S., Ahmad R. R.** A New Blind Signature Scheme Based on Factoring and Discrete Logarithms // Intern. J. of Cryptology Research. 2009. Vol. 1(1). P. 1–9.
5. **Кишмар Р. В., Молдовяну П. А., Новикова Е. С., Сухов Д. К.** Протоколы слепой подписи на основе сложности одновременного решения двух трудных задач // Изв. СПбГЭТУ «ЛЭТИ». 2011. № 4. С. 44–48.
6. **Молдовян Д. Н., Кишмар Р. В., Васильев И. Н.** Двухключевые криптосхемы на основе комбинирования задач факторизации и дискретного логарифмирования // Вопросы защиты информации. 2011. № 4. С. 2–5.
7. **Menezes A. J., Vanstone S. A.** Handbook of Applied Cryptography. — CRC Press, 1996. — 780 p.
8. **Коблиц Н.** Курс теории чисел и криптографии. — М.: ТВП, 2001. — 254 с.
9. **Moldovyan A. A., Moldovyan D. N., Gortinskaya L. V.** Cryptoschemes based on new signature formation mechanism // Computer Science J. of Moldova. 2006. Vol. 14. N 3(42). P. 397–411.
10. **Camenisch J. L., Piveteau J.-M., Stadler M. A.** Blind Signatures Based on the Discrete Logarithm Problem // Advances in Crypology — EUROCRYPT'94. Proc. / Lecture Notes in Computer Science. Springer Verlang, 1995. Vol. 950. P. 428–432.
11. **Молдовян А. А., Молдовян Н. А.** Новые алгоритмы и протоколы для аутентификации информации в АСУ // Автоматика и телемеханика. 2008. № 7. С. 157–169.
12. **Молдовян А. А., Молдовян Н. А.** Коллективная ЭЦП — специальный криптографический протокол на основе новой трудной задачи // Вопросы защиты информации. 2008. № 1. С. 14–18.
13. **Moldovyan N. A., Moldovyan A. A.** Blind Collective Signature Protocol Based on Discrete Logarithm Problem // Intern. J. of Network Security. 2010. Vol. 11. N 2. P. 106–113.
14. **Moldovyan N. A.** Blind Signature Protocols from Digital Signature Standards // Intern. J. of Network Security. 2011. Vol. 13. N 1. P. 22–30.
15. **Молдовян Н. А., Дернова Е. С., Молдовян Д. Н.** Протоколы слепой и коллективной подписи на основе стандарта ЭЦП ДСТУ 4145-2002 // Вопросы защиты информации. 2011. № 2. С. 14–18.
16. **ElGamal T.** A public key cryptosystem and a signature scheme based on discrete logarithms // IEEE Transactions on Information Theory. 1985. Vol. IT-31. N 4. P. 469–472.
17. **Молдовян А. А., Молдовян Д. Н., Васильев И. Н., Головачев Д. А.** Протоколы с нулевым разглашением секрета и обоснование безопасности схем цифровой подписи // Вопросы защиты информации. 2011. № 4. С. 6–11.
18. **Молдовян Д. Н., Васильев И. Н., Латышев Д. М., Сухов Д. К.** Построение схемы 240-битовой цифровой подписи // Вопросы защиты информации. 2011. № 3. С. 6–10.
19. **Васильев И. Н., Краснова А. И., Молдовян Д. Н.** Схема слепой 240-битовой цифровой подписи // Информационно-управляющие системы. 2011. № 6(55). С. 49–53.
20. **Moldovyan A., Moldovyan N., Novikova E.** Blind 384-bit Digital Signature Scheme: 6th Intern. Conf. MMM-ACNS'12. St.-Petersburg, Russia, Oct. 17–20 // Springer LNCS. 2012. Vol. 7531. P. 77–83.

УДК 621.391.01

АЛГОРИТМ РАЗРЕШЕНИЯ НЕИЗВЕСТНОГО ЧИСЛА ЦЕЛЕЙ ПО ДАЛЬНОСТИ

В. В. Акимцев,
канд. техн. наук, доцент
Санкт-Петербургское высшее военное училище радиоэлектроники (военный институт)

Предложена процедура разрешения по дальности неизвестного числа целей, основанная на анализе цепного отображения цифрового принимаемого сигнала импульсной радиолокационной станции обзора. Приведены характеристики ее качества для простейшего, но весьма распространенного случая, когда сигналы наблюдаются на фоне собственного шума приемника.

Ключевые слова *— разрешение целей по дальности, цифровая обработка сигналов, цепное отображение, непараметрическая статистика.*

Введение

Требования к разрешающей способности радиолокационной станции (РЛС) по различным координатам постоянно повышаются. Это вызвано возрастанием интенсивности воздушного движения и разработкой новых процедур радиолокационного наблюдения, связанных с распознаванием целей.

Задачу разрешения целей можно решать одновременно с задачей их обнаружения и измерения координат. Однако при большом количестве целей в зоне обзора РЛС, особенно когда их число N не известно, практическая реализация процедур обнаружения-измерения-разрешения в реальном времени может вызвать значительные трудности из-за больших вычислительных затрат [1]. По этой причине в ряде случаев целесообразно решать задачу разрешения не по всем целям на этапе их обнаружения и измерения координат, а после обнаружения всех целей в зоне обзора РЛС и лишь по тем целям, в отношении которых из каких-либо соображений необходимо получить дополнительную информацию. Применение таких процедур разрешения совместно с цифровой обработкой сигналов позволяет снизить вычислительные затраты и обеспечить требуемые для практических приложений значения разрешающей способности РЛС по дальности [2].

В работе [3] исследована модель оцифрованного в тракте промежуточной частоты сигнала, принимаемого импульсной РЛС. Если сигнал является наложением перекрывающихся во времени сигналов, отраженных от N целей, имеющих одинаковые угловые координаты, то он описывается матрицей

$$\mathbf{Y} = \left[\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_r, \ldots, \mathbf{y}_h\right] =$$

$$= \sum_{j=1}^{N} \left[\underbrace{\mathbf{o}, \ldots, \mathbf{o}}_{r_{0j}}, \underbrace{\mathbf{s}_1^{(j)}, \ldots, \mathbf{s}_r^{(j)}}_{r}, \underbrace{\mathbf{o}, \ldots, \mathbf{o}}_{r_{j0}} \right] + \mathbf{W} + \mathbf{N} =$$

$$= \left[\mathbf{0}_{0j},\ \mathbf{S}_j,\ \mathbf{0}_{j0}\right] + \mathbf{W} + \mathbf{N} = \sum_{j=1}^{N} \boldsymbol{\Xi}_j + \mathbf{W} + \mathbf{N}, \quad (1)$$

где $\mathbf{y}_k = \{y_1[t_1 + (k-1)\Delta_t], y_2[t_2 + (k-1)\Delta_t], \ldots, y_M[t_M + (k-1)\Delta_t]\}^{\mathrm{T}}$, $k = 1, 2, \ldots, h$ — столбцы матрицы $\mathbf{Y}$; T — знак транспонирования; M — число импульсов в пачке отраженных импульсов; $M \times h$ — размеры угломестного строба, в пределах которого анализируется отраженный сигнал; t_i, $i = 1, 2, \ldots, M$ — момент первого отсчета входного процесса приемника в i-м периоде зондирования; $r = \mathrm{Ent}(\tau_{\text{и}}/\Delta_t)$ — число отсчетов каждого импульса пачки отраженных импульсов; $\tau_{\text{и}}$ — длительность зондирующего импульса; Δ_t — шаг временной дискретизации входного процесса приемника; $\mathrm{Ent}(x)$ — целая часть x; $\boldsymbol{\Xi}_j = [\mathbf{0}_{0j}, \mathbf{S}_j, \mathbf{0}_{j0}]$ — блочная матрица; $\mathbf{0}_{0j}$, $\mathbf{0}_{j0}$ — нулевые блоки, состоящие из r_{0j} и r_{j0} нулевых столбцов соответственно; $r_{0j} + r + r_{j0} = h$;

$$\mathbf{S}_j = \left[\mathbf{s}_1^{(j)}, \mathbf{s}_2^{(j)}, \ldots, \mathbf{s}_r^{(j)}\right] = \sqrt{2P_j}\,\mathbf{G}^2\mathbf{Z}_j\left[\mathbf{s}_{01}^{(j)}, \mathbf{s}_{02}^{(j)}, \ldots, \mathbf{s}_{0r}^{(j)}\right] \quad (2)$$

— матрица, состоящая из столбцов $\mathbf{s}_k^{(j)}$ ($k = 1, 2, \ldots, r$), образованных отсчетами с одинаковыми номерами отраженного от j-й цели сигнала в каждом

из M периодов зондирования; $\mathbf{G} = \mathrm{diag}(g_1, ..., g_M)$ и $\mathbf{Z}_j = \mathrm{diag}(z_1^{(j)}, ..., z_M^{(j)})$ — диагональные матрицы, описывающие направленные свойства приемо-передающей антенны РЛС и флюктуации пачки отраженных импульсов соответственно; P_j — мощность отраженного от j-й цели сигнала; $\mathbf{W} = [\mathbf{w}_1, ..., \mathbf{w}_h]$ и $\mathbf{N} = [\mathbf{n}_1, ..., \mathbf{n}_h]$ — матрицы размера $M \times h$ отсчетов шума приемника $w(t)$ и помехи $n(t)$;

$$\mathbf{s}_{0k}^{(j)} = \left\{\cos\left[\omega_0\left(t_1^{(j)} + (k-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_1^{(j)}\right], \right.$$
$$\cos\left[\omega_0\left(t_2^{(j)} + (k-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_2^{(j)}\right], \ldots$$
$$\left. \ldots, \cos\left[\omega_0\left(t_M^{(j)} + (k-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_M^{(j)}\right]\right\}^{\mathrm{T}},$$
$$t_i^{(j)} + (k-1)\Delta_t - t_{\mathrm{d}j} \in [0, \tau_{\mathrm{и}}],$$
$$i = 1, \ldots, M,\ k = 1, \ldots, r \qquad (3)$$

— вектор k-х отсчетов гармонических функций, описывающих форму отраженного от j-й цели сигнала в M периодах зондирования; $t_i^{(j)}$ — момент первого отсчета отраженного от j-й цели импульса в i-м периоде зондирования; $\omega_0 = 2\pi f_0$ — частота заполнения импульсов пачки (промежуточная частота); $\varphi_i^{(j)}$ — начальная фаза отраженного от j-й цели импульса в i-м периоде зондирования; $t_{\mathrm{d}j}$ — время запаздывания отраженного от j-й цели сигнала.

Как видно из (1)–(3), свойства модели зависят от величины выбранного шага дискретизации Δ_t. Путем анализа этой зависимости с позиций использования модели в задаче разрешения сигналов по времени было показано [2, 3], что при

$$\Delta_t = \frac{n}{2f_0} < \tau_{\mathrm{и}},$$

где n — целое число, можно построить алгоритм разрешения-обнаружения импульсных сигналов по времени, основанный на оценке эффективного ранга матрицы $\mathbf{Y}$ (1). Некоторые характеристики качества такого алгоритма приведены в работе [4]. Если же

$$\Delta_t = \frac{n}{f_0} < \tau_{\mathrm{и}}, \qquad (4)$$

то можно построить непараметрические алгоритмы полного разрешения импульсных сигналов по времени, основанные на обнаружении статистической неоднородности некоторого числа столбцов матрицы $\mathbf{Y}$ (1) или какого-либо ее преобразования. Характеристики качества одного из вариантов непараметрического алгоритма разрешения импульсных сигналов по времени подробно исследованы в работе [5].

Однако свойства модели (1)–(3) не исчерпываются возможностью построения только двух указанных типов алгоритмов разрешения импульсных сигналов по времени. Модель позволяет предложить для решения задачи разрешения сигналов по времени и некоторые другие способы, в частности способ цепного отображения [6].

Свойства цепного отображения оцифрованного сигнала, принимаемого импульсной РЛС

Цепное отображение подразумевает отображение кластеров в некотором многомерном пространстве на пространство с меньшим числом измерений. Считая h столбцов $\mathbf{y}_k = [y_{k1}, y_{k2}, ..., y_{kM}]^{\mathrm{T}}$ матрицы $\mathbf{Y}$ векторами в M-мерном пространстве, можно вычислить расстояния между всеми соседними столбцами матрицы $\mathbf{Y}$ [7]

$$d_{k,\,k+1} = d(\mathbf{y}_k, \mathbf{y}_{k+1}) = \sqrt{\sum_{i=1}^{M}\left[y_{k+1,\,i} - y_{ki}\right]^2} = \|\mathbf{y}_{k+1} - \mathbf{y}_k\| = \sqrt{(\mathbf{y}_{k+1} - \mathbf{y}_k)^{\mathrm{T}}(\mathbf{y}_{k+1} - \mathbf{y}_k)},\ k = \overline{1, h-1}, \qquad (5)$$

где $\|\mathbf{y}_{k+1} - \mathbf{y}_k\|$ — норма вектора $\mathbf{y}_{k+1} - \mathbf{y}_k$. Расстояние $d_{k,\,k+1}$ может служить мерой различия векторов $\mathbf{y}_k$ и $\mathbf{y}_{k+1}$. Если векторы $\mathbf{y}_k$ «тесно кластеризированы» (образуют несколько групп, внутри которых различия между векторами мало, а различие между векторами, принадлежащим различным группам, — значительно больше), то кластеры можно распознать как совокупность векторов, расположенных между доминирующими значениями расстояния $d_{k,\,k+1}$.

При определении расстояний между соседними столбцами матрицы $\mathbf{Y}$ возможны следующие ситуации.

1. Пара $(\mathbf{y}_k, \mathbf{y}_{k+1}) = (\mathbf{w}_k + \mathbf{n}_k, \mathbf{w}_{k+1} + \mathbf{n}_{k+1})$ — соседние столбцы матрицы $\mathbf{Y}$ содержат лишь отсчеты шума приемника $w(t)$ и помехи $n(t)$ и не содержат отсчетов отраженных от целей сигналов (для этих векторов $1 \le k \le r_{0j}$ или $r_{0j} + r < k \le h - 1$ для всех $j = 1, 2, ..., N$). Рассматриваемые столбцы в статистическом смысле не различаются и, следовательно, должны находиться в пределах одного кластера. Квадрат расстояния между векторами этого кластера

$$d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right] = (\mathbf{y}_{k+1} - \mathbf{y}_k)^{\mathrm{T}}(\mathbf{y}_{k+1} - \mathbf{y}_k) = (\mathbf{w}_{k+1} + \mathbf{n}_{k+1} - \mathbf{w}_k - \mathbf{n}_k)^{\mathrm{T}}(\mathbf{w}_{k+1} + \mathbf{n}_{k+1} - \mathbf{w}_k - \mathbf{n}_k).$$

Использование правил матричной алгебры [8] в предположении, что помеха $n(t)$ и шум $w(t)$ — взаимно независимые стационарные процессы, дает

$$d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right] =$$
$$= \left(\mathbf{w}_{k+1}^{\mathrm{T}}\mathbf{w}_{k+1} + \mathbf{w}_k^{\mathrm{T}}\mathbf{w}_k + \mathbf{w}_{k+1}^{\mathrm{T}}\mathbf{w}_k + \mathbf{w}_k^{\mathrm{T}}\mathbf{w}_{k+1}\right) +$$
$$+ \left(\mathbf{n}_{k+1}^{\mathrm{T}}\mathbf{n}_{k+1} + \mathbf{n}_k^{\mathrm{T}}\mathbf{n}_k + \mathbf{n}_{k+1}^{\mathrm{T}}\mathbf{n}_k + \mathbf{n}_k^{\mathrm{T}}\mathbf{n}_{k+1}\right) =$$
$$= \left(\sum_{i=1}^{M} w_{k+1,i}^2 + \sum_{i=1}^{M} w_{ki}^2 - \sum_{i=1}^{M} w_{k+1,i}w_{ki} - \sum_{i=1}^{M} w_{ki}w_{k+1,i}\right) +$$
$$+ \left(\sum_{i=1}^{M} n_{k+1,i}^2 + \sum_{i=1}^{M} n_{ki}^2 - \sum_{i=1}^{M} n_{k+1,i}n_{ki} - \sum_{i=1}^{M} n_{ki}n_{k+1,i}\right) =$$
$$= 2M\left\{\left[\left(\sigma_w^*\right)^2\left(1-\rho_w^*(\Delta_t)\right)\right] + \left(\sigma_n^*\right)^2\left[1-\rho_n^*(\Delta_t)\right]\right\}, \quad (6)$$

где

$$\left(\sigma_w^*\right)^2 \approx \frac{1}{M}\sum_{i=1}^{M} w_{ki}^2 \approx \frac{1}{M}\sum_{i=1}^{M} w_{k+1,i}^2,$$

$$\left(\sigma_n^*\right)^2 \approx \frac{1}{M}\sum_{i=1}^{M} n_{ki}^2 \approx \frac{1}{M}\sum_{i=1}^{M} n_{k+1,i}^2$$

— оценки дисперсии процессов $w(t)$ и $n(t)$;

$$\rho_w^*(\Delta_t) \approx \frac{1}{M\left(\sigma_w^*\right)^2}\sum_{i=1}^{M} w_{ki}w_{k+1,i} \approx \frac{1}{M\left(\sigma_w^*\right)^2}\sum_{i=1}^{M} w_{k+1,i}w_{ki},$$

$$\rho_n^*(\Delta_t) \approx \frac{1}{M\left(\sigma_n^*\right)^2}\sum_{i=1}^{M} n_{ki}n_{k+1,i} \approx \frac{1}{M\left(\sigma_n^*\right)^2}\sum_{i=1}^{M} n_{k+1,i}n_{ki}$$

— оценки нормированных корреляционных функций $\rho_w(\tau)$ и $\rho_n(\tau)$ процессов $w(t)$ и $n(t)$ при значении аргумента $\tau = \Delta_t$. Так как $w(t)$ — δ-коррелированный случайный процесс, то $\rho_w^*(\Delta_t) \to 0$, и (6) приводится к окончательному виду

$$d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right] =$$
$$= 2M\left\{\left(\sigma_w^*\right)^2 + \left(\sigma_n^*\right)^2\left[1-\rho_n^*(\Delta_t)\right]\right\} =$$
$$= d^2(\mathbf{w})\left\{1 + q_{nw}^2\left[1-\rho_n^*(\Delta_t)\right]\right\}, \quad (7)$$

где $d^2(\mathbf{w}) = 2M(\sigma_w^*)^2$; $q_{nw}^2 = (\sigma_n^*/\sigma_w^*)^2$ — отношение помеха/шум.

Как видно из (7), расстояния между векторами рассматриваемого кластера статистически однородны (различаются только в силу конечного размера векторов $\mathbf{y}_k$) и определяются лишь статистическими свойствами шума приемника $w(t)$ и помехи $n(t)$. В частном случае, когда помеха отсутствует:

$$d^2(\mathbf{w}_k, \mathbf{w}_{k+1}) = d^2(\mathbf{w}) = 2M\left(\sigma_w^*\right)^2. \quad (8)$$

Отметим, что для различных типов помех $n(t)$, обладающих свойством стационарности, оценки $\rho_n^*(\Delta_t)$ также имеют различные значения и лишь таким образом влияют на величину $d^2[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})]$. Так, для сильно коррелированной помехи $\rho_n^*(\Delta_t) \to 1$ и

$$d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right] \approx$$
$$\approx d^2(\mathbf{w}_k, \mathbf{w}_{k+1}) = d^2(\mathbf{w}) = 2M\hat{\sigma}_w^2.$$

Для некоррелированной помехи $\rho_n^*(\Delta_t) \to 0$ и

$$d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right] \approx$$
$$\approx 2M\left(\hat{\sigma}_w^2 + \hat{\sigma}_n^2\right) = d^2(\mathbf{w})\left(1 + q_{nw}^2\right).$$

Для нестационарных помех оценки $\rho_n^*(\Delta_t)$ также могут быть формально определены, и от их величин будут зависеть значения $d^2[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})]$.

2. Пара $(\mathbf{y}_k, \mathbf{y}_{k+1}) = (\mathbf{s}_l^{(j)} + \mathbf{w}_k + \mathbf{n}_k, \mathbf{s}_{l+1}^{(j)} + \mathbf{w}_{k+1} + \mathbf{n}_{k+1})$ — соседние столбцы матрицы $\mathbf{Y}$ содержат отсчеты шума приемника $w(t)$, помехи $n(t)$, а также l-й и $(l+1)$-й отсчеты $(1 \le l \le r-1)$ отраженного от j-й цели сигнала $s^{(j)}(t)$. Рассматриваемые столбцы также находятся в пределах одного кластера. Квадрат расстояния между векторами этого кластера

$$d^2\left[\left(\mathbf{s}_l^{(j)}, \mathbf{w}_k, \mathbf{n}_k\right), \left(\mathbf{s}_{l+1}^{(j)}, \mathbf{w}_{k+1}, \mathbf{n}_{k+1}\right)\right] =$$
$$= \left(\mathbf{y}_{k+1} - \mathbf{y}_k\right)^{\mathrm{T}}\left(\mathbf{y}_{k+1} - \mathbf{y}_k\right) =$$
$$= \left(\mathbf{s}_{l+1}^{(j)} + \mathbf{w}_{k+1} + \mathbf{n}_{k+1} - \mathbf{s}_l^{(j)} - \mathbf{w}_k - \mathbf{n}_k\right)^{\mathrm{T}} \times$$
$$\times \left(\mathbf{s}_{l+1}^{(j)} + \mathbf{w}_{k+1} + \mathbf{n}_{k+1} - \mathbf{s}_l^{(j)} - \mathbf{w}_k - \mathbf{n}_k\right).$$

Элементарные вычисления с учетом (2), (3), (6) в предположении, что $s^{(j)}(t)$, $w(t)$ и $n(t)$ — взаимно независимые процессы, дают

$$d^2\left[\left(\mathbf{s}_l^{(j)}, \mathbf{w}_k, \mathbf{n}_k\right), \left(\mathbf{s}_{l+1}^{(j)}, \mathbf{w}_{k+1}, \mathbf{n}_{k+1}\right)\right] =$$
$$= 2P_j\sum_{i=1}^{M} g_i^4 z_i^2\left[s_{0l}^{(j)}\right]_i^2 + 2P_j\sum_{i=1}^{M} g_i^4 z_i^2\left[s_{0,l+1}^{(j)}\right]_i^2 -$$
$$- 4P_j\sum_{i=1}^{M} g_i^4 z_i^2 s_{0l}^{(j)} s_{0,l+1}^{(j)} + d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right]. \quad (9)$$

Значения сумм в (9)

$$\sum_{i=1}^{M} g_i^4 z_i^2\left[s_{0l}^{(j)}\right]_i^2 =$$
$$= \sum_{i=1}^{M} g_i^4 z_i^2 \cos^2\left[\omega_0\left(t_i^{(j)} + (l-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_i^{(j)}\right] =$$
$$= \frac{1}{2}\sum_{i=1}^{M} g_i^4 z_i^2\left\{1 + \cos 2\left[\omega_0\left(t_i^{(j)} + (l-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_i^{(j)}\right]\right\} \approx$$
$$\approx \frac{1}{2}\sum_{i=1}^{M} g_i^4 z_i^2,$$

так как

$$\sum_{i=1}^{M} g_i^4 z_i^2 \cos 2\left[\omega_0\left(t_i^{(j)} + (l-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_i^{(j)}\right] \to 0.$$

Аналогичным образом

$$\sum_{i=1}^{M} g_i^4 z_i^2 \left[s_{0,\,l+1}^{(j)}\right]_i^2 \approx \frac{1}{2}\sum_{i=1}^{M} g_i^4 z_i^2;$$

$$\sum_{i=1}^{M} g_i^4 z_i^2 s_{0l}^{(j)} s_{0,\,l+1}^{(j)} =$$

$$= \sum_{i=1}^{M} g_i^4 z_i^2 \cos\left[\omega_0\left(t_i^{(j)} + (l-1)\Delta_t - t_{\mathrm{d}j}\right) + \varphi_i^{(j)}\right] \times$$

$$\times \cos\left[\omega_0\left(t_i^{(j)} + l\Delta_t - t_{\mathrm{d}j}\right) + \varphi_i^{(j)}\right] \approx \frac{\cos\omega_0\Delta_t}{2}\sum_{i=1}^{M} g_i^4 z_i^2.$$

После подстановки значений сумм в (9)

$$d^2\left[\left(\mathbf{s}_l^{(j)}, \mathbf{w}_k, \mathbf{n}_k\right), \left(\mathbf{s}_{l+1}^{(j)}, \mathbf{w}_{k+1}, \mathbf{n}_{k+1}\right)\right] =$$

$$= 2P_j(1-\cos\omega_0\Delta_t)\sum_{i=1}^{M} g_i^4 z_i^2 +$$

$$+ d^2\left[(\mathbf{w}_k, \mathbf{n}_k), (\mathbf{w}_{k+1}, \mathbf{n}_{k+1})\right] =$$

$$= 2P_j(1-\cos\omega_0\Delta_t)\mathrm{tr}\left(\mathbf{G}^4\mathbf{Z}_j^2\right) +$$

$$+ d^2(\mathbf{w})\left\{1 + q_{nw}^2\left[1-\hat{\rho}_n(\Delta_t)\right]\right\} =$$

$$= 2MP_{\mathrm{cp}}^{(j)}(1-\cos\omega_0\Delta_t) +$$

$$+ d^2(\mathbf{w})\left\{1 + q_{nw}^2\left[1-\hat{\rho}_n(\Delta_t)\right]\right\}, \qquad (10)$$

где $\mathrm{tr}(\mathbf{X})$ — след матрицы $\mathbf{X}$; $P_{\mathrm{cp}}^{(j)} = P_j\mathrm{tr}(\mathbf{G}^4\mathbf{Z}_j^2)/M$ — средняя мощность импульсов пачки, отраженной от j-й цели [5].

Значение (10) отличается от (7) на величину $2MP_{\mathrm{cp}}^{(j)}(1 - \cos\omega_0\Delta_t)$, зависящую от выбранного шага дискретизации входного процесса приемника Δ_t. Так как рассматриваемые столбцы находятся в пределах одного кластера, то необходимо потребовать минимального значения расстояния между ними, т. е. минимального неотрицательного значения $2MP_{\mathrm{cp}}^{(j)}(1 - \cos\omega_0\Delta_t)$. Таким образом, величину Δ_t необходимо определить из условия $1 - \cos\omega_0\Delta_t = 0$, откуда

$$\Delta_t = \frac{n}{f_0} < \tau_{\mathrm{и}},$$

что совпадает с (4). При этом

$$d^2\left[\left(\mathbf{s}_l^{(j)}, \mathbf{w}_k, \mathbf{n}_k\right), \left(\mathbf{s}_{l+1}^{(j)}, \mathbf{w}_{k+1}, \mathbf{n}_{k+1}\right)\right] =$$

$$= d^2(\mathbf{w})\left\{1 + q_{nw}^2\left[1-\hat{\rho}_n(\Delta_t)\right]\right\},$$

как и в предыдущем случае.

Отметим, что аналогичный результат получится, если под $\mathbf{s}_l^{(j)}$ и $\mathbf{s}_{l+1}^{(j)}$ подразумевать наложение одного и того же числа одних и тех же сигналов.

3. Столбец $\mathbf{y}_k$ содержит лишь отсчеты шума приемника $w(t)$ и помехи $n(t)$, а столбец $\mathbf{y}_{k+1}$ дополнительно содержит первый отсчет сигнала, отраженного от j-й цели из состава групповой цели. Столбец $\mathbf{y}_{k+1}$, очевидно, связан с границей между двумя кластерами. Причем, если Δ_t удовлетворяет условию (4), то внутрикластерные расстояния $d_{i,\,j+1}(1 \le i \le k)$ и $d_{j,\,j+1}(k+1 \le j \le k+r)$ одинаковы и определяются из (7).

В соответствии с (5) квадрат расстояния между вектором $\mathbf{y}_k$ (последним вектором кластера $\mathbf{w}_k + \mathbf{n}_k$, $1 \le k \le r_{0j}$) и вектором $\mathbf{y}_{k+1}$ (первым вектором кластера $\mathbf{s}_l^{(j)} + \mathbf{w}_{k+1} + \mathbf{n}_{k+1}$, $1 \le l \le r$, $r_{0j} < k \le r_{0j} + r$)

$$d_{k,\,k+1}^2 = \left(\mathbf{y}_{k+1} - \mathbf{y}_k\right)^{\mathrm{T}}\left(\mathbf{y}_{k+1} - \mathbf{y}_k\right) =$$

$$= \left(\mathbf{s}_1^{(j)} + \mathbf{w}_{k+1} + \mathbf{n}_{k+1} - \mathbf{w}_k - \mathbf{n}_k\right)^{\mathrm{T}} \times$$

$$\times\left(\mathbf{s}_1^{(j)} + \mathbf{w}_{k+1} + \mathbf{n}_{k+1} - \mathbf{w}_k - \mathbf{n}_k\right).$$

Элементарные вычисления с учетом (2), (6), принятого предположения о независимости процессов $w(t)$ и $n(t)$ и указанных выше значений сумм дают

$$d_{k,\,k+1}^2 = 2P_j\sum_{i=1}^{M} g_i^4 z_i^2 \cos^2\left[\omega_0\left(t_i^{(j)} - t_{\mathrm{d}j}\right) + \varphi_i^{(j)}\right] +$$

$$+ 2M\left\{\left(\sigma_w^*\right)^2 + \left(\sigma_n^*\right)^2\left[1-\rho_n^*(\Delta_t)\right]\right\} =$$

$$= P_j\mathrm{tr}\left(\mathbf{G}^4\mathbf{Z}_j^2\right) + 2M\left\{\left(\sigma_w^*\right)^2 + \left(\sigma_n^*\right)^2\left[1-\rho_n^*(\Delta_t)\right]\right\} =$$

$$= MP_{\mathrm{cp}}^{(j)} + 2M\left\{\left(\sigma_w^*\right)^2 + \left(\sigma_n^*\right)^2\left[1-\rho_n^*(\Delta_t)\right]\right\} =$$

$$= M\left(\sigma_w^*\right)^2 q_j^2 + d^2(\mathbf{w})\left\{1 + q_{nw}^2\left[1-\rho_n^*(\Delta_t)\right]\right\}, \qquad (11)$$

где $q_j^2 = P_{\mathrm{cp}}^{(j)}/(\sigma_w^*)^2$ — отношение сигнал/шум для j-й цели.

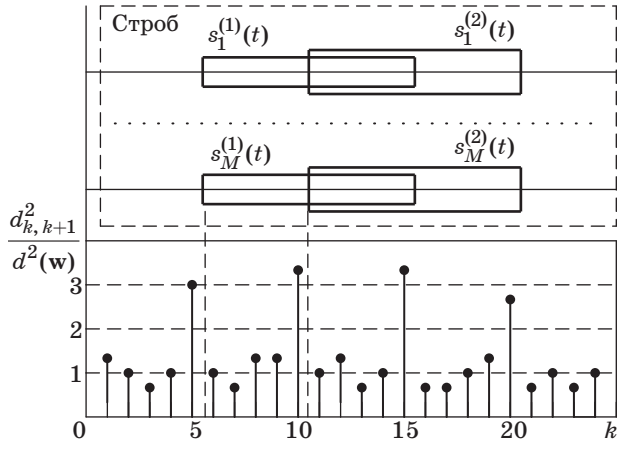
Таким образом, квадрат расстояния между векторами, принадлежащими различным кластерам, отличается от внутрикластерных расстояний на величину

$$\Delta d_{k,\,k+1}^2 = M\left(\sigma_w^*\right)^2 q_j^2. \qquad (12)$$

Отметим, что вектор $\mathbf{y}_{k+1}$ характеризует положение передних фронтов импульсов пачки, отраженной от j-й цели. Аналогичное значение $\Delta d_{k,\,k+1}^2$ получится и в симметричном случае — для задних фронтов импульсов этой пачки. Кроме того,

$$d_{k,\,k+1}^2 > d^2(\mathbf{w})\left\{1 + q_{nw}^2\left[1-\rho_n^*(\Delta_t)\right]\right\}$$

и в случае, когда векторы $\mathbf{y}_k$ и $\mathbf{y}_{k+1}$ содержат отсчеты наложения различного числа перекрывающихся во времени сигналов.



■ ***Рис. 1.*** Пример цепного отображения для двух не разрешаемых по критерию Рэлея пачек отраженных импульсов

В качестве примера приведено (рис. 1) цепное отображение для ситуации, когда в строб размера $M \times h$ попадают две перекрывающиеся во времени пачки отраженных импульсов. Полагалось, что

$$\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_h] = \sum_{j=1}^{2} \Xi_j + \mathbf{W},$$

$M = 60$, $h = 25$, $r = 10$, $q_1 = q_2 = q = 2$. Началу и окончанию отсчетов сигнала $s^{(1)}(t)$ соответствует $k_{н1} = 6$ и $k_{к1} = 15$, а сигнала $s^{(2)}(t)$ — $k_{н2} = 11$ и $k_{к2} = 20$. Временной сдвиг между перекрывающимися сигналами $\delta t_{12} \approx \tau_и/2 = 5\Delta_t$, Δ_t удовлетворяет условию (4). Такие сигналы не разрешаются по критерию Рэлея. На графике явно прослеживаются четыре границы между кластерами, которые связаны с передними и задними фронтами импульсов перекрывающихся пачек, причем, как это следует из (12), с увеличением отношения сигнал/шум q_j эти границы будут проявляться все отчетливее.

Обнаружение аномальных значений величин $d^2_{k,\,k+1}$, $k = 1, 2, \ldots, h - 1$, цепного отображения соседних столбцов матрицы $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_h]$ (1) и анализ их взаимных расположений с тех позиций, что они связаны с моментами появления передних или задних фронтов перекрывающихся импульсов, открывает возможность для построения алгоритма разрешения неизвестного числа целей по дальности.

Структура алгоритма разрешения неизвестного числа целей по дальности

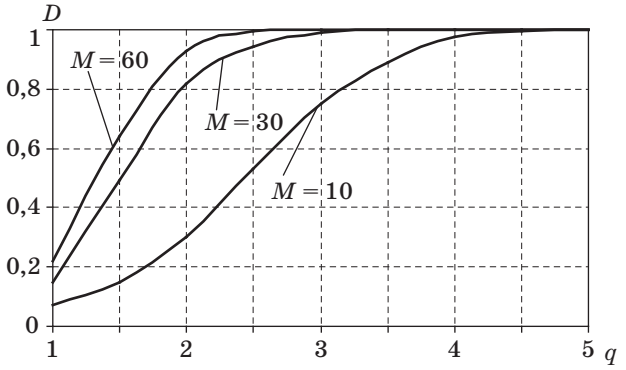
Чтобы обнаружение некоторого числа аномальных величин $d^2(\mathbf{y}_k, \mathbf{y}_{k+1})$ не носило субъективный характер, необходимо воспользоваться каким-либо статистическим критерием проверки значимости различия двух сравниваемых значений $d^2_{m,\,m+1}$ и $d^2_{n,\,n+1}$. Как видно из (7) и (11), величины $d^2(\mathbf{y}_k, \mathbf{y}_{k+1})$ имеют смысл выборочных дисперсий. Поэтому для проверки можно применить критерий Кокрена [9], который в данном случае сводится к сравнению с порогом величин:

$$G_{(m,\,m+1),\,(n,\,n+1)} = \frac{\max\left(d^2_{m,\,m+1}, d^2_{n,\,n+1}\right)}{d^2_{m,\,m+1} + d^2_{n,\,n+1}}. \quad (13)$$

Если $G_{(m,\,m+1),\,(n,\,n+1)} \le G_{кр}(\alpha, к, \chi)$, где $G_{кр}(\alpha, к, \chi)$ — критическая точка, определяемая уровнем значимости критерия α, количеством сравниваемых выборок $\chi = 2$ и числом степеней свободы $к = M - 1$, то принимается решение об однородности величин $d^2_{m,\,m+1}$ и $d^2_{n,\,n+1}$, в противном случае — о значимости различия этих величин (об их неоднородности).

Представим графики (рис. 2) зависимости вероятности D правильного обнаружения неоднородности величин $d^2_{k-1,\,k}$ и $d^2_{k,\,k+1}$ от отношения сигнал/шум q в условиях, когда векторы $\mathbf{y}_{k-1}$ и $\mathbf{y}_k$ содержат отсчеты только шума $w(t)$, а вектор $\mathbf{y}_{k+1}$ — отсчеты шума $w(t)$ и сигнала $s(t)$ (обнаружение переднего фронта импульсов отраженной пачки). Предполагалось, что уровень значимости $\alpha = 5$ %, а шаг дискретизации Δ_t удовлетворяет условию (4). Как видно из графиков, правильное обнаружение неоднородности величин $d^2_{k-1,\,k}$ и $d^2_{k,\,k+1}$ является практически достоверным событием уже при достаточно малых q, причем с увеличением размера пачки M требуется все меньшее значение q для практически достоверного обнаружения неоднородности.

В работе [5] приводятся аналогичные зависимости вероятности D правильного обнаружения неоднородности столбцов матрицы $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_h]$ от отношения сигнал/шум q, полученные с использованием методов непараметрической статистики. Хотя под однородностью там понимается принадлежность рассматриваемых столбцов ма-



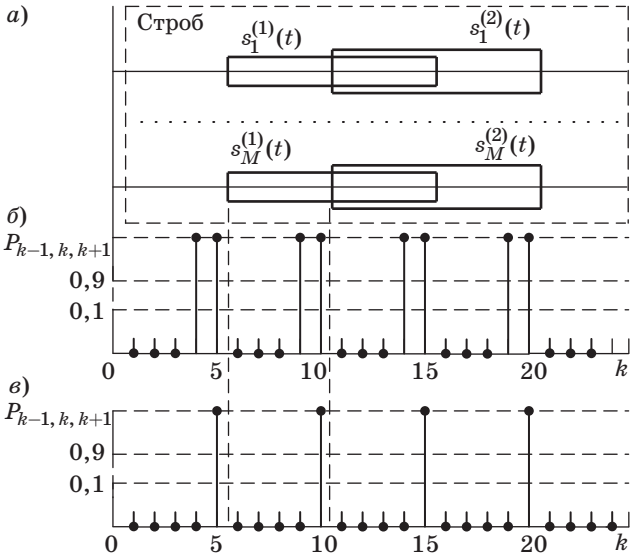
■ ***Рис. 2.*** Характеристики обнаружения аномальных значений цепного отображения для пачки сигналов, отраженных от одиночной цели

трицы **Y** к одному и тому же распределению вероятностей, тем не менее, обнаружение неоднородности соседних столбцов матрицы **Y** также означает обнаружение переднего или заднего фронта одного из перекрывающихся импульсов отраженной пачки. Сопоставление графиков рис. 2 с графиками работы [5] показывает, что оба способа обнаружения фронтов перекрывающихся импульсов отраженной пачки имеют практически одинаковые показатели качества. Однако, поскольку сравниваемые величины $d^2_{k-1,\,k}$ и $d^2_{k,\,k+1}$ (7) или (11) в методе цепного отображения не зависят от таких факторов как модуляция пачки диаграммой направленности приемо-передающей антенны РЛС и флюктуации импульсов отраженной пачки, то и значения вероятности D не зависят от перечисленных факторов, как это имеет место в непараметрическом алгоритме разрешения целей по дальности, предложенном в работе [5]. Это свойство цепного отображения, очевидно, следует отнести к его преимуществам.

Отметим, что при последовательном сравнении всех пар ($d^2_{k-1,\,k}$, $d^2_{k,\,k+1}$), что соответствует сравнению расстояний между столбцами $\mathbf{y}_{k-1}$, $\mathbf{y}_k$ и $\mathbf{y}_k$, $\mathbf{y}_{k+1}$ матрицы **Y** (1), обнаружение границы между кластерами (фронтов импульсов) всегда будет происходить дважды, так как вектор, связанный с границей между двумя кластерами (например, $\mathbf{y}_k$), участвует в формировании как величины $d^2_{k-1,\,k}$, так и величины $d^2_{k,\,k+1}$. Эта особенность, присущая последовательному сравнению пар величин ($d^2_{k-1,\,k}$, $d^2_{k,\,k+1}$), не имеет принципиального значения и может быть учтена при построении алгоритма разрешения целей по дальности на основе цепного отображения столбцов матрицы **Y**.

На рис. 3, *б* показана зависимость вероятностей $P_{k-1,\,k,\,k+1}$ принятия решения об обнаружении переднего и заднего фронтов отраженных импульсов пачки как функции дискретной величины k для приведенного выше примера (рис. 3, *а*). На графике видно, что даже при выбранном сравнительно небольшом отношении сигнал/шум по каждому из сигналов два решения об обнаружении фронтов импульсов принимаются с вероятностью, близкой к единице. Этим решениям соответствуют две оценки положения передних фронтов импульсов, отраженных от j-й цели, связанные с их истинными значениями $k_{н\,j}$ соотношениями $k^*_{н\,j} = k_{н\,j} - 1$ и $k^{**}_{н\,j} = k_{н\,j} - 2$.

Избежать двукратных обнаружений фронтов импульсов можно незначительной корректировкой процедуры сравнения пар ($d^2_{k-1,\,k}$, $d^2_{k,\,k+1}$), состоящей в следующем. Начиная процедуру сравнения, следует положить $\delta_1 = 0$ (δ_k — решение: неоднородность пары связана с k-м отсчетом). Далее, для $k \geq 2$ вычисляются статистики $G_{(k-1,\,k),\,(k,\,k+1)}$ (13).



■ ***Рис. 3.*** Вероятности обнаружения фронтов пачек импульсов, отраженных от двух целей: *а* — не разрешаемые по критерию Рэлея пачки отраженных импульсов; *б* — вероятности обнаружения фронтов импульсов пачки как функции номера отсчета k; *в* — вероятности обнаружения фронтов импульсов пачки как функции параметра k для скорректированной процедуры

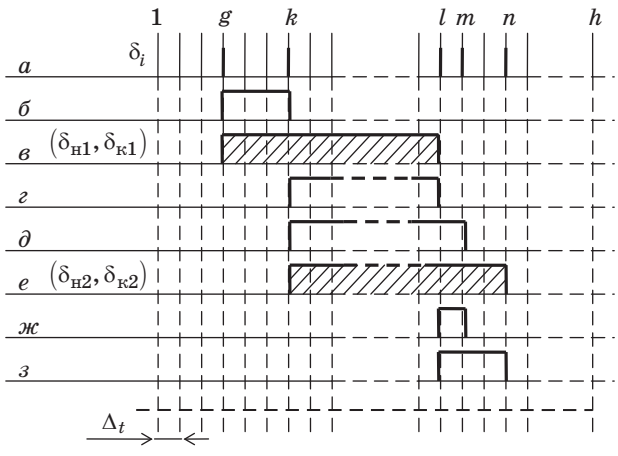
Пока $G_{(k-1,\,k),\,(k,\,k+1)} < G_{кр}(\alpha, к, \chi)$, полагается $\delta_k = 0$ и параметр k увеличивается на единицу. Как только $G_{(k-1,\,k),\,(k,\,k+1)} > G_{кр}(\alpha, к, \chi)$, что свидетельствует о том, что вектор $\mathbf{y}_{k+1}$ связан с моментом появления переднего или заднего фронта очередного импульса пачки, следует положить $\delta_k = 0$, $\delta_{k+1} = 1$, увеличить параметр k на две единицы и продолжить процедуру сравнения пар величин ($d^2_{k-1,\,k}$, $d^2_{k,\,k+1}$).

Зависимость вероятностей $P_{k-1,\,k,\,k+1}$ принятия решения об обнаружении переднего и заднего фронтов отраженных импульсов пачки как функции дискретной величины k в соответствии с описанной процедурой показана на рис. 3, *в*. Видно, что скорректированная процедура позволяет избавиться от двукратного обнаружения фронтов импульсов, а за оценку положения передних и задних фронтов импульсов можно принять величины $k^*_{н\,j} = k_{н\,j} - 1$ и $k^*_{к\,j} = k_{к\,j} - 1 = k^*_{н\,j} + r$.

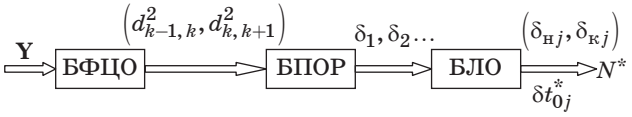
В результате применения описанной выше скорректированной процедуры формируется совокупность решений δ_i ($i = k_{н1} - 1$, $k_{н2} - 1$, ..., $k_{к1} - 1$, $k_{к2} - 1$, ...) об обнаружении передних и задних фронтов импульсов, отраженных от группы не разрешаемых по дальности целей. Для оценки конфигурации такой группы целей необходимо выделить из этой совокупности те пары ($\delta_{н\,j}$, $\delta_{к\,j}$), которые связаны с началом и окончанием отсчетов одного и того же сигнала, отраженного от j-й цели. Ясно, что число пар ($\delta_{н\,j}$, $\delta_{к\,j}$) будет оценкой

N^* числа N перекрывающихся сигналов, а временная задержка момента принятия решений $\delta_{н\,j}$ относительно выбранного начала отсчетов (например, относительно начала строба) определит оценку δt^*_{0j} временной задержки δt_{0j} отраженного от j-й цели сигнала.

Знание числа r отсчетов импульсов отраженной пачки с шагом дискретизации Δ_t позволяет предложить логическую процедуру формирования пар $(\delta_{н\,j}, \delta_{к\,j})$. Пусть в результате анализа стробируемого участка дальности, связанного с некоторой обнаруженной целью, получена последовательность решений δ_i об обнаружении передних и задних фронтов импульсов отраженной пачки (рис. 4, *а*). Момент принятия первого решения δ_g, очевидно, следует связать с моментом обнаружения переднего фронта отраженного от ближайшей цели сигнала, т. е. положить $\delta_{н1} = \delta_g$. Начиная с g-го отсчета, необходимо приступить к формированию последовательности импульсов, задние фронты которых определяются моментами принятия решений δ_k, δ_l, ... (рис. 4, *б*, *в*). Если число отсчетов между началом и окончанием какого-либо из сформированных импульсов равно r (например, $l - g = r$), то принимается, что $\delta_{к1} = \delta_l$. При этом локализуется сигнал, отраженный от ближайшей цели (см. рис. 4, *в*). Если же не найдется такого δ_i, для которого $i - g = r$, то считается, что δ_g — ошибочное решение. Затем аналогичная операция применяется к отсчету, связанному с моментом формирования решения δ_k, в предположении, что $\delta_{н2} = \delta_k$ (рис. 4, *г*, *д*). Если для некоторого решения (например, для δ_n) выполняется условие $n - k = r$, то принимается, что $\delta_{к2} = \delta_n$. При этом локализуется сигнал, отраженный от более удаленной цели (рис. 4, *е*). Если же не найдется решение δ_i, для которого $i - k = r$, то считается, что δ_k — ошибочное решение. Эту процедуру необходимо повторять до тех пор, пока, начиная с некоторого δ_i, не будет выполняться условие $h - i < r$. В результате рассмотренной процедуры будут локализованы N^* отраженных сигналов и определено их относительное расположение во времени, т. е. будет решена задача полного разрешения целей в выделенном стробе с разрешающей способностью по времени $\delta t = \Delta_t$.



■ ***Рис. 4.*** Принцип логической обработки: *а* — последовательность решений об обнаружени передних и задних фронтов импульсов отраженной пачки; *б*, *в* — последовательность импульсов, начало которых связано с моментом принятия решения δ_g; *г* — *е* — последовательность импульсов, начало которых связано с моментом принятия решения δ_k; *ж*, *з* — последовательность импульсов, начало которых связано с моментом принятия решения δ_l



■ ***Рис. 5.*** Структурная схема алгоритма разрешения неизвестного числа целей по дальности

Структура алгоритма разрешения, основанного на цепном отображении, в котором учтены все рассмотренные выше операции, показана на рис. 5. В его состав входят три блока: блок формирования цепного отображения *БФЦО*, в котором для всех соседних столбцов матрицы $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_h]$ вычисляются величины $d^2_{k,\,k+1}$ (5); блок проверки однородности расстояний *БПОР* цепного отображения, в котором применяется скорректированная процедура сравнения пар величин $(d^2_{k-1,\,k}, d^2_{k,\,k+1})$; блок логической обработки *БЛО*, в котором происходит формирование оценок N^* числа перекрывающихся сигналов и их временных задержек δt^*_{0j} относительно некоторого начала отсчетов.

Заключение

Критерий Кокрена относится к числу непараметрических тестов. Как известно [10], непараметрические тесты не конкретизируют распределения, описывающие конкурирующие гипотезы в статистических задачах. Априорная информация, закладываемая в непараметрические тесты, сводится лишь к заданию различий между конкурирующими гипотезами. Очевидно, что расстояние $\Delta d_{k,\,k+1}$ между векторами, принадлежащими соседним кластерам, отличается от внутрикластерных расстояний при любых распределениях конкурирующих гипотез и является мерой различий, необходимых для функционирования непараметрических тестов. По этой причине предлагаемый алгоритм разрешения должен сохранять работоспособность и в присутствии широкого круга помех $n(t)$, действующих совместно с шумом $w(t)$. Для каждого типа помех $n(t)$ ме-

тодом статистического моделирования можно построить характеристики обнаружения аномальных значений цепного отображения, аналогичные приведенным на рис. 2, которые позволяют судить о возможностях алгоритма разрешения при работе в различных помеховых ситуациях.

Литература

1. **Давыдов В. С., Лукошкин А. П., Шаталов А. А., Ястребков А. Б.** Радиолокация сложных целей (разрешение и распознавание). — СПб.: Янис, 1993. — 280 с.
2. **Акимцев В. В.** Разрешающая способность по дальности при цифровой обработке сигналов // Радиотехника. 2004. № 1. С. 3–11.
3. **Акимцев В. В., Мещерин А. Н.** Цифровой принимаемый сигнал импульсных РЛС обзора и сопровождения и его возможности по разрешению целей по дальности // Информационно-управляющие системы. 2008. № 1. С. 43–49.
4. **Акимцев В. В., Гниденко И. Ю.** Алгоритм разрешения-обнаружения целей по дальности в обзорных РЛС // Радиотехника. 2002. № 1. С. 61–66.
5. **Акимцев В. В.** Непараметрический алгоритм разрешения целей по дальности // Радиотехника. 2009. № 9. С. 53–67.
6. **Патрик Э.** Основы теории распознавания образов: пер. с англ. — М.: Сов. радио, 1980. — 408 с.
7. **Колмогоров А. Н., Фомин С. В.** Элементы теории функций и функционального анализа. — М.: Наука, 1976. — 542 с.
8. **Воеводин В. В., Кузнецов Ю. А.** Матрицы и вычисления. — М.: Наука, 1984. — 320 с.
9. **Гмурман В. Е.** Теория вероятностей и математическая статистика. — М.: Высш. шк., 1999. — 479 с.
10. **Тарасенко Ф. П.** Обзор основных понятий и методов непараметрической статистики // Тр. Сибирского физико-технического института им. В. Д. Кузнецова при Томском государственном университете / ТГУ. Томск, 1973. Вып. 63. С. 49–68.

УДК 681.326.74

ВЕРИФИКАЦИЯ, ВАЛИДАЦИЯ И ТЕСТИРОВАНИЕ КОМПЬЮТЕРНЫХ МОДЕЛЕЙ ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Г. С. Бритов,
канд. техн. наук, доцент
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассмотрены задачи верификации, валидации и тестирования компьютерных моделей линейных динамических систем. Показано, что при верификации модели необходимо перед ее запуском выполнить визуальную проверку, а после запуска — использовать средства верификации применяемых математических пакетов. При валидации модели целесообразно построить устройство функционального диагностирования. Оно позволит проверять процесс моделирования при любых исходных данных. Валидацию можно провести и при тестовых данных, используя специальные режимы моделирования. Приведены примеры компьютерных моделей, построенных на основе математических описаний моделируемых динамических объектов.

Ключевые слова *— верификация, валидация, компьютерные модели, функциональное диагностирование, линейные динамические системы, устройство функционального диагностирования, тестирование моделей.*

Введение

Термины верификация, валидация и тестирование в широком смысле связаны с проверкой качества производимой продукции — оборудования, лекарственных препаратов, технологических процессов. Эти термины, особенно два первых, сейчас хорошо известны в области проверки качества программного обеспечения [1].

В соответствии со стандартом ISO 9000:2000 верификация и валидация изготовленного продукта определяются как подтверждения на основе представления объективных свидетельств того, что установленные требования были выполнены (верификация) и что требования для конкретного применения выполнены (валидация). Международный словарь по метрологии [2] определяет верификацию также как предоставление объективных свидетельств того, что данный объект полностью удовлетворяет установленным требованиям. А валидация — это верификация, при которой установленные требования связаны с предполагаемым использованием объекта.

Таким образом, верификация и валидация предполагают проверку правильности компьютерных моделей (КМ) как перед запуском, так и в процессе моделирования. Это означает, что необходимо диагностирование моделей. Тестовое диагностирование позволит убедиться в том, что модель удовлетворяет установленным требованиям. Например, методы тестового диагностирования линейных динамических систем [3, 4] могут быть использованы и для тестирования КМ. В работах [5, 6] приведен расчет тестового режима линейной системы управления, который тоже можно использовать при проверке ее КМ.

Правильность модели проверяется и с помощью функционального диагностирования. В таком случае можно будет сделать вывод о том, что модель удовлетворяет установленным требованиям, уже непосредственно в процессе моделирования системы.

Простейшее функциональное диагностирование может осуществляться методом контроля по модели [3, 6], в котором диагностические признаки получаются как отклонения выходных сигналов КМ от соответствующих сигналов упрощенной модели. Другой метод связан с построением устройства функционального диагностирования [7, 8]. При компьютерном моделировании под устройством следует понимать дополнительную схему моделирования.

В статье будут даны примеры интерпретации и применения верификации и валидации при компьютерном моделировании линейных динамических систем.

Виды компьютерных моделей динамических систем

Обозначим векторы входных и выходных сигналов моделируемой динамической системы через $\mathbf{u}(t)\in R^m$ и $\mathbf{y}(t)\in R^s$, где t — время функционирования системы. Целью компьютерного моделирования системы является получение на основе математического описания ее основных динамических характеристик. К ним, прежде всего, относятся переходная и весовая характеристики, реакция на гармонический входной сигнал заданной амплитуды и частоты и др. Кроме того, при компьютерном моделировании можно проверить диагностические возможности устройства функционального диагностирования *УФД* системы, на выходе которого формируется диагностический признак $\Delta(t)$ (рис. 1).

Диагностирование динамической системы будет осуществляться проверкой равенства нулю диагностического признака. Компьютерная модель позволит исследовать влияние помех $\mathbf{w}(t)$ на величину диагностического признака $\Delta(t)$.

Рассмотрим три варианта задания математического описания динамической системы.

1. *Матричное описание*. Модель динамической системы задана уравнениями состояния

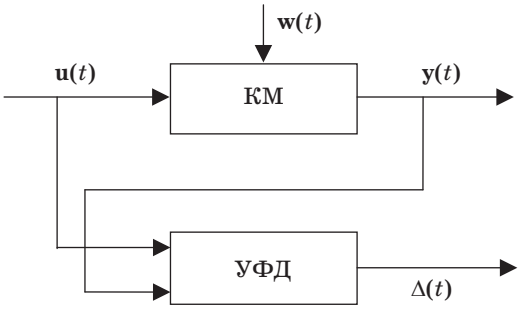
$$\dot{\mathbf{x}}(t)=\mathbf{A}\mathbf{x}(t)+\mathbf{B}\mathbf{u}(t);\ \mathbf{y}(t)=\mathbf{C}\mathbf{x}(t)+\mathbf{D}\mathbf{u}(t), \qquad (1)$$

где $\mathbf{x}(t)\in R^n$ — вектор состояния системы; **A**, **B**, **C**, **D** — постоянные матрицы.

Положим, осуществляется моделирование системы автоматического регулирования 4-го порядка, матричное описание которой имеет вид

$$\dot{\mathbf{x}}(t)=\begin{bmatrix}-3 & 0 & 0 & -0{,}75\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 2 & 0 & 1 & 0\end{bmatrix}\cdot\mathbf{x}(t)+\begin{bmatrix}1\\0\\0\\0\end{bmatrix}\cdot\mathbf{u}(t);$$

$$\mathbf{y}(t)=\begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\cdot\mathbf{x}(t). \qquad (2)$$



■ ***Рис. 1***. Компьютерная модель системы с УФД

Расчет УФД системы осуществляется на основе метода избыточных переменных [7]. В результате КМ системы автоматического регулирования 4-го порядка дополняется УФД с описанием вида

$$\Delta=y_1+y_2+\frac{1}{s}(-y_1+2{,}5y_2)+\frac{1}{s^2}(-2{,}5y_1+1{,}5y_2-2u).$$

Результат компьютерного моделирования системы с УФД представлен на рис. 2, откуда следует, что УФД обеспечивает получение диагностического признака, равного нулю (точнее, близкого к нулю из-за погрешностей системы) при правильной работе КМ.

Таким образом, КМ на основе матричного описания удобна тем, что ее достаточно просто реализовать в любом математическом пакете, например в Simulink.

2. *Операторное описание*. Модель динамической системы задана матричной передаточной функцией (ПФ) $\mathbf{W}(s)$. Используя введенные обозначения, можно записать уравнение системы в виде

$$\mathbf{y}=\mathbf{W}(s)\mathbf{u},$$

где $\mathbf{W}(s)=[W_{ij}(s)]$, $W_{ij}(s)=\dfrac{B_{ij}(s)}{A_{ij}(s)}$ — скалярные дробно-рациональные ПФ.

Матричная ПФ связана с матрицами описания в пространстве состояний уравнением

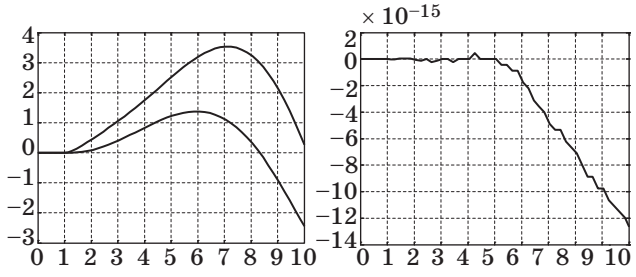
$$\mathbf{W}(s)=\mathbf{C}(s\mathbf{E}-\mathbf{A})^{-1}\mathbf{B},$$

причем переход от $\mathbf{W}(s)$ к описанию в пространстве состояний неоднозначен.

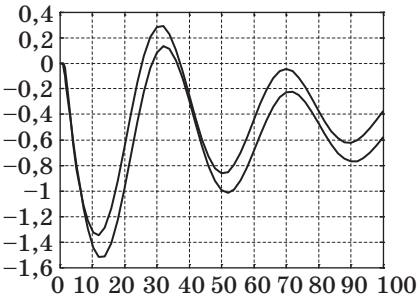
Скалярная ПФ динамики тяжелого транспортного самолета по каналу тангажа имеет следующий вид:

$$W(s)=\frac{-0{,}3s^2-0{,}193s-0{,}016}{s^4+1{,}234s^3+1{,}05s^2+0{,}067s+0{,}025}. \qquad (3)$$

Расчет УФД системы осуществляется на основе метода редукции модели [3]. В результате редуцированная модель динамики тяжелого транс-



■ ***Рис. 2***. Переходная характеристика системы автоматического регулирования 4-го порядка с УФД



■ ***Рис. 3.*** Переходные характеристики динамики тяжелого транспортного самолета по каналу тангажа исходной и редуцированной моделей

портного самолета по каналу тангажа характеризуется ПФ вида

$$W_r(s)=\frac{-0{,}19s-0{,}0144}{s^2+0{,}0379s+0{,}0264}.$$

Переходные характеристики исходной и редуцированной моделей представлены на рис. 3.

Диагностический признак получается как разность реакции исходной и редуцированной моделей. Из графиков на рис. 3 следует, что диагностирование будет эффективным только в начале моделирования. Затем переходные характеристики исходной и редуцированной моделей постепенно расходятся.

Таким образом, КМ на основе операторного описания тоже может быть реализована, например, в пакете Simulink.

3. *Структурное описание.* Математическая модель динамической системы может быть задана блок-схемой, которая состоит из линейных блоков с известными ПФ и сумматоров, связывающих указанные блоки. В этом случае уравнения модели можно записать следующим образом:

$$\mathbf{z}=\mathbf{Q}(s)\mathbf{v},\ \mathbf{v}=\mathbf{F}\mathbf{z}+\mathbf{G}\mathbf{u},\ \mathbf{y}=\mathbf{H}\mathbf{z},$$

где $\mathbf{v}$, $\mathbf{z}$ — входы и выходы линейных блоков; $\mathbf{u}$, $\mathbf{y}$ — входы и выходы системы; $\mathbf{Q}(s)$ — диагональная матрица ПФ блоков; $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$ — матрицы связей.

От структурного описания можно перейти к матричной ПФ с помощью формулы

$$\mathbf{y}=\mathbf{H}(\mathbf{E}-\mathbf{Q}(s)\mathbf{F})^{-1}\cdot\mathbf{Q}(s)\mathbf{G}\mathbf{u}=\mathbf{W}(s)\mathbf{u}. \qquad (4)$$

Положим, матрицы системы имеют вид

$$\mathbf{Q}(s)=\begin{bmatrix}\dfrac{1}{s+1} & 0 & 0 & 0\\ 0 & \dfrac{s}{s^2+s+1} & 0 & 0\\ 0 & 0 & \dfrac{-1}{s^2+s+1} & 0\\ 0 & 0 & 0 & \dfrac{s}{s+1}\end{bmatrix};$$

$$\mathbf{F}=\begin{bmatrix}0&0&1&0\\0&0&1&0\\1&0&1&1\\0&0&1&0\end{bmatrix};\ \mathbf{G}=\begin{bmatrix}1\\1\\0\\1\end{bmatrix};\ \mathbf{H}=\begin{bmatrix}0&1&0&1\end{bmatrix}.$$

Компьютерную модель можно реализовать непосредственным образом в любом математическом пакете. Расчет УФД системы выполним на основе метода диагностирования по модели. ПФ диагностической модели рассчитывается по формуле (4):

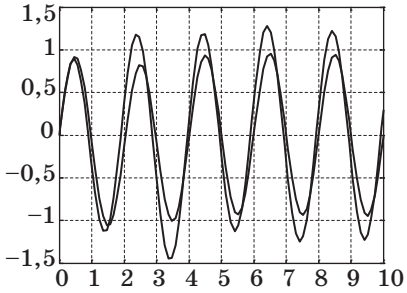
$$\mathbf{W}_{\mathrm{dm}}(s)=\frac{s^5+3s^4+4s^3+8s^2+4s}{s^5+3s^4+7s^3+9s^2+7s+3}.$$

Результаты моделирования системы в виде реакций ее и диагностической модели на гармонический входной сигнал показаны на рис. 4. Видно, что диагностический признак будет возрастать в точках максимума и минимума реакции КМ на гармоническое воздействие.

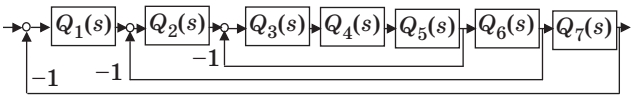
Таким образом, КМ на основе структурного описания тоже диагностируется и реализуется достаточно простым способом.

Все три рассмотренных варианта задания математического описания эквивалентны в том смысле, что, зная один из них, можно перейти к другим. Поэтому для решения задач компьютерного моделирования необходимо ориентироваться на конкретный вид динамической системы. Например, предложенная в работе [9] трехконтурная система подчиненного регулирования электропривода состоит из семи линейных блоков 1-го порядка, охваченных тремя отрицательными обратными связями (рис. 5).

Для этой системы естественно применить структурное описание. Оно будет содержать матрицу $\mathbf{Q}(s)$ 7-го порядка с ПФ 1-го порядка. Ма-



■ ***Рис. 4.*** Результаты компьютерного моделирования системы



■ ***Рис. 5.*** Трехконтурная система подчиненного регулирования электропривода

трицы связей системы будут очень разреженными матрицами. Поэтому сумматоров в модели будет всего три и число входов у них по два. Вид матриц связи следующий:

$$\mathbf{F}=\begin{bmatrix}0&0&0&0&0&0&-1\\1&0&0&0&0&-1&0\\0&1&0&0&-1&0&0\\0&0&1&0&0&0&0\\0&0&0&1&0&0&0\\0&0&0&0&1&0&0\\0&0&0&0&0&1&0\end{bmatrix};$$

$$\mathbf{G}=\begin{bmatrix}1\\0\\0\\0\\0\\0\\0\end{bmatrix};\ \mathbf{H}=\begin{bmatrix}0&0&0&0&0&0&1\end{bmatrix}.$$

Моделирование системы подчиненного регулирования электропривода будет выполнено ниже. Рассмотрим теперь задачи верификации и валидации КМ.

Верификация и валидация компьютерной модели

При решении задач верификации КМ динамической системы необходимо получить представление о том, что установленные математическим описанием требования были выполнены при реализации модели. Верификацию следует осуществлять до и после запуска модели. До запуска модели, реализованной в виде отдельных блоков, необходимо:

— проверить обязательное наличие выходов у используемых блоков;

— проверить наличие входов у используемых блоков;

— обосновать отсутствие входов у некоторых блоков;

— проверить правильность установки параметров блоков;

— проверить связи блоков.

Связи блоков осуществляются в модели с помощью многовходовых сумматоров. Для операторного описания можно предложить следующее правило проверки. Если строка матричной ПФ содержит $k \le m$ ненулевых ПФ, то соответствующий выход модели системы должен иметь сумматор с k входами. Для структурного описания можно предложить следующее правило проверки. Если строка матрицы **F** и соответствующая строка матрицы **G** содержат вместе k единиц, то соответствующий вход блока реализации ПФ $Q_i(s)$ должен иметь сумматор с k входами. Количество выходных сумматоров определяется аналогичным образом для матрицы **H**.

Таким образом, все проверки модели носят визуальный характер. Более информативна верификация после запуска. Положим, для моделирования используется пакет Simulink. Тогда могут быть использованы результаты, получаемые с помощью специальных блоков *Model Verification*. Например, блок *Check Dynamic Gap* проверяет факт попадания в заданные пределы того результата моделирования, который подан на вход блока. Всего *Model Verification* содержит 11 подобных блоков. За счет усложнения схемы включением этих блоков можно будет проверить качество результатов моделирования.

Для верификации модели может быть использована функция *linmod*. С ее помощью получают параметры математического описания модели. Например, для модели (2) получим

```
[A,B,C,D] = linmod('m12')
A =
  -3.0000       0       0  -0.7500
   1.0000       0       0       0
        0  1.0000       0       0
   2.0000       0  1.0000       0
B =
1.0000
        0
        0
        0
C =
        0       0  1.0000       0
        0       0       0  1.0000
D =
     0
     0
```

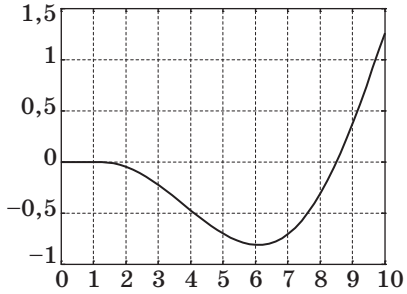
После этого можно проверить устойчивость системы:

```
 >> l=eig(A);
>> re=real(l)
re =
  -2.2781
  -1.1158
   0.1970
   0.1970
```

Следующий шаг верификации — проверка управляемости и наблюдаемости системы:

```
 >> rank(ctrb(A,B))
ans =
     4
>> rank(obsv(A,C))
ans =
     4
```

Видно, что система 4-го порядка (2) неустойчива (правда, степень неустойчивости достаточно мала), управляема и наблюдаема. Все это не должно противоречить полученным ранее результатам исследования системы.



■ ***Рис. 6.*** Диагностический признак КМ (2) при изменении одного из коэффициентов системы на 10 %

Валидация КМ предполагает проверку того, что построена модель, необходимая исследователю. Основной метод валидации базируется на *Model checking*. Он предполагает проверку модели в специальных, тестовых режимах. Однако можно осуществлять проверку и в рабочих режимах, если используются результаты функционального диагностирования.

При валидации КМ может быть использован диагностический признак, получаемый с помощью УФД. Так, если по каким-то причинам коэффициент A_{41} матрицы **A** матричного описания (2) изменился на 10 %, то переходная характеристика меняется незначительно, а результат работы УФД показан на рис. 6.

Диагностический признак КМ говорит о том, что требования для конкретного применения ее, строго говоря, не выполнены, так как значение диагностического признака не равно нулю. Аналогичные результаты получаются и для других приведенных выше КМ.

Тестирование компьютерной модели

В работе [3] приведен большой список специальных, тестовых режимов для организации проверки систем автоматического управления. Все они могут быть использованы и для проверки КМ динамических систем [10]. Рассмотрим некоторые из них.

1. *Нулевой режим* связан с понятием передаточного нуля динамической системы. Подробно расчет нулевого режима, основанного на использовании передаточного нуля, приведен в работе [3]. Для того чтобы обеспечить нулевой режим модели, необходимо найти вещественный корень λ числителя ПФ операторного описания модели (передаточный нуль). Входной сигнал будет иметь вид $\mathbf{u}(t) = e^{\lambda t}$. Затем по формуле

$$\mathbf{x_0} = (\lambda\mathbf{E} - \mathbf{A})^{-1} \cdot \mathbf{B} \qquad (5)$$

рассчитываются специальные начальные условия модели. Здесь **A**, **B** — матрицы соответствующего матричного описания скалярной системы.

Ниже показан этот расчет для модели динамики тяжелого транспортного самолета по каналу тангажа с операторным описанием (3). Передаточный нуль здесь является корнем числителя (получен с помощью функции *roots*), так как при $s = -0{,}5456$ ПФ $\mathbf{W}(-0{,}5456) = 0$. Значит, в случае, когда входной сигнал модели $\mathbf{u}(t) = e^{-0{,}5456t}$, вынужденное движение модели будет равно 0. Для того чтобы сделать нулевым собственное движение модели, вызванное скачком входного воздействия в нулевой момент времени, рассчитываем специальные начальные условия по формуле (5). Соответствующие системные матрицы можно получить с помощью функции *canon*. Расчет и моделирование выполнено в пакете MatLab. Результаты тестирования модели в нулевом режиме представлены на рис. 7, где хорошо видно, что экспоненциальный входной сигнал приводит к нулевой реакции правильной модели и ненулевой реакции неправильной модели.

2. *Модальный режим* предполагает исследование собственного движения модели, т. е. входной сигнал $\mathbf{u}(t) = 0$. Рассмотрим матричное описание модели в собственном движении:

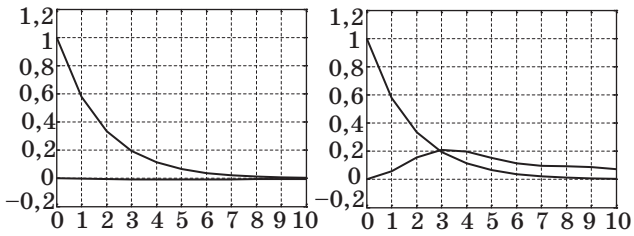
$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t);\ \mathbf{y}(t) = \mathbf{C},\ x(0) = x_0.$$

Зададим начальные условия как собственный вектор **v** матрицы **A** для вещественного собственного числа λ. Тогда выходной сигнал модели будет следующим:

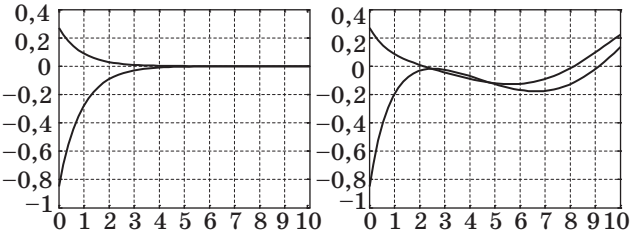
$$\mathbf{y}(t) = e^{\lambda t}\mathbf{C}\mathbf{v},$$

где $e^{\lambda t}$ — мода системы. Следует обратить внимание, что собственное движение модели в модальном режиме пропорционально этой моде. Это обстоятельство может быть использовано для проверки модели.

Выполним расчеты в пакете MatLab для матриц **A** и **C** системы автоматического регулирования 4-го порядка (2). Получим собственное число $\lambda = -1{,}1158$ и собственный вектор $\mathbf{v} = [0{,}3379; -0{,}3028; 0{,}2714; -0{,}8488]$. Если использовать его как начальные условия модели, то при нулевом входном сигнале получим выходные сигналы $y_1(t) = 0{,}2714e^{-1{,}1158t}$, $y_2(t) = -0{,}8488e^{-1158t}$. На



■ ***Рис. 7.*** Результаты нулевого режима модели динамики тяжелого транспортного самолета по каналу тангажа



■ ***Рис. 8.*** Результаты модального режима модели системы автоматического регулирования 4-го порядка

рис. 8 показаны результаты проверки модели в модальном режиме.

Левая осциллограмма получена для расчетной матрицы **A**, а правая — для неправильной матрицы **A**. В первом случае модель правильная, во втором — нет.

3. *Режим комплементарного сигнала*. Предлагается подать на входы КМ системы последовательность импульсов, ширина и амплитуды которых рассчитываются так, чтобы модель из нулевых начальных условий за расчетное время вернулась опять в нулевое состояние. Такой сигнал называется комплементарным [11]. Соответственно получается и тестовый режим, который приводит к естественному нулевому диагностическому признаку. Расчет комплементарного сигнала осуществляется на основе матрицы системы **A** [11]. Пусть h — ширина импульса в секундах. Тогда составляющие вектора амплитуд комплементарного сигнала рассчитываются по формулам

$$\alpha_1 = -\sum_{i=1}^{n}\mu_i,\ \alpha_2 = \sum_{i,j=1}^{n}\mu_i\mu_j, \ldots, \alpha_n = (-1)^n\mu_1\ldots\mu_n,$$

где $\mu_i, i = 1, ..., n$ — собственные числа матрицы $e^{\mathbf{A}h}$. Расчетное время тестирования модели $t_k \geq nh$.

Расчет и моделирование выполним ниже для системы подчиненного регулирования электропривода.

Компьютерная модель системы подчиненного регулирования электропривода

Рассмотрим компьютерное моделирование системы подчиненного регулирования электропривода, структурное описание которой было предложено выше. Согласно работе [9], примем следующие ПФ:

— силовой части

$$Q_i(s) = \frac{1}{s+1},\ i = 5, 6, 7;$$

— фильтра

$$Q_4(s) = \frac{1}{s+1};$$

— регулирующей части

$$Q_3(s) = \frac{s+1}{2s+1},\ Q_2(s) = \frac{s+1}{4s+1},\ Q_1(s) = \frac{s+1}{8s+1}.$$

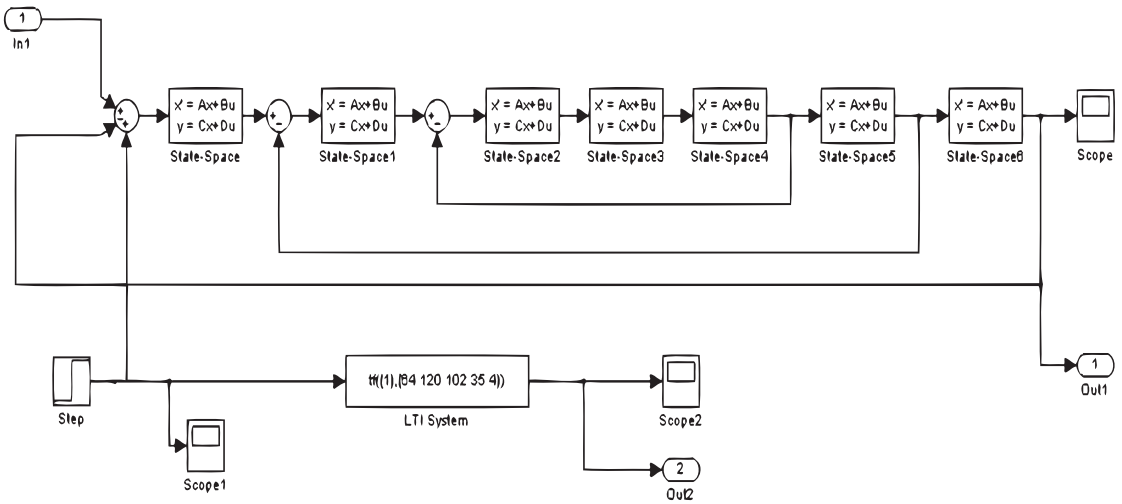
Предположим, что КМ уже построена в пакете Simulink из блоков *State Space* (рис. 9).

В качестве УФД используем блок *LTI System*, в котором реализована ПФ из расчета по формуле (4):

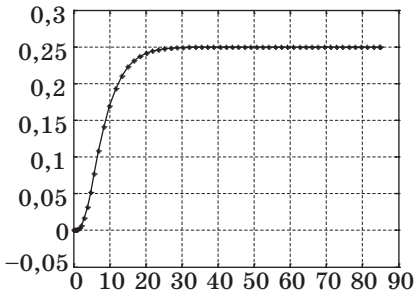
$$W(s) = \frac{1}{64s^4 + 120s^3 + 102s^2 + 35s + 4}.$$

Необходимо обратить внимание на то, что получилась ПФ системы 4-го порядка. В то же время модель, показанная на рис. 9, по внешнему виду должна иметь 7-й порядок.

Результаты моделирования системы с УФД представлены на рис. 10 (сплошная линия соответствует модели, а точечная — УФД).



■ ***Рис. 9.*** Схема КМ системы подчиненного регулирования электропривода



■ ***Рис. 10.*** Переходные характеристики модели и УФД

Следовательно, модель пока правильная, и можно доверять результатам ее исследования. В частности, если необходимо не только построить модель, но и выполнить верификацию и валидацию, то следует провести подготовительную работу.

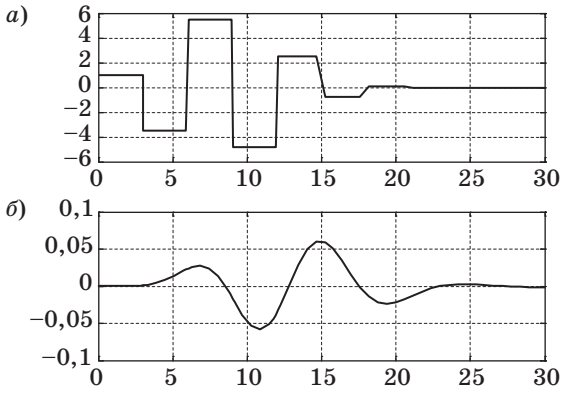
С помощью функции *linmod* определяются (7 × 7)-матрица **A** коэффициентов системы и (7 × × 1)-матрица **B** ее входов:

```
>> [A,B,C,D]=linmod('m11')
A =
  -1.0000       0   1.0000       0       0       0       0
  -0.1094  -0.1250       0       0       0       0       0
        0       0  -1.0000       0   1.0000       0       0
  -0.0234   0.1875  -0.1875  -0.2500       0       0       0
        0       0       0       0  -1.0000       0   1.0000
  -0.0078   0.0625  -0.0625   0.2500  -0.2500  -0.5000       0
  -0.0156   0.1250  -0.1250   0.5000  -0.5000   1.0000  -1.0000
B =
        0
   0.1094
        0
   0.0234
        0
   0.0078
   0.0156
C =    1.0000       0       0       0       0       0       0
D =    0
```

Теперь можно рассчитать необходимые данные для нулевого и модального режимов подобно тому, как это было сделано для предыдущих примеров. Результаты тестирования КМ аналогичны результатам, полученным на рис. 7, 8. Следует отметить, что при расчете начальных условий нулевого режима возникают трудности с обращением матрицы. Поэтому результаты нулевого режима оказываются неточными, и показывать их не имеет смысла.

Расчет комплементарного сигнала приводится полностью:

```
>> chm=eig(A)
chm =
  -1.0000
  -0.6485 + 0.5844i
  -0.6485 - 0.5844i
  -0.2500
  -0.3281
   -1.0000
   -1.0000
>> h=1;
>> mu=exp(chm*h);
>> alf=poly(mu)
alf = 1.0000  -3.4749   5.1649
 -4.2754   2.1333  -0.6409   0.107   -0.0076
```



■ ***Рис. 11.*** Комплементарный сигнал (*а*) и реакция на него модели (*б*)

Результаты тестирования системы с помощью комплементарного сигнала представлены на рис. 11, *а*, *б*. Понятно, что модель, показанная на рис. 9, действительно правильная.

При компьютерном моделировании системы подчиненного регулирования электропривода с указанными выше параметрами теперь можно использовать для верификации построенное УФД, а для валидации — рассчитанный комплементарный сигнал.

Таким образом, рассмотренные подходы к верификации и валидации линейных динамических систем применены для различных примеров их КМ.

Заключение

Изложен подход к организации процедур верификации, тестирования и валидации КМ динамических систем, описываемых матрицами уравнений состояния, матричными ПФ или структурной схемой. В частном случае моделироваться могут линейные системы автоматического управления. Приведены примеры использования указанных процедур для простых моделей, наглядно демонстрирующих и стандартные приемы верификации, и специальные методы, основанные на применении устройств функционального диагностирования. Все модели реализованы в пакете MatLab с использованием пакета Simulink. Показано, что для линейных динамических систем всегда может быть синтезировано УФД сравнительно небольшой размерности, вырабатывающее диагностический признак для применения при валидации модели системы. Результаты верификации, валидации и тестирования разработанных моделей показали их работоспособность и целесообразность применения при анализе характеристик динамических систем.

Работа выполнена по гранту № 11-08-00240.

Литература

1. **Baier C., Katoen J. P.** Principles of Model Checking. — Boston: The MIT Press, 2008. — 975 p.
2. **Международный** словарь по метрологии. Основные и общие понятия и соответствующие термины. ISO/IEC Guid 99: 2007.
3. **Мироновский Л. А.** Функциональное диагностирование динамических систем. — М.: Изд-во МГУ, 1998. — 340 с.
4. **Мироновский Л. А., Соловьева Т. Н.** Диагностирование систем с фазовращательными и бисингулярными передаточными функциями // Информационно-управляющие системы. 2012. № 6. С. 60–66.
5. **Бритов Г. С., Мироновский Л. А.** Расчет тестового режима линейных систем управления // Приборы и системы. Управление, контроль, диагностика. 2006. № 11. С. 44–49.
6. **Лоскутов А. И., Вечеркин В. Б., Шестопалова О. Л.** Автоматизация контроля состояния сложных технических систем на основе использования конечно-автоматной модели и нейросетевых структур // Информационно-управляющие системы. 2012. № 2. С. 74–81.
7. **Бритов Г. С., Мироновский Л. А.** Автоматизированное проектирование устройств функционального диагностирования // Информационно-управляющие системы. 2010. № 2. С. 55–61.
8. **Безмен Г. В., Колесов Н. В.** Функциональное диагностирование линейных динамических систем с использованием нечеткого анализа // Информационно-управляющие системы. 2009. № 5. С. 67–73.
9. **Шрейнер Р. Т.** Системы подчиненного регулирования электроприводов / УрГППУ. — Екатеринбург, 2008. — 279 с.
10. **Кириллов А. Н.** Моделирование динамики структур гибридных систем // Информационно-управляющие системы. 2011. № 4. С. 42–46.
11. **Мироновский Л. А.** Диагностирование линейных систем методом комплементарного сигнала // Приборы и системы. Управление, контроль, диагностика. 2002. № 5. С. 52–57.

УДК 156.6

РЕШЕНИЕ ЗАДАЧИ РАСПРЕДЕЛЕНИЯ РЕСУРСОВ ПРИ ВЫПОЛНЕНИИ АДМИНИСТРАТИВНЫХ РЕГЛАМЕНТОВ

В. В. Науменко,
аспирант
В. В. Копытов,
доктор техн. наук, профессор
Северо-Кавказский федеральный университет, г. Ставрополь

Рассматриваются проблемы повышения эффективности системы государственного управления, одной из которых является распределение ресурсов в процессе исполнения государственных функций и предоставления государственных услуг. Для решения данной проблемы предлагается применение оптимизационной модели, направленной на эффективное распределение выполняемых работ между исполнителями. При этом используются методы теории массового обслуживания и алгоритмы поиска оптимального значения целевой функции.

Ключевые слова *— административный регламент, функциональная безопасность, распределение ресурсов.*

Введение

Современная нормативно-правовая база [1, 2] требует от органов власти исполнения государственных функций и предоставления государственных услуг строго в соответствии с требованиями административных регламентов (АР), делая тем самым АР ключевым элементом системы государственного управления, от эффективности выполнения которого зависит эффективность всей системы государственного управления. В настоящее время информационная система государственного управления является критической социотехнической системой, неэффективная работа которой связана с невозможностью выполнять функции, закрепленные нормативно-правовыми актами и законами, что ухудшает социально-экономическое развитие. Поэтому на функциональную безопасность информационной подсистемы системы государственного и муниципального управления напрямую влияет выполнимость АР, которая в свою очередь выражается в невыполнении требований АР.

Анализ структуры административных регламентов

Рассмотрим структуру АР. Выделяются следующие составляющие АР [3]:

— область применения, которая ограничивается набором соответствующих нормативных регуляторов;

— функциональные цели;

— субъекты выполнения регламента, в интересах которых осуществляется регламент и (или) которые охватываются функциональной целью;

— объекты действия регламента;

— процессы и операции регламента.

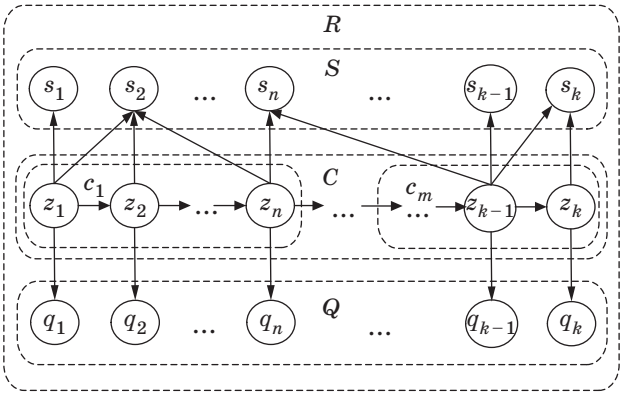
Главное назначение АР заключается в интеграции совокупности процессов и операций, реализуемых субъектами действия над существующими объектами в интересах, определенных нормативными регуляторами и инструкциями, для достижения заданной цели [3]. Поэтому если рассматривать АР как алгоритмический процесс, реализующий государственную услугу, то его основу будут составлять процессы и операции, закрепленные в нормативно-правовом акте АР (рис. 1). Типовая структура АР предполагает наличие описания административных процедур c, каждая из которых достигается путем выполнения определенной последовательности задач z:

$$P(c_m) \to P\left((z_n)_{n=1}^{k}\right),\ k > 0,$$

где P — предикат наличия элемента системы ($P(e)$ принимает значение «истина», если e существует в системе).

При этом сам АР R будет характеризоваться последовательностью процедур:

$$P(R) \to P\left((c_n)_{n=1}^{k}\right),\ k > 0.$$



■ ***Рис. 1.*** Типовая структура административного регламента

Каждая задача предполагает совершение определенных операций субъектов над объектами, т. е.

$$P(z_n) \to P(s_m),\ P(z_n) \to P(q_m).$$

Субъект-пользователь инициирует выполнение первой задачи последовательности, в результате чего запускается процесс выполнения АР. При этом субъекты и объекты выступают в качестве ресурсов, необходимых для перехода от одной задачи к другой [4].

Основные проблемы, препятствующие выполнению требований АР, можно разделить на две группы:

— связанные с логической противоречивостью описания процессов АР;

— возникающие в процессе выполнения АР.

Решение проблем логической противоречивости связано с правильным выбором методов и средств моделирования информационных процессов, подходящих для описания АР. Построение адекватной информационной модели АР позволит выявить противоречия в описании до его реализации в виде совокупности информационных процессов, исключая появление в системе ряда угроз функциональной безопасности. Однако проблемы, возникающие в процессе выполнения, являются следствием либо отсутствия, либо неправильного распределения ресурсов системы — субъектов и объектов и требуют решения оптимизационных задач. Решение одной из таких задач предлагается ниже.

Постановка задачи распределения ресурсов

Как уже было отмечено, выполнение некоторой задачи z_n связано с наличием в системе соответствующего субъекта s'_m и объекта q'_k, которые можно определить как необходимые требования для z_n. Однако стоит отметить, что система государственного управления является динамичной системой, и поэтому соблюдение условий $P(s'_m) \to \to P(s_m)$ и $P(q'_k) \to P(q_k)$ зависит от того, заняты ли s_m и q_k в момент времени t_i, где t_i — время начала выполнения задачи z_n:

$$\tau(t_i) \to \left(P(s'_m) \to P(s_m)\right) \wedge \left(P(q'_k) \to P(q_k)\right),$$

где $\tau(t_i)$ — предикат, принимающий значение «истина» в момент времени t_i.

При этом, как показывает практика, каждый объект АР уникален и соответствует только одной задаче. Поэтому решение задачи обеспечения функциональной безопасности выполнения АР состоит в согласовании множества объектов — необходимых Q' и существующих в системе Q: $Q' \subset Q$.

Иная ситуация обстоит с субъектами. Каждый субъект в системе государственного управления функционально может выполнять несколько задач АР. *Поэтому для обеспечения функциональной безопасности выполнения АР необходимо решение оптимизационной задачи, направленной на эффективное распределение задач между субъектами-исполнителями.*

Рассмотрим процедуру решения задачи оптимального распределения субъектов-исполнителей, входящих в группу S_o (с общим количеством субъектов-исполнителей a), где каждый из субъектов-исполнителей $s_n \in S_o$ функционально может быть задействован в выполнении задач из группы Z_l (с общим количеством задач b). Учитывая, что процесс выполнения задач непрерывный, а процесс поступления задач на выполнение — случайный, то АР можно рассматривать как сеть узлов массового обслуживания, где каждый узел — это субъект-исполнитель s_n, который представляет собой систему массового обслуживания типа $M/M/1$, выполняющую поток заявок λ_k, соответствующий задаче $z_k \in Z_l$. Каждая выполненная задача z_k в свою очередь инициирует выполнение следующей задачи z_{k+1}, поэтому, согласно теореме Бёрке, исходящий поток с входящим потоком с параметром λ и показательным распределением интенсивности обслуживания μ является пуассоновским потоком с тем же параметром λ [5]. Для определения интенсивности обслуживания задачи z_k необходимо в качестве исходных данных использовать среднее время выполнения для каждой из выполняемых задач отдельно взятой организации $t(z_k)$:

$$\mu_k = \frac{1}{t(z_k)}. \qquad (1)$$

В случае если субъект s_n выполняет несколько задач, то интенсивность обслуживания отдельной задачи определяется с учетом всех входящих потоков, т. е.

$$\mu_k = \frac{p_k}{t(z_k)}, \quad (2)$$

где p_k — вероятность появления заявки из потока λ_k, определяемая отношением потока λ_k к сумме всех потоков, выполняемых субъектом:

$$p_k = \frac{\lambda_k}{\sum_{i=k}^{k+p} \lambda_i}. \quad (3)$$

Среднее время обслуживания для $M/M/1$ будет определяться следующим образом [5]:

$$T_k = \frac{1/\mu_k}{1-\lambda_k/\mu_k}. \quad (4)$$

В случае если одну задачу (один поток) выполняет несколько субъектов, то система массового обслуживания будет $M/M/m$, где m — число субъектов. Среднее время обслуживания

$$T_k = \frac{(\rho_k m)^{m+1}}{\lambda_k (m-1)! \sum_{n=0}^{m} \left[\frac{(\rho_k m)^n}{n!}\right]\left[(m-L)^2 - L\right]}, \quad (5)$$

где L — число заявок, ожидающих в очереди; среднее значение определяется по формуле

$$L_q = \frac{\rho_k (m\rho_k)^m}{m!(1-\rho_k)^2} \frac{1}{\left(\sum_{n=0}^{m-1} \frac{(m\rho_k)^n}{n!} + \frac{(m\rho_k)^m}{m!(1-\rho_k)}\right)}, \quad (6)$$

где $\rho_k = \lambda_k/(m\mu_k)$.

Просчитав для каждой задачи среднее время выполнения T_k и сравнив его с максимальным временем $T_{k\,\max}$, определенным нормативными регуляторами АР, можно сказать, что если $T_k > T_{k\,\max}$, то задача не выполняется с заданными условиями, т. е. не соблюдаются требования функциональной безопасности. Результат представим в виде коэффициента

$$\alpha_k = \frac{T_k}{T_{k\max}}. \quad (7)$$

Данный показатель характеризует загруженность субъекта-исполнителя при выполнении задачи z_k и позволяет оценить запас времени, который доступен для решения других задач из группы Z_l.

Решение любых задач распределения ресурсов требует выбора определенного критерия в качестве целевой функции, на получение оптимальных значений которого и будет направлена задача.

Необходимо отметить, что государственное управление не имеет формализуемой целевой функции. С этим и связано отсутствие интегрального показателя качества государственного управления, на повышение которого могла быть направлена задача распределения ресурсов. Поэтому в качестве целевой функции необходимо использовать показатель, который характеризует, насколько сбалансировано распределение исполнителей между задачами АР. В качестве такого показателя целесообразно использовать дисперсию величины α_k каждой из задач АР:

$$D(\alpha) = \frac{1}{b}\sum_{k=1}^{b} (\alpha_k - M[\alpha_k])^2. \quad (8)$$

Таким образом, данная задача распределения ресурсов (субъектов между задачами АР) схожа с классической задачей распределения ресурсов (между поставщиками и потребителями) и состоит в определении величин x_{nk}, определяющих поставку n-субъекта k-задаче и принимающих значение 1 либо 0. При этом для каждой из задач АР выполняется условие $\alpha_k \le 1$, а величина $D(\alpha)$ минимальна.

Анализ методов решения задачи распределения ресурсов

Стоит отметить, что данная задача имеет сложную целевую функцию, которая зависит не только от входных данных (λ_k, $t(z_k)$), но и от сложности взаимосвязей «субъект-задача» внутри системы. Поэтому для вычисления оптимального значения целевой функции необходимо отказаться от градиентных методов (методов, использующих значения градиентов или высших производных непрерывных функций, предназначенных для оценки шага и скорости приближения к точке оптимума в итерационных процессах) и применить один из безградиентных методов либо алгоритм открытого поиска. Подобные алгоритмы основаны на оценках критерия качества по множеству точек, расположенных вокруг текущей точки, и выборе одной из них, где критерий принимает наименьшее из всех оцениваемых значений [6].

Теперь рассмотрим процедуру распределения задач между субъектами-исполнителями. Приведем технологическую матрицу, характеризующую распределение ресурсов в системе (табл. 1). В левом столбце приведены все задачи группы Z_l, в верхней строке — все субъекты группы S_o.

■ ***Таблица 1.*** Матрица распределения ресурсов в системе

Z_l	S_o					
	s_1	s_2	...	s_n	...	s_a
$z_1\,\lambda_1,\ t(z_1),\ T_{1\max}$	x_{11}	x_{21}	...	x_{n1}	...	x_{a1}
$z_2\,\lambda_2,\ t(z_2),\ T_{2\max}$	x_{12}	x_{22}	...	x_{n2}	...	x_{a2}
...	...	...	...	...	...	...
$z_k\,\lambda_k,\ t(z_k),\ T_{k\max}$	x_{1k}	x_{2k}	...	x_{nk}	...	x_{ak}
...	...	...	...	...	...	
$z_b\,\lambda_b,\ t(z_b),\ T_{b\max}$	x_{1b}	x_{2b}	...	x_{nb}	...	x_{ab}

Для каждой задачи указывается поток заявок λ_k и среднее время выполнения $t(z_k)$. При этом сумма каждой строки должна быть больше 0, так как для любой из задач должен быть определен исполнитель:

$$\begin{cases} x_{11}+x_{21}+\ldots+x_{n1}+\ldots+x_{a1}>0; \\ x_{12}+x_{22}+\ldots+x_{n2}+\ldots+x_{a2}>0; \\ \ldots \\ x_{1k}+x_{2k}+\ldots+x_{nk}+\ldots+x_{ak}>0; \\ \ldots \\ x_{1b}+x_{2b}+\ldots+x_{nb}+\ldots+x_{ab}>0. \end{cases} \quad (9)$$

Решение алгоритма открытого поиска заключается в поиске значения вектора $\mathbf{x} = [x_{11}, \ldots, x_{ab}]$, которое обеспечивало бы минимальное значение целевой функции $D(\alpha)$ за 2^{ab} шагов (размещение с повторениями). На каждом шаге определяется соответствие системе неравенств (9). Если вектор $\mathbf{x}$ не удовлетворяет условиям неравенства, то идет переход к следующему шагу, в противном случае вычисляется вектор $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_b]$ по выражениям (1)–(3), далее вычисляется вектор $\mathbf{T} = [T_1, \ldots, T_b]$ [выражения (4)–(6)], затем вектор $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_b]$ и проводится проверка на соответствие условию $\alpha_k \le 1$ [(выражение (7)]. Если любое из $\alpha_k > 1$, то идет переход к следующему шагу, иначе вычисляется значение $D(\alpha)$ [(выражение (8)]. После этого выбирается значение вектора $\mathbf{x}$, соответствующее минимальному значению $D(\alpha)$:

$$\mathbf{x} = \arg\min\left(D_1(\alpha), \ldots, D_{2^{ab}}(\alpha)\right). \quad (10)$$

Алгоритм открытого поиска позволяет получить все оптимальные значения вектора $\mathbf{x}$, однако из-за большого количества шагов решения данный метод является слишком трудоемким.

В работе [7] проведен подробный анализ безградиентных методов оптимизации для решения задачи подбора технических средств охраны. Рассматриваемая в работе задача, так же как и решаемая авторами данной статьи, имеет множество дискретных решений и сложную целевую функцию, в связи с чем был выбран генетический алгоритм (ГА).

Применение генетического алгоритма для решения задачи распределения ресурсов при выполнении АР

Рассмотрим применяемые в настоящее время различные модели ГА. Классическим (каноническим) ГА принято считать алгоритм Джона Холланда [8], который имеет следующие характеристики:

— целочисленное кодирование особей (хромосом);

— одинаковая длина всех хромосом в популяции;

— постоянный размер популяции;

— пропорциональный отбор;

— одноточечный кроссовер;

— битовая мутация;

— формирование следующего поколения из потомков текущего поколения.

Помимо классического ГА существуют и другие модели: гибридный, Genitor, СНС ГА и др. Они различаются принципами отбора и формирования нового поколения особей, операторами мутации, кодированием генов и т. д. В Genitor-модели (созданной Д. Уитли [9]) используется специфичная стратегия отбора. Вначале производится оценка особей начальной популяции. Далее выбираются случайным образом две особи, которые скрещиваются, причем получается только один потомок, который оценивается и занимает место наименее приспособленной особи. После этого снова случайным образом выбираются две особи, и их потомок занимает место особи с самой низкой приспособленностью. Таким образом, на каждом шаге в популяции обновляется только одна особь. В работе [10] утверждается, что при помощи Genitor-модели поиск гиперплоскостей происходит лучше, а сходимость быстрее, чем у классического ГА.

В СНС-модели (*Cross generational elitist selection*, *Heterogeneous recombination*, *Cataclysmic mutation*), характеризующейся отсутствием мутаций, используется популяция небольшого размера, отбор особей в следующее поколение ведется и между родительскими особями, и между их потомками. После нахождения некоторого решения алгоритм перезапускается, и лучшая особь копируется в новую популяцию, а оставшиеся особи являются сильной мутацией (мутирует примерно треть битов в хромосоме) существующих, и поиск повторяется. Другой специфичной чертой является стратегия скрещивания: все особи разбиваются на пары, причем скрещиваются только те пары, в которых хромосомы особей существенно различны. При скрещивании используется так называемый HUX-оператор (*Half Uniform Crossover*) — разновидность однородного кроссовера, но в нем к каждому потомку попадает ровно половина битов хромосомы от каждого родителя.

Использование гибридного алгоритма (*Hybrid Algorithms*) позволяет объединить преимущества ГА и классических методов. На каждом поколении каждый полученный потомок оптимизируется этим выбранным классическим методом, после чего продолжаются обычные для ГА действия. Стоит отметить, что такой метод ухудшает способность алгоритма к поиску решения с помощью отбора гиперплоскостей, однако на практике гибридные алгоритмы показывают успешные результаты [11]. Это связано с тем, что обычно вели-

ка вероятность того, что одна из особей попадет в область глобального максимума и после проведенной оптимизации окажется решением задачи.

Параллельные ГА предназначены для снижения преждевременной сходимости к локальному оптимуму, стимуляции разнообразия и поиска альтернативных решений той же проблемы. Они основаны на разбиении популяции на несколько отдельных подпопуляций, каждая из которых будет обрабатываться ГА независимо от других. Кроме того, разнообразные миграции индивидов порождают обмен генетическим материалом среди популяций, которые обычно улучшают точность и эффективность алгоритма. Наиболее распространенными из параллельных ГА являются островная модель (*Island model*) и ячеистый алгоритм (*Cellular Genetic Algorithm*).

Несмотря на очевидные преимущества параллельных ГА, для решения поставленной задачи будем использовать классический алгоритм, так как в рамках данной работы необходимо проверить адекватность разработанной математической модели распределения ресурсов при выполнении АР, которая в свою очередь заключается в наличии сходимости целевой функции. Получение приемлемых результатов с помощью алгоритма Холланда позволяет применять и другие модели ГА для решения задачи распределения ресурсов АР.

Этапы и пример решения задачи распределения ресурсов при выполнении АР

Теперь рассмотрим каждый из этапов решения задачи. Вначале случайным образом выбираются N ($N \geq b$) значений вектора $\mathbf{x}$ (начальная популяция из N особей), каждое из которых соответствует системе неравенств (9) и условию $\alpha_k > 1$. Далее генерируется промежуточная популяция — набор особей, получивших право размножаться. Для генерации промежуточной популяции используется *принцип пропорционального отбора*, заключающийся в том, что каждая особь попадает в популяцию с вероятностью, пропорциональной ее приспособленности. Для данной задачи чем меньше значение $D(\alpha)$, тем больше вероятность попадания в промежуточную популяцию:

$$P_p = 1 - \frac{D_p(\alpha)}{D_1(\alpha) + \ldots + D_N(\alpha)}. \qquad (11)$$

Далее особи в случайном порядке разбиваются на пары, и производится *скрещивание* (обмен случайными отсеченными частями):

$$\mathbf{x}_p[01011.01\ldots101] \leftrightarrow \mathbf{x}_{p+1}[11001.11\ldots001] = \mathbf{x}'_p[11001.01\ldots101].$$

К полученному в результате отбора и скрещивания новому поколению применяется оператор мутации, который инвертирует каждый бит популяции с вероятностью $1/N$. Далее из полученных особей выбираются только те, что соответствуют системе неравенств (9) и условию $\alpha_k > 1$, после чего цикл повторяется снова. *Процесс эволюции* (цикл отбор—скрещивание—мутация) может продолжаться бесконечное число шагов, поэтому критерием останова является получение сходимости целевой функции за n число шагов. При этом оптимальному значению вектора $\mathbf{x}$ будет соответствовать наиболее приспособленная особь из последнего поколения:

$$\mathbf{x} = \arg\min\left(\left(D_1(\alpha), \ldots, D_N(\alpha)\right)_n\right). \qquad (12)$$

В случае если ни на одном из шагов не обеспечивается выполнение любого из условий, то распределение ресурсов в системе невозможно ввиду их недостаточности, и решение данной проблемы возможно только путем перераспределения функциональных обязанностей, расширяя при этом группу S_o.

Приведем пример работы ГА с использованием скрипта «ga.m», входящего в пакет среды MatLab, для распределения семи задач между пятью субъектами-исполнителями. В качестве исходных данных будем использовать значения λ_k, $t(z_k)$, $T_{k\max}$, удовлетворяющие условию $l/\mu < 1$:

$$\lambda_k = [1;\ 2;\ 1;\ 0.5;\ 0.4;\ 2;\ 1];$$
$$t(z_k) = [0.2;\ 0.4;\ 0.1;\ 0.5;\ 0.05;\ 0.3;\ 0.8];$$
$$T_{k\max} = [0.8;\ 3;\ 1.2;\ 5.6;\ 1.9;\ 3.3;\ 3.5].$$

Для подсчета дисперсии величины α_k для среды MatLab написан специальный скрипт «ga_example.m», который также проверяет полученные промежуточные значения T_k на соответствие условию $\alpha_k > 1$, и в случае отсутствия такого соответствия присваивает целевой функции значение 100, в противном случае выводит значение дисперсии $D(\alpha)$.

В качестве начальной популяции использовалось значение вектора, при котором целевая функция отлична от 100:

$$\mathbf{x} = [10;\ 30;\ 30;\ 18;\ 30;\ 12;\ 7];$$
$$D(\alpha) = 0.082326984378338.$$

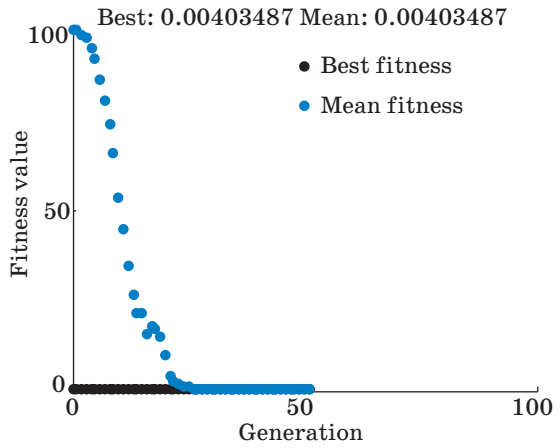
В результате расчета получены следующие численные значения:

Optimization running.
Objective function value: 0.004034869320015893
$\mathbf{x}$ = [10 19 2 25 2 18 7].

Полученное значение вектора $\mathbf{x}$ соответствует матрице (табл. 2).

■ ***Таблица 2.*** Результат распределения ресурсов в системе

Z_l	S_o				
	s_1	s_2	s_3	s_4	s_5
z_1 1, 0.2, 0.8	0	1	0	1	0
z_2 2, 0.4, 3	1	0	0	1	1
z_3 1, 0.1, 1.2	0	0	0	1	0
z_4 0.5, 0.5, 5.6	1	1	0	0	1
z_5 0.4, 0.05, 1.9	0	0	0	1	0
z_6 2, 0.3, 3.3	1	0	0	1	0
z_7 1, 0.8, 3.5	0	0	1	1	1



■ ***Рис. 2.*** Процесс оптимизации целевой функции при помощи ГА

Для наблюдения схождения целевой функции в среде MatLab использовался инструмент psearchtool (рис. 2).

Из графика видно, что процесс сходится к установившемуся значению

$$D(\alpha) = 0.004034869320015893$$

приблизительно за 30 итераций.

Заключение

Внедрение данного метода в виде программного решения в автоматизированные системы исполнения государственных функций и предоставления государственных услуг позволит вывести систему государственного управления на качественно новый уровень путем непрерывного решения задач выполнимости АР.

Литература

1. **Федеральный закон** «Об организации предоставления государственных и муниципальных услуг» № 210-ФЗ от 27 июля 2010 года. http://www.rg.ru/2010/07/30/gosusl-dok.html (дата обращения: 01.03.2013).
2. **Постановление** Правительства РФ от 16.05.2011 № 373 «О разработке и утверждении административных регламентов исполнения государственных функций и административных регламентов предоставления государственных услуг». http://www.rg.ru/2011/05/31/gosuslugi-site-dok.html (дата обращения: 01.03.2013).
3. **Региональное** электронное правительство: стратегия создания, архитектура, типовые решения / под ред. В. И. Дрожжинова, А. А. Лучина. — М.: Эко-Трендз, 2004. — 288 с.
4. **Копытов В. В., Науменко В. В., Минин В. А., Зайцев А. А.** Анализ проблем обеспечения бесконфликтного выполнения электронных административных регламентов // Сб. науч. ст. / Ставропольский филиал ИГУТИ. Ставрополь, 2012. Вып. XII. С. 72–78.
5. **Клейнрок Л.** Теория массового обслуживания / пер. с англ. И. И. Грушко; ред. В. И. Нейман. — М.: Машиностроение, 1979. — 432 с.
6. **Жерновков В. А., Дмитриенко Д.В.** Распределение ресурсов на основе алгоритмов открытого поиска // Журнал университета водных коммуникаций. 2009. № 3. С. 153–156.
7. **Давидюк Н. В.** Разработка системы поддержки принятия решений для обеспечения физической безопасности объектов [электронный ресурс]: дис. ... канд. техн. наук: 05.13.01, 05.13.19 — Астрахань: РГБ, 2010.
8. **Holland J. H.** Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. — Massachusetts Institute of Technology, 1992. — 328 p.
9. **Уитли Д.** Учебник по генетическим алгоритмам // Статистика и компьютинг. 1994. № 4. С. 65–85.
10. **Syswerda G.** Schedule optimization using genetic algorithms // Handbook of genetic algorithms / ed. By L. Davis. — N. Y.: Van Nostrand Reinhold, 1991. P. 332–349.
11. **Газизов Р. К., Гагарин А. В.** Гибридный генетический нейросетевой алгоритм в задаче идентификации параметров цифровых моделей // Вестник УГАТУ. Уфа, 2009. Т. 13. № 2. С. 246–256.

УДК 519.614

О СУЩЕСТВОВАНИИ МАТРИЦ МЕРСЕННА 11-ГО И 19-ГО ПОРЯДКОВ

Н. А. Балонин,
доктор техн. наук, профессор
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Приведено определение обобщенных матриц Мерсенна. Показаны примеры таких матриц порядков, отличающихся от порядков матриц, соответствующих последовательности Сильвестра, сформулирована гипотеза об их существовании.

Ключевые слова *— ортогональные матрицы, матрицы Адамара, матрицы Адамара — Мерсенна, числа Мерсенна.*

В работах [1, 2] предложены версии малоуровневых ортогональных матриц, нечетных порядков, равных числам Мерсенна и Ферма. В процессе исследований выяснено, что последовательность матриц Адамара — Мерсенна, в отличие от последовательности матриц Адамара — Ферма, сходна с матрицами Адамара в том, что ассоциированные с ними матрицы встречаются чаще. Такие общие матрицы для большей простоты будем называть матрицами Мерсенна.

Определение 1. Значения, которым равны элементы матрицы, будем называть ее уровнями.

Значения уровней позволяют формировать графические портреты матриц.

Определение 2. Матрица Мерсенна — это квадратная двухуровневая матрица $\mathbf{M}_n$ порядка n, состоящая из чисел $\{a = 1, -b\}$, столбцы которой ортогональны:

$$\mathbf{M}_n^{\mathrm{T}}\mathbf{M}_n = \mu\mathbf{I},$$

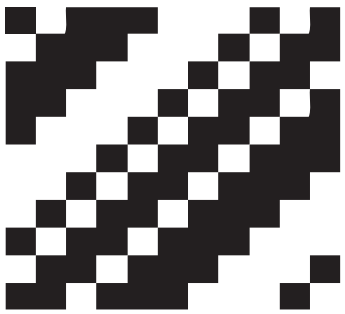
где $b = \frac{1}{2}$ при $n = 3$, в остальных случаях $b = \frac{q-\sqrt{4q}}{q-4}$, $q = n + 1$ (порядок матрицы Адамара); вес $\mu = \frac{(n+1)+(n-1)b^2}{2}$ учитывает, что $\frac{q}{2}$ элементов каждого столбца такой матрицы составляют $a = 1$, остальные элементы равны $-b$.

Портреты двух таких матриц $\mathbf{M}_{11}$ $\left(\text{уровень } b = \frac{3-\sqrt{3}}{2}\right)$ и $\mathbf{M}_{19}$ $\left(\text{уровень } b = \frac{5-\sqrt{5}}{2}\right)$ приведены на рис. 1 и 2, элементу $a = 1$ соответствует белый цвет, элементу $-b$ — черный.

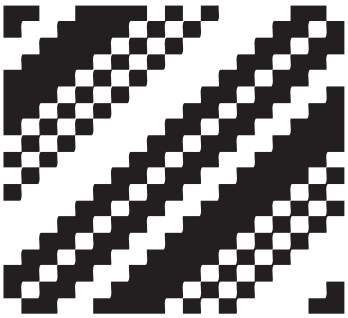
Существование таких матриц позволяет высказать следующее немаловажное предположение, эквивалентное гипотезе Адамара [3].

Гипотеза. Матрицы Мерсенна порядков $4k - 1$ существуют.

В отличие от матриц Адамара матрицы Мерсенна имеют нечетные значения порядков. Первые две матрицы Адамара $\mathbf{H}_{12}$ и $\mathbf{H}_{20}$ соответственно 12-го и 20-го порядков, послужившие основой его гипотезы, приведены в работе [3].



■ ***Рис. 1.*** Портрет матрицы $\mathbf{M}_{11}$



■ ***Рис. 2.*** Портрет матрицы $\mathbf{M}_{19}$

Сформулированная выше гипотеза позволяет ожидать существенного расширения множества четырехуровневых матриц Адамара — Эйлера [4], включая особые порядки 22, 34, 58 и т. п., на которых альтернативных им трехуровневых матриц Белевича не существует. Такие версии пополняют список особых матриц [5], что имеет принципиальное значение для теории *M*-матриц [6, 7] и теории кодирования информации, расширяя ортогональный базис преобразований.

Литература

1. **Балонин Н. А., Сергеев М. Б., Мироновский Л. А.** Вычисление матриц Адамара — Мерсенна // Информационно-управляющие системы. 2012. № 5. С. 92–94.
2. **Балонин Н. А., Сергеев М. Б., Мироновский Л. А.** Вычисление матриц Адамара — Ферма // Информационно-управляющие системы. 2012. № 6. С. 90–93.
3. **Hadamard J.** Résolution d'une question relative aux determinants // Bulletin des Sciences Mathématiques. 1893. Vol. 17. P. 240–246.
4. **Балонин Н. А., Сергеев М. Б.** О двух способах построения матриц Адамара — Эйлера // Информационно-управляющие системы. 2013. № 1. С. 7–10.
5. **Балонин Ю. Н., Сергеев М. Б.** М-матрица 22-го порядка // Информационно-управляющие системы. 2011. № 5. С. 87–90.
6. **Балонин Н. А., Сергеев М. Б.** М-матрицы // Информационно-управляющие системы. 2011. № 1. С. 14–21.
7. **Балонин Н. А., Мироновский Л. А.** Матрицы Адамара нечетного порядка // Информационно-управляющие системы. 2006. № 3. С. 46–50.

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (80x@mail.ru).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию. Рукописи не возвращаются.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

РЕЦЕНЗИЯ НА МОНОГРАФИЮ К. В. ГРИГОРЬЕВОЙ «АППРОКСИМАЦИЯ КРИТЕРИАЛЬНОГО ФУНКЦИОНАЛА В ЗАДАЧАХ МАТЕМАТИЧЕСКОЙ ДИАГНОСТИКИ»

LAP LAMBERT Academic Publishing, Saarbrucken, Germany, 2011. — 232 с. ISBN 978-3-8454-1577-2

Монография посвящена разработке методик решения задач математической диагностики, основанных на математическом программировании в рамках оптимизационного подхода. К задачам математической диагностики относятся задачи распознавания образов, обработки наблюдений, анализа экспериментальных данных, идентификации и др.

Все эти задачи можно свести к исследованию модели, в которой требуется разделить два или более множества точек в многомерных пространствах. Если выпуклые оболочки этих множеств не пересекаются, то их можно разделить с помощью гиперплоскости на два полупространства, в каждом из которых лежит одно из рассматриваемых множеств. Тогда при решении вопроса о принадлежности некоторой произвольной точки с заданными координатами одному из множеств достаточно подставить координаты этой точки в левую часть уравнения гиперплоскости и определить знак результата. Эта идея реализована в монографии в классе плохо разделяемых множеств, т. е. множеств, которые, вообще говоря, гиперплоскостью разделить нельзя. Для таких множеств, какова бы ни была гиперплоскость, хотя бы в одном из образованных ею подпространств найдутся точки того и другого множества.

Автор в своем исследовании ищет такую гиперплоскость, для которой суммарное число точек, попавших не в то полупространство, минимально. Это требование приводит к задаче о минимуме функционала, определяемого параметрами гиперплоскости (ее нормалью и расстоянием от начала координат), значениями которого является число «ошибочно идентифицируемых» точек. Поскольку этот функционал, называемый автором «натуральным», является существенно разрывной функцией, то применение к решению задачи минимизации классических методов, разработанных для гладких или непрерывных негладких функций, затруднено. Эта трудность обходится построением достаточно хорошей аппроксимации «натурального» функционала, называемой «суррогатным» функционалом. Автор рассматривает в качестве «суррогатных» два функционала, один из которых является субдифференцируемым, а второй — непрерывно дифференцируемым. В работе подробно изучены свойства этих функционалов и предложены методы для их оптимизации, являющиеся обобщением и (или) комбинацией известных методов безусловной и условной минимизации.

В качестве методов численного решения упомянутых задач рассматриваются три группы методов построения направления наискорейшего спуска: методы нормирования, методы проектирования и комбинированные методы нормирования и проектирования вектора гиперплоскости в задаче условной минимизации. В каждой группе представлено три метода: «релаксационный» метод, построенный для обоих суррогатных функционалов; «сходящийся» метод, построенный для непрерывно дифференцируемого суррогатного функционала, и «смешанный» метод, построенный с использованием обоих суррогатных функционалов. Разработанные численные методы позволяют использовать информацию о направлениях наискорейшего спуска суррогатных функционалов для решения задачи минимизации натурального функционала.

На основе проведенных в монографии теоретических исследований автором создано программное обеспечение для решения поставленных оптимизационных задач. В результате удается найти оптимальное положение гиперплоскости, благодаря которому можно эффективно решать практические задачи о принадлежности имеющихся (и вновь появляющихся) объектов тому или другому множеству.

Наиболее интересную часть работы представляет выполненное автором исследование нескольких известных (часто используемых в литературе для сравнения) баз данных, среди которых имелись пять медицинских баз данных и одна база данных, связанная с денежным обращением. Эти эксперименты подтверждают эффективность предложенных методов. В монографии разработаны рекомендации для практического использования теоретических результатов (в частности, для решения задачи прогнозирования эффективности применения химиотерапии при лечении онкологических заболеваний).

Монография состоит из введения, трех глав, заключения, приложения и списка литературы,

занимая в общей сложности 232 страницы. Во введении проведен анализ работ по математической диагностике и проблемам идентификации, изложены краткое содержание работы, цели и задачи исследования, перечислены полученные результаты. В первой главе содержится постановка задачи, исследуются свойства суррогатных функционалов, доказывается возможность применения точных штрафных функций к решению поставленной задачи. Вторая глава посвящена численным методам. В ней приведены алгоритм метода условного градиента, методы нормирования и проектирования для минимизации функционала, рассмотрены построение направления наискорейшего спуска и использование полученных результатов при построении методов минимизации натурального функционала. Третья глава содержит результаты применения разработанных методов на перечисленных выше базах данных, имеющие теоретическое и прикладное значение.

Оценивая монографию в целом, можно утверждать, что она представляет собой завершенное научное исследование, свидетельствующее о личном вкладе ее автора в развитие негладкого дискриминантного анализа. Работа выполнена на хорошем математическом уровне. В монографии предложены новые математические методы и вычислительные алгоритмы для решения задач математической диагностики, которые могут использоваться в различных областях знаний, в частности в задачах медицинской и технической диагностики.

Доктор технических наук, профессор,
заведующий кафедрой прикладной информатики
Санкт-Петербургского государственного
университета технологии и дизайна
В. И. Пименов

АКИМЦЕВ Владимир Васильевич

Доцент, старший научный сотрудник научно-исследовательского центра проблем федеральной системы разведки и контроля воздушного пространства Санкт-Петербургского высшего военного училища радиоэлектроники (военного института).
В 1973 году окончил Ленинградский институт авиационного приборостроения по специальности «Радиотехника».
В 1991 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 60 научных публикаций и одного патента на изобретение.
Область научных интересов — статистическая радиолокация, обработка радиолокационной информации.
Эл. адрес: vvznak@mail.ru

БРИТОВ Георгий Семенович

Доцент кафедры информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1963 году окончил Ленинградский институт авиационного приборостроения по специальности «Авиационное приборостроение».
В 1968 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 100 научных публикаций и 13 патентов на изобретения.
Область научных интересов — теория надежности и техническая диагностика информационных систем.
Эл. адрес: bgs@ibi.metrocom.ru

ГОРОДЕЦКИЙ Андрей Емельянович

Доктор технических наук, профессор, заведующий лабораторией методов и средств автоматизации Института проблем машиноведения РАН, г. Санкт-Петербург, заслуженный деятель науки и техники.
В 1965 году окончил Ленинградский политехнический институт им. М. И. Калинина по специальности «Автоматика и телемеханика».
В 1993 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 150 научных публикаций и 70 изобретений.
Область научных интересов — математическое моделирование, оптимальное управление, идентификация и диагностика.
Эл. адрес: gorodetsky@mail23.ipme.ru

БАЛОНИН Николай Алексеевич

Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1982 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Автоматика и телемеханика».
В 2008 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 80 научных публикаций, в том числе трех монографий.
Область научных интересов — теория динамических систем, теория идентификации, теория операторов, теория матриц, вычислительные методы, интернет-робототехника, интернет-книги с исполняемыми алгоритмами, научные социальные сети.
Эл. адрес: korbendfs@mail.ru

ВЕРШИНИНА Анна Сергеевна

Магистрант кафедры электроники и оптической связи Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2011 году с отличием окончила бакалавриат Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Оптотехника».
Является автором восьми научных публикаций.
Область научных интересов — частотно-временные характеристики фемтосекундных импульсов, поляризационные характеристики электромагнитного поля.
Эл. адрес: avershinina1203@gmail.com

ГУРЬЯНОВ Денис Юрьевич

Ассистент кафедры автоматизированных систем управления и обработки информации Санкт-Петербургского государственного электротехнического университета «ЛЭТИ».
В 2007 году окончил Санкт-Петербургский государственный университет водных коммуникаций по специальности «Комплексное обеспечение информационной безопасности автоматизированных систем».
В 2011 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 18 научных публикаций.
Область научных интересов — информационная безопасность, криптография, практическая реализация криптографических протоколов.
Эл. адрес: rightx@gmail.com

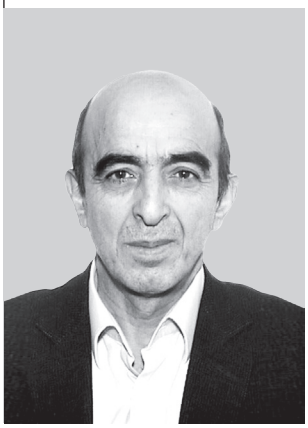
ДЕМЬЯНЧУК Анна Алексеевна

Младший научный сотрудник научно-исследовательского отдела проблем информационной безопасности Санкт-Петербургского института информатики и автоматизации РАН, аспирант Санкт-Петербургского государственного электротехнического университета «ЛЭТИ».
В 2011 году окончила Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина) по специальности «Компьютерная безопасность».
Является автором двух научных публикаций.
Область научных интересов — информационная безопасность, математические основы криптографии, двухключевая криптография, электронные цифровые подписи, протоколы с нулевым разглашением.
Эл. адрес: anonimkina@gmail.com

КОПЫТОВ Владимир Вячеславович

Профессор кафедры организации и технологии защиты информации Северо-Кавказского федерального университета, г. Ставрополь.
В 1983 году окончил Ставропольское высшее военное инженерное училище связи по специальности «Инженер по радиосвязи».
В 2004 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 140 научных публикаций и четырех патентов на изобретения.
Область научных интересов — стохастическая динамика нелинейных динамических систем, методы прогнозирования поведения динамических систем, проектирование информационных систем, информационное противоборство.
Эл. адрес: kopytov@stavsu.ru

КУРБАНОВ Вугар Гариб оглы



Старший научный сотрудник лаборатории методов и средств автоматизации Института проблем машиноведения РАН, г. Санкт-Петербург.
В 1976 году окончил Азербайджанский государственный университет им. С. М. Кирова по специальности «Прикладная математика».
В 1983 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором 50 научных публикаций.
Область научных интересов — математическое моделирование процессов управления, методы логического анализа систем, логико-вероятностные методы.
Эл. адрес:
vugar_borchali@yahoo.com

КЛЕЙМЕНОВА Елена Михайловна

Руководитель НТЦ «Корпоративные информационные технологии» ОАО «РКК «Энергия», г. Королев, аспирант Института проблем управления сложными системами РАН, г. Самара.
В 1992 году окончила Московский институт электронного машиностроения по специальности «Системы автоматизированного проектирования».
Является автором двух научных публикаций.
Область научных интересов — информационные технологии интеллектуальной поддержки принятия решений, средства создания и поддержки проблемно-ориентированных систем, основанных на знаниях.
Эл. адрес:
elena.kleimenova@rsce.ru

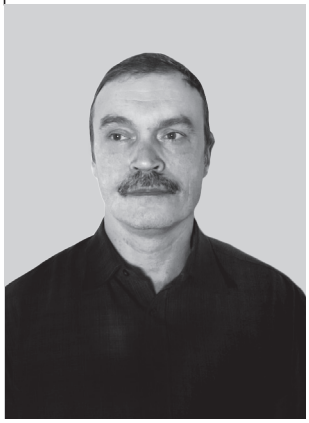
КУЛАКОВ Сергей Викторович

Профессор, заведующий кафедрой электроники и оптической связи Санкт-Петербургского государственного университета аэрокосмического приборостроения, лауреат Государственной премии СССР, заслуженный деятель науки и техники РФ, академик Международной академии наук высшей школы.
В 1953 году окончил Ленинградский институт авиационного приборостроения по специальности «Радиотехника».
В 1980 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 160 научных публикаций и более 30 авторских свидетельств на изобретения.
Область научных интересов — акустооптика, акустоэлектроника, радиотехника.
Эл. адрес: svk25@mail.ru

ЛАРЮХИН Владимир Борисович

Директор по разработкам ООО «НПК «Разумные решения», г. Самара, аспирант кафедры прикладной математики и вычислительной техники Самарского государственного архитектурно-строительного университета.
В 2010 году окончил факультет информационных систем и технологий Самарского государственного архитектурно-строительного университета по специальности «Информационные системы и технологии».
Является автором десяти научных публикаций.
Область научных интересов — мультиагентные системы, прикладные онтологии, управление развитием компетенций и способностей людей и др.
Эл. адрес:
Vladimir.larukhin@live.ru

МАЙОРОВ Игорь Владимирович



Ведущий специалист научно-исследовательского отдела научно-исследовательской группы лаборатории интеллектуальных технологий ООО «НПК «Разумные решения», г. Самара.
В 1988 году окончил Куйбышевский государственный университет по специальности «Теоретическая физика».
Является автором 20 научных публикаций.
Область научных интересов — мультиагентные системы, искусственный интеллект, системы планирования в реальном времени, методы оптимизации.
Эл. адрес:
imayorov@smartsolutions-123.ru

МАШЕВСКИЙ Глеб Алексеевич

Ассистент кафедры биотехнических систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ».
В 2009 году окончил магистратуру Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» по специальности «Биотехнические системы и технологии».
В 2012 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 28 научных публикаций.
Область научных интересов — потенциометрия, биомедицинская электроника, моделирование биологических объектов.
Эл. адрес: Aniket@list.ru

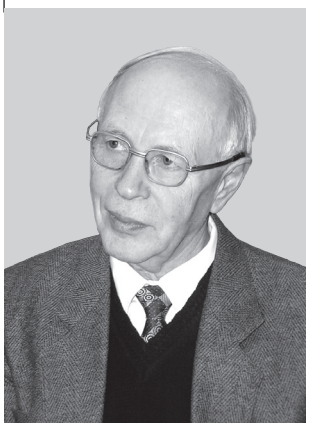
МОЛДОВЯН Дмитрий Николаевич

Младший научный сотрудник научно-исследовательского отдела проблем информационной безопасности Санкт-Петербургского института информатики и автоматизации РАН.
В 2009 году окончил Санкт-Петербургский электротехнический университет «ЛЭТИ» по специальности «Компьютерная безопасность».
В 2012 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 38 научных публикаций и пяти патентов на изобретения.
Область научных интересов — криптография, аутентификация информации, протоколы электронной цифровой подписи, схемы открытого распределения ключей, компьютерная безопасность.
Эл. адрес: mdn.spectr@mail.ru

МАКСИМЕНКО Сергей Леонидович

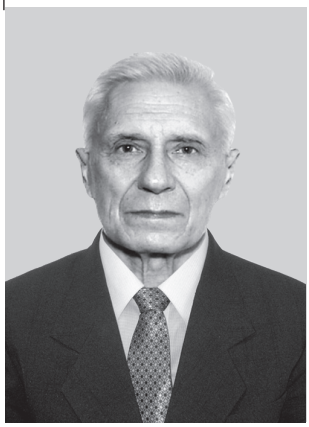
Старший преподаватель кафедры компьютерных систем и программных технологий Санкт-Петербургского государственного политехнического университета.
В 1998 году окончил с отличием Санкт-Петербургский государственный технический университет по специальности «Вычислительные машины, комплексы, системы и сети».
Является автором более 20 научных публикаций.
Область научных интересов — технологии проектирования аппаратуры вычислительных систем, системы на кристалле, цифровая обработка сигналов.
Эл. адрес:
sl_max@kspt.ftk.spbstu.ru

МЕЛЕХИН Виктор Федорович



Профессор, заведующий кафедрой компьютерных систем и программных технологий Санкт-Петербургского государственного политехнического университета, почетный работник высшего профессионального образования РФ.
В 1960 году окончил Ленинградский политехнический институт.
В 1984 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 230 научных публикаций, в том числе четырех монографий и 75 изобретений.
Область научных интересов — теория и технология проектирования вычислительных систем и устройств.
Эл. адрес:
melekhin@kspt.ftk.spbstu.ru

МОСКАЛЕЦ Олег Дмитриевич



Доцент кафедры электроники и оптической связи Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1961 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Радиотехника».
В 1970 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 120 научных публикаций и пяти патентов на изобретения.
Область научных интересов — теория сигналов, теория линейных систем, спектрально-корреляционный анализ сигналов, квантовая физика.
Эл. адрес: molegd@mail.ru

НАУМЕНКО Владимир Викторович

Аспирант кафедры организации и технологии защиты информации Северо-Кавказского федерального университета, г. Ставрополь.
В 2010 году окончил Ставропольский государственный университет по специальности «Организация и технология защиты информации».
Является автором 15 научных публикаций и двух свидетельств о регистрации программы для ЭВМ.
Область научных интересов — теория систем и системный анализ, имитационное моделирование, проектирование информационных систем.
Эл. адрес: ssu-2008@mail.ru

НОСАЛЬ Ирина Алексеевна

Аспирант Санкт-Петербургского института информатики и автоматизации РАН.
В 2010 году окончила с отличием Сыктывкарский государственный университет по специальности «Информационная безопасность».
Является автором трех научных публикаций.
Область научных интересов — информационная безопасность.
Эл. адрес: ironia.i@gmail.com

ПИМЕНОВ Виктор Игоревич

Профессор, заведующий кафедрой прикладной информатики Санкт-Петербургского государственного университета технологии и дизайна.
В 1983 году окончил Ленинградский механический институт «Военмех» по специальности «Системы автоматического управления».
В 2009 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 120 научных публикаций.
Область научных интересов — анализ данных, распознавание образов, мультимедиа технологии, 3D-моделирование.
Эл. адрес: v_pim@mail.ru

НОВИКОВА Евгения Сергеевна

Ассистент кафедры автоматизированных систем управления и обработки информации Санкт-Петербургского государственного электротехнического университета «ЛЭТИ».
В 2007 году с отличием окончила Санкт-Петербургский электротехнический университет «ЛЭТИ» по специальности «Компьютерная безопасность».
В 2009 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором 45 научных публикаций и трех патентов на изобретения.
Область научных интересов — информационная безопасность, визуализация событий безопасности, двухключевая криптография, электронные цифровые подписи и др.
Эл. адрес:
novikova@comsec.spb.ru

ОСИПОВ Василий Юрьевич

Профессор, ведущий научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН.
В 1981 году окончил Высшее военно-морское училище радиоэлектроники им. А. С. Попова по специальности «Радиотехнические средства».
В 2000 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 100 научных публикаций.
Область научных интересов — интеллектуальные системы, моделирование, информационная безопасность.
Эл. адрес: osipov_vasiliy@mail.ru

ПОЛОНЧУК Евгений Владимирович

Начальник отдела корпоративных информационных технологий ОАО «РКК «Энергия», г. Королев.
В 2003 году окончил Московский государственный университет леса по специальности «Информационно-измерительная техника и технологии».
Является автором одной научной публикации.
Область научных интересов — информационные технологии интеллектуальной поддержки принятия решений, средства создания и поддержки проблемно-ориентированных систем, основанных на знаниях.
Эл. адрес:
Evgeny.Polonchuk@rsce.ru

САВЧЕНКО Андрей Владимирович

Доцент кафедры информационных систем и технологий Национального исследовательского университета Высшая школа экономики, г. Нижний Новгород.
В 2008 году окончил Нижегородский государственный технический университет по специальности «Прикладная математика».
В 2010 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 40 научных публикаций и двух патентов на полезную модель.
Область научных интересов — статистическое распознавание образов, анализ и понимание изображений, распознавание речи.
Эл. адрес: avsavchenko@hse.ru

САВЧЕНКО Владимир Васильевич

Профессор, заведующий кафедрой математики и информатики Нижегородского государственного лингвистического университета.
В 1977 году окончил Горьковский политехнический институт по специальности «Автоматизированные системы управления».
В 1994 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 150 научных публикаций и 26 патентов на изобретения.
Область научных интересов — статистические методы обработки информации, распознавание образов и прогнозирование случайных сигналов.
Эл. адрес: svv@lunn.ru

СИМОНОВА Елена Витальевна

Ведущий аналитик ООО «НПК «Разумные решения», доцент кафедры информационных систем и технологий Самарского государственного аэрокосмического университета им. акад. С. П. Королева (национального исследовательского университета).
В 1985 году окончила Куйбышевский авиационный институт им. акад. С. П. Королева по специальности «Автоматизированные системы управления».
В 1994 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 100 научных публикаций.
Область научных интересов — разработка интеллектуальных систем управления мобильными ресурсами на основе мультиагентных технологий и др.
Эл. адрес:
simonova.elena.v@gmail.com

СКИДИН Антон Сергеевич

Научный сотрудник Института вычислительных технологий Сибирского отделения РАН, г. Новосибирск.
В 2005 году окончил Курганский государственный университет по специальности «Автоматизация технологических процессов и производств».
В 2011 году защитил диссертацию на соискание ученой степени кандидата физико-математических наук.
Является автором пяти научных публикаций.
Область научных интересов — теория кодирования, теория обработки сигналов, телекоммуникации.
Эл. адрес: ask@skidin.org

СКОБЕЛЕВ Петр Олегович

Ведущий научный сотрудник Института проблем управления сложными системами РАН, профессор кафедры инженерии знаний Поволжского государственного университета телекоммуникаций и информатики, г. Самара.
В 1983 году окончил Куйбышевский авиационный институт им. акад. С. П. Королева.
В 2003 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 150 научных публикаций и трех патентов на изобретения.
Область научных интересов — мультиагентные технологии для создания интеллектуальных систем управления ресурсами в реальном времени и др.
Эл. адрес:
petr.skobelev@gmail.com

СМИРНОВ Владимир Александрович

Аспирант кафедры микро- и нанотехнологий аэрокосмического приборостроения Санкт-Петербургского государственного университета аэрокосмического приборостроения, ведущий инженер-электроник отдела новой техники НПЦ «Аквамарин», г. Санкт-Петербург.
В 1988 году окончил Ленинградский институт авиационного приборостроения по специальности «Робототехнические системы».
Является автором пяти научных публикаций и двух авторских свидетельств на изобретения.
Область научных интересов — диагностика сложных технических систем.
Эл. адрес:
vlad.sm2010@yandex.ru

ТАРАСОВА Ирина Леонидовна

Доцент, старший научный сотрудник Института проблем машиноведения РАН, г. Санкт-Петербург.
В 1978 году окончила Ленинградский политехнический институт им. М. И. Калинина.
В 1998 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 50 научных публикаций.
Область научных интересов — математическое моделирование, оптимальное управление, идентификация и диагностика.
Эл. адрес: til@msa2.ipme.ru

ФЕОКТИСТОВ Александр Леонидович

Заместитель Генерального конструктора по информационным технологиям НТЦ «Корпоративные информационные технологии», ОАО «РКК «Энергия», г. Королев.
В 1977 году окончил Московский авиационный институт по специальности «Производство летательных аппаратов».
Является автором трех научных публикаций.
Область научных интересов — информационные технологии интеллектуальной поддержки принятия решений, технологии и системы, основанные на знаниях.
Эл. адрес:
Alexander.Feoktistov@rsce.ru

ШОКИН Юрий Иванович

Академик РАН, директор Института вычислительных технологий, член Президиума Сибирского отделения РАН, кавалер ордена Дружбы, ордена «Знак Почета», ордена Почета.
В 1966 году окончил Новосибирский государственный университет по специальности «Математика».
В 1981 году защитил диссертацию на соискание ученой степени доктора физико-математических наук.
Является автором 250 научных публикаций.
Область научных интересов — информатика, математическое моделирование, прикладная математика, телекоммуникации.
Эл. адрес: dir@ict.nsc.ru

ФЕДОРУК Михаил Петрович

Профессор, ректор Новосибирского государственного университета, заведующий лабораторией вычислительных технологий Института вычислительных технологий Сибирского отделения РАН.
В 1982 году окончил Новосибирский государственный университет по специальности «Физика».
В 1999 году защитил диссертацию на соискание ученой степени доктора физико-математических наук.
Является автором 200 научных публикаций.
Область научных интересов — математическое моделирование, телекоммуникации, лазерные системы.
Эл. адрес: mife@ict.nsc.ru

ФИЛЬЧЕНКОВ Андрей Александрович

Аспирант математико-механического факультета Санкт-Петербургского государственного университета, младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики Санкт-Петербургского института информатики и автоматизации РАН.
В 2010 году окончил Санкт-Петербургский государственный университет по специальности «Математическое обеспечение и администрирование информационных систем», в 2011 году — Международный банковский институт по специальности «Финансы и кредит».
Является автором 79 научных публикаций.
Область научных интересов — вероятностные графические модели, алгебраические байесовские сети, автоматическое обучение.
Эл. адрес: aaafil@mail.ru

ЮЛДАШЕВ Зафар Мухамедович

Профессор, заведующий кафедрой биотехнических систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ», член академии медико-технических наук, метрологической академии, почетный работник высшего профессионального образования.
В 1978 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Электронно-медицинская аппаратура».
В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 170 научных публикаций, двух монографий и 17 авторских свидетельств и патентов на изобретения.
Область научных интересов — биомедицинская инженерия, медицинская метрология.
Эл. адрес: yuld@mail.ru

УДК 621.396.96

Поляризационные преобразования зондирующих и отраженных сигналов радиочастотной идентификации

Вершинина А. С., Кулаков С. В., Москалец О. Д. Информационно-управляющие системы, 2013. № 2. С. 2–6.

Исследуются поляризационные преобразования сигналов систем радиочастотной идентификации в среде распространения и приемной антенне. Введены поляризационные спектры векторных сигналов. Метод исследования базируется на представлении поляризационных характеристик сигнала в форме вектора Джонса. Свойства среды распространения и приемной антенны, преобразующие состояние поляризации, описываются частотно-зависимой матрицей Джонса, при этом исходная матрица Джонса представлена в форме матричного ряда.

Ключевые слова — радиочастотная идентификация, поляризационный спектр, вектор Джонса, матрица Джонса, ряд матрицы.

Список лит.: 13 назв.

UDC 621.396.96

Polarization Transformations of Probing and Echo Signals of Radio Frequency Identification

Vershinina A. S., Kulakov S. V., Moskaletz O. D. IUS, 2013. N 2. P. 2–6.

There have been studied polarization transformations of signals of radio-frequency identification systems in a propagation medium and receiving antenna. There have been entered polarization spectra of vector signals. The method of researching is based on representation of polarization characteristics of a signal in the form of Jones vector. Properties of a propagation medium and receiving antenna which transform a polarization state are described with the help of a frequency dependent Jones matrix, meanwhile the parent Jones matrix is presented in the form of a matrix series.

Keywords — Frequency Identification, Polarization Spectrum, Jones Vector, Jones Matrix, Matrix Series.

Refs: 13 titles.

УДК 004.934

Метод фонетического декодирования слов в информационной метрике Кульбака — Лейблера для систем автоматического анализа и распознавания речи с повышенным быстродействием

Савченко В. В., Савченко А. В. Информационно-управляющие системы, 2013. № 2. С. 7–12.

Предложена новая разновидность метода фонетического декодирования слов в расчете на ограниченное множество минимальных звуковых единиц типа отдельных фонем как альтернатива большинству известных методов распознавания речи, основанных на скрытых марковских моделях речевых сигналов. В ее основе используется идея многократного (на порядок и более) сжатия данных за счет того, что слова и фразы из словаря отображаются на последовательность фонетических кодов. Достигаемый эффект, подтвержденный результатами экспериментальных исследований, состоит в увеличении скорости автоматической обработки речевого сигнала при сохранении достаточной точности и надежности распознавания речи.

Ключевые слова — автоматическое распознавание речи, распознавание образов, распознавание с обучением, критерий минимума информационного рассогласования.

Список лит.: 22 назв.

UDC 004.934

The Method of Words Phonetic Decoding Using Kullback-Leibler Information Discrimination for High-Speed Performance Systems of Automatic Speech Analysis and Recognition

Savchenko V. V., Savchenko A. V. IUS, 2013. N 2. P. 7–12.

There has been presented a new kind of words phonetic decoding for a limited set of minimal speech units (separate phonemes) as an alternative to known speech recognition methods based on hidden Markov models of speech signals. It is based on an idea of multiple (by an order of magnitude or more) data compression due to mapping of words and phrases from a vocabulary to a sequence of phonetic codes. The achieved effect confirmed by results of experimental researches is an increase in computational speed of speech signals while preserving adequate speech recognition accuracy and reliability.

Keywords — Automatic Speech Recognition, Pattern Recognition, Supervised Learning, Minimum Information Discrimination Criterion.

Refs: 22 titles.

УДК 004.8

Субоптимальная звездчатая структура алгебраической байесовской сети

Фильченков А. А. Информационно-управляющие системы, 2013. № 2. С. 13–17.

Выделен подкласс минимальных графов смежности — звездчатые графы смежности. Доказано, что множество таких графов содержит оптимальные вторичные структуры, т. е. графы смежности с минимальным диаметром. Представлен алгоритм синтеза такого множества. На основе этого предложен способ ускорения синтеза оптимальной вторичной структуры алгебраической байесовской сети.

Ключевые слова — алгебраическая байесовская сеть, звездчатый граф, обучение глобальной структуры, граф смежности, машинное обучение, субоптимальная вторичная структура.

Список лит.: 20 назв.

UDC 004.8

A Suboptimal Stellate Structure of an Algebraic Bayesian Network

Filchenkov A. A. IUS, 2013. N 2. P. 13–17.

There has been pointed out a subclass of minimal adjacency graphs — stellate adjacency graphs. It is proved that a set of such graphs contains optimal secondary structures, i. e. adjacency graphs with a minimal diameter. There has been presented an algorithm for synthesizing this set.

There has been suggested a method to accelerate synthesis of the optimal secondary structure of an algebraic Bayesian network.

Keywords — Algebraic Bayesian Network, Global Structure Learning, Join Graph, Machine Learning, Stellate Graph, Suboptimal Secondary Structure.

Refs: 20 titles.

УДК 681.3

Анализ надежности функциональных узлов цифровых СБИС со структурным резервированием и периодическим восстановлением работоспособного состояния

Максименко С. Л., Мелехин В. Ф. Информационно-управляющие системы, 2013. № 2. С. 18–23.

Проводится анализ влияния радиационных воздействий на цифровые устройства со структурным резервированием на уровне функциональных узлов интегральных схем в составе информационно-управляющих систем. Предлагается математическая модель, позволяющая оценить надежность узла, представленного на уровне регистровых передач, с учетом цикличности вычислительных процессов и периодического восстановления информации при сбоях в элементах. Показано, что при организации цикличности работы узлов с периодическим восстановлением информации достигается существенное улучшение показателей надежности.

Ключевые слова — информационно-управляющие системы, радиационные эффекты, интегральные схемы, сбои, отказы, восстановление, надежность, модель, структура, троирование, мажоритар.

Список лит.: 7 назв.

UDC 681.3

Analysis of Reliability of Functional Nodes of Digital VLSI Circuits with Structural Redundancy and Periodic Operational State Recovery

Maximenko S. L., Melekhin V. F. IUS, 2013. N 2. P. 18–23.

There has been presented an analysis of radiation effects impact on digital electronic devices with structural redundancy at a level of functional nodes of integrated circuits included in information-management systems. There has been suggested a mathematical model which allows estimating reliability of a node presented at the register transfer level with account of cyclicity of computational processes and periodic information recovery in case of elements faults. The fact has been shown that if cyclic operation of nodes with periodical information backup is arranged significant improvement of reliability factors is reached.

Keywords — Control Systems, Radiation Effects, Integrated Circuit, Fault, Soft Error, Recovery, Reliability, Model, Structure, Triplication, Voter.

Refs: 7 titles.

УДК 519.71

Поиск неисправностей в бортовых системах управления в процессе приемочного контроля

Смирнов В. А. Информационно-управляющие системы, 2013. № 2. С. 24–28.

Рассмотрена методика выбора оптимальной последовательности процедур тестирования электронных блоков, входящих в бортовую систему управления. Определены случаи целесообразного использования предлагаемой методики. На простых примерах показана работа алгоритма поиска неисправного электронного блока.

Ключевые слова — сложная техническая система, диагностирование, байесовская сеть доверия, тестирование, методика выбора, апостериорный вывод, пример решения задачи поиска.

Список лит.: 7 назв.

UDC 519.71

Malfunction Searching in Onboard Control Systems during Acceptance Control

Smirnov V. A. IUS, 2013. N 2. P. 24–28.

There has been considered a procedure of choosing optimal sequence of testing procedures of electronic modules of onboard control systems. There have been defined cases of its expedient and advisable use. There has been demonstrated operation of an algorithm searching for defective electronic blocks.

Keywords — Complex Technical System, Diagnosing, Bayesian Belief Networks, Testing, Technique of Choice, Posterior Output, Example of Solving the Problem of Finding.

Refs: 7 titles.

УДК 005.8:615.478

Метод оценки рисков в мультиагентной системе управления проектами НИР и ОКР в реальном времени

Клейменова Е. М., Феоктистов А. Л., Скобелев П. О., Ларюхин В. Б., Майоров И. В., Симонова Е. В., Полончук Е. В. Информационно-управляющие системы, 2013. № 2. С. 29–37.

Предлагаются удобные для практики количественная модель и метод оценки рисков проектной деятельности по срокам выполнения, позволяющие интерактивно учитывать ход выполнения проектов НИР и ОКР в мультиагентной системе управления проектами в реальном времени. Агенты представляют подразделения, проекты, задачи и сотрудников, причем задачи каждого проекта являются связанными и распределяются на общем поле ресурсов подразделений. Модель основана на линейной аппроксимации вероятностных распределений и вычислении рисков по задачам, связанным отношениями следования. Метод представляет собой порядок расчетов, позволяющий оценивать риск и перепланировать цепочки связанных задач непосредственно в реальном времени, когда распределение задач по сотрудникам постоянно меняется в связи с непредвиденными событиями. Разработанный метод предназначен для снижения рисков при планировании проектов НИР и ОКР в аэрокосмических приложениях.

Ключевые слова — управление проектами, оценка рисков, вероятностный подход, мультиагентные системы, адаптивное планирование, реальное время.

Список лит.: 19 назв.

UDC 005.8:615.478

The Method of Risk Assessment for a Multi-Agent System of Real-Time Management of Research and Development Projects

Kleymenova E. M., Feoktistov A. L., Skobelev P. O., Larukhin V. B., Mayorov I. V., Simonova E. V., Polonchuk E. V. IUS, 2013. N 2. P. 29–37.

There has been suggested a practical quantitative model and method of risk assessment of project for terms of works execution allowing interactive control of progress for research and development projects using a multi-agent system project management in real time. Agents represent departments, projects, tasks and employees while tasks of every project are linked and assigned at the common field of departmental resources. The model is based on a linear approximation of probabilistic distributions and risks calculation for tasks linked by consequence relation. The method is an order of calculations allowing risk assessment and rescheduling chains of the linked tasks in real time when tasks assignment for every employee is constantly changing due to unforeseen events. The developed method is designed for risk reduction during planning of research and development projects in aerospace applications.

Keywords — Project Management, Risks Assessment, Probabilistic Approach, Multi-Agent Systems, Real-Time Planning.

Refs: 19 titles.

УДК 681.5

Имитационное моделирование развития аварийных ситуаций в энергетических установках

Городецкий А. Е., Курбанов В. Г., Тарасова И. Л. Информационно-управляющие системы, 2013. № 2. С. 38–42.

Предложена имитационная модель, которая на основе комбинации логико-вероятностного и логико-лингвистического моделирования позволяет прогнозировать аварийные ситуации в энергетических установках большой единичной мощности.

Ключевые слова — имитационное моделирование, логико-вероятностные переменные, логико-лингвистические переменные, функция принадлежности, вероятность безотказной работы, база данных.

Список лит.: 13 назв.

UDC 681.5

Simulation Modeling of Emergencies Development in Power Installations

Gorodetsky A. E., Kurbanov V. G., Tarasova I. L. IUS, 2013. N 2. P. 38–42.

There has been presented a simulation model based on logical-and-probabilistic and logical-and-linguistic modeling which allows predicting emergencies in power installations of high unit capacity.

Keywords — Simulation Modeling, Logical-and-Probabilistic Variables, Logical-and-Linguistic Variables, Membership Function, Probability of Trouble-Free Work, Database.

Refs: 13 titles.

УДК 681.52

Модель принятия решений при диагностике воспалительных процессов организма по виду интоксикации ионами HS^- и Fe^{2+}

Машевский Г. А., Юлдашев З. М. Информационно-управляющие системы, 2013. № 2. С. 43–47.

Рассматривается возможность использования Pt- и Ag_2S-электродов для контроля развития воспалительных процессов в организме человека, связанных с интоксикацией организма ионами HS^- и Fe^{2+}, а также электрохимическая модель работы данных электродов в присутствии сульфгидрильных соединений. Предложена модель принятия решения по виду интоксикаций, а также алгоритм распознавания данных патологий.

Ключевые слова — система мониторинга, ионометрия, диагностика воспалительного процесса, математическая модель.

Список лит.: 6 назв.

UDC 681.52

A Decision Model for Diagnosing Inflammatory Processes in a Human Body due to Intoxication by HS^- and Fe^{2+} Ions

Mashevskiy G. A., Yuldashev Z. M. IUS, 2013. N 2. P. 43–47.

There has been analyzed a possibility of using Pt- and Ag_2S electrodes for monitoring inflammatory processes connected with HS^- and Fe^{2+} ions intoxication in a human body as well as an electrochemical model of these electrodes over sulfhydryl compounds. There has been presented a decision model relating to intoxication forms as well as an algorithm identifying such pathologies.

Keywords — Monitoring System, Ionometry, Inflammatory Process Diagnostics, Mathematical Model.

Refs: 6 titles.

УДК 681.3.067

Обоснование мероприятий информационной безопасности

Осипов В. Ю., Носаль И. А. Информационно-управляющие системы, 2013. № 2. С. 48–53.

Предложен подход к обоснованию мероприятий информационной безопасности с учетом ценности защищаемых информационных ресурсов. Рассмотрена модель ценности этих ресурсов. Приведены математическая формулировка и алгоритм решения задачи поиска целесообразных мероприятий информационной безопасности, предусматривающие синтез и анализ возможных программ деструктивных воздействий на защищаемые информационные ресурсы. Отражены результаты моделирования.

Ключевые слова — информационная безопасность, оптимизация, методы, программы.

Список лит.: 13 назв.

UDC 681.3.067

Substantiation of Information Security Measures

Osipov V. Yu., Nosal I. A. IUS, 2013. N 2. P. 48–53.

There has been presented an approach to substantiate information security measures with account of value of protected information resources. There has been considered a model of value of these resources. There has been presented mathematical formulation and an algorithm for solving the problem of searching for expedient information security measures providing for synthesis and analysis of possible programs of destructive influence on protected information resources. The modeling results are provided.

Keywords — Information Security, Optimization, Methods, Programs.

Refs: 13 titles.

УДК 519.727, 621.391

Особенности передачи и обработки информации в сверхскоростных волоконно-оптических линиях связи

Шокин Ю. И., Скидин А. С., Федорук М. П. Информационно-управляющие системы, 2013. № 2. С. 54–59.

Проведен анализ особенностей искажения сигнала в высокоскоростных волоконно-оптических линиях связи. На основе анализа предложены методы кодирования и обработки оптического сигнала, учитывающие специфику воздействия на сигнал в волоконном световоде при передаче данных на высокой скорости.

Ключевые слова — волоконная оптика, математическое моделирование, теория кодирования, нелинейные эффекты.

Список лит.: 15 назв.

UDC 519.727, 621.391

Aspects of Information Transmission and Processing in Ultra High-Rate Fiber Optic Communication Lines

Shokin Yu. I., Skidin A. S., Fedoruk M. P. IUS, 2013. N 2. P. 54–59.

There has been conducted an analysis of signal distortion aspects in ultra high-hate fiber optic communication lines. According to the analysis there have been presented methods of optical signal processing and encoding with account of specificity of influence on a signal in a fiber optic conduit during high-rate information transmission.

Keywords — Fiber Optics, Mathematical Modeling, Coding Theory, Nonlinear Effects.

Refs: 15 titles.

УДК 681.3

Подход к построению криптосхем на основе нескольких вычислительно трудных задач

Демьянчук А. А., Молдовян Д. Н., Новикова Е. С., Гурьянов Д. Ю. Информационно-управляющие системы, 2013. № 2. С. 60–66.

Предлагается подход к построению криптосхем, основанных на двух вычислительно трудных задачах, который обеспечивает формирование подписи небольшой длины. Определены требования к выбору системных параметров криптосхем и личных ключей пользователя. Разработанный способ построения криптосхем обладает свойством универсальности и может быть применим для построения протоколов различного типа, таких как протокол открытого распределения ключей, протокол аутентификации с нулевым разглашением секрета.

Ключевые слова — электронная цифровая подпись, протокол открытого шифрования, протокол обмена ключами, задача дискретного логарифмирования, задача факторизации.

Список лит.: 20 назв.

UDC 681.3

An Approach to Crypto Integrated Circuits Construction Based on Several Challenging Computation Problems

Demyanchuk A. A., Moldovyan D. N., Novikova E. S., Gurianov D. U. IUS, 2013. N 2. P. 60–66.

There has been presented an approach to crypto integrated circuits construction based on two challenging computation problems which allows generating short length signatures. There have been defined requirements for choosing system parameters and user's personal keys. The developed approach to constructing crypto integrated circuits is of universal application and can be used for designing cryptographic protocols of different types, such as the public key distribution protocol, the authentication protocol with zero-knowledge.

Keywords — Digital Signature Scheme, Encryption Protocol, Key Agreement Protocol, Discrete Logarithm Problem, Factorization Problem.

Refs: 20 titles.

УДК 621.391.01

Алгоритм разрешения неизвестного числа целей по дальности

Акимцев В. В. Информационно-управляющие системы, 2013. № 2. С. 67–74.

Предложена процедура разрешения по дальности неизвестного числа целей, основанная на анализе цепного отображения цифрового принимаемого сигнала импульсной радиолокационной станции обзора. Приведены характеристики ее качества для простейшего, но весьма распространенного случая, когда сигналы наблюдаются на фоне собственного шума приемника.

Ключевые слова — разрешение целей по дальности, цифровая обработка сигналов, цепное отображение, непараметрическая статистика.

Список лит.: 10 назв.

UDC 621.391.01

Distance Range Resolution Algorithm for Unknown Number of Targets

Akimtsev V. V. IUS, 2013. N 2. P. 67–74.

There has been presented a procedure of distance range resolution for unknown number of targets based on the analysis of received signals chain mapping of a pulse surveillance radar. There have been provided quality characteristics of the procedure for an elementary widespread case when signals are observed against the background of a receiver noise.

Keywords — Distance Resolution of Targets, Digital Signal Processing, Chain Displaying, Nonparametric Statistic.

Refs: 10 titles.

УДК 681.326.74

Верификация, валидация и тестирование компьютерных моделей линейных динамических систем

Бритов Г. С. Информационно-управляющие системы, 2013. № 2. С. 75–82.

Рассмотрены задачи верификации, валидации и тестирования компьютерных моделей линейных динамических систем. Показано, что при верификации модели необходимо перед ее запуском выполнить визуальную проверку, а после запуска — использовать средства верификации применяемых математических пакетов. При валидации модели целесообразно построить устройство функционального диагностирования. Оно позволит проверять процесс моделирования при любых исходных данных. Валидацию можно провести и при тестовых данных, используя специальные режимы моделирования. Приведены примеры компьютерных моделей, построенных на основе математических описаний моделируемых динамических объектов.

Ключевые слова — верификация, валидация, компьютерные модели, функциональное диагностирование, линейные динамические системы, устройство функционального диагностирования, тестирование моделей.

Список лит.: 11 назв.

UDC 681.326.74

Verification, Validation and Testing of Computer Models of Linear Dynamic Systems

Britov G. S. IUS, 2013. N 2. P. 75–82.

There have been analyzed verification, validation and testing of computer models of linear dynamic systems. There has been demonstrated that during verification of a model it is necessary to run visual checking before its launch and verify mathematic software packages after the launch. During validation it is advisable to construct a functional diagnosis device. It allows checking modeling process using whatever initial data. Validation can be run with testing data using special modeling modes. There have been provided examples of computational models based on various mathematical formulations of modeled dynamic objects.

Keywords — Verification, Validation, Computer Models, Functional Diagnose, Linear Dynamic Systems, Functional Diagnose Device, Model Checking.

Refs: 11 titles.

УДК 156.6

Решение задачи распределения ресурсов при выполнении административных регламентов

Науменко В. В., Копытов В. В. Информационно-управляющие системы, 2013. № 2. С. 83–88.

Рассматриваются проблемы повышения эффективности системы государственного управления, одной из которых является распределение ресурсов в процессе исполнения государственных функций и предоставления государственных услуг. Для решения данной проблемы предлагается применение оптимизационной модели, направленной на эффективное распределение выполняемых работ между исполнителями. При этом используются методы теории массового обслуживания и алгоритмы поиска оптимального значения целевой функции.

Ключевые слова — административный регламент, функциональная безопасность, распределение ресурсов.

Список лит.: 11 назв.

UDC 156.6

Solution to the Problem of Resource Allocation at Implementation of Administrative Regulations

Naumenko V. V., Kopytov V. V. IUS, 2013. N 2. P. 83–88.

There have been considered problems of improving efficiency of public administration, one of them is distribution of resources in the course of implementation of public functions and provision of public services. To solve this problem it has been proposed to use an optimization model ensuring efficient distribution of works performed between contractors. Methods of queuing theory and algorithms for finding the optimal value of an objective function have been applied.

Keywords — Administrative Regulation, Functional Safety, Resource Allocation.

Refs: 11 titles.

УДК 519.614

О существовании матриц Мерсенна 11-го и 19-го порядков

Балонин Н. А. Информационно-управляющие системы, 2013. № 2. С. 89–90.

Приведено определение обобщенных матриц Мерсенна. Показаны примеры таких матриц порядков, отличающихся от порядков матриц, соответствующих последовательности Сильвестра, сформулирована гипотеза об их существовании.

Ключевые слова — ортогональные матрицы, матрицы Адамара, матрицы Адамара — Мерсенна, числа Мерсенна.

Список лит.: 7 назв.

UDC 519.614

Existence of Mersenne Matrices of 11th and 19th Orders

Balonin N. A. IUS, 2013. N 2. P. 89–90.

A definition of generalized Mersenne matrices has been given. Examples of such matrices of orders different from orders of matrices corresponding to Sylvester's sequence have been provided; a hypothesis on their existence has been formulated.

Keywords — Orthogonal Matrices, Hadamard Matrices, Hadamard-Mersenne Matrices, Mersenne Numbers.

Refs: 7 titles.

УВАЖАЕМЫЕ АВТОРЫ!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, электронные адреса авторов, которые по требованию ВАК должны быть опубликованы на страницах журнала. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени — эта информация будет опубликована в ссылке на первой странице.

Формулы набирайте в Word, не используя формульный редактор (Mathtype или Equation), при необходимости можно использовать формульный редактор; для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта никогда не пользуйтесь вкладкой Other..., используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; в формулах не отделяйте пробелами знаки: + = –.

Для набора формул в Word никогда не используйте Конструктор (на верхней панели: «Работа с формулами» — «Конструктор»), т. к. этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Иллюстрации в текст не заверстываются и предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы изготавливаются в векторных программах: Visio 4, 5, 2002–2003 (*.vsd); Coreldraw (*.cdr); Excel; Word; AdobeIllustrator; AutoCad (*.dxf); Компас; Matlab (*.ps, *.pdf или экспорт в формат *.ai);

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисуночных подписей обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40 × 55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Оформление статей».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: 80x@mail.ru
Сайт: www.i-us.ru

ISSN 1684-8853

5-я Международная выставка
ОХРАНА. БЕЗОПАСНОСТЬ. ПРОТИВОПОЖАРНАЯ ЗАЩИТА

SIPS OUTH RUSSIA

3 –5 сентября 2013

КРАСНОДАР
ул. Зиповская, 5

Одновременно с выставками:

IDES SOUTH OF RUSSIA — Развитие инфраструктуры Юга России

SOUTH RUSSIA OIL & GAS — Нефть и газ Юга России*

* Ранее выставки GAS RUSSIA и PETROLEUM

Организатор:

ITE

КРАСНОДАРЭКСПО
В составе группы компаний ITE

T +7 (861) 200–12–34, 200–12–81
F +7 (861) 200–12–54
E ides@krasnodarexpo.ru

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

www.SIPS-EXPO.ru