

ИНФОРМАТИКА И ЕЁ ПРИМЕНЕНИЯ

**Научный журнал Отделения нанотехнологий
и информационных технологий Российской академии наук**

Издается с 2007 года
Журнал выходит ежеквартально

Учредители:
Российская академия наук
Институт проблем информатики Российской академии наук

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик С. В. Емельянов (главный редактор, член Редакционного совета)
академик Ю. И. Журавлев (председатель Редакционного совета)
академик С. К. Коровин
академик Г. И. Савин
академик И. А. Соколов (зам. главного редактора, член Редакционного совета)
академик А. Л. Стемпковский
академик Ю. И. Шокин (член Редакционного совета)
член-корреспондент РАН В. Л. Арлазаров
член-корреспондент РАН А. Б. Жижченко
член-корреспондент РАН И. А. Каляев
член-корреспондент РАН Ю. С. Попков
член-корреспондент РАН К. В. Рудаков
член-корреспондент РАН Ю. А. Флеров
член-корреспондент РАН Б. Н. Четверушкин
член-корреспондент РАН Р. М. Юсупов
профессор, д.т.н. В. И. Будзко
профессор, д.т.н. А. А. Зацаринный
профессор, д.ф.-м.н. В. Ю. Королёв
профессор, д.ф.-м.н. А. В. Печинкин
профессор, д.т.н. И. Н. Синицын
профессор, д.ф.-м.н. С. Я. Шоргин (ответственный секретарь)

Редакция

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков;
к.ф.-м.н. Е. Н. Арутюнов;
О. В. Ломакина

© Институт проблем информатики Российской академии наук, 2009

Адрес редакции:

Москва 119333, ул. Вавилова 44, корп. 2, ИПИ РАН,
редакция журнала «Информатика и её применения»
Тел. 8(499)135-86-92, e-mail rust@ipiran.ru

Журнал «Информатика и её применения» включен 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», утвержденный ВАК

Подписной индекс журнала в каталоге «Пресса России» 88018 (годовая подписка)

Информатика и её применения

Том 3 Выпуск 1 Год 2009

СОДЕРЖАНИЕ

Методы построения информационных моделей эредитарных флуктуаций неравномерности вращения Земли И. Н. Сеницын	2
Проблемы распараллеливания биометрических вычислений в крупномасштабных идентификационных системах О. С. Урмаев	8
Развитие методов повышения качества изображений лиц в видеопотоке А. В. Насонов, А. С. Крылов, О. С. Урмаев	19
Восстановление вероятностных характеристик случайных функций в задачах однофотонной эмиссионной томографии В. Г. Ушаков, О. В. Шестаков	29
Концепция создания отечественных интегрированных коммуникационных микроконтроллеров для пакетной коммутации В. Б. Егоров	34
О нестационарных системах обслуживания с катастрофами А. И. Зейфман, Я. А. Сатин, А. В. Чегодаев	47
Байесовские модели массового обслуживания и надежности: экспоненциально-эрланговский случай А. А. Кудрявцев, С. Я. Шоргин	55
Об одном подходе к формированию изображений без использования экрана А. В. Торчигин	60
Оценки скорости сходимости распределений экстремумов обобщенных процессов Кокса с ненулевым средним к сдвиговым смесям нормальных законов С. В. Артюхов	69
Abstracts	74
Об авторах	77
About Authors	78

Выпускающий редактор *Л. Кокушкина*

Технический редактор *Т. Торжкова*

Художественный редактор *М. Седакова*

Сдано в набор 12.01.09. Подписано в печать 16.02.09. Формат 60 x 84 / 8
Бумага офсетная. Печать офсетная. Усл.-печ. л. 10,0. Уч.-изд. л. 10,8. Тираж 200 экз.
Заказ №

Издательство «ТОРУС ПРЕСС», Москва 121614, ул. Крылатская, 29-1-43

torus@torus-press.ru; <http://www.torus-press.ru>

Отпечатано в ППП «Типография «Наука» с готовых файлов
Москва 121099, Шубинский пер., д. 6.

МЕТОДЫ ПОСТРОЕНИЯ ИНФОРМАЦИОННЫХ МОДЕЛЕЙ ЭРЕДИТАРНЫХ ФЛУКТУАЦИЙ НЕРАВНОМЕРНОСТИ ВРАЩЕНИЯ ЗЕМЛИ*

И. Н. Синицын¹

Аннотация: Рассматриваются методы построения информационных моделей флуктуаций неравномерности вращения Земли вследствие эредитарных гравитационно-приливных диссипативных возмущений. Изучены вопросы эквивалентности различных эредитарных возмущений. Созданное методическое и экспериментальное программное обеспечение включено в состав информационных ресурсов по проблеме «статистическая динамика вращения Земли».

Ключевые слова: априорные и апостериорные данные; информационная модель; информационные ресурсы; квазилинейные методы; спектрально-корреляционные характеристики; эредитарные флуктуации неравномерности вращения Земли; эредитарное ядро

1 Введение

Как известно [1–7], одной из важных современных задач статистической динамики вращения Земли является построение информационных моделей флуктуаций неравномерности вращения Земли. Как показывают результаты астрометрических измерений, такие возмущения на внутригодовых интервалах времени оказываются существенными. Разработанные в [7] модели учитывают только нелинейные стохастические механизмы гравитационно-приливных диссипативных возмущений. Дадим обобщение результатов [7] на случай эредитарных стохастических гравитационно-приливных диссипативных возмущений.

2 Эредитарные стохастические интегродифференциальные уравнения флуктуаций неравномерности вращения Земли

Обобщая математические модели [1–7] флуктуаций угловой скорости собственного вращения Земли на случай воздействия как аддитивных, так и параметрических гармонических и широкополосных случайных гравитационно-приливных и эреди-

тарных флуктуационно-диссипативных возмущений на внутригодовых интервалах времени, представим дифференциальное уравнение изменения угла собственного вращения Земли $\delta\varphi$ в следующем виде:

$$\begin{aligned} \delta\ddot{\varphi}(t) = & M_0(t) + X^S(t) + X^L(t) - \\ & - \mu_1 \left[1 + \tilde{\mu}_1(t) + X_1^{\mu S}(t) + X^{\mu L}(t) \right] \delta\dot{\varphi}(t) - \\ & - \int_{t-T_0}^t \mu_1^H(t, \tau) \delta\dot{\varphi}(\tau) d\tau + \mu_n \left[1 + \tilde{\mu}_n + \right. \\ & \left. + X_n^{\mu S}(t) + X_n^{\mu L}(t) \right] F_n(\delta\varphi(t), \delta\dot{\varphi}(t)), \quad (1) \end{aligned}$$

где

$$\begin{aligned} \tilde{M}_{0t} = & M_{10}^S \cos(2\pi f_{\Gamma} t + \chi_1^S) + \\ & + M_{20}^S \cos(4\pi f_{\Gamma} t + \chi_2^S) + \\ & + M_{m0}^L \cos(2\pi\nu_m t + \chi_m^L) + \\ & + M_{f0}^L \cos(2\pi\nu_f t + \chi_f^L); \quad (2) \end{aligned}$$

$$\begin{aligned} \tilde{\mu}_{lt} = & \mu_l \left[\pi_{l1}^{\mu S} \cos(2\pi f_{\Gamma} t + \chi_{l1}^{\mu S}) + \right. \\ & + \pi_{l2}^{\mu S} \cos(4\pi f_{\Gamma} t + \chi_{l2}^{\mu S}) + \\ & + \pi_{lm}^{\mu L} \cos(2\pi\nu_m t + \chi_{lm}^{\mu L}) + \\ & \left. + \pi_{lf}^{\mu L} \cos(2\pi\nu_f t + \chi_{lf}^{\mu L}) \right] \quad (l = 1, n). \quad (3) \end{aligned}$$

* Работа выполнена при финансовой поддержке РФФИ (проект № 07-07-00031) и программы ОНИТ РАН «Информационные технологии и анализ сложных систем» (проект 1.5).

¹ Институт проблем информатики Российской академии наук, sinitsin@iit.ru

Здесь введены следующие обозначения: $\delta\varphi = \varphi - r_*\omega_*^{-1}t$ — изменение угла собственного вращения Земли; φ — угол собственного вращения Земли, значения которого определены на дату t (он является параметром, характеризующим вращение земной системы координат по отношению к небесной); $r_* = 7,292115 \cdot 10^{-5}$ рад/с — постоянная средняя составляющая угловой скорости собственного вращения Земли; ω_* — угловая скорость обращения Земли по орбите; t — безразмерное время, измеряемое стандартными годами; $f_\Gamma = 1, 2f_\Gamma$, $\nu_m = 13,28$, $\nu_f = 26,28$, M_{10}^S , M_{20}^S , M_{m0}^L , M_{f0}^L и χ_1^S , χ_2^S , χ_m^L , χ_f^L — частоты, амплитуды и начальные фазы аддитивных гармонических возмущений от Солнца (S) и Луны (L), соответствующие годовому, полугодовому, месячному и двухнедельному циклам; μ_h ($h = 1, n$) — коэффициенты диссипативных моментов сил, обусловленных разнообразием геофизических процессов (приливное трение океанических и земных приливов, атмосферные воздействия, океанические течения, перераспределение водных масс и т.п.); $\pi_{h1}^{\mu S}$, $\pi_{h2}^{\mu S}$, $\pi_{hm}^{\mu L}$, $\pi_{hf}^{\mu L}$ и $\chi_{h1}^{\mu S}$, $\chi_{h2}^{\mu S}$, $\chi_{hm}^{\mu L}$, $\chi_{hf}^{\mu L}$ ($h = 1, n$) — амплитуды и начальные фазы параметрических гармонических диссипативных моментов сил на частотах f_Γ , $2f_\Gamma$, ν_m , ν_f ; $F_n = F_n(\delta\varphi, \delta\dot{\varphi})$ — нелинейная составляющая диссипации (в случае рэлеевского механизма диссипации $F_n(\delta\dot{\varphi}) = \delta\dot{\varphi}^3$); $X^S(t)$, $X^L(t)$ и $X_h^{\mu S}(t)$, $X_h^{\mu L}(t)$ — нормальные (гауссовские) широкополосные аддитивные и параметрические случайные возмущения с известными математическими ожиданиями и ковариационными характеристиками; $\mu_1^H = \mu_1^H(t, \tau)$ — нестационарные эредитарные ядра, определяющие эредитарность («память») гравитационно-приливных диссипативных сил.

Примерами стационарных эредитарных ядер служат функции вида $\mu_1^H(t, \tau) = \mu_1^H(\rho)$, $\rho = t - \tau$, где

$$\mu_1^H(\rho) = \mu_{10}^H e^{-\alpha|\rho|} \mathbf{1}(\rho); \quad (4)$$

$$\mu_1^H(\rho) = \mu_{10}^H e^{-\alpha|\rho|} (1 + \alpha|\rho|) \mathbf{1}(\rho); \quad (5)$$

$$\mu_1^H(\rho) = \mu_{10}^H e^{-\alpha|\rho|} \left(1 + \alpha|\rho| + \frac{1}{3} \alpha^2 \rho^2 \right) \mathbf{1}(\rho); \quad (6)$$

$$\mu_1^H(\rho) = \mu_{10}^H e^{-\alpha|\rho|} \cos \omega\rho \cdot \mathbf{1}(\rho); \quad (7)$$

$$\mu_1^H(\rho) = \mu_{10}^H e^{-\alpha|\rho|} (\cos \omega\rho + \gamma \sin \omega|\rho|) \mathbf{1}(\rho), \quad (8)$$

допускающие производные по ρ различных порядков. Ядра (4)–(6) определяют чисто затухающую эредитарность, а ядра (7) и (8) — затухающую колебательную эредитарность на частоте ω .

3 Априорные методы построения информационных корреляционных эредитарных моделей

Основываясь на уравнениях (1)–(4), составим приближенные уравнения для математических ожиданий, дисперсий, ковариаций и ковариационных функций основных переменных при следующих основных допущениях.

- 1° Аддитивный возмущающий момент допускает представление $X^S(t) + X^L(t) = m_0^{SL} + V_1$, где V_1 является скалярным нормальным белым шумом интенсивности $\nu_1 = \nu_1(t)$, происходящим от первого источника [7].
- 2° Возмущающие моменты $X_3 = X_1^{\mu S}(t) + X_1^{\mu L}(t)$ и $X_4 = X_n^{\mu S}(t) + X_n^{\mu L}(t)$ происходят от второго источника нормального белого шума V_2 единичной интенсивности ($\nu_2 = 1$). Они имеют конечные математические ожидания m_3 и m_4 и дисперсии σ_3^2, σ_4^2 и удовлетворяют следующим скалярным уравнениям формирующего фильтра [8]:

$$\dot{X}_l^0 = -\alpha_l X_l^0 + \sigma_l \sqrt{2\alpha_l} \cdot V_2;$$

$$X_l^0 = X_l - m_l^*, \quad l = 3, 4, \quad \alpha_l > 0,$$

где m_l^* — постоянные значения m_l .

- 3° Нелинейные функции $X_2 X_3, F_n = F_n(X_1, X_2)$, $F'_n = F'_n(X_1, X_2, X_4) = F_n(X_1, X_2) X_4$, входящие в (1), допускают статистическую линеаризацию нелинейностей по формулам [8]:

$$X_2 X_3 \approx m_2 m_3 + k_{23} + m_2 X_3^0 + m_3 X_2^0;$$

$$F_n = F_n(X_1, X_2) \approx F_{n0} + F_{n1} X_1^0 + F_{n2} X_2^0;$$

$$F'_2 = F_1(X_1, X_2) X_4 \approx F'_{n0} + F'_{n1} X_1^0 + F'_{n2} X_2^0 + F'_{n4} X_4^0.$$

Здесь m_i и k_{ij} — математические ожидания и ковариационные моменты переменных X_i и $X_i^0 = X_i - m_i$;

$$F_{n0} = F_{n0}(m_1, m_2, k_{11}, k_{12}, k_{22}) = \mathbf{M}_N^{(1,2)} [F_n(X_1, X_2)];$$

$$F_{n1} = \frac{\partial F_{n0}}{\partial m_1}; \quad F_{n2} = \frac{\partial F_{n0}}{\partial m_2};$$

$$F'_{n0} = F'_{n0}(m_1, m_2, m_4, k_{11}, k_{12}, k_{22}, k_{14}, k_{24}, k_{44}) = \mathbf{M}_N^{(1,2,4)} [F'_n(X_1, X_2, X_4)];$$

$$F'_{n1} = \frac{\partial F'_{n0}}{\partial m_1}; \quad F'_{n2} = \frac{\partial F'_{n0}}{\partial m_2}; \quad F'_{n4} = \frac{\partial F'_{n0}}{\partial m_4},$$

где $M_N^{(1,2)}[\cdot]$ и $M_N^{(1,2,4)}[\cdot]$ — символы вероятностного осреднения для двумерного и трехмерного нормального распределения переменных X_1, X_2 и X_1, X_2, X_4 .

4° Эредитарные ядра $\mu_1^H(t, \tau)$ имеют вид (4), причем при $t < t_0 + T_0$ влиянием значений $X_2(\tau)$ при $\tau \in (t - T_0, t_0)$ можно пренебречь, положив $t - T_0 \approx t_0$. Тогда, следуя [8–10] и приняв

$$\int_{t_0}^t \mu_{10}^H e^{-\alpha_5 |t-\tau|} X_2(\tau) d\tau = X_5,$$

будем иметь следующие дифференциальные уравнения, связывающие вспомогательную переменную X_5 с исходными переменными:

$$\dot{X}_5 = -\alpha_5 X_5 - \mu_{10}^H X_2.$$

В результате стохастическое нелинейное дифференциальное уравнение второго порядка (1) будет эквивалентно нелинейной системе для математических ожиданий $m_i = M X_i$ ($i = \overline{1, 5}$):

$$\dot{m} = A^m m, \quad m = [m_1 m_2 m_3 m_4 m_5]^T; \quad (9)$$

$$A^m = [A_1^m A_2^m A_3^m A_4^m A_5^m]^T, \quad (10)$$

где

$$A_1^m = m_2; \quad A_2^m = M_0^3;$$

$$M_0^3 = \tilde{M}_{0t} + m_0^{SL} - (\mu_1 + \tilde{\mu}_{1t})m_2 - \mu_1(m_2 m_3 + k_{23}) + (\mu_n + \tilde{\mu}_{nt})F_{n0} + \mu_n F'_{n0} - m_5 + m_5;$$

$$A_3^m = -\alpha_3(m_3 - m_3^*), \quad A_4^m = -\alpha_4(m_4 - m_4^*);$$

$$A_5^m = -\alpha_5 m_5 + \mu_{10}^H m_2,$$

и линейной дифференциальной системе уравнений для центрированных составляющих:

$$\dot{X}_1^0 = X_2^0;$$

$$\dot{X}_2^0 = V_1 + [(\mu_n + \tilde{\mu}_{nt})F_{n1} + \mu_n F'_{n1}] X_1^0 - [(\mu_1 + \tilde{\mu}_{1t}) + \mu_1 m_3 - (\mu_n + \tilde{\mu}_{nt})F_{n2} - \mu_n F'_{n2}] X_2^0 - \mu_1 m_2 X_3^0 + \mu_n F'_{n4} X_4^0 - X_5^0;$$

$$\dot{X}_l^0 = -\alpha_l X_l^0 + \sigma_l \sqrt{2\alpha_l} V_2 \quad (l = 3, 4).$$

Обозначая через ε и β матрицы размерности (5×5) и (2×5)

$$\varepsilon = \varepsilon(t) = [\varepsilon_{ij}]; \quad \beta = [\beta_{ij}],$$

где

$$\varepsilon_{21} = (\mu_n + \tilde{\mu}_{nt})F_{n1} + \mu_n F'_{n1};$$

$$\varepsilon_{22} = -(\mu_1 + \tilde{\mu}_{1t}) - \mu_1 m_3 + (\mu_n + \tilde{\mu}_{nt})F_{n2} + \mu_n F'_{n2};$$

$$\varepsilon_{23} = -\mu_1 m_2; \quad \varepsilon_{24} = \mu_n F'_{n1};$$

$$\varepsilon_{25} = -1; \quad \varepsilon_{33} = -\alpha_3;$$

$$\varepsilon_{44} = -\alpha_4; \quad \varepsilon_{52} = -\mu_{10}^H; \quad \varepsilon_{55} = -\alpha_5;$$

$$\beta_{21} = 1; \quad \beta_{32} = \sigma_3 \sqrt{2\alpha_3}; \quad \beta_{42} = \sigma_4 \sqrt{2\alpha_4},$$

представим уравнения для ковариационной матрицы $K(t) = [k_{ij}(t)]$, $k_{ij}(t) = M [X_i^0(t) X_j^0(t)]$ и матрицы для ковариационных функций $K(t_1, t_2) = [K_{ij}(t_1, t_2)]$, $K_{ij}(t_1, t_2) = M [X_i^0(t_1) X_j^0(t_2)]$ ($i, j = \overline{1, 5}$) в следующем виде [8]:

$$\dot{K}(t) = \varepsilon(t)K(t) + K(t)\varepsilon^T(t) + \beta\nu\beta^T; \quad (11)$$

$$k_{ij}(0) = k_{ij0}; \quad (12)$$

$$\frac{\partial K(t_1, t_2)}{\partial t_2} = K(t_1, t_2)\varepsilon^T(t_2); \quad (13)$$

$$K_{ij}(t, t') = k_{ij}(t), \quad (14)$$

где $\nu = [\nu_{ij}]$ ($\nu_{11} = \nu_1(t)$, $\nu_{12} = \nu_{12}(t)$, $\nu_{22} = 1$) — матрица интенсивностей белых шумов V_1 и V_2 .

Совокупность уравнений (9)–(14) определяет аналитическую квазилинейную гауссовскую корреляционную модель флуктуаций приливной внутригодовой неравномерности вращения Земли при экспоненциальной затухающей эредитарности.

Аналогично выписываются уравнения для эредитарных ядер (5)–(8), а также модулированных частотами f_Γ, ν_m, ν_t эредитарных ядер $\mu_1^H(t, \tau)$.

Усредняя по периодам основных гармоник, отвечающих частотам $f_\Gamma, 2f_\Gamma, \nu_m$ и ν_f для стационарных регулярных и нерегулярных колебаний и постоянных $m_h = m_h^*$ ($h = 2, 3, 4$), из уравнений (9)–(14) находим соответствующие уравнения для математических ожиданий, ковариационной матрицы, матрицы ковариационных функций и матрицы спектральных плотностей.

Уравнения (9)–(14) применимы для нелинейных функций $F_n = F_n(X_1, X_2)$, допускающих угловые точки и даже разрывы. В случае гладких функций F_n рассмотренный квазилинейный метод переходит в метод непосредственной линеаризации в окрестностях математических ожиданий m_1 и m_2 .

Для негауссовских возмущений в (1), как показано в [8], можно воспользоваться методом эквивалентной линеаризации, взяв в качестве осредняющего вероятностного распределения отрезок параметризованного (моментами, квазимоментами, семиинвариантами и др.) разложения одно- и двумерных плотностей.

4 Апостериорные информационные модели эредитарных флуктуаций

Примем за информационные переменные Z_1 и Z_2 , допускающие измерение переменных $X_1 = \delta\varphi$ и $X_2 = \delta\dot{\varphi}$. Положим

$$Z_1 = X_1 + V_3, \quad Z_2 = X_2 + V_4, \quad (15)$$

где V_3 и V_4 — независимые нормальные белые шумы с интенсивностями ν_5 и ν_6 , соответственно. Тогда совокупность уравнений (1)–(4) и (15) будет представлять собой исходную систему уравнений для синтеза фильтра для обработки информации о флуктуациях неравномерности вращения Земли по апостериорным данным, т. е. результатам измерения Z_1 и Z_2 .

Для построения квазилинейного нормального фильтра, согласно [11, 12], перепишем уравнения (1)–(4) и (15) в следующем стандартном виде:

$$\dot{X} = a(X, t) + b(t)\bar{V}_1; \quad Z = a_1(X, t) + \bar{V}_2. \quad (16)$$

Здесь

$$X = [X_1 \dots X_6]^T; \quad Z = [Z_1 Z_2]^T;$$

$$\bar{V}_1 = [V_1 V_2]^T; \quad \bar{V}_2 = [V_3 V_4]^T;$$

$$a = [a_1 \dots a_5]^T$$

$$a_1 = [a_{11} \ a_{12}]^T; \quad a_{11} = X_1; \quad a_{12} = X_2;$$

$$a_2 = \tilde{M}_{0t} - [(\mu_1 + \tilde{\mu}_{1t}) + \mu_1 X_3] X_2 + [(\mu_n + \tilde{\mu}_{nt}) + \mu_n X_4] F_n(X_1, X_2) - X_5;$$

$$a_3 = -\alpha_3(X_3 - m_3^*); \quad a_4 = -\alpha_4(X_4 - m_4^*); \quad (17)$$

$$a_5 = -\alpha_5 X_5 + \mu_{10}^H X_2; \quad (18)$$

$$b = b(t) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \sigma_3 \sqrt{2\alpha_3} \\ 0 & \sigma_4 \sqrt{2\alpha_4} \end{bmatrix};$$

$$\bar{\nu}_1 = \begin{bmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{bmatrix};$$

$$\bar{\nu}_2 = \begin{bmatrix} \nu_3 & 0 \\ 0 & \nu_4 \end{bmatrix}.$$

Заменим (16) статистически линеаризованной системой уравнений, нелинейной относительно математических ожиданий $m^x = [m_1^x \dots m_6^x]^T$, $m^z = [m_1^z \ m_2^z]$ и линейной относительно центрированных составляющих $X^0 = X - m^x$:

$$\dot{m}^x = a_{00}; \quad m^z = a_{10}; \quad (19)$$

$$\dot{X}^0 = a_{01}X^0 + \psi(t)\bar{V}_1; \quad Z^0 = a_{11}X^0 + V_2. \quad (20)$$

Здесь $a_{ij} = a_{ij}(m^x, K^x, t)$ ($i, j = 0, 1$) — коэффициенты статистической линеаризации функций $a = a(X, t)$ и $a_1 = a_1(X, t)$. При этом ковариационная матрица K^x определяется уравнением вида (11):

$$\dot{K}^x = a_{01}K^x + K^x a_{01}^T + b\bar{\nu}_1 b^T. \quad (21)$$

Применяя к (19) и (20) уравнения фильтра Калмана–Бьюси [11], получим

$$\begin{aligned} \dot{\hat{X}} &= a_{00} - a_{01}m^x + a_{01}\hat{X} + \\ &+ Ra_{11}\bar{\nu}_2^{-1}(Z - a_{11}\hat{X} - a_{10} + a_{11}m^x); \end{aligned} \quad (22)$$

$$\hat{X}_0 = MX(t_0); \quad (23)$$

$$\begin{aligned} \dot{R} &= a_{01}R + Ra_{01}^T - Ra_{11}^T\bar{\nu}_2^{-1}a_{11}R + \\ &+ b\bar{\nu}_1 b^T; \end{aligned} \quad (24)$$

$$R_0 = M \left[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T \right]. \quad (25)$$

Совокупность фильтрационных уравнений (22)–(25) при условиях (19) и (21) определяет искомым квазилинейный нормальный фильтр для обработки информации о флуктуациях неравномерности вращения Земли по апостериорным данным, в том числе в темпе получения результатов наблюдений. Коэффициенты статистической линеаризации a_{00} и a_{01} приведены в разд. 3, а $a_{10} = 0$ и $a_{11} = I_2$ в силу линейности второго уравнения (30).

Фильтрационные уравнения (22)–(25) для произвольного линейного наблюдения, когда $a_1(X, t) = b_1(t)X + b_0$, $a_{11} = b_1(t)$, если через β обозначить коэффициент усиления фильтра $\beta = Rb_1(t)^T\bar{\nu}_2^{-1}$, принимают следующий вид:

$$\begin{aligned} \dot{\hat{X}} &= a_{00} - a_{01}m^x + a_{01}\hat{X} + \beta[Z - b_1(t)\hat{X} - \\ &- b_0(t) + b_1(t)m^x], \quad \hat{X}_0 = MX_0; \end{aligned}$$

$$\dot{R} = a_{01}R + Ra_{01}^T - \beta b_1(t)R + b(t)\bar{\nu}_1(t)b(t)^T;$$

$$R_0 = M \left[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T \right].$$

Важно отметить, что коэффициенты статистической линеаризации a_{00} , a_{01} и вспомогательная инструментальная матрица ошибки фильтрации R не содержат результатов наблюдений и могут быть определены отдельно (до получения результатов наблюдений). Таким образом, возможна априорная оценка точности квазилинейного фильтра.

5 Об эквивалентности моделей эредитарных флуктуаций неравномерности вращения Земли

Будем считать известными совместные канонические разложения (КР) возмущений

$$\begin{aligned} X_{0t}^{SL} &= \tilde{M}_0(t) + X^s(t) + X^L(t); \\ X_{1t}^{SL} &= \tilde{\mu}_1(t) + X_1^{\mu S}(t) + X_1^{\mu L}(t). \end{aligned}$$

Запишем их в виде

$$X_{ht}^{SL} = m_h^{SL}(t) + \sum_{\nu} V_{\nu} x_{\nu h}^{SL}(t) \quad (h = \overline{0, n}). \quad (26)$$

Здесь $m_h^{SL}(t)$ — математические ожидания возмущений; V_{ν} — случайные независимые (некоррелированные) коэффициенты с известными распределениями (дисперсиями D_{ν}); $x_{\nu h}^{SL}(t)$ — детерминированные координатные функции. Тогда уравнения (1)–(3) совместно с (4)–(8) для различных эредитарных ядер с помощью методов [9, 10] приводятся к уравнениям вида

$$\dot{X}_1 = X_2; \quad (27)$$

$$\dot{X}_2 = -\mu_1 X_2 + \mathcal{F}(X_1, X_2, X_{ht}^{SL}, X_{lt}^I), \quad (28)$$

где входные возмущения $X_{ht}^{SL}(t)$ определяются (26); $X_{lh}^{SL}(t)$ ($l = 1, 2, \dots$) — инструментальные вспомогательные переменные, удовлетворяющие дифференциальным уравнениям, задаваемым типом эредитарного ядра; $\mathcal{F}(\bullet)$ — в общем случае нелинейная функция отмеченных переменных.

Эффективное значение линейного коэффициента диссипации μ_1^{\exists} (по математическому ожиданию) определяется по формуле

$$\mu_1^{\exists} = \mu_1 + \frac{\partial}{\partial m_2^x} M\mathcal{F}.$$

Эффективное значение линейного коэффициента диссипации (по случайной составляющей) находится из рассмотрений уравнений (27) и (28) для центрированных переменных.

При использовании КР (26) аналогично вычисляются эффективные коэффициенты диссипации по случайным величинам V_{ν} или координатным функциям входных возмущений $x_{\nu h}^{SL}(t)$.

Принцип эквивалентной замены исходной эредитарной стохастической системы на систему с эквивалентной линейной диссипацией позволяет вычислять μ_1^{\exists} не только для различных типов эредитарных ядер, но и для эредитарных нелинейностей

вида $\int_{t-T_0}^t \mu_n^H(t, \tau) F_n(\delta\varphi(\tau), \delta\dot{\varphi}(\tau)) d\tau$. В некоторых практических задачах применяется принцип эквивалентности, связанный с критериями совпадения параметров систематического и флуктуационного дрейфа по переменной $\delta\varphi$ [7].

6 Тестовые примеры

Пример 1. Сохраняя в уравнениях (9)–(12) постоянное возмущение m_0^{SL} , вязкое и эредитарное вязкое трение ($\mu_1 \neq 0$, $\mu_{10}^H \neq 0$, $\alpha_5 \neq 0$), получим уравнения:

$$\dot{m}_1 = m_2; \quad (29)$$

$$\dot{m}_2 = m_0^{SL} - \mu_1 m_2 - m_5; \quad (30)$$

$$\dot{m}_5 = -\alpha_5 m_5 + \mu_{10}^H m_2; \quad (31)$$

$$\dot{K}_{11} = 2K_{12}; \quad \dot{K}_{12} = K_{22} - \mu_1 K_{12} - K_{15}; \quad (32)$$

$$\dot{K}_{15} = K_{25} - \alpha_5 K_{15} + \mu_{10}^H K_{12}; \quad (33)$$

$$\dot{K}_{22} = \nu_1 - 2(\mu_1 K_{22} + K_{25});$$

$$\dot{K}_{25} = -(\mu_1 + \alpha_5) K_{25} - K_{55} + \mu_{10}^H K_{22};$$

$$\dot{K}_{55} = -2(\alpha_5 K_{55} - \mu_{10}^H K_{25}). \quad (34)$$

Отсюда, во-первых, следует, что для переменных X_2 и X_5 имеют место стационарные решения, отвечающие постоянным значениям математических ожиданий

$$\begin{aligned} m_2^* &= m_0^{SL} \mu_{11}^{\exists}; \quad m_5^* = m_2^* \mu_{10}^H \alpha_5^{-1}; \\ \mu_{11}^{\exists} &= \mu_1 + \mu_{10}^H \alpha_5^{-1} \end{aligned} \quad (35)$$

и постоянным значениям дисперсий и ковариаций

$$\begin{aligned} K_{22}^* &= \frac{\nu_1}{2\mu_{12}^{\exists}}; \quad K_{25}^* = \frac{\mu_{10}^H}{(\mu_1 + \alpha_5) + \mu_{10}^H \alpha_5^{-1}}; \\ K_{55}^* &= \mu_{10}^H \alpha_5^{-1} K_{25}^*; \\ \mu_{12}^{\exists} &= \frac{\mu_1(\mu_1 + \alpha_5) + \mu_{10}^H(1 + \mu_1 \alpha_5^{-1})}{(\mu_1 + \alpha_5) + \mu_{10}^H \alpha_5^{-1}}. \end{aligned} \quad (36)$$

Во-вторых, имеет место систематический и флуктуационный дрейф по $X_1 = \delta\varphi$, определяемый уравнениями (29), (32) и (33). Наличие эредитарного вязкого трения приводит к увеличению эффективного вязкого трения как по математическому ожиданию μ_{11}^{\exists} , так и по случайной составляющей μ_{12}^{\exists} . Момент m_0^{SL} определяет величину смещенных колебаний и не влияет на корреляционные характеристики случайных колебаний. При чисто эредитарном вязком трении его эффективные коэффициенты будут определяться формулами

$$\mu_{11}^{\exists} = \frac{\mu_{10}^H}{\alpha_5}, \quad \mu_{12}^{\exists} = \frac{\mu_{10}^H}{\alpha_5 + \mu_{10}^H \alpha_5^{-1}}. \quad (37)$$

Пример 2. Если в условиях примера 1 дополнительно учесть аддитивный полигармонический возмущающий момент \tilde{M}_{0t} , определяемый формулой (2), то уравнения (29)–(31) примут следующий вид:

$$\begin{aligned} \dot{m}_1 &= m_2; \\ \dot{m}_2 &= -\mu_1 m_2 - m_5 + \tilde{M}_{0t}; \\ \dot{m}_5 &= -\alpha_5 m_5 + \mu_{10}^H m_2 \end{aligned}$$

а уравнения (32)–(34) останутся без изменения. Амплитудно-частотные характеристики вынужденных полигармонических колебаний находятся по известным формулам стационарных линейных дифференциальных систем [8]. По переменной $X_1 = \delta\varphi$ возможно накопление систематической ошибки, зависящее от амплитудно-частотных характеристик. При этом могут быть использованы формулы (37) и результаты [7].

Пример 3. Разработанные в [7] тестовые примеры 3–10 были использованы для моделирования переменных с учетом (35) и (36).

7 Заключение

Разработанные методы построения информационных моделей эредитарных флуктуаций неравномерности вращения Земли по априорным и апостериорным данным реализованы в виде экспериментального программного обеспечения в среде MATLAB. Проведено тестирование программного обеспечения на десяти примерах. Разработаны критерии эквивалентности эредитарных и неэредитарных моделей.

Методы, алгоритмы, программное обеспечение и тестовые примеры включены в состав информационных ресурсов по фундаментальной проблеме «Статистическая динамика вращения Земли».

Квазилинейные методы, как показали вычислительные эксперименты и сравнение с результатами статистического моделирования, обеспечивают высокую точность фильтрации скорости $\delta\dot{\varphi}$ (примерно 2%–3% для априорной стационарной информации и 0,5%–1% при апостериорной информации). Из-за отсутствия возвращающей силы по $\delta\varphi$ появляются дрейфы и накапливающиеся ошибки. Поэтому так необходимы точные измерения $\delta\varphi$.

Направления дальнейших исследований:

- (1) учет влияния автокоррелированности эредитарных ядер различных возмущений в уравнениях (1);
- (2) оценка негауссовости распределений возмущения для оценок больших уклонений по $\delta\varphi$;

- (3) оценивание и распознавание возмущений в уравнениях (1) на основе апостериорной информации от нескольких нелинейных измерительных систем различной точности;
- (4) разработка комплексных статистических моделей вращения Земли, учитывающих эредитарные флуктуации полюса и неравномерности вращения Земли.

Автор благодарен Н. Н. Семендяеву за выполненные вычислительные эксперименты.

Литература

1. Марков Ю. Г., Сеницын И. Н. Спектрально-корреляционные модели флуктуаций вращательного движения Земли // ДАН, 2003. Т. 393. № 5. С. 618–623.
2. Марков Ю. Г., Сеницын И. Н. Спектрально-корреляционные и кинетические модели движения Земли // Астрон. журнал, 2004. Т. 81, № 2. С. 184–192.
3. Сеницын И. Н. Стохастические модели флуктуаций движения Земли в условиях пуассоновских возмущений // Системы и средства информатики. Спец. вып. «Геоинформационные технологии». — М.: ИПИ РАН, 2004. С. 39–55.
4. Марков Ю. Г., Дасаев Р. Р., Перепёлкин В. В., Сеницын И. Н., Сеницын В. И. Стохастические модели вращения Земли с учетом влияния Луны и планет // Космические исследования, 2005. Т. 43. № 1. С. 54–66.
5. Акуленко Л. Д., Марков Ю. Г., Перепёлкин В. В. Внутригодовые неравномерности вращения Земли // ДАН, 2007. Т. 417. № 2. С. 118–121.
6. Марков Ю. Г., Сеницын И. Н. Корреляционная модель приливной неравномерности вращения Земли // ДАН, 2008. Т. 49. № 3. С. 338–341.
7. Сеницын И. Н. Квазилинейные методы построения информационных моделей флуктуаций неравномерности вращения Земли // Информатика и её применения, 2008. Т. 2. Вып. 1. С. 35–43.
8. Пугачёв В. С., Сеницын И. Н. Теория стохастических систем. 2-е изд. — М.: Логос, 2004.
9. Sinitsyn I. N. Stochastic hereditary control systems // Пробл. упр. и теория информации, 1986. Vol. 15. No. 4. P. 287–298.
10. Пугачёв В. С., Сеницын И. Н. Стохастические дифференциальные системы. Анализ и фильтрация. 2-е изд. — М.: Наука, 1990.
11. Сеницын И. Н. Фильтры Калмана и Пугачёва. 2-е изд. — М.: Логос, 2007.
12. Сеницын И. Н. Развитие теории фильтров Пугачёва для оперативной обработки информации в стохастических системах // Информатика и её применения, 2007. Т. 1. Вып. 1. С. 3–13.

ПРОБЛЕМЫ РАСПАРАЛЛЕЛИВАНИЯ БИОМЕТРИЧЕСКИХ ВЫЧИСЛЕНИЙ В КРУПНОМАСШТАБНЫХ ИДЕНТИФИКАЦИОННЫХ СИСТЕМАХ

О. С. Ушмаев¹

Аннотация: Рассматриваются проблемы распараллеливания биометрических вычислений в крупномасштабных идентификационных системах. Предлагается методика вычисления производительности биометрического кластера, методы определения необходимой мощности вычислительных средств. Разрабатываются методы оптимизации и распараллеливания вычислений в задаче мультибиометрической идентификации.

Ключевые слова: мультибиометрическая идентификация; параллельные вычисления

1 Введение

К настоящему времени накоплен значительный опыт создания крупномасштабных биометрических систем. Большинство из них являются автоматизированными дактилоскопическими системами (АДИС). Выбор отпечатков пальцев обусловлен множеством факторов. Основным из них является традиционное использование отпечатков пальцев в криминальном учете и высокий потенциал дактилоскопии с точки зрения ошибок распознавания. Для практически безошибочного распознавания людей в масштабах населения страны достаточно использования всех десяти отпечатков пальцев. При использовании меньшего числа отпечатков соотношение ошибок 1-го и 2-го родов является удовлетворительным во многих приложениях [1, 2]. Однако использование дактилоскопии имеет недостатки. Накопленный опыт реализации АДИС для криминалистической и гражданской идентификации позволяет выделить следующие наиболее критичные направления развития [3–6]:

- идентификация людей, не обладающих пригодными для распознавания отпечатками пальцев (инвалиды, плохое состояние кожи);
- увеличение производительности.

Первая задача возникает из непосредственных требований к биометрическим системам, а именно: она должна автоматически идентифицировать личность по предъявляемым биометрическим образцам. Соответственно люди с плохим качеством отпечатков пальцев не могут надежно идентифицироваться средствами АДИС.

Вторая задача — увеличение производительности — связана с тем, что при создании крупномасштабных биометрических систем более 75% средств затрачиваются на аппаратные средства вычислительных узлов [7]. Уже в настоящее время потребность в вычислительных мощностях современных АДИС достигает тысяч серверов, что требует параллельных вычислений и суперкомпьютерных мощностей. При полномасштабном внедрении систем гражданской идентификации, таких как паспортно-визовые документы нового поколения, биометрические массивы вырастут многократно, поэтому проблемы увеличения производительности будут становиться все более актуальными. В такой ситуации первоочередной проблемой становится организация вычислительного кластера биометрической системы и параллельных вычислений [2, 7–9].

Проблемы организации параллельных вычислений усугубляются тем, что многие перспективные крупномасштабные биометрические системы планируются как мультибиометрические с привлечением одновременно нескольких биометриков, отличных по своим эксплуатационным показателям и используемым вычислительным средствам. В [2] показано, что привлечение нескольких биометриков может увеличить скорость идентификации при правильной организации вычислений. В то же время это создает дополнительные трудности при проектировании. В частности, определение необходимой производительности вычислительных средств мультибиометрической системы становится довольно сложной методологической проблемой.

¹Институт проблем информатики Российской академии наук, oushmaev@ipiran.ru

Решение проблемы параллельных вычислений в мультибиометрических идентификационных системах требует решения следующих задач:

- выработка методологии оценки производительности биометрического кластера;
- определение потребности в вычислительных средствах в зависимости от состава биометрического программного обеспечения;
- организация параллельных вычислений.

В статье рассмотрена проблема параллельных вычислений в мультибиометрической системе. В настоящем разделе поставлена задача. В разд. 2 дана методология оценки производительности биометрической идентификации. Особенности оценки производительности при мультибиометрической идентификации представлены в разд. 3. Раздел 4 посвящен определению потребности в вычислительных средствах. В разд. 5 изложены подходы к организации параллельных вычислений. Пример применения разработанных методов на примере идентификационной системы, использующей отпечатки пальцев и изображение лица, представлены в разд. 6. Основные выводы и рекомендации приведены в заключении.

2 Оценка производительности биометрического кластера

Основным показателем производительности биометрической системы является проектное время ожидания отклика. Для крупномасштабных систем, таких как паспортно-визовые системы или системы криминального учета, время отклика обычно устанавливается в 24 ч для суточного цикла (в редких случаях 168 ч для недельного цикла) функционирования системы [10]. Биометрическая

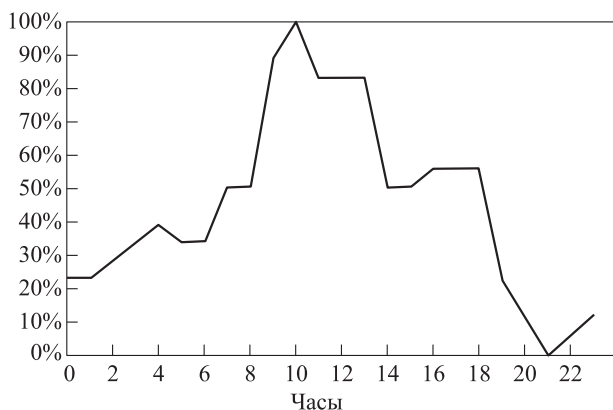


Рис. 1 Пример графика потока заявок в биометрическую систему

система должна успевать обрабатывать заявки на идентификацию, поступающие в течение суток или недели соответственно. Отказ от реального времени в пользу суточного цикла обычно связан с тем, что поток заявок имеет прогнозируемую неравномерную структуру. Примерный вид графика интенсивности запроса приведен на рис. 1. Максимальный участок соответствует времени, когда функционирует большинство пунктов сбора биометрической информации во всех часовых поясах. С точки зрения экономии ресурсов целесообразно производить вычисления равномерно в течение суток, а не согласно графику поступления.

В таком случае проектное время обработки заявок, поступающих в течение суток, рассчитывается по следующей формуле:

$$T_{\text{сут}} = r_{\text{сут}} t, \tag{1}$$

где $r_{\text{сут}}$ — максимальный проектный поток заявок в течение суток; t — среднее время идентификации по биометрической базе.

В случае АДИС (или другой однофакторной биометрии) время t обычно линейно пропорционально количеству записей в базе данных (БД), поскольку в ходе идентификации предъявляемые образцы последовательно сравниваются с каждым хранимым, т. е. в большинстве приложений

$$t = \frac{N t_{\text{ср}}}{W}, \tag{2}$$

где N — число записей в БД; $t_{\text{ср}}$ — среднее время сравнения пары биометрических образцов на единицу мощности вычислительных средств (с); W — мощность вычислительных средств, нормированная на единичную номинальную мощность (например, на один процессор с тактовой частотой 1 ГГц).

Формула (2) верна, если пренебречь потерями, связанными с распараллеливанием вычислений.

Чтобы система идентификации справлялась с плановым потоком заявок в течение суток, накладывается ограничение $T_{\text{сут}} < 24$ ч. Резерв R системы (избыточность) определяется как

$$R = \frac{24 - T_{\text{сут}}}{24} = 1 - \frac{N t_{\text{ср}}}{W} \frac{r_{\text{сут}}}{24} = 1 - r \frac{N t_{\text{ср}}}{W} = 1 - r t,$$

где r — интенсивность потока заявок.

Избыточность в основном необходима в следующих случаях:

- сезонные колебания и резкие скачки нагрузки на биометрические серверы;

- выход из строя части вычислительных мощностей;
- плановый профилактический вывод из эксплуатации части вычислительных мощностей;
- сбой системы, приводящий к необходимости повторной обработки запросов.

В случае систем оперативной идентификации, где время ожидания ограничено минутами, помимо средней способности системы обеспечить обработку потока заявок требуется ограничить дисперсию времени ожидания, чтобы в моменты пиковой загрузки вычислительных мощностей проектное время ожидания не было превышено. В таком случае требуется проводить специальное исследование структуры потока заявок, чтобы определить максимально возможную интенсивность.

Как видно из приведенной формулы (2), скорость сравнения биометрических образцов является значимым фактором при определении производительности.

Вторым ограничивающим фактором является объем адресуемой памяти. Как показывает анализ современных тенденций развития аппаратных средств, темпы прироста мощности вычислительных средств опережают темпы прироста скорости доступа к оперативной памяти и жестким дискам. В такой ситуации становится критичным организация доступа к памяти. В крупномасштабных системах, где востребованы отпечатки пальцев, радужная оболочка глаза или изображение лица, единственным способом быстрого сравнения является кэширование биометрической базы в оперативной памяти. Хранение на жестком диске нежелательно из-за низкой скорости доступа, которая многократно превышает время сравнения t_{cp} . Таким образом,

получаем, что на одном сервере биометрического кластера можно разместить V/s_{tpl} биометрических шаблонов, где V — размер доступной оперативной памяти, s_{tpl} — размер биометрического шаблона.

Для 32-битной версии максимальное значение V составляет 4 Гб. Реально за вычетом памяти, занятой вычислительным процессом и операционной системой (ОС), остается порядка 3,5–3,7 Гб. В случае 64-битной ОС максимальное адресуемое пространство составляет 8 Гб.

В случае одномодальной системы единственно возможной конфигурацией кластера, реализующей эффективное распараллеливание, является распределение биометрической базы на нескольких узлах.

3 Мультибиометрическая идентификация

При мультибиометрической идентификации оценка времени сравнения (2) является отдельной задачей. При привлечении нескольких технологий биометрической идентификации возможны несколько вариантов реализации мультибиометрической системы, позволяющих корректировать эксплуатационные показатели.

В качестве базового примера рассмотрим бимодальные технологии, использующие две биометрики. В таком случае возможны две основные схемы интеграции:

- (1) процессы сравнения независимы (рис. 2);
- (2) процессы сравнения зависимы (рис. 3).

В [10–15] разработана методология интеграции биометрических систем в случае независи-

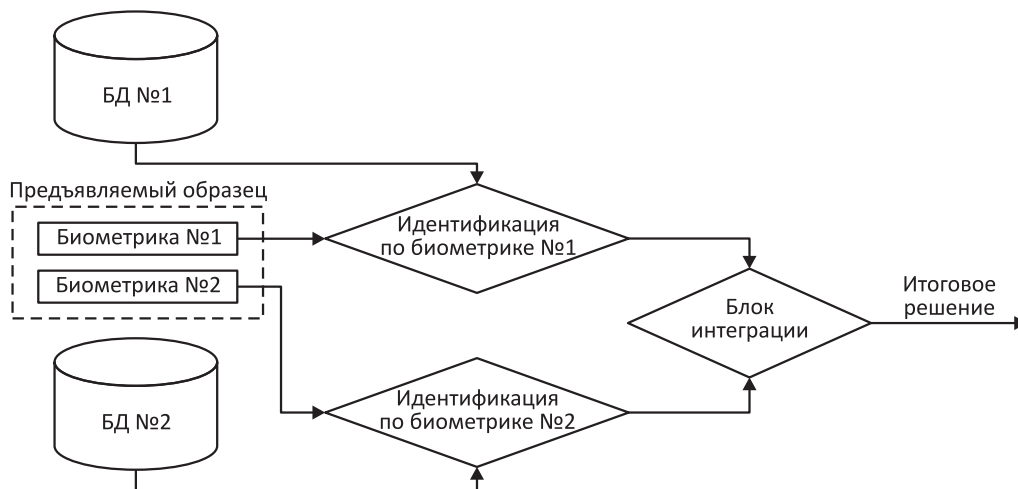


Рис. 2 Независимые процессы идентификации

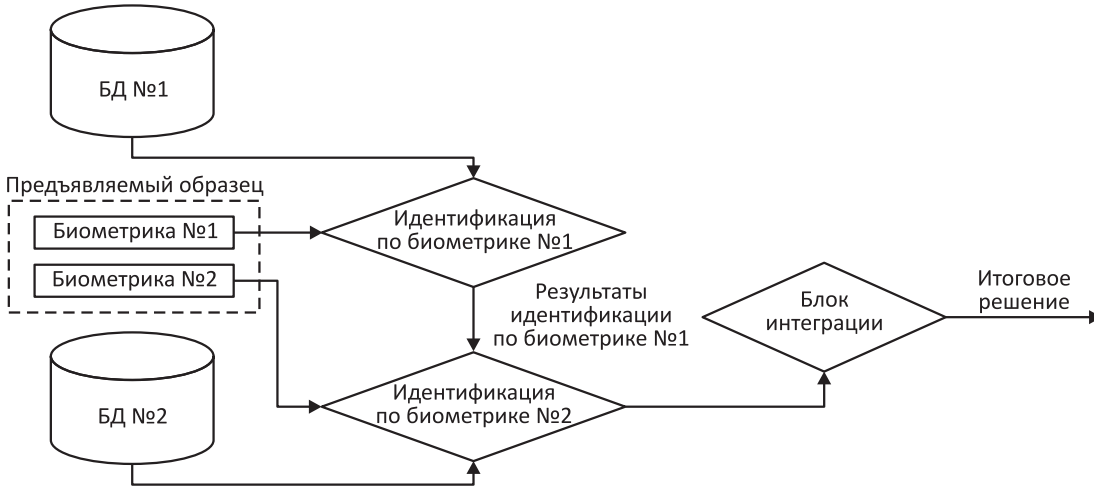


Рис. 3 Зависимые процессы идентификации

мого сравнения. Тогда достигаются минимально возможные ошибки распознавания. Однако это приводит к потерям производительности. Несложно показать, что время сравнения такой системы получается суммированием времени сравнения по каждой биометрике.

Во многих задачах, решаемых современными АДИС, качество идентификации является приемлемым в отличие от производительности, которая остается достаточно низкой. Поэтому далее в разделе мы сосредоточим основное внимание на схеме интеграции с зависимыми процессами идентификации, которая позволяет достичь прироста в производительности с приемлемыми потерями в качестве распознавания относительно независимого процесса сравнения.

Рассмотрим реализацию идентификации с зависимыми процессами сравнения более детально. На вход функции одномодальной биометрической идентификации поступают предъявляемая биометрическая запись или образец и биометрическая БД, или линейный список биометрических записей. На выходе мы получаем меры сходства предъявляемой и хранимых в БД записей. На основе этой информации принимается решение, принадлежат ли записи одному человеку или нет. Большинство систем идентификации по отпечаткам пальцев и изображению лица используют пороговые методы принятия решения.

При идентификации по двум биометрикам увеличения производительности можно достичь, если по результатам идентификации по более быстрой из двух технологий принимать решение о целесообразности дальнейшего поиска по второй биометрике (схема реализации функции сравнения приведена на рис. 4).

Как видно из рис. 4, в такой функции сравнения биометрических образцов есть четыре терминальных состояния:

- 2 — при сравнении по первой биометрике принято решение об идентичности образцов (Accept), так как результат сравнения m_1 превышает определенный порог A_1 ;
- 4 — при сравнении по первой биометрике принято решение о различности образцов (Reject), так как мера сходства m_1 меньше некоторого минимального порога R_1 ;
- 5 — после сравнения по второй биометрике принято решение об идентичности образцов, что суммарная мера сходства m_2 больше порога заданного порога A_2 , проблемы построения интегральной меры сходства при мультибиометрической идентификации изложены в [11–15];
- 6 — по результатам сравнения принято решение о различности образцов.

Оценим статистические характеристики временных показателей выполнения функции мультибиометрического сравнения. Следует разделить следующие два случая:

- (1) образцы принадлежат одному человеку (обозначим время сравнения t^g);
- (2) образцы принадлежат разным людям (t^i).

В первом случае вероятности события $m_1 \geq A_1$ и вероятность $m_1 < R_1$ являются стандартными показателями качества распознавания и обозначаются $TAR(A_1)$, True Acceptance Rate, и $FRR(R_1) = 1 - TAR(R_1)$, False Rejection Rate, ошибка первого рода. Среднее время сравнения «своих» будет

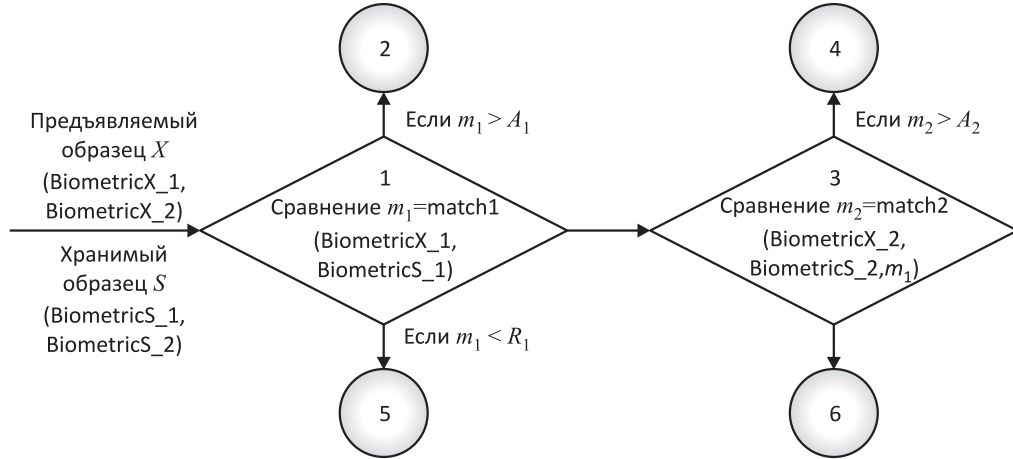


Рис. 4 Реализации функции мультибиометрического сравнения

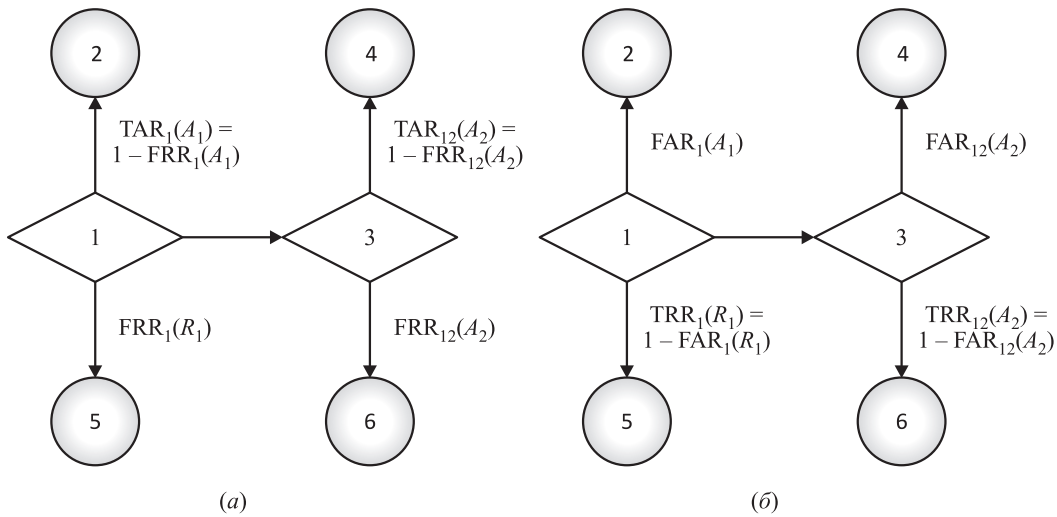


Рис. 5 Вероятности перехода: (а) свои сравнения; (б) чужие сравнения

складываться из двух слагаемых: времени выполнения сравнения по первой биометрике и времени сравнения по второй биометрике. Причем сравнение отпечатков пальцев будет проводиться только в случае $R_1 \leq m_1 < A_1$. Вероятности переходов схематично представлены на рис. 5.

Соответственно среднее суммарное время определяется по формуле:

$$t^g = t_1^g + (1 - TAR_1(A_1) - FRR_1(R_1))t_2^g = t_1^g + (FRR_1(A_1) - FRR_1(R_1))t_2^g, \quad (3)$$

где t_1^g — среднее время выполнения «своих» сравнений для первой биометрики; t_2^g — среднее время выполнения «своих» сравнений для второй биометрики.

Во втором случае вероятности $m_1 \geq A_1$ и $m_1 < R_1$ выражаются аналогичным образом через

ошибку второго рода $FAR(A_1)$, False Acceptance Rate — ошибка второго рода, и $TRR(R_1) = 1 - FAR(R_1)$, True Rejection Rate. Среднее суммарное время определяется как

$$t^i = t_1^i + (1 - FAR_1(A_1) - TRR_1(R_1))t_2^i = t_1^i + (FAR_1(R_1) - FAR_1(A_1))t_2^i, \quad (4)$$

где t_1^i — среднее время выполнения «чужих» сравнений для изображения лица; t_2^i — среднее время выполнения «чужих» сравнений для отпечатков пальцев.

Таким образом, получаем, что время сравнения (4) очевидно меньше времени сравнения при независимых процессах сравнения. Более того, при определенных соотношениях t_1^i и t_2^i время (4) меньше времени t_2^i сравнения по одной биометрике. Данный результат может применяться, например,

для ускорения дактилоскопической идентификации [2].

4 Определение потребности в вычислительных средствах

В задаче массовой идентификации, когда размер базы достаточно велик, одного сервера или ПЭВМ недостаточно. Поэтому требуется решение задачи определения вычислительной мощности.

Массовость сравнения позволяет при оценке производительности в значительной степени ориентироваться на формулы (3) и (4), так как при сравнении по большой базе входящие в формулу вероятности дают достаточно точную оценку времени идентификации. При выполнении операции массовой идентификации доминируют операции сравнения «чужих». Так, если в базе зарегистрированы по одному образцу для N субъектов, то в процессе идентификации будет выполнено N «чужих» и 1 «свое» сравнение. При большом N доля времени t^g в процессе идентификации ничтожно мала, ей можно пренебречь. Таким образом из формулы (1) получаем, что требуемая мощность вычислительных средств составляет

$$W = \frac{Nt_{cp}}{t}. \quad (5)$$

С точки зрения производительности схема интеграции с зависимыми процессами сравнения заведомо медленнее одномодальной биометрической системы на базе первой технологии и заведомо быстрее бимодальной с независимыми процессами сравнения. Как было отмечено выше, при определенных условиях бимодальная система с зависимыми процессами сравнения может быть производительней одномодальной с использованием второй биометрики (важный пример: интеграция в АДИС изображения лица). Если дополнительная биометрика дает существенный выигрыш в производительности, то ее использование целесообразно. Если выигрыша нет, то требуются дополнительные аргументы в пользу мультибиометрии.

Для оценки выигрыша в производительности введем обозначения для относительного изменения времени сравнения $a(A_1, R_1)$. На одном вычислительном узле относительное изменение производительности при зависимых процессах идентификации с выбранными порогами A_1 и R_1 можно вычислить как отношение (4) к времени идентификации $t_1^i + t_2^i$ при независимых процессах идентификации:

$$a(A_1, R_1) = \frac{t^i}{t_1^i + t_2^i} = \frac{t_1^i + (\text{FAR}_1(R_1) - \text{FAR}_1(A_1))t_2^i}{t_1^i + t_2^i}. \quad (6)$$

Значение $a(A_1, R_1)$ изменяется в пределах от $t_1^i/(t_1^i + t_2^i)$ до 1.

При измерении ускорения относительно одномодальной системы выражение (6) модифицируется следующим образом:

$$a(A_1, R_1) = \frac{t^i}{t_2^i} = \frac{t_1^i}{t_2^i} + (\text{FAR}_1(R_1) - \text{FAR}_1(A_1)). \quad (7)$$

Выражения (6) и (7) применимы, когда процессы физически функционируют на одних и тех же вычислительных средствах и выполняются последовательно. Но во многих случаях для различных биометрик потребуются отдельные вычислительные средства (обозначим их долю через g_i). В таком случае можно считать, что в продолжительном интервале времени процессы идентификации исполняются параллельно со средним временем сравнения t_1^i/g_1 для первой биометрики и t_2^i/g_2 для второй в случае независимых процессов сравнения. Среднее время выполнения определяется самым медленным звеном, т.е. итоговое среднее время сравнения равно $\max(t_1^i/g_1, t_2^i/g_2)$. При зависимых процессах на вторую биометрику приходится меньшая вычислительная нагрузка, поскольку до сравнения по второй биометрике доходит в среднем только $(\text{FAR}_1(R_1) - \text{FAR}_1(A_1))$ образцов.

Распределение вычислительных мощностей g_1 и g_2 (балансировка кластера) для каждой биометрики может производиться из следующего основного соображения. Суммарное время выполнения идентификации по каждой из биометрик должно быть одинаковым на достаточно длинном интервале (например, сутки). Если одна из биометрик систематически обрабатывает быстрее, то часть ее мощностей целесообразно отдать второй.

При независимых процессах сравнения вычислительные мощности находятся из соотношений (первое уравнение является условием равной загрузки биометрик, второе — условием того, что все мощности распределены):

$$\begin{cases} \frac{t_1^i}{g_1} - \frac{t_2^i}{g_2} = 0; \\ g_1 + g_2 = 1. \end{cases} \quad (8)$$

В таком случае получаем

$$g_1 = \frac{t_1^i}{t_1^i + t_2^i}; \quad g_2 = \frac{t_2^i}{t_1^i + t_2^i}. \quad (9)$$

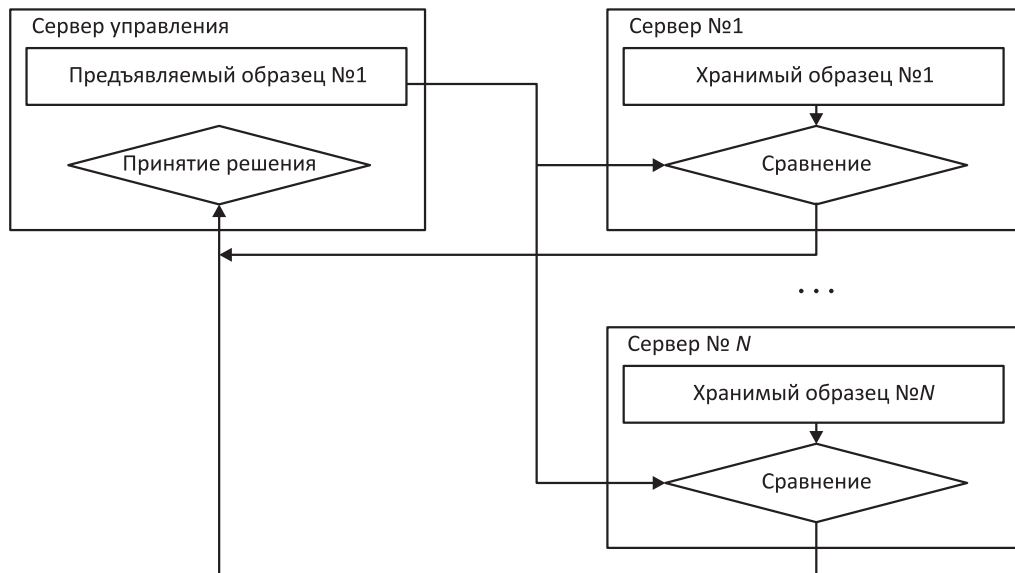


Рис. 6 Распараллеливание биометрических вычислений

При зависимых процессах распределение вычислительных мощностей определяется как

$$\begin{cases} \frac{t_1^i}{g_1} - (\text{FAR}_1(R_1) - \text{FAR}_1(A_1)) \frac{t_2^i}{g_2} = 0; \\ g_1 + g_2 = 1. \end{cases} \quad (10)$$

Следовательно,

$$g_1 = \frac{t_1^i}{t_1^i + (\text{FAR}_1(R_1) - \text{FAR}_1(A_1))t_2^i};$$

$$g_2 = \frac{(\text{FAR}_1(R_1) - \text{FAR}_1(A_1))t_2^i}{t_1^i + (\text{FAR}_1(R_1) - \text{FAR}_1(A_1))t_2^i}.$$

Изменение производительности при параллельной организации вычислений в точности соответствует формулам (6) и (7), выведенных в предположении о функционировании двух биометрик на одних аппаратных средствах. При дефиците памяти распределение различных биометрик по различным аппаратным средствам является предпочтительным. Соответственно при определении вычислительной нагрузки следует руководствоваться следующим алгоритмом:

- вычисление необходимого времени идентификации на основе размера базы, интенсивности потока заявок и требуемой избыточности методами разд. 2 с учетом поправок (5) и (6) на производительность;
- вычисление доли серверов, специализирующихся на каждом из биометрических методов, по формулам (7)–(10).

При большем числе биометрик следует руководствоваться тем же принципом равной нагрузки.

5 Конфигурация кластера

При значительном числе зарегистрированных пользователей биометрической системе необходимо более одного сервера. Это требует решения проблемы распараллеливания вычислений.

Возможны два принципиально методологически отличных подхода к данной проблеме. Первый заключается в распараллеливании операций биометрической идентификации. В таком случае организацию параллельных вычислений можно переложить на штатные средства кластера. Однако такой подход практически не используется. Во-первых, качество распараллеливания отдельных операций вряд ли будет высоким. Во-вторых, биометрическая идентификация является процессом с высокой внутренней степенью параллелизма, который следует максимально использовать.

Второй подход к организации параллельных вычислений заключается в комплектовании биометрической системы специализированными средствами управления параллельными вычислениями.

Базовая идея организации параллельных вычислений приведена на рис. 6. В самой простой ситуации каждый сервер можно ассоциировать с одним хранимым биометрическим образцом. Тогда идентификация будет осуществляться следующим образом. Предъявляемый образец сравнивается с каждым хранимым на отдельном сервере. В таком случае без учета потерь на передачу время идентификации равно времени сравнения пары образцов.

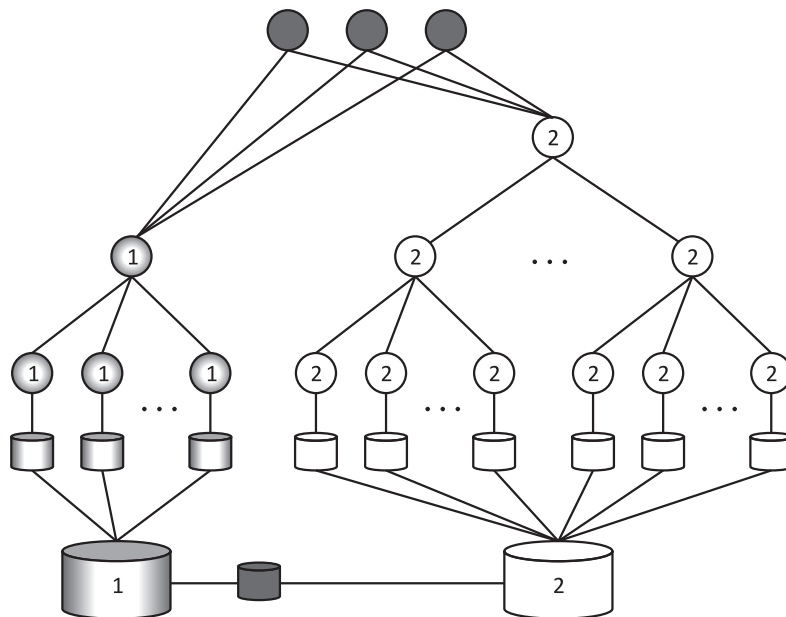


Рис. 7 Топология мультибиометрического кластера (вариант отдельных вертикалей)

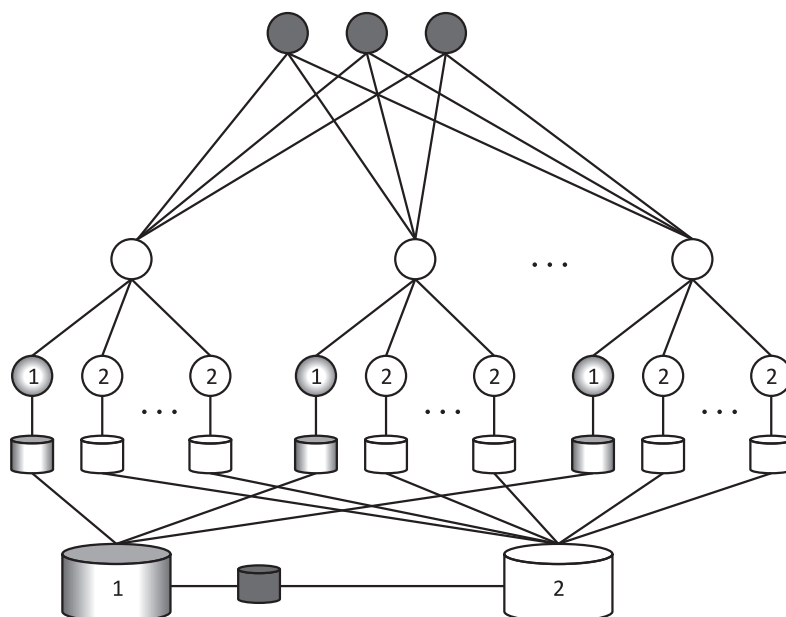


Рис. 8 Топология мультибиометрического кластера (вариант репликации)

В общем случае при n серверах на каждом можно размещать $1/n$ всей базы.

Помимо вычислительных серверов и серверов управления в конфигурации кластера требуется предусмотреть специализированные аппаратные средства для мониторинга и управления функционированием, хранения биометрической базы, технологические серверы (доменные контроллеры,

серверы авторизации) и пр. Определение требований и потребностей в данных категориях выходят за рамки данной статьи.

С точки зрения организации вычислительного кластера можно выделить две основные топологии кластера: организация нескольких независимых вертикалей (рис. 7) и репликация мультибиометрических кластеров меньшего размера (рис. 8).

Таблица 1 Эксплуатационные показатели современных биометрических технологий

Технология	Скорость сравнения t_{cp}^{-1*} , сравнений/с	Размер шаблона, байт	Размер образца, КБ	Время создания шаблона*, с
Отпечатки пальцев				
Biolink ATK	9 000	600–3000	100	0,2
Neurotechnologija MegaMatcher	10 000	600–3000		0,3
Изображение лица				
Neurotechnologija VeriLook	60 000	3000	200	0,6
Радужная оболочка глаза				
Biolink Idenium	200 000	3400	150	0,1
Голос				
ЦРТ Трал	1	55 296	N/A	5

* В расчете на одно ядро тактовой частотой 1 ГГц.

С точки зрения требований к функционированию каждый из вариантов имеет свои преимущества и недостатки. Отдельные вертикали более удобны в профилактике и модернизации и больше подходят для систем с независимыми процессами сравнения. Объединение более аналогичных систем в рамках большей более перспективно в задачах, связанных с интеграцией биометрической информации из различных ведомств или территорий.

При применении обеих перечисленных топологий в задаче идентификации с зависимыми процессами сравнения возникает значимая проблема с балансировкой нагрузки, так как отобранные при сравнении первой биометрики образцы необязательно будут равномерно распределены по аппаратным средствам второй биометрики. В таком случае может потребоваться дополнительная избыточность.

6 Пример использования

В качестве примера рассмотрим часть процесса проектирования системы мультибиометрической идентификации по отпечаткам пальцев и изображению лица. Данный пример имеет важное практическое значение, потому что такая комбинация биометрических идентификаторов является основой многих современных высокопроизводительных биометрических систем. Как видно из табл. 1, идентификация по изображению лица является более быстрой по сравнению даже с одним отпечатком пальца, не говоря уже о более трудоемких четырех- и десятипальцевых дактокартах. В то же время отпечатки пальцев являются более надежной биометрией с точки зрения качества идентификации. Та-

кое соотношение производительность–надежность идентификации делает привлекательным совместное использование этих биометрик. При этом распознавание по изображению лица будет первой технологией предварительного поиска, окончательное решение принимается на основе идентификации по отпечаткам пальцев.

Схематично идентификация представлена на рис. 9. Размеры A_1, R_1, A_3, R_3 определяются вероятностями перехода в соответствующие состояния графа (рис. 9). Априори, в связи с высокой надежностью бимодальной биометрии по дактокарте и изображению лица можно считать, что переходы из узла 3 выполняются без ошибок, и основной целью усиления АДИС лицевой биометрией является увеличение производительности.

Учитывая реальные требования к современным системам биометрической идентификации вероятность перехода в узел 2 (рис. 9) будет очень редким событием, поскольку на ошибку второго рода накладываются серьезные ограничения $FAR_{face}(A_1) < 10^{-3} \div 10^{-8}$, из которого определяется допустимое значение порога A_1 . Соответственно в чужих сравнениях вероятность перехода практически нулевая и не влияет на среднее время. Поэтому следует отметить, что в большинстве приложений при оптимизации производительности придется отказаться от возможности принятия положительного решения на основе идентификации только по изображению лица, так как прирост минимальный, а риски неправильной оценки FAR_{face} высоки. В то же время порог R_1 будет регулироваться другим ограничением: максимально допустимым уровнем ошибки 1-го рода, $FRR_{face}(R_1) \approx 0$.

В качестве конкретного примера определения конфигурации серверов выберем следующие пока-

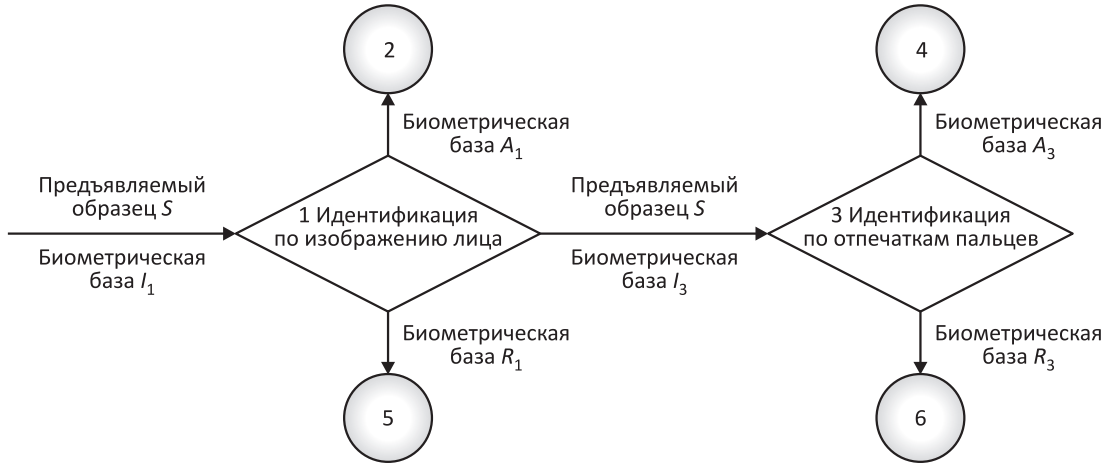


Рис. 9 Идентификация по лицу и отпечаткам пальцев

Таблица 2 Модельные параметры мультибиометрического кластера

Параметр	Значение
Размер шаблона дактокарты	20 Кб
Скорость сравнения дактокарт	5000 сравнений/с
Размер шаблона изображения лица	2 Кб
Скорость сравнения изображения лица	250 000 сравнений/с
Размер базы (записей)	20 000 000
Суточный поток заявок с учетом избыточности	60 000
$FAR_{face}(R_1) - FAR_{face}(A_1)$	10%

затели стационарного режима¹ (табл. 2), с параметрами российских биометрических технологий Vi-olink (отпечатки пальцев) и ИТ (изображение лица).

Получаем, что всего за сутки требуется произвести $60 \cdot 10^3 \cdot 20 \cdot 10^6 = 1,2 \cdot 10^{12}$ сравнений мультибиометрических образцов. Среднее время сравнения

$$t_{face}^i + t_{finger}^i (FAR_{face}(R_1) - FAR_{face}(A_1)) = \frac{1}{250\,000} + 0,10 \frac{1}{5000} = \frac{1 + 0,10 \cdot 50}{250\,000} = \frac{6}{250\,000}.$$

Скорость сравнения равна примерно 41 600 сравнений/с (или $3,6 \cdot 10^9$ сравнений/сут), что значительно выше скорости сравнения по отпечаткам пальцев. Суммарная потребность в вычислительных мощностях равна: $1,2 \cdot 10^{12} / (3,6 \cdot 10^9) = 0,333 \cdot 10^3 \sim 333$ ядра, при четырехъядерном процессоре — 84 сервера.

На лицо приходится 1/6 вычислительных ресурсов, на отпечаток пальца — 5/6. Всего 14 лицевых серверов и 70 дактилоскопических. Суммарный объем адресуемой памяти лицевых серверов не менее 100 Гб, дакто — 504 Гб (при 64-битной адре-

сацией). Общая потребность в памяти: 40 Гб на лицевую биометрию и 400 Гб на дактилоскопию (в частности, лицевая биометрия допускает использование 32-битных версий).

При таком составе вычислительного узла возможны две крайние конфигурации: репликация минимальной конфигурации 1 лицевой сервер и 5 дактосерверов (всего 14 комплектов) или две вертикали: 14 лицевых серверов и 70 дактосерверов. Данные конфигурации различаются с точки зрения организации вычислительного кластера. При этом с точки зрения балансировки нагрузки они идентичны. Плановая загрузка вычислительных мощностей при максимальной нагрузке в 60 000 запросов/сут составляет 99,2%. Данное значение получается как соотношение плановой нагрузки в $1,2 \cdot 10^{12}$ сравнений/сут и максимальной производительности лицевого кластера ($1,2096 \cdot 10^{12}$ сравнений/с) и дактилоскопического кластера ($1,2096 \times 10^{11}$). Дактилоскопический кластер в приведенной конфигурации планово выполняет только 10% сравнений, оставшихся после лицевой биометрии, т. е. всего $1,2 \cdot 10^{11}$ сравнений/сут.

¹Режим, в котором размер базы и поток заявок достигли своего максимального штатного значения.

Столь высокая нагрузка в модельном примере определяется оптимальным соотношением памяти и производительности. При снижении плановой загрузки до 30 000 запросов/сут произойдет эффект избытка памяти. При размере базы в 20 млн записей необходимо иметь минимум 56 серверов дактобиометрии для поддержки необходимых 400 ГБ адресуемой памяти. В таком случае конфигурация кластера будет далека от оптимальной (7 лицевых серверов и 56 дактилоскопических). Плановая нагрузка составит 99% для лицевых серверов и 62% для дактилоскопических соответственно.

7 Заключение

В статье выработан подход к распараллеливанию вычислений в мультибиометрической системе. В качестве преимуществ данного подхода можно отметить следующее:

- эффективное распараллеливание биометрических вычислений;
- разделение вычислительных мощностей между различными биометриками;
- оптимизация производительности мультибиометрического кластера;
- возможность регулирования соотношения производительность — ошибка первого рода;
- возможность оптимального выбора конфигурации кластера в случае распределенных вычислений.

С практической точки зрения разработанный подход позволяет проводить быструю оценку конфигурации мультибиометрической системы на этапе эскизного проектирования.

Литература

1. *Ushmaev O. S., Novikov S. O.* Integral criteria for large-scale multiple fingerprint solutions // *Biometric technology for human identification* / Edited by A. K. Jain, N. K. Ratha. SPIE Proceedings. Berlingham, WA: SPIE, 2004. Vol. 5404. P. 534–543.
2. *Урмаев О. С.* Информационная технология интеграции идентификации по изображению лица для ускорения автоматической дактилоскопической идентификации // *Информатика и её применения*, 2008. Т. 2. Вып. 4. С. 66–73.
3. *Болл Р. М., Коннел Дж. Х., Панканти Ш., Ратха Н. К., Сеньор Э. У.* Руководство по биометрии. — М.: Техносфера, 2007.
4. *Урмаев О. С.* Концепция мультибиометрической идентификации в информационно-аналитических системах // *Паспортные и правоохранительные системы 2008, Интерполитех-2008*. <http://www.dancom.ru/rus/AIA/Archive/RUXIX-IPIRAN-Ushmaev-MultimodalBiometricsFramework.ppt>.
5. *Урмаев О. С., Босов А. В.* Реализация концепции многофакторной биометрической идентификации в интегрированных аналитических системах // *Бизнес и безопасность в России*, 2008. № 49. С. 104–105.
6. *Синицын И. Н., Губин А. В., Урмаев О. С.* Метрологические и биометрические технологии и системы // *История науки и техники*, 2008. № 7. С. 41–44.
7. *Dizard III, Wilson P.* FBI plans major database upgrade // *Government Computer News*. Available at http://www.gcn.com/print/25_26/41792-1.html?page=1.
8. *Урмаев О. С., Синицын И. Н.* Опыт проектирования многофакторных биометрических систем, Труды VIII международной научно-технической конференции «Кибернетика и высокие технологии XXI века», 2007. Т. 1. С. 17–28.
9. *Урмаев О. С.* Сервисно-ориентированный подход к разработке мультибиометрических технологий // *Информатика и её применения*, 2008. Т. 2. Вып. 3. С. 41–53.
10. FBI — next generation identification. <http://www.fbi.gov/hq/cjisd/ngi.htm>.
11. *Синицын И. Н., Новиков С. О., Урмаев О. С.* Развитие технологий интеграции биометрической информации // *Системы и средства информатики*, 2004. Вып. 14. С. 5–36.
12. *Ushmaev O., Novikov S.* Biometric fusion: Robust approach // *MMUA 06 Proceedings*. Toulouse, France. May 11–12, 2006.
13. *Biometric fusion: Robust approach* // *MMUA 06 Proceedings*. Toulouse, France. May 11–12, 2006.
14. *Урмаев О. С., Босов А. В.* Реализация концепции многофакторной биометрической идентификации в интегрированных аналитических системах // *Системы высокой доступности*, 4, 2007. Т. 3. С. 13–23.
15. *Урмаев О. С., Синицын И. Н.* Программная реализация мультибиометрической идентификации в интегрированных аналитических приложениях // *Труды IX международной научно-технической конференции «Кибернетика и высокие технологии XXI века»*. Воронеж, 13–15 мая 2008. Т. 2. С. 735–746.

РАЗВИТИЕ МЕТОДОВ ПОВЫШЕНИЯ КАЧЕСТВА ИЗОБРАЖЕНИЙ ЛИЦ В ВИДЕОПОТОКЕ*

А. В. Насонов¹, А. С. Крылов², О. С. Ушмаев³

Аннотация: В работе предложен метод улучшения качества изображений лиц на видео, основанный на использовании общего метода суперразрешения. Суперразрешение моделируется как обратная задача к задаче понижения разрешения, т. е. осуществляется поиск изображения, которое, будучи уменьшенным с учетом движения, дает минимальное суммарное квадратичное отклонение от исходных изображений низкого разрешения. Рассмотрены качественный и быстрый варианты метода суперразрешения. Предложен специальный метод подавления эффекта Гиббса для метода быстрого суперразрешения. Предложен новый многомасштабный метод оценки движения для последовательных кадров.

Ключевые слова: суперразрешение; подавление эффекта Гиббса; изображения лиц; многомасштабный метод оценки движения; быстрый метод

1 Введение

В настоящее время активно развиваются информационные технологии, связанные с интеллектуальной обработкой видеoinформации [1–8]. Одним из наиболее динамично развивающихся направлений интеллектуальной обработки является анализ динамических последовательностей лица (видеоряда) [9, 10] с целью последующей идентификации.

Существенной особенностью идентификации по видеоряду является то, что на этапе регистрации используется один источник (статичное фото, фоторобот и т. д.), в то время как при идентификации доступна информация из принципиально другого источника (видеокамера). Последний отличается, с одной стороны, худшим качеством изображения и потенциальной некооперативностью человека. С другой стороны, при относительно длительном нахождении в кадре лицо человека меняет ракурс, что теоретически устраняет значимую проблему идентификации человека по изображению лица, а именно: стремительную деградацию качества распознавания при различных ракурсах и низком разрешении.

На практике при анализе видео используют два принципиально различных подхода. Первый заключается в использовании длины видеоряда. А именно, современные библиотеки распознава-

ния по изображению лица позволяют сравнивать изображения низкого разрешения, но с значительными ошибкам распознавания. Однако длина видеоряда позволяет в определенной степени компенсировать негативные факторы за счет массового сравнения (рис. 1).

Второй подход заключается в интеграции до этапа сравнения (рис. 2) с целью приблизить качество видеокadra к уровню статической фотографии. А именно, интеграция происходит на уровне выделения признаков, влияющих на качество распознавания. Например, по длинной серии изображений можно определить положение зрачков с шагом, значительно меньшим размер пикселя камеры.

Наиболее перспективным направлением в рамках второго подхода является применение методов суперразрешения, которые позволяют получить изображение высокого разрешения на основе серии изображений низкого разрешения. Такой подход имеет следующие преимущества.

Во-первых, он решает ряд проблем автоматизированной идентификации. Например, около 90% данных видеозаписи не пригодны для фотопортретной экспертизы [11]. Технологии offline и online повышения разрешения видеоизображения могут обеспечить определенный прогресс в этом направлении.

* Работа поддержана грантами РФФИ (проекты 09-07-00173 и 09-01-00703) и Программой ОНИТ РАН «Информационные технологии и методы анализа сложных систем».

¹Московский государственный университет им. М. В. Ломоносова, факультет ВМиК, nasonov@cs.msu.ru

²Московский государственный университет им. М. В. Ломоносова, факультет ВМиК, kryl@cs.msu.ru

³Институт проблем информатики Российской академии наук, oushmaev@ipiran.ru

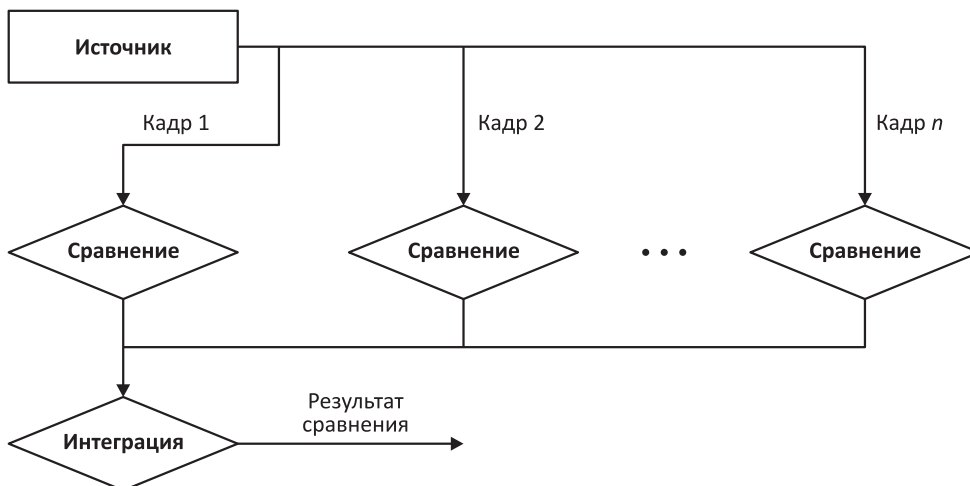


Рис. 1 Первый подход в использовании нескольких кадров в задаче распознавания: интеграция результатов сравнений

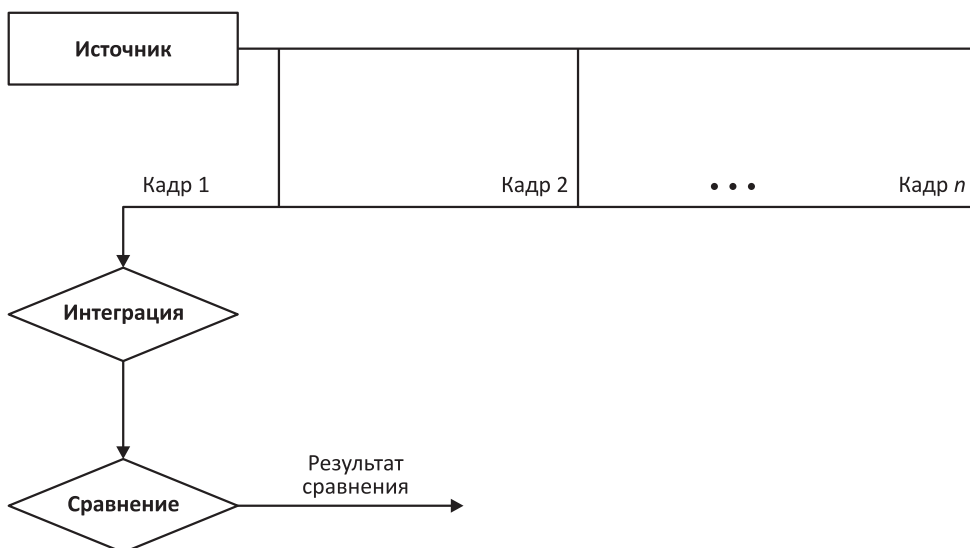


Рис. 2 Второй подход в использовании нескольких кадров в задаче распознавания: сравнение результата интеграции

Во-вторых, он имеет самостоятельную ценность и может быть интегрирован в произвольные видеосистемы.

В-третьих, использование такого подхода позволяет улучшить качество идентификации по изображению лица.

В разд. 2 дано общее описание методов суперразрешения. В разд. 3 описывается математическая модель задачи суперразрешения. В данной работе предлагаются методы качественного и быстрого суперразрешения, описанные соответственно в разд. 4 и 5. Метод подавления эффекта Гиббса для быстрого суперразрешения описан в разд. 6. Новый алгоритм определения векторов движения при

решении задачи суперразрешения по видеоданным приведен в разд. 7. Результаты и сравнение качества предложенных методов проиллюстрированы в разд. 8. В заключении представлены основные выводы.

2 Суперразрешение

В отличие от интерполяции и увеличения (resampling) изображений, повышающих разрешение, но не вносящих в изображение новой информации, суперразрешение использует информацию сразу с нескольких изображений, поэтому результиру-

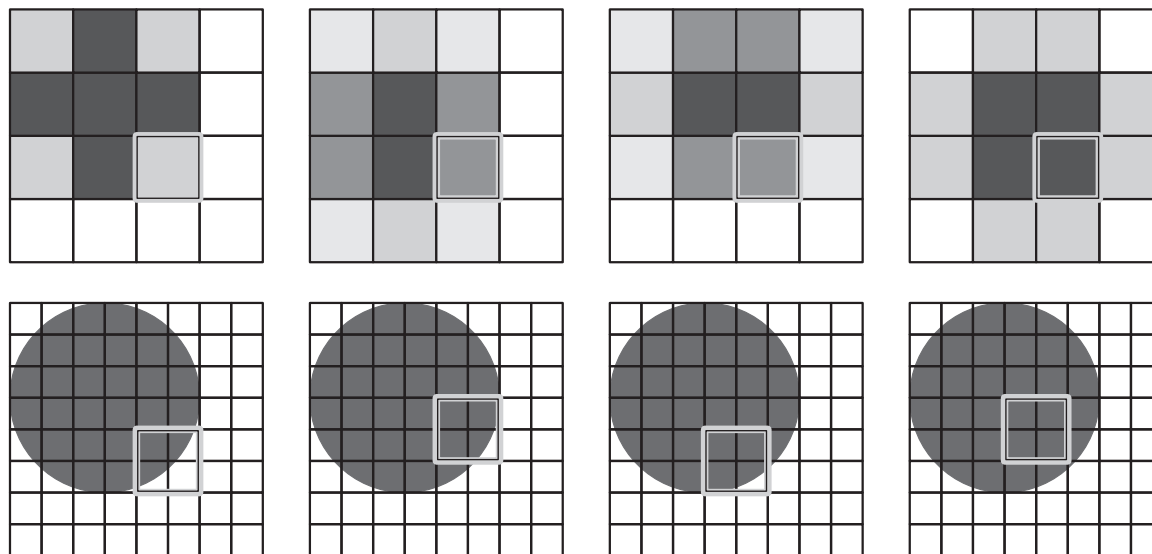


Рис. 3 Соответствие между пикселями изображений низкого разрешения (вверху) и изображения высокого разрешения (внизу)

ющее изображение высокого разрешения содержит в себе больше полезной информации. Это положительно влияет на качество методов идентификации.

Основной источник информации для суперразрешения — это изображения одного и того же объекта, незначительно движущегося на последовательных кадрах. Пиксели камеры, получающей изображение, имеют ненулевой размер, поэтому наблюдаемое значение пикселя соответствует не значению в конкретной точке на реальном изображении, а является усреднением по некоторой окрестности точки. Объект смещается, как правило, на нецелое число пикселей, поэтому на разных кадрах усреднение производится по разным окрестностям (рис. 3). Если движение объекта известно, то можно использовать информацию со всех кадров для построения одного изображения высокого разрешения.

В реальных ситуациях движение объекта неизвестно, поэтому его необходимо сначала вычислить. Близкие друг к другу точки обычно движутся одинаково, поэтому обычно используется предположение о том, что движение в небольшой окрестности точки, в которой вычисляется движение, является плоскопараллельным. В этом случае повышается точность методов оценки движения и устойчивость к шуму, но в случае резкого изменения движения, например в случае движущегося объекта на неподвижном фоне, метод будет давать неточные результаты.

Задача суперразрешения обычно ставится в виде задачи минимизации: найти такое изображение,

которое, будучи уменьшенным с учетом движения, даст минимальное суммарное квадратичное отклонение от исходных изображений низкого разрешения. Эта задача является некорректно поставленной. Для ее решения применяется регуляризирующий метод, основанный на методе регуляризации Тихонова [12]. При этом некорректно поставленная задача заменяется на близкую к ней корректно поставленную задачу путем добавления ограничений.

3 Математическая модель

Перед исследованием методов суперразрешения рассмотрим обратную задачу построения кадров низкого разрешения при известном изображении высокого разрешения z (двумерная матрица интенсивности пикселей) и известных операторах движения F_k (отображение одного изображения на последующее) для каждого кадра. Эти операторы задают соответствия между точками первого и k -го кадров. При этом кадры низкого разрешения w_k можно представить в следующем виде:

$$w_k = A_k z = DHF_k z, \quad k = \overline{1, N}, \quad (1)$$

где DH является оператором понижения разрешения. Этот оператор осуществляет последовательное применение фильтра высоких частот H , представляющего собой свертку с фильтром Гаусса и оператора прореживания D . Схематично это может быть проиллюстрировано на рис. 4.

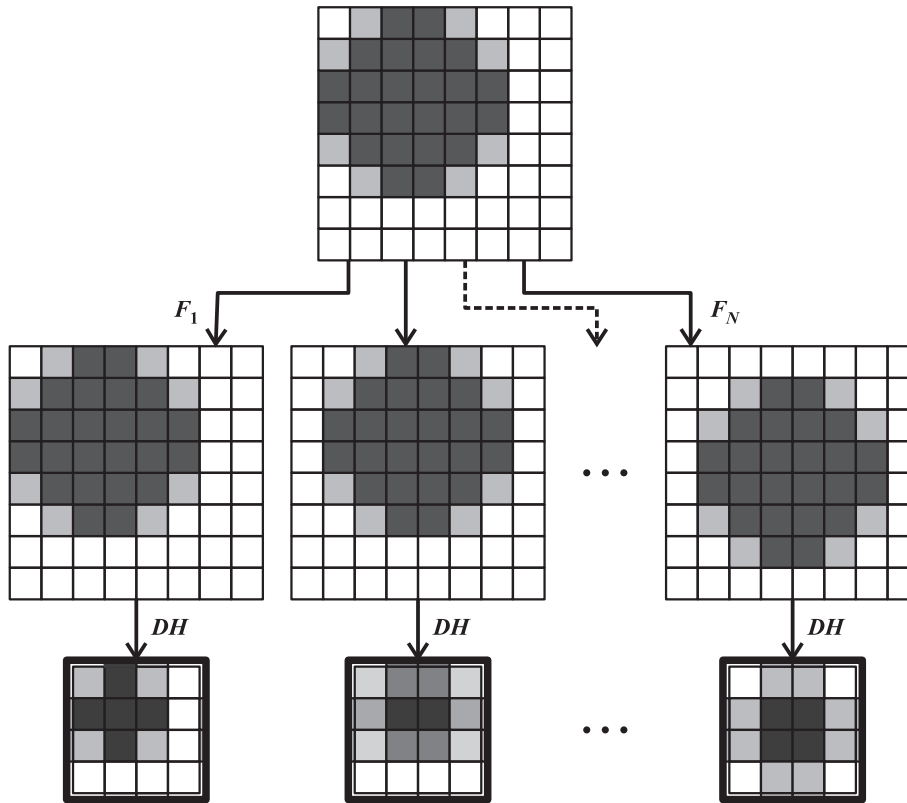


Рис. 4 Схема построения кадров низкого разрешения (1)

Задача суперразрешения ставится в виде обратной задачи к задаче понижения разрешения (1). Таким образом, происходит поиск изображение высокого разрешения z при известных w_k и F_k . Задача качественного суперразрешения формулируется в виде задачи минимизации ошибки z_R

$$z_R = \arg \min_z \sum_{k=1}^N \| A_k z - w_k \|_1, \quad (2)$$

где $\| z \|_1 = \sum_{i,j} |z_{i,j}|$.

Решением задачи является изображение высокого разрешения, которое наиболее близко ко всей последовательности w_k одновременно.

В качестве оценки ошибки вместо (2) может быть использована стандартная евклидова норма:

$$Z_R = \arg \min_z \sum_{k=1}^N \| A_k - w_k \|_2^2. \quad (3)$$

4 Качественное суперразрешение

Задачи минимизации (2) и (3) являются некорректно поставленными, поэтому для их решения

необходимо использовать регуляризирующие алгоритмы. Для задачи качественного суперразрешения (2) мы минимизируем функционал вида

$$z_R = \arg \min_z \left(\sum_{k=1}^N \| A_k z - w_k \|_1 + \alpha \Omega(z) \right). \quad (4)$$

В качестве стабилизатора $\Omega(z)$ используется функционал Bilateral Total Variation [13]

$$\Omega(z) = \sum_{-p \leq x, y \leq p} \gamma^{|x|+|y|} \| S_{x,y} z - z \|_1,$$

где $S_{x,y}$ — операторы сдвига по горизонтали на x и по вертикали на y пикселей соответственно, $\gamma = 0,8, p = 1$.

Минимизация функционала (4) производится с помощью субградиентного метода [14, 15].

5 Быстрое суперразрешение

Решение задачи в постановке (4) требует больших вычислительных затрат. Часто является важным быстрое построение приближенного решения задачи суперразрешения. Мы используем подход,

близкий к [16]. В рамках такого подхода мы заменяем регуляризирующий алгоритм обращения оператора DH из формулы (1) на прямой оператор повышения разрешения U . В качестве этого оператора мы в данной работе используем оператор повышения разрешения, основанный на гауссовской фильтрации [17].

Данная реализация алгоритма быстрого суперразрешения включает в себя следующие шаги:

1. Взять первый кадр w_1 и вычислить движение между ним и остальными кадрами w_k , $k = 2, 3, \dots, N$. Движение между w_1 и w_1 принимается равным нулю, т. е. $z = F_1 z$.
2. Увеличить каждый кадр w_k и применить оператор движения, переводящий k -й кадр в первый:

$$W_k = F_k U w_k, \quad (5)$$

где U — оператор повышения разрешения, основанный на гауссовской фильтрации. Быстрое вычисление (5) описано в [17].

3. Произвести усреднение $z = (1/N) \sum_{k=1}^N W_k$.
4. Поднять резкость изображения при помощи нерезкой маски и адаптивного фильтра [18].

6 Подавление эффекта Гиббса

При поднятии резкости изображения в быстром суперразрешении возникает побочный эффект, называемый эффектом Гиббса и выражающийся в появлении ореолов возле резких границ.

Мы используем метод, аналогичный методу подавления эффекта Гиббса после увеличения изображений [19]. При этом коэффициент увеличения s для полученного методом быстрого суперразрешения изображения определяется выбором оператора U в (5). Основная идея метода заключается в ограничении максимального значения функционала суммарной вариации, в дискретном двумерном случае задаваемом в виде

$$TV(z) = \sum_{i,j} |z_{i+1,j} - z_{i,j}| + \sum_{i,j} |z_{i,j+1} - z_{i,j}|.$$

Ограничение заключается в том, что при увеличении в s раз суммарная вариация не может увеличиться больше, чем в s раз. Если это условие не выполняется, то считается проекция на множество изображений M с суммарной вариацией, удовлетворяющей этому ограничению:

$$z_R = \arg \min_{z \in M} \|z - z^*\|_2^2,$$

где z^* — результат интерполяции или быстрого суперразрешения. Для задачи суперразрешения это множество задается в виде

$$M = \left\{ z \mid TV(z) \leq \frac{1}{N} s \sum_{k=1}^N TV(w_k) \right\}.$$

Иллюстрация применения метода приведена ниже в разд. 8.

7 Оценка оператора движения

Одной из подзадач в задаче суперразрешения является вычисление оператора — вектора движения точек F_k для пар изображений. Результаты качественного суперразрешения существенно зависят от его точности.

Существует много методов вычисления векторов движения. Среди них можно выделить группу методов, вычисляющих движение на основе совмещения особых точек. В качестве особых точек могут использоваться, например, угловые точки [20], элементы лиц [21]. Эти методы не дают результата с субпиксельной точностью, поэтому его приходится уточнять. Другая группа методов — методы, использующие вариационный подход.

В качестве базового метода используется метод Канаде—Лукаса [22], относящийся к группе вариационных методов. Этот метод использует представление изображения $J(x, y)$ через $I(x, y)$ в виде

$$I(x + u(x, y), y + v(x, y)) = J(x, y), \quad (6)$$

где $u(x, y)$ и $v(x, y)$ — компоненты вектора движения $\vec{V} = (u, v)$. Изображения $I(x, y)$ и $J(x, y)$ считаются достаточно гладкими и дифференцируемыми, $I(x + u, y + v)$ раскладывается в ряд Тейлора с отбрасыванием старших производных

$$I(x + u, y + v) \approx I(x, y) + \frac{\partial I(x, y)}{\partial x} u + \frac{\partial I(x, y)}{\partial y} v.$$

При этом (6) принимает вид

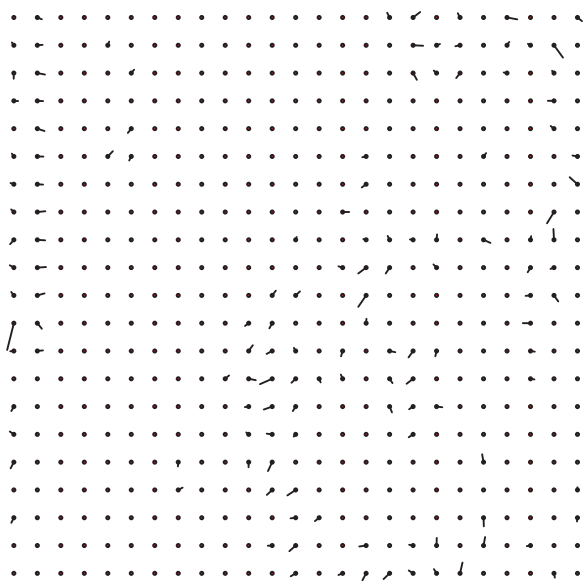
$$I_x(x, y)u(x, y) + I_y(x, y)v(x, y) = I_t(x, y),$$

где $I_t(x, y) = J(x, y) - I(x, y)$. Так как векторы движения соседних точек близки друг к другу, то в методе Канаде—Лукаса при вычислении вектора \vec{V} в конкретной точке (x_0, y_0) предполагается, что в окрестности этой точки

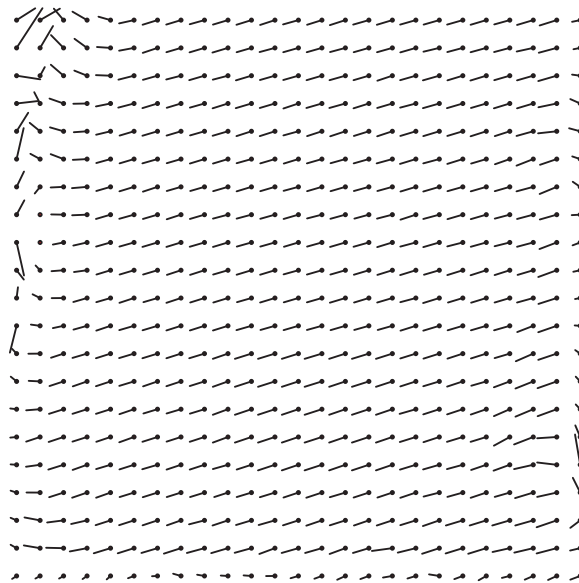
$$\Omega(x_0, y_0) = \{(x, y) : (x_0 - x)^2 + (y_0 - y)^2 \leq \varepsilon^2\}$$



(a)



(б)



(в)

Рис. 5 Иллюстрация простого и многомасштабного методов оценки движения: (а) пара изображений, для которых вычислялся оператор движения; (б) обычный метод Канаде–Лукаса; (в) многомасштабный метод

вектор движения постоянен:

$$I_x(x, y)u(x_0, y_0) + I_y(x, y)v(x_0, y_0) = I_t(x, y), \quad (x, y) \in \Omega(x_0, y_0). \quad (7)$$

Это повышает устойчивость метода, но делает невозможным резкие смены векторов движения. Далее используется модификация метода Канаде–Лукаса, использующая помимо самих изображений и их производные [17]. Это увеличивает точность

работы метода в случае переменной освещенности кадров видеопоследовательности.

Метод Канаде–Лукаса дает относительно точные результаты только в случае небольших векторов движения. Для преодоления этой проблемы нами предлагается многомасштабный подход, заключающийся в последовательном уточнении векторов движения, предварительно найденных для уменьшенных изображений.

Предположим, что известны векторы движения $\vec{V}_2 = (u_2, v_2)$ для пары изображений вдвое меньшего разрешения. Тогда для изображений оригинального разрешения эти вектора движения можно представить в виде $\vec{V}^* = (u^*, v^*)$, где $u^*(x, y) = 2u_2(x/2, y/2)$, $v^*(x, y) = 2v_2(x/2, y/2)$. Заменяя $u(x, y) = u^*(x, y) + \delta u(x, y)$ и $v(x, y) = v^*(x, y) + \delta v(x, y)$, переписываем (6) в виде

$$I(x + u^*(x, y) + \delta u(x, y), y + v^*(x, y) + \delta v(x, y)) = J(x, y). \quad (8)$$

Пользуясь тем, что метод Канаде–Лукаса предполагает постоянность вектора движения в окрестности точки, в которой считается вектор движения, заменяем (8) на

$$I(x + \delta u(x, y), y + \delta v(x, y)) = J(x - u^*(x, y), y - v^*(x, y))$$

и вычисляем $\delta u(x, y)$ и $\delta v(x, y)$ через (7). Затем подставляем эти значения в $u(x, y) = u^*(x, y) + \delta u(x, y)$ и $v(x, y) = v^*(x, y) + \delta v(x, y)$.

8 Эксперименты и обсуждение результатов

В этой работе приводятся результаты следующих экспериментов:

1. *Сравнение простого и многомасштабного методов Канаде–Лукаса.* Были взяты два изображения, сдвинутые друг относительно друга примерно на 20 пикселей по горизонтали и 5 по вертикали. Для этих изображений были применены обычный и многомасштабный методы Канаде–Лукаса. Результаты их работы приведены на рис. 5. Многомасштабный метод довольно точно определил смещение в отличие от простого метода, который такое смещение не обнаружил.
2. *Синтетический тест.* Было взято изображение лица, к нему применены несколько произвольных сдвигов, затем сдвинутые изображения были уменьшены в 4 раза. Всего было построено 16 таких изображений. К уменьшенным изображениям были применены методы интерполяции, использующие информацию только об одном изображении, и предложенные методы суперразрешения. Результаты приведены на рис. 6. Для вычисления среднеквадратичного отклонения использовались полученное и исходное изображения высокого разрешения.

3. *Применение метода подавления эффекта Гиббса* для метода быстрого суперразрешения. Результаты представлены на рис. 7.
4. *Обработка реального видеопотока.* Результаты приведены на рис. 8.

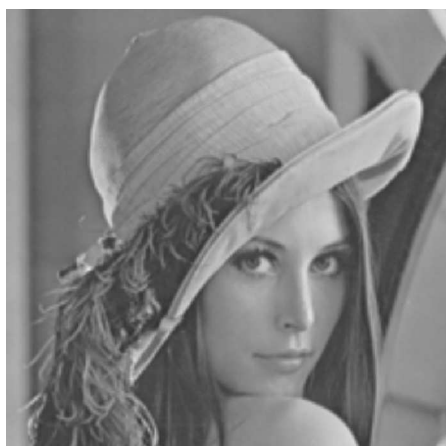
9 Заключение

Таким образом, в работе рассмотрено применение общего подхода суперразрешения видеопоследовательности к задаче улучшения качества изображений лиц. Анализ предложенных качественного и быстрого методов суперразрешения и нового многомасштабного метода оценки движения показал практическую применимость разработанных алгоритмов для задач видеонаблюдения. Рассмотрена задача постобработки результата суперразрешения видеоданных, получаемого быстрым методом с целью подавления эффекта Гиббса.

Основными направлениями дальнейших исследований являются создание комбинированного метода оценки движения, использующего, в том числе, характерные точки лица, и разработка алгоритмов для случая зашумленных видеоданных.

Литература

1. *McNeese M. D.* How video informs cognitive systems engineering: Making experience count // *Cognition, Technology & Work*, 2004. Vol. 6. No. 3. P. 186–194.
2. *Samanich N.* Analytics driving transition in video // *Security Technology & Design*, 2006. September. P. 96–98.
3. *Ушмаев О. С.* Реализации концепции многофакторной биометрической идентификации в правоохранительных системах. Интерполитех-2007. http://www.dancom.ru/rus/AIA/Archive/RUVI_BioLinkSolutions_MultimodalBiometricsConcept.pdf.
4. *Обработка и анализ цифровых изображений с примерами на LabVIEW и IMAQ Vision: [учебный курс] / Ю. В. Визильтер, С. Ю. Желтов, В. А. Князь и др. — М.: ДМК Пресс, 2007. 464 с.*
5. *Синицын И. Н., Губин А. В., Ушмаев О. С.* Метрологические и биометрические технологии и системы // *История науки и техники*, 2008. № 7. С. 41–44.
6. *Ушмаев О. С., Босов А. В.* Реализация концепции многофакторной биометрической идентификации в интегрированных аналитических системах // *Бизнес и безопасность в России*, 2008. № 49, январь. С. 104–105.
7. *Ушмаев О. С., Синицын И. Н.* Информационные технологии распознавания лиц в потоке // *Опτικο-электронные приборы и устройства в системах распознавания образов. Ч. 2. — Курск, 2008. С. 138–140.*



(a)



(б)



(в)



(г)



(д)



(е)

Рис. 6 Результаты работы методов увеличения изображений, использующих только один кадр, и предложенных методов суперразрешения для синтетической видеопоследовательности: (a) оригинал; (б) примеры уменьшенных изображений; (в) бикубическая интерполяция, $MSE = 608,63$; (г) интерполяция, основанная на регуляризации [23], $MSE = 480,07$; (д) быстрое суперразрешение, $MSE = 232,76$; (е) качественное суперразрешение, $MSE = 137,88$

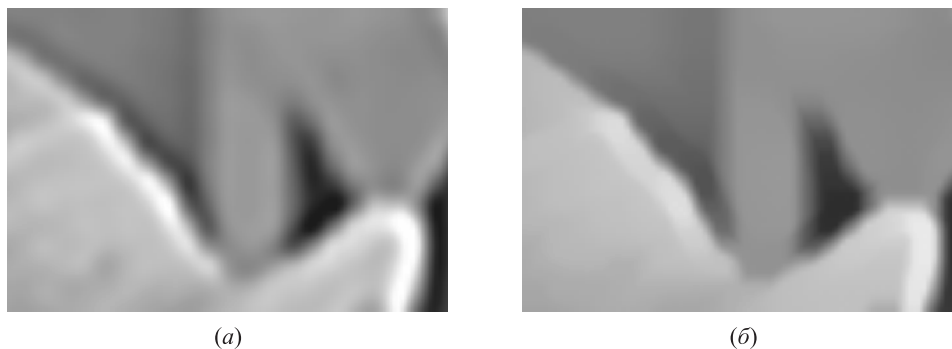


Рис. 7 Подавление эффекта Гиббса для быстрого суперразрешения: (а) быстрое суперразрешение, $MSE = 232,76$; (б) быстрое суперразрешение с подавлением эффекта Гиббса, $MSE = 197,81$



Рис. 8 Применение метода суперразрешения к реальным видеоданным: (а) примеры исходных кадров; (б) метод «ближайшего соседа»; (в) билинейная интерполяция; (г) интерполяция, основанная на регуляризации; (д) быстрое суперразрешение; (е) качественное суперразрешение

8. *Niels H., Venetiaren P.L., Lipton A.* The evolution of video surveillance: An overview // *Machine Vision and Applications*, 2008. Vol. 19. No. 5–6. P. 279–290.
9. Face Recognition Vendor Test. <http://www.frvt.org>.
10. Face Recognition Vendor Test 2006. <http://frvt.org/FRVT2006/>.
11. Портретная экспертиза / Под ред. Зинина А. М. — М.: Право и Закон, Экзамен, 2004. 160 с.
12. *Тихонов А. Н., Арсенин В. Я.* Методы решения некорректных задач. М., 1979.
13. *Farsiu S., Robinson D., Elad M., Milanfar P.* Fast and robust multi-frame super-resolution // *IEEE Trans. On Image Processing*, 2004. Vol. 13. No. 10. P. 1327–1344.
14. *Boyd S., Xiao L., Mutapcic A.* Subgradient methods. Lecture notes of EE392. — Stanford University, 2003.
15. *Krylov A. S., Nasonov A. V., Sorokin D. V.* Face image super-resolution from video data with non-uniform illumination // *Conference (International) Graphicon 2008 Proceedings*, 2008. P. 150–155.
16. *Lin F., Fookes C., Chandran V., Sridharan S.* Investigation into optical flow super-resolution for surveillance applications // *APRS Workshop on Digital Image Computing Proceedings*, 2005. P. 73–78.
17. *Krylov A., Nasonov A.* Fast super-resolution from video data using optical flow estimation // *IEEE ICSP'08 Proceedings*. Beijing, 2008. P. 853–856.
18. *Krylov A., Nasonov A., Ushmaev O.* Image super-resolution using fast deconvolution // *9th Conference on Pattern Recognition and Image Analysis: New Information Technologies Proceedings*. Nizhni Novgorod, 2008. Vol. 1. No. 2. P. 362–364.
19. *Krylov A., Nasonov A.* Adaptive total variation deringing method for image interpolation // *ICIP'08 Proceedings*. San Diego, 2008. P. 2608–2611.
20. *Sung Won Park, Savvides M.* Breaking the limitation of manifold analysis for super-resolution of facial images // *IEEE Conference (International) on Acoustics, Speech and Signal Processing*, 2007. Vol. 1. P. 573–576.
21. *Jiangang Yu.* Super-resolution and facial expressions for face recognition in video. PhD. Thesis. Riverside: University of California, 2007.
22. *Lucas B. D., Kanade T.* An iterative image registration technique with an application to stereo vision // *Imaging Understanding Workshop Proceedings*, 1981. P. 121–130.
23. *Lukin S., Krylov A. S., Nasonov A. V.* Image interpolation by super-resolution // *Conference (International) Graphicon 2006 Proceedings*, 2006. P. 239–242.

ВОССТАНОВЛЕНИЕ ВЕРОЯТНОСТНЫХ ХАРАКТЕРИСТИК СЛУЧАЙНЫХ ФУНКЦИЙ В ЗАДАЧАХ ОДНОФОТОННОЙ ЭМИССИОННОЙ ТОМОГРАФИИ

В. Г. Ушаков¹, О. В. Шестаков²

Аннотация: В работе рассматривается задача восстановления вероятностных характеристик объекта, случайным образом меняющего свою структуру в процессе регистрации проекционных данных. В рамках предложенной модели томографического эксперимента разрабатывается метод восстановления распределений случайной функции по распределениям проекций в случае, когда случайная функция имеет не более чем счетное число состояний.

Ключевые слова: однофотонная эмиссионная томография; стохастическая томография; экспоненциальное преобразование Радона; случайные функции; проекционные данные

1 Введение

В задачах однофотонной эмиссионной томографии возникает проблема обращения обобщенного преобразования Радона, имеющего вид

$$R_{\mu}(s, \theta) = \int_{x\theta=s} f(x) e^{D\mu(x, \theta^{\perp})} dl, \quad \theta \in S^1, s \in \mathbf{R}, \quad (1)$$

где $f(x)$ — непрерывная функция с компактным носителем, имеющая смысл интенсивности излучения, интеграл берется вдоль прямой $x\theta = s$, S^1 — множество направлений, задаваемых единичными векторами в \mathbf{R}^2 с центром в начале координат, а $D\mu(x, \theta^{\perp})$ — весовая функция, равная

$$D\mu(x, \theta^{\perp}) = \int_0^{\infty} \mu(x + t\theta^{\perp}) dt.$$

Здесь $\mu(x)$ — известная функция с компактным носителем, $\theta = (\cos \alpha, \sin \alpha)$, а $\theta^{\perp} = (-\sin \alpha, \cos \alpha)$. Функция $-\mu(x)$ имеет смысл коэффициента поглощения. Если $\mu(x) = 0$, то $R_{\mu}(s, \theta)$ превращается в классическое преобразование Радона, а если $\mu(x)$ равна константе на носителе $f(x)$, $R_{\mu}(s, \theta)$ превращается в экспоненциальное преобразование Радона. Формулы обращения для классического и экспоненциального преобразований Радона получены довольно давно, однако вопрос возможности

обращения преобразования (1) оставался открытым более 20 лет. В последние годы проблема была успешно решена, и были получены различные формулы обращения [1–3].

Часто в подобных задачах естественно считать функцию, описывающую интенсивность излучения, случайной. При этом состояния (реализации) этой функции меняются во время процесса получения проекций. Это приводит к тому, что восстановление даже одной реализации случайной функции обычными томографическими методами становится невозможно.

В работах [4–8] рассматривается задача определения вероятностных характеристик двумерных случайных функций по характеристикам одномерных проекций без учета поглощения и в предположении, что коэффициент поглощения равен константе. Показывается, что в общем эта задача характеризуется сильной неоднозначностью, и если не накладывать ограничений на вид реализаций случайной функции, то содержательные результаты удастся получить лишь в том случае, когда случайная функция имеет не более чем счетное число состояний. В работах [5, 8] для класса таких функций разрабатывается метод восстановления распределений двумерных случайных функций.

В данной работе задача определения вероятностных характеристик случайных функций рассматривается в предположении, что коэффициент поглощения равен известной функции $-\mu(x)$,

* Работа выполнена при финансовой поддержке РФФИ, грант 08-01-00567.

¹Московский государственный университет им. М. В. Ломоносова, кафедра математической статистики факультета ВМиК; Институт проблем информатики Российской академии наук, vgushakov@mail.ru

²Московский государственный университет им. М. В. Ломоносова, кафедра математической статистики факультета ВМиК, oshestakov@cs.msu.su

имеющей компактный носитель и принадлежащей классу Гёльдера (с некоторым параметром).

2 Постановка задачи

Пусть стохастический объект описывается двумерной случайной функцией $\xi(x)$. Предполагается выполнение следующих условий:

- (1) почти все значения $\xi(x)$ непрерывны и интегрируемы,
- (2) $\xi(x)$ имеет компактный носитель (без потери общности будем считать, что этим носителем является единичный круг: $U = \{x \in \mathbf{R}^2 : x_1^2 + x_2^2 \leq 1\}$).

Функции, совпадающие всюду за исключением множеств нулевой лебеговой меры, будем считать эквивалентными.

При выполнении этих условий определены проекции функции $\xi(x)$ с учетом поглощения — одномерные случайные функции вида

$$R_\mu \xi_\theta(s) = \int_{x\theta=s} \xi(x) e^{D\mu(x, \theta^\perp)} dl, \quad \theta \in S^1, s \in \mathbf{R}. \quad (2)$$

В задачах стохастической томографии предполагается, что имеется некоторая информация о вероятностных характеристиках учитывающих поглощение проекций (всех или некоторого множества). Задача состоит в нахождении определенных вероятностных характеристик случайной функции $\xi(x)$. Первый вопрос касается однозначности соответствия между рассматриваемыми характеристиками двумерной случайной функции и ее проекций. Как и в случае отсутствия поглощения, этот вопрос может возникать, например, в следующих формах:

1. Можно ли однозначно определить совместные распределения $\xi(x_1), \dots, \xi(x_n)$, если известны совместные распределения $R_\mu \xi_\theta(s_1), \dots, R_\mu \xi_\theta(s_m)$ для всех $m = 1, 2, \dots$ и всех $\theta \in S^1$?
2. Можно ли однозначно определить дисперсии $\text{var}(\xi(x))$, если известны дисперсии $\text{var}(R_\mu \xi_\theta(s))$ для всех $\theta \in S^1$?
3. Можно ли однозначно восстановить дисперсии $\text{var}(\xi(x))$, если известны все совместные распределения величин $R_\mu \xi_\theta(s_1), \dots, R_\mu \xi_\theta(s_m)$ для всех $m = 1, 2, \dots$ и всех $\theta \in S^1$?
4. Имеется ли связь между «изменчивостью» (величиной дисперсии) двумерной случайной функции в некоторой точке и изменчивостью проекций в точках, являющихся проекциями этой точки?

Проблема восстановления математического ожидания двумерной случайной функции по математическим ожиданиям учитывающих поглощение проекций не рассматривается, поскольку, как легко видеть, эта проблема эквивалентна обычной (нестохастической) томографии:

$$\begin{aligned} ER_\mu \xi_\theta(s) &= \mathbf{E} \int_{x\theta=s} \xi(x) e^{D\mu(x, \theta^\perp)} dl = \\ &= \int_{x\theta=s} \mathbf{E} \xi(x) e^{D\mu(x, \theta^\perp)} dl. \end{aligned}$$

3 Общий случай

В данном разделе будет приведено несколько утверждений, относящихся к сформулированным выше вопросам 1–4 и показывающих, что в общем случае ответы на, по крайней мере, вопросы 2 и 4 отрицательны. По-видимому, в общем случае отрицательными будут ответы и на вопросы 1 и 3, хотя пока не удалось построить соответствующие примеры.

Утверждение 1. *Существуют две случайные функции $\xi(x)$ и $\eta(x)$, определенные в единичном круге $x_1^2 + x_2^2 \leq 1$, такие, что $\text{var}(\xi(x)) \neq \text{var}(\eta(x))$ для всех x , удовлетворяющих неравенству $x_1^2 + x_2^2 < 1$, в то время как $\text{var}(R_\mu \xi_\theta(s)) = \text{var}(R_\mu \eta_\theta(s))$ для всех $\theta \in S^1$ и $s \in \mathbf{R}^1$.*

В ряде задач бывает достаточно установить, в каких точках значения реализаций двумерной случайной функции имеют большой разброс по сравнению с другими точками. Другими словами, нет необходимости восстанавливать численные значения дисперсии, а достаточно на качественном уровне указать те точки, в которых значения дисперсии относительно велики. В связи с этим возникает вопрос о связи величины разброса значений реализаций в некоторой точке и величинами разброса реализаций проекций с учетом поглощения в проекциях этой точки. Следующие два утверждения показывают, что в общем случае такая связь отсутствует.

Утверждение 2. *Существует случайная функция $\xi(x)$, $x_1^2 + x_2^2 \leq 1$, такая, что все учитывающие поглощение проекции наиболее изменчивой точки (точки, имеющей максимальную дисперсию) являются наименее изменчивыми точками соответствующих одномерных учитывающих поглощение проекций функции $\xi(x)$.*

Утверждение 3. *Существует случайная функция $\xi(x)$, $x_1^2 + x_2^2 \leq 1$, такая, что все учитывающие поглощение*

проекции наименее изменчивой точки (точки, имеющей минимальную дисперсию) являются наиболее изменчивыми точками соответствующих одномерных учитывающих поглощение проекций функции $\xi(x)$.

Сформулированные утверждения справедливы даже в том случае, когда коэффициент поглощения равен константе. Соответствующие доказательства можно найти в работе [7].

4 Класс T случайных функций

Как следует из предыдущего раздела, в случае, когда случайная функция $\xi(x)$ произвольна, по-видимому, невозможно получить какие-либо содержательные результаты. Однако ситуация изменится, если наложить на нее некоторые ограничения. В данном разделе будет введен класс случайных функций, с одной стороны, достаточный для многих приложений, а с другой — позволяющий единственным образом восстанавливать вероятностные характеристики стохастических объектов по характеристикам проекций.

Пусть T — множество всех двумерных случайных функций $\xi(x)$ вида

$$\xi(x) = f_\nu(x),$$

где $f_1(x), f_2(x), \dots$ — последовательность непрерывных интегрируемых функций, определенных в единичном круге $U = \{(x) \in \mathbf{R}^2 : x_1^2 + x_2^2 \leq 1\}$, а ν — случайная величина, принимающая целые положительные значения.

Вероятностная структура случайных функций из класса T полностью определяется распределением, т. е. набором $(f_1(x), f_2(x), \dots; p_1, p_2, \dots)$, где $p_i = P(\xi(x) = f_i(x))$, $i = 1, 2, \dots$, $\sum_{i=1}^{\infty} p_i = 1$.

Распределение $\xi(x)$ будем обозначать P_ξ .

Оказывается, что в рамках этой модели, как и в случае постоянного коэффициента поглощения, рассмотренного в работе [7], распределение двумерной случайной функции полностью определяется распределениями проекций. Однако, если для модели с постоянным коэффициентом поглощения удалось разработать метод реконструкции распределения двумерной случайной функции по распределениям проекций [8], то в модели с неравномерным поглощением пока удалось лишь установить принципиальную возможность такой реконструкции.

Теорема. Пусть $\xi(x) \in T$, $\eta(x) \in T$ и $P_{R_\mu \xi_\theta} = P_{R_\mu \eta_\theta}$ для всех $\theta \in \Lambda$, где Λ — подмножество S^1 , имеющее положительную меру. Тогда

$$P_\xi = P_\eta.$$

Другими словами, в классе T распределение любой двумерной случайной функции однозначно определяется распределениями проекций, зарегистрированных в любом сколь угодно узком диапазоне углов обзора.

Доказательство. Покажем, что если функция $f(x)$ непрерывна и имеет компактный носитель, функция $\mu(x)$ принадлежит классу Гёльдера (с некоторым параметром) и также имеет компактный носитель, а $R_\mu f(s, \theta) = 0$ для всех $s \in \mathbf{R}$ и всех $\theta \in E$, где E — подмножество S^1 , имеющее положительную меру, то $f(x)$ тождественно равна нулю. (При доказательстве этого факта будем следовать идеям работы [9].)

Разложим функцию

$$D\mu(x, \theta^\perp) = \int_0^\infty \mu(x + t\theta) dt$$

на четную и нечетную компоненты (по θ): $D\mu(x, \theta^\perp) = u(x, \theta) + v(x, \theta)$, где

$$u(x, \theta) = \frac{1}{2}(D\mu(x, \theta^\perp) + D\mu(x, -\theta^\perp));$$

$$v(x, \theta) = \frac{1}{2}(D\mu(x, \theta^\perp) - D\mu(x, -\theta^\perp)).$$

Легко видеть, что функция $2u(x, \theta)$ равна преобразованию Радона функции $\mu(x)$, вычисленному в точке $(s, \theta) = (x \cdot \theta, \theta)$. Поэтому эта функция постоянна на прямых $x \cdot \theta = s$, и если $\mu(x)$ принадлежит классу Гёльдера и имеет компактный носитель, то $u(x, \theta)$ также принадлежит классу Гёльдера. Для данной непрерывной функции $\psi(\theta)$ обозначим за $\tilde{\psi}(\theta)$ такую функцию от θ с нулевым средним, для которой функция $\psi(\theta) + i\tilde{\psi}(\theta)$ является граничным значением аналитической функции, определенной в единичном круге. Функция $\tilde{\psi}(\theta)$ называется функцией, сопряженной к $\psi(\theta)$, и ее можно вычислить с помощью свертки

$$\tilde{\psi}(\theta) = \frac{1}{2\pi} \int_0^{2\pi} \psi(\theta(\alpha - \beta)) \operatorname{ctg} \frac{\beta}{2} d\beta,$$

где $\theta(\alpha) = (\cos \alpha, \sin \alpha)$. По теореме Привалова [10], если $\psi(\theta)$ принадлежит классу Гёльдера, то $\tilde{\psi}(\theta)$ также принадлежит классу Гёльдера. Найдем сопряженную функцию для нечетной компоненты $v(x, \theta)$:

$$\begin{aligned} \tilde{v}(x, \theta) &= \frac{1}{2} \left(\tilde{D}\mu(x, \theta^\perp) - \tilde{D}\mu(x, -\theta^\perp) \right) = \\ &= \frac{1}{4\pi} \int_0^{2\pi} D\mu(x, \theta^\perp(\alpha - \beta)) \operatorname{ctg} \frac{\beta}{2} d\beta - \end{aligned}$$

$$-\frac{1}{4\pi} \int_0^{2\pi} D\mu(x, -\theta^\perp(\alpha - \beta)) \operatorname{ctg} \frac{\beta}{2} d\beta,$$

где $\theta^\perp(\alpha) = (-\sin \alpha, \cos \alpha)$. Далее, воспользовавшись формулой

$$\frac{1}{\cos \beta} = \frac{1}{2} \left(\operatorname{ctg} \frac{\beta + \pi/2}{2} - \operatorname{ctg} \frac{\beta - \pi/2}{2} \right),$$

получаем

$$\begin{aligned} \tilde{v}(x, \theta) &= \frac{1}{2\pi} \int_{S^1} \frac{1}{\theta\omega} D\mu(x, \omega) d\omega = \\ &= \frac{1}{2\pi} \int_{S^1} \frac{1}{\theta\omega} \int_0^\infty \mu(x + t\omega) dt. \end{aligned}$$

Переходя в последнем интеграле к декартовым координатам $x + t\omega = y$, имеем

$$\begin{aligned} \tilde{v}(x, \theta) &= -\frac{1}{2\pi} \int_{\mathbf{R}^2} \frac{\mu(y)}{(x-y)\theta} dy = \\ &= -\frac{1}{2\pi} \int_{-\infty}^\infty \frac{1}{x\theta - s} \left(\int_{y\theta=s} \mu(y) dl \right) ds = \\ &= -\frac{1}{2} HR\mu(x\theta, \theta), \end{aligned}$$

где $HR\mu(x\theta, \theta)$ — преобразование Гильберта функции $R\mu(\theta, s)$ по первой переменной, вычисленное в точке $(x\theta)$. (Преобразование Гильберта функции $q(s)$ задается формулой

$$Hq(t) = \frac{1}{\pi} \int_{-\infty}^\infty \frac{1}{t-s} q(s) ds,$$

где интеграл понимается в смысле главного значения.)

Следовательно, $\tilde{v}(x, \theta)$ постоянна на прямых $x\theta = s$. Функции $u(x, \theta)$ и $v(x, \theta)$ принадлежат классу Гельдера, и по теореме Привалова $\tilde{v}(x, \theta)$ также принадлежит классу Гельдера, а функция $v(x, \theta) + i\tilde{v}(x, \theta)$ является граничным значением аналитической функции, определенной в единичном круге.

Введем функцию

$$a(x, \theta) = e^{D\mu(x, \theta^\perp)} b(x, \theta),$$

где

$$b(x, \theta) = e^{(1/2)(-R\mu(\theta, x\theta) - iHR\mu(\theta, x\theta))}.$$

Таким образом,

$$a(x, \theta) = e^{v(x, \theta) + i\tilde{v}(x, \theta)}.$$

Поскольку $b(x, \theta)$ постоянна на прямых $x\theta = s$, из того, что $R_\mu f(s, \theta) = 0$ для $\theta \in E$, следует

$$R_a f(s, \theta) = \int_{x\theta=s} f(x) a(x, \theta) dx = 0, \quad \theta \in E, s \in \mathbf{R}.$$

Возьмем произвольный многочлен $p(s)$ и рассмотрим функцию

$$\begin{aligned} W_p(\theta) &= \int_{-\infty}^\infty p(s) R_a f(s, \theta) ds = \\ &= \int_{\mathbf{R}^2} p(x\theta) f(x) a(x, \theta) dx. \end{aligned}$$

Функция $W_p(\theta) = 0$ при $\theta \in E$, и, поскольку подынтегральное выражение при каждом фиксированном x является граничным значением аналитической функции, а функция $f(x)$ имеет компактный носитель, $W_p(\theta) = 0$ при всех $\theta \in S^1$. Функция $R_a f(s, \theta)$ для любого θ имеет носитель в некотором интервале $|s| < A$, и, так как множество многочленов плотно в пространстве непрерывных на интервале функций, $R_a f(s, \theta) = 0$ для всех θ и всех s . Так как $b(x, \theta)$ постоянна (и не равна нулю) на прямых $x\theta = s$, должно также выполняться равенство $R_\mu f(s, \theta) = 0$ для всех θ и всех s . Функция $f(x)$ однозначно определяется значениями $R_\mu f(s, \theta)$ для $\theta \in S^1$ и $s \in \mathbf{R}$ по формуле Новикова [3] или Наттерера [2] и, следовательно, $f(x) \equiv 0$.

Предположим теперь, что $P_\xi \neq P_\eta$. Это значит, что существует непрерывная функция $f(x)$ с носителем в круге U такая, что

$$P(\xi(x) = f(x)) \neq P(\eta(x) = f(x)).$$

Обозначим через $f_1(x), f_2(x), \dots$ значения случайной функции $\xi(x)$, отличные от $f(x)$, и через $g_1(x), g_2(x), \dots$ обозначим аналогичные значения $\eta(x)$ (таким образом, $f(x) \neq f_i(x)$ и $f(x) \neq g_i(x)$, $i = 1, 2, \dots$). Для каждого фиксированного $i = 1, 2, \dots$ пусть A_i обозначает множество всех $\theta \in \Lambda$, для которых

$$R_\mu f(s, \theta) \equiv R_\mu f_i(s, \theta),$$

и, соответственно, B_i — множество всех $\theta \in \Lambda$, для которых

$$R_\mu f(s, \theta) \equiv R_\mu g_i(s, \theta).$$

Каждое из множеств A_i и B_i имеет меру нуль. Если для некоторого i это было бы не так, то $R_\mu f(s, \theta)$ совпала бы с $R_\mu f_i(s, \theta)$ или с $R_\mu g_i(s, \theta)$ на множестве положительной меры и в силу доказанного

выше это влекло бы за собой совпадение функции $f(x)$ с функцией $f_i(x)$ или с функцией $g_i(x)$.

Таким образом, множество

$$C = \bigcup_{i=1}^{\infty} (A_i \cup B_i)$$

имеет меру нуль, а следовательно, множество $\Lambda \setminus C$ непусто. Возьмем произвольное $\theta \in \Lambda \setminus C$. Поскольку для этого θ выполнено условие $R_\mu f(s, \theta) \neq R_\mu f_i(s, \theta), i = 1, 2, \dots$, следовательно,

$$P(R_\mu \xi_\theta(s) = R_\mu f(s, \theta)) = P(\xi(x) = f(x))$$

и аналогично

$$P(R_\mu \eta_\theta(s) = R_\mu f(s, \theta)) = P(\eta(x) = f(x)).$$

Полученное противоречие доказывает теорему.

Литература

1. *Arbuzov E. V., Bukhgeim A. L., Kazantsev S. G.* Two-dimensional tomography problems and the theory of A -analytic functions // *Siberian Adv. Math.*, 1998. Vol. 8. P. 1–20.
2. *Natterer F.* Inversion of the attenuated Radon transform // *Inverse Problems*, 2001. Vol. 17. P. 113–119.
3. *Novikov R. G.* An inversion formula for the attenuated X-ray transformation // *Ark. Mat.*, 2002. Vol. 40. P. 145–167.
4. *Ушаков В. Г., Ушаков Н. Г.* Восстановление вероятностных характеристик многомерных случайных функций по проекциям // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2001. № 4. С. 32–39.
5. *Shestakov O. V.* An algorithm to reconstruct probabilistic distributions of multivariate random functions from the distributions of their projections // *J. of Mathematical Sciences*, 2002. Vol. 112. No. 2. P. 4198–4204.
6. *Шестаков О. В.* О единственности восстановления вероятностных характеристик многомерных случайных функций по вероятностным характеристикам их проекций // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2003. № 3. С. 37–41.
7. *Ушаков В. Г., Шестаков О. В.* Экспоненциальное преобразование Радона случайных функций // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2005. № 1. С. 49–55.
8. *Shestakov O. V.* Inversion of exponential Radon transform of random functions // *Transactions of XXV Seminar on Stability Problems for Stochastic Models*, 2005. P. 264–269.
9. *Boman J., Stromberg J.-O.* Novikov's inversion formula for the attenuated radon transform – a new approach // *J. Geom. Anal.*, 2004. Vol. 14. P. 185–198.
10. *Привалов И. И.* *Граничные свойства аналитических функций.* — М.: ГИТТЛ, 1950.

КОНЦЕПЦИЯ СОЗДАНИЯ ОТЕЧЕСТВЕННЫХ ИНТЕГРИРОВАННЫХ КОММУНИКАЦИОННЫХ МИКРОКОНТРОЛЛЕРОВ ДЛЯ ПАКЕТНОЙ КОММУТАЦИИ

В. Б. Егоров¹

Аннотация: В статье предлагается концепция упрощенных интегрированных коммуникационных микроконтроллеров (ИКМ), которые могли бы найти применение в различных отечественных устройствах пакетной коммутации и маршрутизации, расширяя их функциональные возможности и упрощая разработку.

Ключевые слова: интегрированный коммуникационный микроконтроллер; PowerQUICC; децентрализованная коммутация; маршрутизирующий коммутатор

1 Современные интегрированные коммуникационные микроконтроллеры

Интегрированные коммуникационные микроконтроллеры — один из наиболее динамично развивающихся секторов рынка коммуникационной микроэлектроники [1, 2]. Лидеры этого направления — компания *Motorola* и ее преемница *Freescale Semiconductor* — уже выпустили на рынок шесть поколений ИКМ, а общее число предлагаемых модификаций таких приборов перевалило за сотню [3]. Популярность ИКМ объясняется как высокой степенью интеграции электронных компонентов, так и удачной интеграцией функций, в совокупности обеспечивающими ИКМ уникальные свойства и эффективное применение в самых разнообразных приложениях, связанных с телекоммуникацией вообще и коммутацией в частности, причем как каналов, так и пакетов.

Типичный ИКМ компании *Freescale* представляет собой «систему на кристалле», объединяющую в единое целое современный мощный универсальный микроконтроллер с различными сопроцессорами и коммуникационными модулями (рис. 1).

Микроконтроллерная часть ИКМ включает в себя свободно программируемое процессорное ядро и модуль системной интеграции.

Программируемое ядро, на котором выполняются пользовательские программы, представляет собой мощный универсальный процессор с архи-

тектурой PowerPC. Ядра комплектуются диспетчером памяти, отдельными кэшами инструкций и данных, дополняемыми в некоторых моделях кэшем второго уровня и процессором плавающей запятой. Процессорные ядра самых производительных ИКМ способны работать на частотах до полутора гигагерц, обеспечивая производительность до трех с половиной миллиардов инструкций в секунду по эталонному тесту Dhrystone (DMIPS).

Модуль системной интеграции объединяет разнообразные средства организации системы: контроллеры различных памятей, включая DDR SDRAM (double data rate synchronous dynamic random access memory) и NAND-flash; адаптеры внешних системных шин, в частности шины PCI (peripheral component interconnect), и стандартных последовательных интерфейсов ввода-вывода I²C (inter integrated circuits) и SPI (serial peripheral interface); контроллеры современных высокоскоростных последовательных интерфейсов PCIe (PCI express), SRIO (serial RapidIO) и SATA (serial advanced technology attachment).

Большинство ИКМ имеет встроенный сопроцессор безопасности, выполняющий шифрование пользовательских данных «на лету» по всем основным принятым в мировой практике стандартам. Некоторые модели дополнительно комплектуются сопроцессорами, упрощающими анализ контекста заголовков пакетов и/или табличный поиск, в частности с использованием хеширования адресов.

Однако основной изюминкой ИКМ являются встроенные коммуникационные модули двух видов: универсальные и специализированные.

¹Институт проблем информатики Российской академии наук, vegorov@ipiran.ru

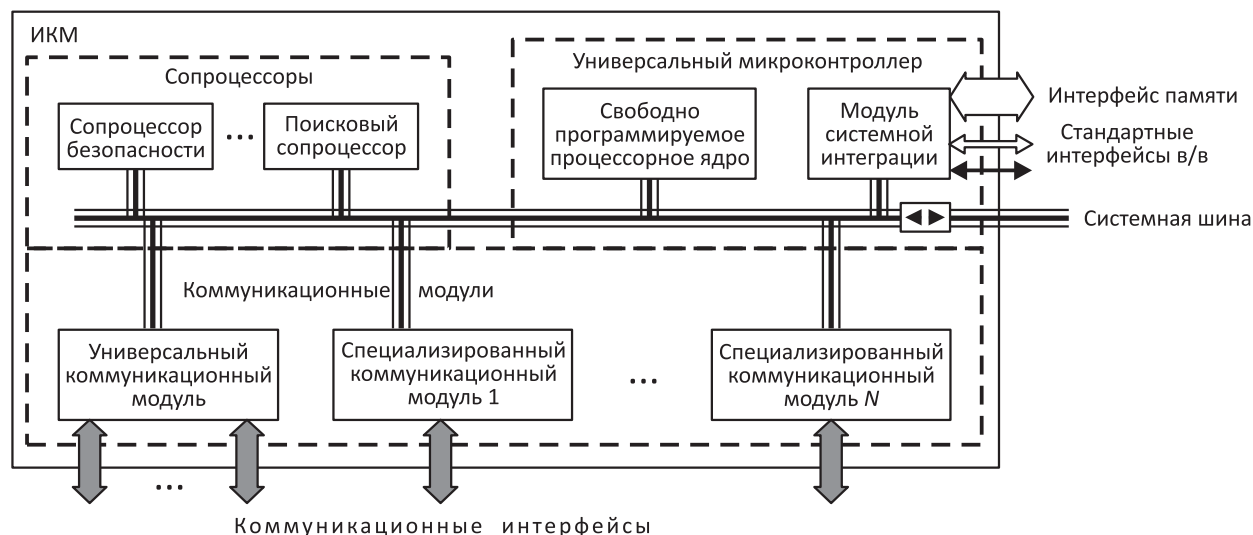


Рис. 1 Обобщенная структура ИКМ

Универсальный коммуникационный модуль, как правило, единственный в ИКМ, обслуживает несколько программно конфигурируемых среднескоростных коммуникационных интерфейсов и поддерживает целый ряд соответствующих им широко распространенных в мире коммуникационных протоколов канального уровня благодаря тому, что включает отдельный RISC-процессор со своим встроенным фирменным программным обеспечением (ФПО). Специализированные коммуникационные модули обслуживают один–два однотипных высокоскоростных интерфейса и тоже могут содержать процессоры и ФПО.

Так как ИКМ долгое время отслеживали эволюцию телекоммуникационных технологий, все историческое многообразие интерфейсов и протоколов нашло в них отражение различными типами коммуникационных модулей и широким ассортиментом конфигурируемых интерфейсов универсальных модулей. На разных этапах развития в номенклатуру поддерживаемых ИКМ коммуникационных интерфейсов и протоколов вошли:

- асинхронные интерфейсы типа UART (universal asynchronous receiver/transmitter);
- синхронные интерфейсы на основе протоколов SDLC/HDLC (synchronous data link control/high-level data link control) и BISYNC (binary synchronous communications);
- иерархия интерфейсов ISDN (integrated services digital network), включая как basic rate (IDL, GCI), так и primary rate (T1/E1, T3/E3);
- интерфейсы ATM (asynchronous transfer mode), в том числе UTOPIA (universal test and operations physical interface for ATM);

– интерфейс Ethernet, причем по мере совершенствования коммуникационных модулей скорость передачи данных последовательно росла с 10 Мбит/с до 1 Гбит/с.

Обогащение номенклатуры интерфейсов и протоколов в ИКМ параллельно сопровождалось расширением перечня коммуникационных технологий, поддерживаемых аппаратурой и ФПО коммуникационных модулей. В результате современные ИКМ находят применение и в сетях цифровой телефонии, в частности ISDN, и в пакетных сетях Token ring, X.25, Frame relay, ATM, Ethernet, Packet over SONET (synchronous optical networking) и других.

Для ИКМ характерны универсальность и «дружелюбие» к пользователю. Непросто найти область телекоммуникации и пакетной коммутации, где ИКМ оказался бы не к месту. Для любого не слишком экзотического приложения найдутся приборы с соответствующими внешними интерфейсами и поддержкой требуемых коммуникационных протоколов, которые в совокупности с исчерпывающими возможностями универсального микроконтроллера действительно оказываются эффективной законченной телекоммуникационной «системой на кристалле», позволяющей решать самые разные по характеру и сложности задачи исключительно внутренними ресурсами одного единственного прибора, а следовательно, предельно быстро и с минимальными затратами.

Но обратная сторона универсальности и «дружелюбия» ИКМ — его сложность. В ИКМ заложен целый ряд технических решений, представляющих собой интеллектуальную собственность компании

Freescale самой высокой пробы в различных областях электроники, вычислительной техники и телекоммуникации. Ее главные составляющие [4]:

- современный высокопроизводительный универсальный процессор, оснащенный эффективными средствами разработки программного обеспечения;
- сверхреактивный RISC-процессор и особенно его уникальное многоплановое ФПО, поддерживающее все основные современные телекоммуникационные технологии и их эффективное взаимодействие (interworking) между собой;
- многофункциональные адаптеры практически всех интерфейсов, нашедших применение в компьютерах и телекоммуникации.

2 Концепции пакетных коммутаторов на основе интегрированных коммутационных микроконтроллеров

С учетом имеющейся на мировом рынке номенклатуры ИКМ и опыта их применения [5, 6] можно выделить три концептуальных подхода к созданию устройств пакетной коммутации и маршрутизации на основе некоего обобщенного ИКМ [7, 8]:

- (1) централизованный коммутатор как «система на кристалле», предполагающий мощный высокоинтегрированный и высокопроизводительный многопортовый ИКМ;
- (2) децентрализованный коммутатор на основе гомогенной «сверхлокальной» сети тесно связанных малопортовых ИКМ средней производительности;
- (3) децентрализованный коммутатор на основе гетерогенной «сверхлокальной» сети с общей разделяемой буферной памятью и множеством ИКМ ограниченной производительности.

Ярким примером высокопроизводительной «системы на кристалле» является ИКМ MPC8568E компании *Freescale* [9]. В рамках общей структуры, показанной на рис. 1, он включает универсальный микроконтроллер, сопроцессоры безопасности и табличного поиска, а также три коммуникационных модуля. Микроконтроллер содержит свободно программируемое ядро с рабочей частотой до 1,33 ГГц, 64-разрядный интерфейс памяти DDR2

SDRAM с частотой до 533 МГц, 32-разрядную параллельную локальную шину и 32-разрядный интерфейс PCI, работающий в режимах host и agent на частоте до 66 МГц, а также последовательные интерфейсы PCIe, SRIО, два I²C и два UART.

Универсальный коммуникационный модуль MPC8568E с двумя RISC-процессорами и мощным ФПО обслуживает 8 одинаковых коммуникационных портов, 8 портов ISDN и два порта SPI. Универсальные порты поддерживают технологии ISDN, X.25, Frame relay, ATM, Ethernet и другие. В режиме ATM максимальная скорость передачи данных через два 16-разрядных интерфейса UTOPIA составляет 622 Мбит/с (OC-12); в режиме Ethernet все 8 портов способны работать на скоростях 10 и 100 Мбит/с, а три порта — на скорости 1 Гбит/с. Фирменное программное обеспечение модуля поддерживает все современные телекоммуникационные технологии и обеспечивает эффективное взаимодействие между ними. Благодаря этому универсальный коммуникационный модуль реализует исключительно внутренними ресурсами восьмипортовый коммутатор L2 и обеспечивает инкапсуляцию трафика Ethernet в трафик ATM.

Помимо универсального коммуникационного модуля в ИКМ MPC8568E входят два одинаковых специализированных «интеллектуальных» коммуникационных модуля трехскоростного (10, 100 и 1000 Мбит/с) Ethernet с аппаратной поддержкой протокола IP, что дает возможность прикладному программному обеспечению эффективно реализовывать функции коммутатора L3, граничного маршрутизатора и устройства сетевой безопасности.

В целом ИКМ MPC8568E позволяет создать с минимальной дополнительной аппаратурой (необходимые внешние памяти и трансиверы Ethernet) законченный централизованный пакетный коммутатор максимально с 10 внешними портами (из них до пяти гигабитных), который также может работать в качестве граничного маршрутизатора и устройства безопасности региональных и глобальных сетей, в том числе сетей Metro Ethernet и MPLS (multiple protocol label switching) [1, 2]. Этот подход весьма привлекателен [4], но здесь он не обсуждается, поскольку, к сожалению, появление в обозримом будущем отечественного ИКМ, подобного MPC8568E, нереально [10]. Далее внимание уделено альтернативным подходам к созданию устройств пакетных коммутаторов и маршрутизаторов, а именно основанным на децентрализованной архитектуре пакетной коммутации [11] с использованием более простых ИКМ.

Первой альтернативой является децентрализованная коммутация в гомогенной «сверхлокаль-

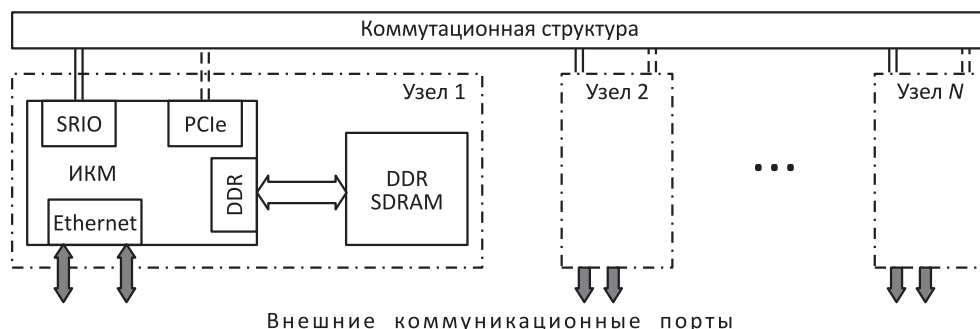


Рис. 2 Гомогенный коммутатор с узлами на основе ИКМ

ной» сети из множества тесно связанных одинаковых малопортовых ИКМ средней производительности, оснащенных небольшим числом коммуникационных портов. Примерами таких ИКМ могут служить приборы *Freescale MPC8542* (2 коммуникационных порта трехскоростного Ethernet) и *MPC8548* (4 порта Ethernet) [12], снабженные также адаптерами интерфейсов SRIO и PCIe. На рис. 2 показана примерная организация гомогенного децентрализованного коммутатора на базе неких малопортовых ИКМ средней производительности, функционально близких, например, к прибору *MPC8542*.

В гомогенном коммутаторе «сверхлокальная» сеть реализуется на высокоскоростном интерфейсе, рассчитанном на объединение множества ИКМ в пределах одной платы или одного шасси и допускающем скорости передачи данных, не уступающие суммарной скорости передачи на внешних портах одного ИКМ. Этим условиям в полной мере отвечает интерфейс SRIO, скорость передачи данных на котором может достигать 10 Гбит/с. Принципиально важно, что спецификация SRIO включает механизм инкапсуляции чужих протоколов для потокового транзита инкапсулированных блоков данных через коммутационные структуры с учетом их приоритетов и классов обслуживания, причем с минимальными накладными расходами инкапсуляции и задержками транзита [13].

Данная концепция предполагает децентрализованную коммутацию, при которой решение о продвижении каждого входящего в коммутатор пакета принимает тот узел «сверхлокальной» сети, в нашем случае ИКМ, в который входит этот пакет. Система гомогенна в том смысле, что все узлы распределенного коммутатора одинаковы и равноправны. Оценочно, на основании опыта компании *Freescale*, для собственно продвижения принятых пакетов с выполнением сопутствующих действий по их фильтрации и квалификации, а также других функций вплоть до распределенной маршрутиза-

ции требуется производительность свободно программируемого ядра ИКМ не ниже 1000 DMIPS. Коммутация пакетов в реальном времени может быть обеспечена таким ядром вплоть до суммарно гигабитной скорости передачи данных на внешних коммуникационных портах. В свою очередь, при тех же условиях пропускная способность интерфейса SRIO достаточна не только для доставки скомутированных пакетов из входного узла «сверхлокальной» сети в выходной, но и для обмена маршрутной и служебной информацией между ИКМ.

Концепция реализуется всего двумя базовыми элементами: ИКМ средней производительности в качестве узлов «сверхлокальной» сети и одним коммутатором, реализующим коммутационную структуру. Требования к ИКМ такого рода подробно обсуждаются ниже. Что же касается коммутатора, то, например, коммутаторы SRIO выпускаются несколькими западными фирмами, например компаниями *IDT* и *Tundra Semiconductor* [14, 15]. Однако типичный коммутатор для интерфейса SRIO или ему подобного, функционально достаточный для реализации концепции «сверхлокальной» сети, может быть реализован на программируемых логических интегральных схемах (ПЛИС) умеренной сложности или, в перспективе, в виде отечественной заказной СБИС (сверхбольшой интегральной схемы).

Другая альтернатива коммутации внутри «системы на кристалле» — концепция централизованной коммутации в гетерогенном распределенном коммутаторе с общей буферной памятью. Она также базируется на «сверхлокальной» сети и предполагает в коммуникационных узлах коммутатора множество ИКМ, но более простых, в пределах однопортовых и относительно небольшой производительности. Суть этой концепции поясняет рис. 3.

Гетерогенный коммутатор предполагает общую для всего коммутатора буферную память, множество коммуникационных узлов, обеспечива-

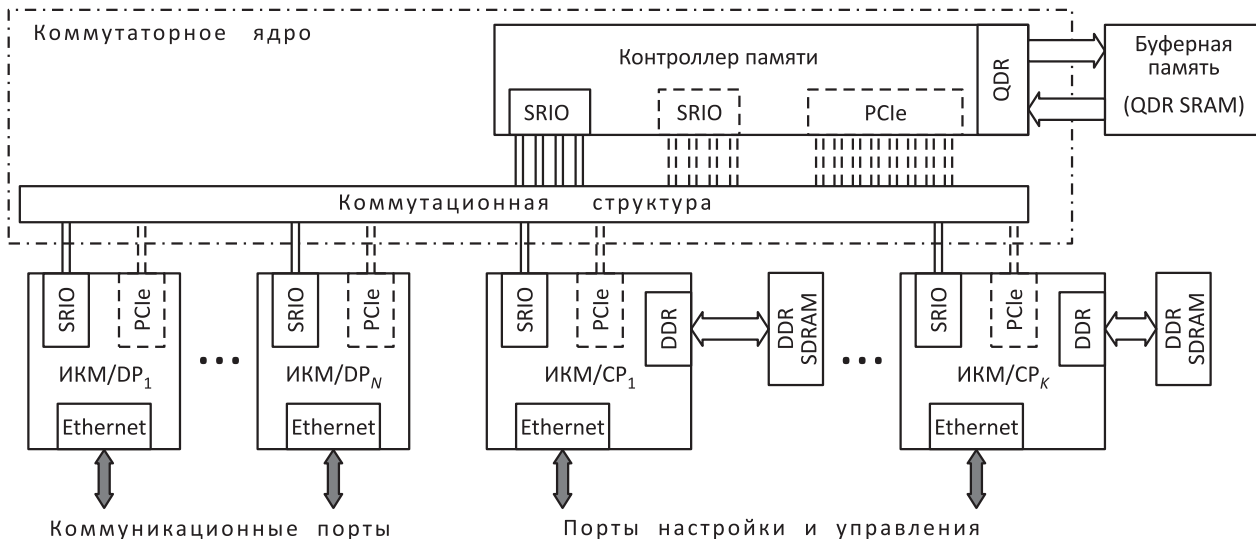


Рис. 3 Гетерогенный коммутатор с общей буферной памятью

ющих коммуникационные порты, и выделенные узлы для выполнения коммутации, маршрутизации и управления коммутатором, причем все узлы также объединяются некоей «сверхлокальной» сетью. Коммуникационные узлы, непосредственно участвующие в продвижении пакетов, — узлы data plane — реализуются на относительно простых ИКМ (на рис. 3 ИКМ/DP_{1...N}) с процессорным ядром умеренной производительности 300...500 DMIPS. Все решения по коммутации и маршрутизации, а также общему управлению коммутатором, т.е. функции control plane, принимаются выделенными ИКМ повышенной производительности (на рис. 3 ИКМ/CP_{1...K}).

Принципиально новым элементом архитектуры коммутатора является общая буферная память, например типа QDR SRAM (quadruple data rate static random access memory), со своим контроллером, который подключается к коммутационной структуре звеном интерфейса «сверхлокальной» сети с высокой пропускной способностью. В частности, в случае интерфейсов SRIO и/или PCie должны быть использованы варианты звеньев максимальной «ширины» и наибольших скоростей передачи данных. Контроллер общей буферной памяти может быть реализован на ПЛИС средней сложности, предоставляющей возможность подключения современных синхронных памяти. В частном, но, как будет показано ниже, важном случае контроллер памяти может быть совмещен с коммутационной структурой в виде некоторого коммутаторного ядра (на рис. 3 оно обведено штрих-пунктиром), для реализации которого, разумеется, потребуется ПЛИС большей сложности.

3 Современные тенденции в области пакетной коммутации

Во второй половине прошлого века телекоммуникация в высоком темпе преодолела несколько знаковых рубежей. Сначала переход от аналоговой технологии связи к цифровой увенчался объединением различных услуг, таких как телефон, факс и передача данных между компьютерами, в единой цифровой сети интегрированных сервисов ISDN, основанной на идее временного разделения традиционных телефонных каналов связи между несколькими соединениями. Однако распространение технологии ISDN на высокоскоростные каналы связи, вылившееся в сложные синхронные и плездохронные иерархии цифровых каналов, не успев толком завершиться, показало свою относительную неэффективность вследствие плохого использования дорогих широкополосных линий связи. Осознание этого принципиального недостатка технологии коммутации каналов стимулировало в телекоммуникации решительный переход к технологиям коммутации пакетов.

Пакетные сети первых поколений, изначально ориентированные только на передачу данных между компьютерами, отличались широким многообразием. В области локальных сетей конкурировало множество технологий, постепенно полностью вытесненных, однако, технологией Ethernet. Несколько меньшее исходное разнообразие региональных пакетных сетей, реализуемых на традиционных телефонных каналах, также постепенно свелось к одной доминирующей технологии X.25. Однако

идея интеграции услуг, уже успешно опробованная в сетях ISDN, плохо ложилась на пакетные сети X.25 из-за ненадежности каналов связи и низких скоростей передачи данных в них. Поэтому с внедрением в региональные и глобальные сети более быстродействующих и надежных каналов, в частности оптоволоконных, появились и более гибкие телекоммуникационные технологии, адекватные требованиям интеграции услуг. Отличительной особенностью новых пакетных технологий стала возможность организации виртуальных соединений, аналогичных аналоговым, т. е., по существу, наложения на пакетные сети виртуальных сетей с коммутацией каналов. В таких сетях передача компьютерных данных могла по-прежнему выполняться традиционными методами, в частности дейтаграммами с динамической маршрутизацией, но для критических сервисов строились прямые виртуальные тракты «точка–точка» от одного конечного абонента к другому и резервировались требуемые полосы пропускания во всех звеньях этих трактов.

Первая такого рода гибридная технология Frame relay, поддерживающая организацию как постоянных, так и коммутируемых виртуальных соединений, наглядно продемонстрировала правильность подхода к пакетным сетям и эффективно решила проблему передачи, в частности, человеческой речи. Затем для передачи через пакетные сети в реальном времени изображения, «живой» телевизионной картинки, потребовались не только линии связи с более широкой полосой пропускания, но и новые более гибкие коммуникационные технологии, обеспечивающие доставку последовательности пакетов изображения получателю в нужном порядке и с постоянным ритмом. Основной в этом классе стала технология АТМ. Еще совсем недавно успешное решение с ее помощью практических проблем организации разнородных трафиков в пакетных сетях регионального масштаба позволило экспертам предсказывать всеобщую «АТМизацию» глобальных пакетных сетей, несмотря на сложности реализации и дороговизну этой технологии. Однако жизнь нашла другое решение и заставила переоценить ценности под влиянием трех мощных факторов технического прогресса:

- (1) широчайшего распространения персональных компьютеров и объединяющих их локальных сетей Ethernet;
- (2) впечатляющего рывка техники передачи данных с мегабитных на гигабитные скорости, т. е. на три порядка (!) менее чем за четверть века;
- (3) огромной популярности Всемирной паутины, сумевшей за два десятилетия охватить практически всю планету.

Персональные компьютеры и локальные сети превратили технологию Ethernet, по первоначальным прогнозам самую дорогую из конкурировавших, в самую дешевую и массово доступную на практике технологию физического и канального уровней. Прогресс в технике передачи данных позволил в исторически короткие сроки создать целую гамму стандартов Ethernet, покрывших диапазон скоростей передачи данных от 10 Мбит/с до 10 Гбит/с и позволивших вывести Ethernet из области сугубо локальной на региональный и даже глобальный уровень. С другой стороны, популярность Интернета сделала его стандарты сетевого и транспортного уровня, «стек» TCP/IP (transport control protocol / Internet protocol), абсолютно доминирующими в мире. В результате сегодня связка Ethernet + IP становится преобладающей в пакетных сетях любого масштаба, от локальных до глобальных, практически лишая шансов все конкурирующие технологии.

Однако отказ от технологий, ориентированных на виртуальные соединения «точка–точка», таких как Frame relay и АТМ, потребовал для поддержки критических трафиков существенной доработки протоколов Интернета. Так, например, в протокол IP были введены понятия обслуживания (в версии 4) или класса трафика (в версии 6). Однако это не решило всех проблем. Для быстрого продвижения пакетов в глобальных дейтаграммных сетях через множество коммутаторов все же пришлось вернуться к идее виртуальных «туннелей» и добавить к связке Ethernet + IP еще один промежуточный уровень (уровень $2^{1/2}$) — технологию MPLS, совместившую в себе простоту Frame relay и гарантированное качество обслуживания QoS (quality of service) АТМ. В результате минимально необходимый и практически достаточный комплект телекоммуникационных технологий, которые необходимо поддерживать современным пакетным коммутаторам, свелся к триаде Ethernet, MPLS и IP. Соответственно, разработка любого нового ИКМ, особенно упрощенного и не предполагающего неременной обратной совместимости с длинным рядом предшественников, могла бы быть ориентирована только на эту триаду.

4 Внутренние интерфейсы «сверхлокальной» сети

Важным элементом обеих концепций «распределенного» коммутатора является его внутренний интерфейс «сверхлокальной» сети, характеристики

которого в значительной степени определяют возможности и параметры всего коммутатора. Исходная концепция децентрализованной коммутации основывалась на некоей шине или системе параллельных шин, слотовых или блочных [11], но современные скорости передачи данных смещают акценты в сторону высокоскоростных последовательных интерфейсов. На сегодняшний день в качестве такого рода интерфейсов конкурируют интерфейсы SRIO и PCIe, в целом очень похожие и близкие по своим характеристикам. Оба интерфейса основаны на пакетной передаче данных по последовательным высокоскоростным каналам. У них сходная физическая реализация: одна или несколько дорожек (lanes), по которым передаются низковольтные дифференциальные сигналы — с той лишь небольшой разницей, что в SRIO принята стандартная сигнализация для витой пары 10-gigabit Ethernet, а PCIe использует собственную оригинальную. Возможности PCIe и SRIO практически равны в части адресации памяти: PCIe допускает длину адреса 32 и 64 разряда, а SRIO — 34, 50 и 66 разрядов. Схожи у обоих интерфейсов механизмы контроля целостности передаваемых данных и подтверждения передачи пакета на звене.

Но есть у этих интерфейсов и существенные различия, вытекающие из их назначения. Интерфейс PCIe предназначен, в первую очередь, для замены традиционной шины PCI, вследствие чего ориентирован на подключение быстродействующей периферии к центральному процессору и подразумевает древовидную топологию с центральным процессором в корне дерева. При этом подключение других процессоров представляет определенные сложности и возможно только через дополнительные нестандартные и непрозрачные порты промежуточных коммутаторов. В противовес этому SRIO изначально создавался как интерфейс взаимодействия множества равноправных процессоров и допускает самые разнообразные топологические воплощения объединяющих такие процессоры коммутационных структур. Ниже более подробно рассмотрены вытекающие из этих основополагающих различий особенности обоих интерфейсов.

Безусловным преимуществом интерфейса PCIe является существенно большая максимальная пропускная способность звена. Интегральная пропускная способность обоих интерфейсов зависит от их ширины, т. е. числа задействованных дорожек, и скорости передачи данных по дорожке. В самом широком варианте с 16 дорожками (принятое обозначение $\times 16$) и на максимальной скорости 5 Гбод PCIe способен обеспечить на звене «чистую» пользовательскую скорость 64 Гбит/с. Интерфейс SRIO в самом широком варианте с четырьмя дорожками

(принятое обозначение $4\times$) и на максимальной скорости 3,125 Гбод предоставляет пользовательскую скорость только 10 Гбит/с.

Формально интерфейс PCIe, использующий общую адресную сетку для адресации памяти и подключаемых устройств ввода-вывода, может адресовать практически бесконечное число абонентов. Однако вряд ли это следует рассматривать как преимущество, поскольку 8- или 16-разрядные идентификаторы SRIO покрывают все реальные потребности адресации конечных абонентов «сверхлокальной» сети и значительно упрощают реализацию коммутационных структур.

Аналогичным образом как относительное можно оценить преимущество механизма реализации QoS на интерфейсе PCIe. Формально восемь классов трафика (traffic classes), отражаемые в восемь виртуальных каналов, и гибкая система арбитража между ними выглядят гораздо более внушительно, чем простые три уровня приоритета пакетов в SRIO. Однако эта система сложна в реализации и требует больших объемов буферов, сепаратно выделяемых в конечном пользователе каждому задействованному сочетанию класса трафика и виртуального канала. В результате большинство практических реализаций интерфейса PCIe поддерживают только единственный класс трафика и один виртуальный канал. Между тем простая приоритетная схема SRIO, дополняемая механизмом CoS (classes of service), допускает в системе тысячи одновременных потоков и десятки тысяч виртуальных каналов для каждого конечного абонента. При этом она ориентирована на общий для пакетов любых приоритетов и потому меньший по суммарному объему буфер в конечном абоненте.

Интерфейс SRIO обладает безусловным преимуществом при взаимодействии множества равноправных ИКМ между собой и в использовании ими общей памяти. Адресный доступ к общей памяти как разделяемому ресурсу поддерживается «неделимыми» (atomic) операциями. Но интерфейс SRIO помимо традиционного адресного доступа предоставляет два дополнительных гибких механизма: обмен сообщениями, исключающий насильственные вторжения в память других конечных абонентов, и инкапсуляцию сквозных потоков (streaming flows), позволяющую осуществлять эффективный транзит высокоскоростных трафиков с произвольными протоколами от одного конечного абонента к другому. Оба этих механизма задействуют процедуру SAR (segmentation and reassembly) как для потоков данных, так и длинных сообщений, что позволяет ограничить длину пакетов 256 байтами и тем самым упростить их буферирование внутри коммутационных структур. Протокол звена интерфейса

Таблица 1 Сравнение интерфейсов PCIe и SRIO

Параметр	PCIe	SRIO
Топология	древовидная	любая
Число конечных абонентов	очень большое	256 или 65 536
Подтверждение операции записи	нет	есть
Поддержка обмена сообщениями	весьма ограниченная	полная и многоадресная
Поддержка режима дейтаграмм	нет	есть
Поля, заменяемые коммутатором	3 поля	нет
Число сигнальных пар	1, 2, 4, 8, 12, 16	1, 4
Скорость передачи данных	2,5; 5,0 Гбод	1,25; 2,5; 3,125 Гбод
Пропускная способность звена	до 64 Гбит/с	до 10 Гбит/с
Механизм QoS	8 классов трафика	3 приоритета пакетов
Аппаратура реализации QoS	более сложная	более простая
Механизмы управления потоком	только на уровне звена	на уровне звена и абонента

SRIO настолько прост, что компонентам коммутационной структуры не нужно менять, в отличие от коммутаторов PCIe, поля в заголовках пакетов и, соответственно, пересчитывать контрольные суммы. Гораздо богаче возможности SRIO в управлении потоком: помимо схожих с интерфейсом PCIe механизмов синхронизации на звене интерфейс SRIO дополнительно включает протокольные средства управления потоком между конечными абонентами, в том числе традиционный механизм *xon/xoff*.

Сводные характеристики обоих интерфейсов представлены в табл. 1, из которой видно, что в целом интерфейс SRIO обладает абсолютными преимуществами для реализации «сверхлокальной» сети гомогенных коммутаторов, но в случае гетерогенного коммутатора он проигрывает интерфейсу PCIe в одном-единственном пункте — пропускной способности звена подключения к коммутационной структуре контроллера общей буферной памяти.

Картину может изменить недавно анонсированная архитектура «сверхлокальной» сети ASI (*advanced switching interconnect*), представляющая собой транспортную надстройку на интерфейс PCIe. Сохраняя тем самым основные достоинства физического уровня PCIe, в первую очередь высокую потенциальную пропускную способность звена, ASI одновременно предоставляет пользователю целый ряд возможностей, ранее свойственных только логическому уровню интерфейса SRIO, в числе которых многоадресная рассылка, инкапсуляция произвольных протоколов, в том числе с механизмом SAR, туннелирование потоков инкапсулированных данных и управление коммутационными структурами.

Кроме того, в гипотетических отечественных ИКМ возможно появление оригинальных высо-

коскоростных последовательных интерфейсов, совмещающих в себе достоинства SRIO и PCIe, что позволит им успешно конкурировать с интерфейсами уровня ASI, по крайней мере, в плане реализации рассматриваемых концепций децентрализованной коммутации. Но такие решения имеют право на жизнь, если только не иметь в виду возможность объединения в одной «сверхлокальной» сети этих гипотетических отечественных ИКМ с реально существующими и будущими приборами компании *Freescale*. В любом случае рассмотрение подобных интерфейсов выходит за рамки настоящей статьи.

5 Упрощенные интегрированные коммутационные микроконтроллеры для пакетной коммутации

В рамках двух концепций децентрализованного пакетного коммутатора — гомогенной и гетерогенной — были, по существу, обозначены два варианта упрощенных ИКМ для децентрализованного коммутатора: большей (свыше 1000 DMIPS) и меньшей (порядка 500 DMIPS) производительности. Они различаются процессорным ядром и некоторыми системными компонентами, но должны иметь минимум два одинаковых функциональных блока: коммуникационный модуль внешнего интерфейса и модуль внутреннего интерфейса «сверхлокальной» сети. С учетом выводов предыдущих разделов в качестве первого необходимо и достаточно иметь трехскоростной (10, 100 и 1000 Мбит/с) Ethernet, а в качестве второго могут выступать интерфейсы SRIO или PCIe. Поскольку интерфейс SRIO обладает рядом существенных преимуществ

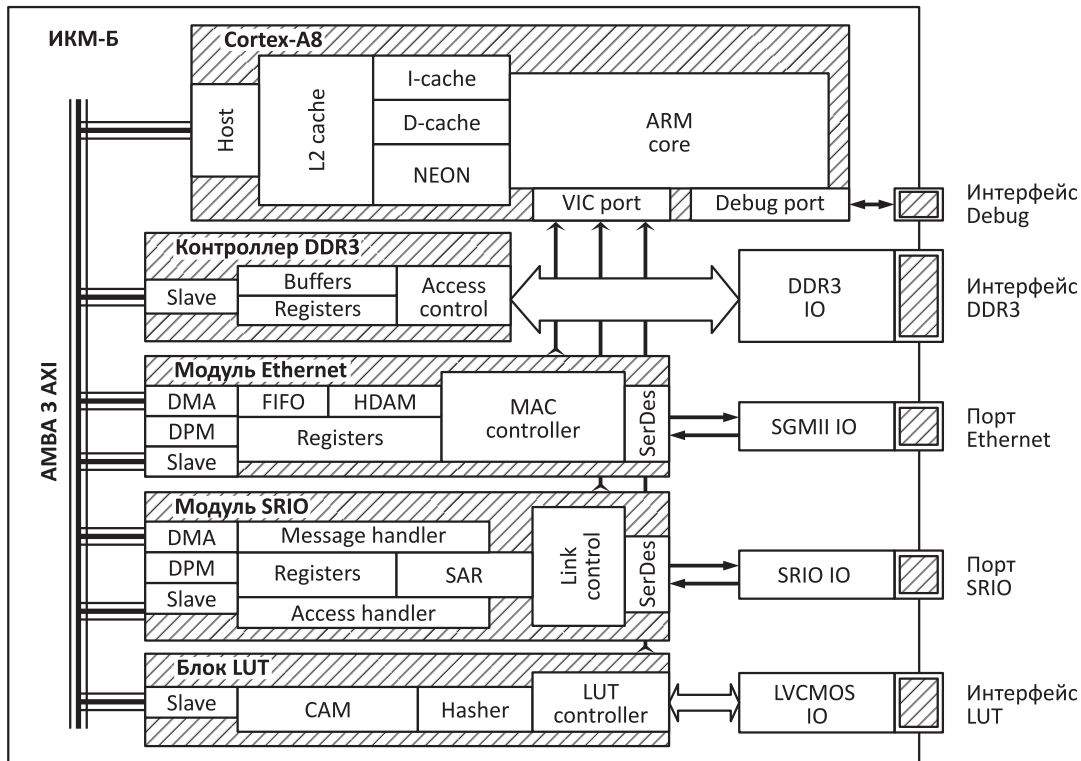


Рис. 4 Структура ИКМ-Б

перед PCIe, далее для определенности он принят в рассматриваемых упрощенных ИКМ в качестве интерфейса «сверхлокальной» сети.

Возможная структура упрощенного ИКМ большей производительности на 1000 DMIPS и более (далее ИКМ-Б) приведена на рис. 4. В целом она не противоречит традиционному подходу (см. рис. 1), но отличается существенно меньшей сложностью, сохраняя при этом достаточно высокую степень универсальности, позволяющую применять такой ИКМ в различных приложениях, в том числе в обеих рассмотренных выше концепциях: в качестве универсального узла «сверхлокальной» сети в гомогенном коммутаторе либо узлов коммутации, маршрутизации и управления гетерогенного коммутатора, т. е. ИКМ/СР на рис. 3.

В качестве процессорного ядра в ИКМ-Б может быть задействован процессор любой архитектуры с суммарной производительностью не хуже 1000 DMIPS. Хорошим решением может стать процессорное ядро с открытой архитектурой ARM, например, как показано на рис. 4, Cortex-A8 [16]. Лицензию и технологию производства различных ядер ARM можно приобретать, в том числе на условиях *royalty free*, у компании *ARM*.

Обязательным элементом любого ИКМ, в том числе ИКМ-Б, должен быть модуль трехскоростно-

го Ethernet, включающий MAC-контроллер, контроллер прямого доступа в память DMA (direct memory access) и специфический блок отделения и слияния заголовков HDAM (headers detach and merge). Внешним интерфейсом может служить любой стандартный MAC-интерфейс гигабитного Ethernet, в частности GMII (gigabit media independent interface), RGMII (reduced GMII) или, в целях экономии числа выводов прибора, последовательный интерфейс SGMII (serial GMII), как показано на рис. 4.

Второй обязательный элемент всех упрощенных ИКМ для пакетной коммутации — это модуль высокоскоростного последовательного интерфейса «сверхлокальной» сети, в частности модуль SRIO. Он обязательно должен включать контроллеры DMA и звена интерфейса (link control), обработчики доступа к общей буферной памяти (access handler) и обмена сообщениями (message handler). Также весьма желательно для улучшения условий доставки инкапсулированных блоков данных через коммутационные структуры, чтобы в него входил блок SAR. Сегментация, с одной стороны, упростит коммутационную структуру, а с другой — ускорит доставку пакетов между узлами «сверхлокальной» сети и облегчит реализацию сопутствующих требований QoS.

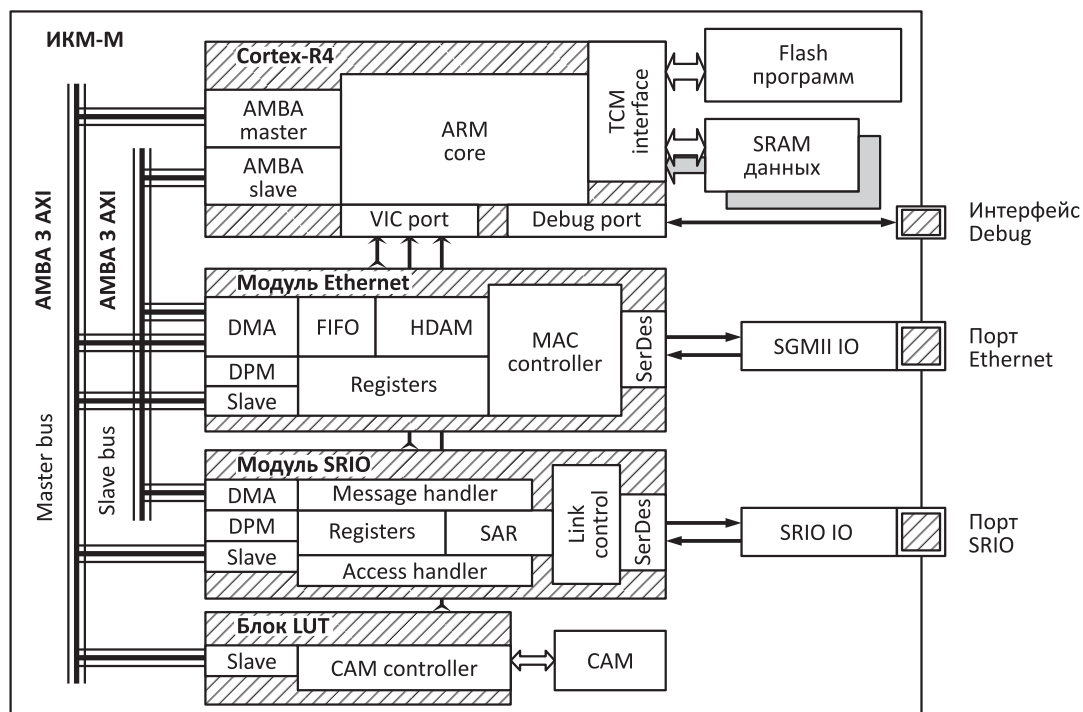


Рис. 5 Структура ИКМ-М

В предположении, что контроллеры DMA умеют работать в свойственном классическим ИКМ активном режиме, используя механизм цепочек описателей буферов [4], модули Ethernet и SRIO могут включать небольшого объема двупортовые памяти DPM (dual-port memories) для циркулярных цепочек такого рода описателей.

Специфическим именно для ИКМ-Б блоком должен быть контроллер внешней памяти большого объема, например DDR3, необходимой для буферирования блоков данных при реализации узла гомогенного коммутатора или хранения маршрутной и служебной информации при реализации коммутирующего процессора гетерогенного коммутатора.

Для ускорения адресной фильтрации и классификации входящих кадров в ИКМ-Б может быть предусмотрен блок быстрого доступа к адресным таблицам LUT (look-up table) и/или другие блоки, выполняющие функции, аналогичные функциям сопроцессоров табличного и контекстного поиска в классических ИКМ (см. рис. 1). Блок LUT может также включать небольшую, размером в несколько элементов, ассоциативную память CAM (context addressable memory) для хранения текущих MAC-адресов и как возможное расширение — интерфейс для подключения внешней табличной памяти.

Общая организация гомогенного коммутатора и место в нем ИКМ-Б достаточно очевидны. Каждый

ИКМ работает здесь как сетевой узел, используя любую известную сетевую технологию продвижения пакетов по «сверхлокальной» сети, возможно, адаптированную к особенностям конкретного интерфейса — SRIO или другого. Все ИКМ коммутатора совокупно выполняют маршрутизацию, поддерживая соответствующие маршрутизирующие протоколы на своих коммуникационных портах и обмениваясь между собой маршрутной информацией в форме сообщений.

На рис. 5 показана структура ИКМ меньшей производительности на 500...1000 DMIPS (далее ИКМ-М), которые могут использоваться в узлах продвижения пакетов, т.е. в качестве ИКМ/DP на рис. 3, и, возможно, некоторых узлах менеджмента гетерогенных коммутаторов, т.е. в качестве ИКМ/CP на рис. 3.

В ИКМ-М может быть задействовано более простое процессорное ядро, чем в ИКМ-Б, например ядро Cortex-R4 [17] с производительностью порядка 750 DMIPS. Сравнительные характеристики обоих ядер ARM, Cortex-A8 и Cortex-R4, приведены в табл. 2.

Еще одно отличие ИКМ-М от ИКМ-Б может заключаться в отказе от использования внешней памяти и, соответственно, исключении аппаратуры ее интерфейсов. Разумеется, такой шаг предполагает наличие в ИКМ-М достаточных ресурсов внутрикристалльной памяти. Это могут быть памяти

Таблица 2 Сравнение ядер Cortex-A8 и Cortex-R4

Параметр	Cortex-A8	Cortex-R4
L1 кэш	переменный	0..64 К
L2 кэш	0..1 М	нет
Тесно связанная память	нет	есть
Внешняя шина	AMBA 3 AXI	2 × AMBA 3 AXI
Технологии NEON и Jazelle RCT	есть	нет
Процесс	65 нм (GP)	90 нм (HS)
Максимальная частота	1,1 ГГц	475 МГц
Максимальное быстродействие	2000 DMIPS	760 DMIPS
Напряжение питания	1,0 В	1,0 В
Занимаемая площадь кристалла	4 мм ²	1,74 мм ²
Потребляемая мощность	0,5 Вт	150 мВт

программ (ROM или Flash объемом до 512 Кбайт) и данных (SRAM до 64 Кбайт). Память данных может дополняться блоками DPM модулей Ethernet и SRIO, а также блоком LUT с небольшого объема памятью CAM для текущих MAC-адресов.

Необходимо отметить, что исключение из ИКМ интерфейсов внешних памяти резко сокращает число выводов корпуса ИКМ-М и, следовательно, в значительной степени стоимость прибора. В структуре, показанной на рис. 5, ИКМ-М требуется всего полтора десятка (!) сигнальных выводов. Для сравнения современные ИКМ компании *Freescale* выпускаются в корпусах с числом выводов, переваливающим за тысячу.

Если организация гомогенного коммутатора и использование в ней ИКМ были тривиальны, то работа гетерогенного коммутатора требует пояснений.

Дезинтегрированная архитектура пакетной коммутации [11] предполагает, что входные порты расчлениют входящие Ethernet-кадры: заголовки кадра остаются в микроконтроллере порта (в случае ИКМ-М — в его внутренней памяти данных), а «обезглавленное» тело кадра сохраняется в общей буферной памяти коммутатора. В тех случаях, когда адрес назначения в оторванном заголовке принятого кадра имеется во внутренней CAM, ИКМ-М входного узла способен сам принять решение по классификации и продвижению этого кадра. После принятия решения о продвижении ИКМ-М отправляет в соответствующий выходной узел *сообщение продвижения* с модифицированным заголовком скоммутированного кадра и ссылкой на буфер в общей памяти, где сохранено тело этого кадра. В остальных случаях ИКМ-М входного узла отправляет *сообщение коммутации*, включающее оторванный заголовок и ссылку на тело соотнесенного кадра, коммутирующему процессору. Коммутирующий процессор, в типичном случае ИКМ-Б,

после определения порта назначения высылает в соответствующий выходной узел *сообщение продвижения* с новым заголовком и ссылкой на тело соотнесенного кадра, а во входной узел — *сообщение обновления*, которое позволит ИКМ-М входного узла обновить содержимое своей CAM и в дальнейшем коммутировать кадры с теми же MAC-адресами самостоятельно. ИКМ-М выходного узла, получив *сообщение продвижения*, ставит полученный заголовок со ссылкой на соотнесенное тело в одну из очередей на отправку в соответствии с требованиями QoS. Когда очередь доходит до этого заголовка, ИКМ-М склеивает его с телом кадра из буферной памяти и включает ссылку на готовый кадр в выходную цепочку описателей буферов блока Ethernet. После отправки кадра (или в зависимости от протокола после подтверждения его получения адресатом) ИКМ выходного узла отправляет *сообщение освобождения* буфера в общей буферной памяти входному узлу, после чего процесс коммутации завершается.

Если гетерогенный коммутатор включает отдельный узел или даже несколько узлов общего управления и менеджмента, то в таких узлах могут использоваться в зависимости от сложности решаемых задач как ИКМ-М, так и ИКМ-Б.

Таким образом, и децентрализованная коммутация в гомогенном коммутаторе, и централизованная коммутация в гетерогенном коммутаторе предполагают на интерфейсе «сверхлокальной» сети:

- потоковые пересылки больших инкапсулированных блоков данных (кадров между ИКМ или тел кадров между ИКМ и общей буферной памятью);
- передачу между ИКМ коротких высокоприоритетных сообщений.

Длина инкапсулированных блоков данных не имеет принципиального значения при условии их

фрагментации, что характерно, например, для интерфейса SRIO. Длина сообщений определяется в основном размером оторванных заголовков кадров. В зависимости от уровня коммутации в коммутирующем процессоре это могут быть заголовки только L2 (кадра Ethernet) или также заголовки L3 и даже L4 (пакетов IP и TCP/UDP). Соответственно, длина сообщений может варьироваться от 32 байт в случае коммутации L2 до 128 байт при коммутации L4 и версии 6 протокола IP. Все это вполне соответствует возможностям интерфейса SRIO, который поддерживает эффективный обмен сообщениями размером до 256 байт, а с использованием механизма SAR — до 4096 байт.

Суммарная производительность гетерогенного коммутатора определяется двумя узкими местами: быстродействием общей буферной памяти и пропускной способностью звена интерфейса между контроллером буферной памяти с коммутационной структурой.

Современная синхронная память способна пропустить через себя десятки гигабайт информации в секунду. Например, 64-разрядная память QDR SRAM (4 прибора IDT71P74604), работающая на частоте 250 МГц, пропускает через себя дуплексный поток 32 Гбит/с и, следовательно, может обслужить до 32 дуплексных гигабитных портов Ethernet. Однако максимальная пропускная способность звена SRIO $4\times$ составляет всего 10 Гбит/с, что ограничивает число гигабитных портов гетерогенного коммутатора десятью. Увеличивать пропускную способность интерфейса буферной памяти можно, добавляя звенья интерфейса SRIO, как показано пунктиром на рис. 3. Однако такое решение предполагает динамическую маршрутизацию коммутатором SRIO потоков между параллельными звеньями, что нетипично для существующих коммутаторов SRIO.

Другое решение заключается в замене для передачи тел блоков данных интерфейса SRIO интерфейсом PCIe, который обладает гораздо большей пропускной способностью: на звене $\times 16$ она достигает 64 Гбит/с, что позволяет обслуживать до 64 гигабитных портов. Но PCIe не поддерживает необходимый в децентрализованном пакетном коммутаторе обмен сообщениями, поэтому такое решение заставляет вводить в ИКМ-М в дополнение к модулю SRIO еще и блок PCIe, что выглядит неоправданно избыточным.

Проблема может быть автоматически решена в будущем новыми версиями спецификации интерфейса SRIO, в которых пропускная способность звена будет существенно повышена до конкурентных с PCIe величин, либо переходом на архитектуру ASI. Но возможно и другое решение *ad hoc* разра-

боткой в виде отдельной микросхемы контроллера общей буферной памяти гетерогенного коммутатора, совмещенного с объединяющим ИКМ коммутатором SRIO, т. е. коммутаторного ядра на рис. 3.

6 Заключение

Интегрированные коммутационные микроконтроллеры — общепризнанный строительный материал для средств телекоммуникации и пакетной коммутации, который особенно удобен для типового «крупноблочного» строительства. Устройства на основе ИКМ могут быть широко востребованы в самых разных приложениях. Но вряд ли устройства, реализованные на ИКМ, достигнут рекордных показателей в своем классе. Характеристики таких устройств, скорее всего, окажутся на приличном среднем уровне, который достаточен для большинства типичных применений. Зато разработка на базе ИКМ может быть завершена быстро и с минимальными затратами.

Хотя приборы ведущего поставщика ИКМ компании *Freescale* становятся все более популярными и в России, сдерживающим фактором их активного использования в отечественных разработках остается цена, которая в зависимости от сложности микросхемы колеблется от десятков до сотен долларов США. Кроме того, есть области, связанные с телекоммуникацией и пакетной коммутацией, в которых применение импортных комплектующих, в том числе ИКМ, нежелательно или невозможно по различным причинам, в том числе нетехнического характера. В этих областях были бы активно востребованы приборы отечественного производства. Однако разработка в России современных ИКМ, эквивалентных по своим возможностям лучшим образцам от компании *Freescale*, сегодня совершенно нереальна. Настоящая статья указывает один из путей решения проблемы разработкой ограниченной номенклатуры отечественных ИКМ, существенно более простых, чем приборы компании *Freescale*, но позволяющих, тем не менее, решать серьезные задачи на современном техническом уровне как минимум в области пакетной коммутации.

Таким решением могли бы стать два ИКМ, меньшей и большей производительности, с программируемыми процессорными ядрами семейства Cortex, лицензионно поставляемых компанией *ARM*, и оснащенных по минимуму всего-навсего двумя специализированными коммуникационными модулями: трехскоростным Ethernet и интерфейсом (типа) SRIO. С использованием этих двух

вариантов ИКМ можно было бы строить высокопроизводительные «распределенные» маршрутизирующие коммутаторы с гомогенной и гетерогенной организацией. Такие коммутаторы, способные предоставить пользователю до нескольких десятков гигабитных портов Ethernet, могли бы, помимо коммутации на уровнях 2..4, также выполнять функции граничных маршрутизаторов в сетях Metro Ethernet и MPLS, обеспечивать сетевую защиту и эффективно решать целый ряд других смежных задач по мере развития собственного прикладного программного обеспечения.

Можно ожидать, что предлагаемые упрощенные ИКМ при нормально организованном крупносерийном производстве могли бы оказаться существенно дешевле более сложных ИКМ компании *Freescale* и вполне конкурентоспособны во многих приложениях, причем не только на внутреннем, но и на мировом рынке.

Литература

1. *Егоров В. Б.* Новое поколение коммуникационных микроконтроллеров компании Freescale Semiconductor // *Chip News*, 2007. No. 3. С. 14–18.
2. *Егоров В. Б.* Интегрированные коммуникационные процессоры компании Freescale Semiconductor // *Электронные компоненты*, 2007. № 8. С. 85–89.
3. PowerQUICC Communications Processors // Сайт компании Freescale Semiconductor: <http://www.freescale.com/webapp/sps/site/homepage.jsp?nodeId=0162468rH3bTdGJk19>.
4. *Егоров В. Б.* Интегрированные коммуникационные микроконтроллеры компании «Freescale Semiconductor» и их применение в пакетной коммутации. — М.: ИПИ РАН, 2008. 129 с.
5. *Соколов И. А., Гребенищikov В. И., Егоров В. Б. и др.* Прототип многофункционального устройства пакетной коммутации // *Системы и средства автоматизи.* Вып. 16. — М.: Наука, 2006. С. 449–462.
6. *Егоров В. Б.* Опыт разработки пакетных коммутаторов и маршрутизаторов на основе интегрированных коммуникационных процессоров // *Материалы Международной научно-технической конференции “Intermatic-2007”*. Москва. Ч. 2. — М.: МИРЭА, 2007. С. 148–151.
7. *Егоров В. Б.* Принципы создания коммутационной аппаратуры на основе специализированных микроконтроллеров // *Системы и средства автоматизи.* Вып. 9. — М.: Наука, 1999. С. 44–55.
8. *Соколов И. А., Егоров В. Б.* Дезинтеграционный подход к архитектуре универсального процессора коммутации пакетов // *Информационные технологии и вычислительные системы*, 2005. № 2. С. 76–85.
9. MPC8568E PowerQUICC III Processor // Сайт компании Freescale Semiconductor: http://www.freescale.com/webapp/sps/site/prod_summary.jsp?code=MPC8568E&nodeId=0162468rH3bTdGJk196465.
10. *Егоров В. Б.* О возможном подходе к созданию отечественных технических средств пакетной коммутации двойного назначения // *Информационные технологии управления информационными ресурсами двойного применения (ИП)* / Препринт (ДСП). — М.: ИПИ РАН, 2007. С. 26–34.
11. *Соколов И. А., Егоров В. Б.* Дезинтегрированная архитектура пакетной коммутации // *Информатика и её применения*, 2008. Т. 2. Вып. 4. С. 2–11.
12. MPC8548E PowerQUICC III Processors // Сайт компании Freescale Semiconductor: http://www.freescale.com/webapp/sps/site/prod_summary.jsp?code=MPC8548E&nodeId=0162468rH3bTdGJk196465.
13. RapidIO Specifications // RapidIO WEB site: <http://www.rapidio.org/specs/current>.
14. Serial RapidIO Preprocessing Switches // Сайт компании IDT: <http://www.idt.com/?catID=18388262>.
15. RapidIO Switches // Сайт компании Tundra: <http://www.tundra.com/products/rapidio-switches>.
16. Application Processors // Сайт компании ARM: <http://www.arm.com/products/CPUs/application.html>.
17. Embedded Processors // Сайт компании ARM: <http://www.arm.com/products/CPUs/embedded.html>.

О НЕСТАЦИОНАРНЫХ СИСТЕМАХ ОБСЛУЖИВАНИЯ С КАТАСТРОФАМИ*

А. И. Зейфман¹, Я. А. Сатин², А. В. Чегодаев³

Аннотация: Рассматриваются модели обслуживания, описываемые процессами рождения и гибели (ПРГ) с катастрофами. Получены как оценки скорости сходимости к предельному режиму, так и оценки различных характеристик этого предельного режима. Рассмотрены также вопросы построения предельных характеристик и пример конкретной системы обслуживания.

Ключевые слова: нестационарные системы обслуживания; марковские модели с катастрофами; слабая эргодичность; оценки; предельные характеристики; аппроксимация

1 Введение

Простейшие модели систем массового обслуживания с катастрофами начали изучаться несколько лет назад (см., например, подробную мотивацию и первые результаты в работах [1, 2]). Несколько другой подход для изучения близких моделей был применен в [3]. Обсуждение современных исследований в этой области и некоторые новые результаты для общих стационарных ПРГ с катастрофами приведены в [4]. Нестационарная система обслуживания типа $M(t)/M(t)/S$ с катастрофами была изучена недавно в [5]. В настоящей работе будет рассмотрен более общий класс моделей, описываемых нестационарными ПРГ с катастрофами.

Предлагаемый подход базируется на методе, возможность применения которого впервые была отмечена в заметке Б. В. Гнеденко и И. П. Макарова [6], а развитие его было проведено в работах одного из авторов настоящей статьи [7, 8].

Пусть $X = X(t)$, $t \geq 0$ — ПРГ с катастрофами, а $\lambda_n(t)$, $\mu_n(t)$ и $\xi(t)$ — интенсивности рождения, гибели и катастрофы соответственно.

Обозначим через

$$p_{ij}(s, t) = Pr \{X(t) = j | X(s) = i\}, \\ i, j \geq 0, 0 \leq s \leq t,$$

переходные вероятности процесса $X = X(t)$, а через $p_i(t) = Pr \{X(t) = i\}$ — его вероятности состояний.

Тогда вероятности состояний (при выполнении некоторых естественных дополнительных условий)

удовлетворяют прямой системе дифференциальных уравнений Колмогорова

$$\begin{cases} \frac{dp_0}{dt} = -(\lambda_0(t) + \xi(t))p_0 + \mu_1(t)p_1 + \xi(t), \\ \frac{dp_k}{dt} = \lambda_{k-1}(t)p_{k-1} - (\lambda_k(t) + \mu_k(t) + \xi(t))p_k + \mu_{k+1}(t)p_{k+1}, \quad k \geq 1. \end{cases} \quad (1)$$

Обозначим через

$$\mathbf{p}(t) = (p_0(t), p_1(t), \dots)^T, \quad t > 0,$$

вектор-столбец вероятностей состояний, а через $\mathbf{A}(t) = \{a_{ij}(t), t \geq 0\}$ — матрицу, порождаемую системой (1), при этом элементы матрицы $\mathbf{A}(t)$ определяются по формулам

$$a_{ij}(t) = \begin{cases} \lambda_{i-1}(t) & \text{при } j = i - 1, \\ \mu_{i+1}(t) & \text{при } j = i + 1, \\ -(\lambda_i(t) + \mu_i(t) + \xi(t)) & \text{при } j = i, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Ограничимся здесь рассмотрением только таких процессов, интенсивности которых можно представить в следующем виде:

$$\lambda_n(t) = \nu_n \lambda(t), \quad \mu_n(t) = \eta_n \mu(t), \quad t \geq 0, \quad n \in E,$$

предполагая, что числовые множители ограничены, т. е. $0 \leq \eta_n \leq M$, $0 \leq \nu_n \leq M$ (см. подробное рассмотрение в [9]).

* Исследование поддержано грантом РФФИ 06-01-00111 и научным грантом Вологодской области.

¹ Вологодский государственный педагогический университет, Институт проблем информатики РАН и ВНКЦ ЦЭМИ РАН, a.zeifman@mail.ru

² Вологодский государственный педагогический университет, yacovi@mail.ru

³ Вологодский государственный педагогический университет, cheg_al@mail.ru

Теперь можно систему (1) рассмотреть как дифференциальное уравнение

$$\frac{d\mathbf{p}}{dt} = \mathbf{A}(t)\mathbf{p} + \mathbf{g}(t), \quad t \geq 0, \quad (2)$$

в пространстве последовательностей l_1 , где

$$\mathbf{g}(t) = (\xi(t), 0, 0, \dots)^T.$$

Обозначим через $\Omega = \{\mathbf{x} : \mathbf{x} \geq 0, \|\mathbf{x}\|_1 = 1\}$ множество всех стохастических векторов.

Далее всюду будем считать «базисные» функции $\lambda(t)$, $\mu(t)$ и $\xi(t)$ локально интегрируемыми на $[0; \infty)$. Более того, для простоты вычислений будем предполагать эти функции ограниченными, т. е. будем считать, что при некотором L выполняется неравенство

$$\lambda(t) + \mu(t) + \xi(t) \leq L < \infty$$

почти при всех $t \geq 0$.

Тогда

$$\|A(t)\|_1 = \sup_j \sum_i |a_{ij}(t)| \leq 2ML$$

почти при всех $t \geq 0$, а значит, задача Коши для уравнения (2) с начальным условием $\mathbf{p}(0)$ имеет единственное решение

$$\mathbf{p}(t) = U(t)\mathbf{p}(0) + \int_0^t U(t, \tau)\mathbf{g}(\tau) d\tau,$$

где $U(t, s)$ — оператор Коши уравнения (2). При этом если $\mathbf{p}(s) \in \Omega$, то и $\mathbf{p}(t) \in \Omega$ при любом $t \geq s$.

2 Оценки, случай большой интенсивности катастроф

Теорема 1. Пусть

$$\int_0^\infty \xi(t) dt = \infty. \quad (3)$$

Тогда $X(t)$ слабо эргодичен в равномерной операторной топологии. При этом справедлива оценка

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 2e^{-\int_0^t \xi(\tau) d\tau}$$

для любых начальных условий $\mathbf{p}^*(0)$, $\mathbf{p}^{**}(0)$.

Доказательство. Здесь и далее будет использоваться понятие логарифмической нормы операторной функции и соответствующие оценки

(см. [9, 10]. В условиях теоремы 1 имеем (через $\gamma(A(t))_1$ обозначена логарифмическая норма в пространстве l_1):

$$\gamma(A(t))_1 = \sup_i \left(a_{ii}(t) + \sum_{j \neq i} a_{ji}(t) \right) = -\xi(t).$$

А тогда справедлива следующая (точная!) оценка:

$$\|U(t, s)\| \leq e^{-\int_s^t \xi(\tau) d\tau}$$

для всех $0 \leq s \leq t$, а значит, справедливо и требуемое утверждение.

Замечание 1. В общем случае можно рассматривать любой режим распределения вероятностей состояний $\mathbf{p}^*(t)$ в качестве предельного. Однако если все интенсивности ($\lambda(t)$, $\mu(t)$ и $\xi(t)$) 1-периодичны, то существует 1-периодический предельный режим, скажем $\pi(t) = (\pi_0(t), \pi_1(t), \dots)^T$.

Далее будем изучать математические ожидания типа

$$E_{\mathbf{p}(0)}(t) = E_{\mathbf{p}(0)}\{X(t)\} = E\{X(t) | \mathbf{p}(0)\}$$

и, в частности,

$$E_k(t) = E\{X(t) | X(0) = k\}.$$

Выпишем первые простейшие оценки.

Теорема 2. Пусть выполнено условие (3). Тогда

$$\begin{aligned} e^{-\int_0^t \xi(\tau) d\tau} E_{\mathbf{p}(0)}(0) - M \int_0^t \mu(\tau) e^{-\int_\tau^t \xi(s) ds} d\tau &\leq \\ &\leq E_{\mathbf{p}(0)}(t) \leq e^{-\int_0^t \xi(\tau) d\tau} E_{\mathbf{p}(0)}(0) + \\ &+ M \int_0^t \lambda(\tau) e^{-\int_\tau^t \xi(s) ds} d\tau \end{aligned} \quad (4)$$

при всех $t \geq 0$ и любом $\mathbf{p}(0)$.

Доказательство. Из (1) получаем:

$$\frac{dE_{\mathbf{p}(0)}(t)}{dt} = \sum_{k \geq 0} (\lambda_k(t) - \mu_k(t) - k\xi(t)) p_k(t).$$

Отсюда

$$\frac{dE_{\mathbf{p}(0)}(t)}{dt} \leq M\lambda(t) - \xi(t)E_{\mathbf{p}(0)}(t) \quad (5)$$

и

$$\frac{dE_{\mathbf{p}(0)}(t)}{dt} \geq -M\mu(t) - \xi(t)E_{\mathbf{p}(0)}(t). \quad (6)$$

Теперь (4) вытекает из (5) и (6).

Определение 1. Будем говорить, что марковская цепь $X(t)$ имеет предельное среднее $\varphi(t)$, если

$$\lim_{t \rightarrow \infty} (\varphi(t) - E_k(t)) = 0$$

при любом k .

Теорема 3. Пусть интенсивности процесса $\lambda(t)$, $\mu(t)$, $\xi(t)$ 1-периодичны. Пусть, далее,

$$\int_0^1 \xi(t) dt > 0. \quad (7)$$

Тогда $X(t)$ имеет 1-периодическое предельное среднее $\varphi(t)$.

Доказательство. Прежде всего отметим, что в условиях теоремы найдется $\delta > 1$, при котором

$$\int_0^1 (\xi(t) - M(\delta - 1)\lambda(t)) dt > 0. \quad (8)$$

Рассмотрим матрицу

$$D = \text{diag}(1, \delta, \delta^2, \dots)$$

и пространство последовательностей \mathcal{B} таких, что

$$\|\mathbf{x}\|_{\mathcal{B}} = \sum_{i=0}^{\infty} \delta^i |x_i| < \infty.$$

Будем исследовать теперь прямую систему Колмогорова (2) как уравнение в пространстве \mathcal{B} . Тогда логарифмическую норму $\gamma(A(t))$ в \mathcal{B} можно оценить следующим образом:

$$\begin{aligned} \gamma(A)_{\mathcal{B}} &= \sup_{i \geq 0} (\delta \lambda_i(t) - (\lambda_i(t) + \mu_i(t) + \xi(t)) + \\ &+ \delta^{-1} \mu_i(t)) \leq \sup_{i \geq 0} ((\delta - 1) \lambda_i(t) - \xi(t)) \leq \\ &\leq -(\xi(t) - M(\delta - 1)\lambda(t)). \end{aligned}$$

А тогда имеем

$$\|U(t, s)\|_{\mathcal{B}} \leq e^{-\int_s^t (\xi(\tau) - M(\delta - 1)\lambda(\tau)) d\tau}$$

при всех $0 \leq s \leq t$.

Следовательно,

$$\begin{aligned} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{\mathcal{B}} &\leq \\ &\leq e^{-\int_0^t (\xi(\tau) - M(\delta - 1)\lambda(\tau)) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{\mathcal{B}} \quad (9) \end{aligned}$$

для любых допустимых начальных условий $\mathbf{p}^*(0)$, $\mathbf{p}^{**}(0)$.

С другой стороны, используя условия теоремы, получаем следующую оценку:

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\mathbf{p}(t)\|_{\mathcal{B}} &\leq \limsup_{t \rightarrow \infty} (\|U(t)\mathbf{p}(0)\|_{\mathcal{B}} + \\ &+ \int_0^t \|U(t, \tau)\mathbf{g}(\tau)\|_{\mathcal{B}} d\tau) \leq \\ &\leq \limsup_{t \rightarrow \infty} \int_0^t \xi(\tau) e^{-\int_{\tau}^t (\xi(u) - M(\delta - 1)\lambda(u)) du} d\tau = \\ &= M < \infty \end{aligned}$$

при любом $\mathbf{p}(0)$.

Положим

$$W = \sup_n \frac{n}{\delta^n} < \infty.$$

Тогда получаем

$$\begin{aligned} \limsup_{t \rightarrow \infty} E_{\mathbf{p}(0)}(t) &= \limsup_{t \rightarrow \infty} \sum_{k=0}^{\infty} k p_k(t) \leq \\ &\leq W \limsup_{t \rightarrow \infty} \|\mathbf{p}(t)\|_{\mathcal{B}} \leq WM \end{aligned}$$

при любом допустимом $\mathbf{p}(0)$.

Выберем теперь $\mathbf{p}^*(0) = \pi(0)$, $\varphi(t) = \sum_{k=0}^{\infty} k \pi_k(t)$ и $\mathbf{p}^{**}(0) = \mathbf{p}(0) = \mathbf{e}_0$. Тогда в (9) имеем:

$$\begin{aligned} \|\pi(t) - \mathbf{p}(t)\|_{\mathcal{B}} &\leq \\ &\leq e^{-\int_0^t (\xi(\tau) - M(\delta - 1)\lambda(\tau)) d\tau} \sum_{k=0}^{\infty} \delta^k \pi_k(0) \leq \\ &\leq M e^{-\int_0^t (\xi(\tau) - M(\delta - 1)\lambda(\tau)) d\tau}. \end{aligned}$$

Окончательно получаем следующую оценку:

$$|\varphi(t) - E_0(t)| \leq WM e^{-\int_0^t (\xi(\tau) - M(\delta - 1)\lambda(\tau)) d\tau}. \quad (10)$$

Правая часть (10) стремится к нулю при $t \rightarrow \infty$ в соответствии с (8). Легко убедиться, что тогда и $|\varphi(t) - E_k(t)| \rightarrow 0$ при $t \rightarrow \infty$ для *любого* k . А значит, $\varphi(t)$ — 1-периодическое предельное среднее для $X(t)$.

Следствие 1. Пусть выполнены условия теоремы 3. Тогда справедлива оценка скорости сходимости к предельному среднему (10).

Замечание 2. Предельное среднее существует, в принципе, независимо от свойств интенсивностей как функций времени. Однако в общем случае в качестве предельного среднего можно выбирать *любое* среднее, поскольку никаких его особых свойств гарантировать нельзя. Достаточным условием существования в этом общем случае будет (8) вместо (7).

3 Оценки, общий случай

Пусть d_i — некоторые положительные числа. Положим

$$\alpha_k(t) = \lambda_k(t) + \mu_{k+1}(t) - \frac{d_{k+1}}{d_k} \lambda_{k+1}(t) - \frac{d_{k-1}}{d_k} \mu_k(t), \quad k \geq 0, \quad (11)$$

и

$$\alpha(t) = \inf_{k \geq 0} \alpha_k(t).$$

Теорема 4. Пусть заданы интенсивности $\lambda_k(t)$ и $\mu_k(t)$ для процесса рождения и гибели с катастрофами $X(t)$.

Пусть существует последовательность положительных чисел $\{d_j\}$ такая, что $1 \leq d_1 \leq d_2 \leq \dots$, и

при этом $\int_0^\infty \alpha(t) dt = +\infty$. Тогда $X(t)$ слабо эргодичен для любой $\xi(t)$ и справедлива следующая оценка:

$$\begin{aligned} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_B &\leq \\ &\leq 2e^{-\int_s^t \alpha(\tau) d\tau} \|\mathbf{p}^*(s) - \mathbf{p}^{**}(s)\|_{1D} \end{aligned}$$

при всех $s, t, 0 \leq s \leq t$, и любых допустимых начальных условиях $\mathbf{p}^*(s), \mathbf{p}^{**}(s)$.

Доказательство. Поскольку $\mathbf{p}(t) \in \Omega$ при всех $t \geq s$, можно положить $p_0(t) = 1 - \sum_{i \geq 1} p_i(t)$

(для обычных ПРГ этот подход подробно описан, например, в [9]), и тогда из (2) получим систему

$$\begin{pmatrix} \frac{dp_1}{dt} \\ \frac{dp_2}{dt} \\ \vdots \\ \frac{dp_n}{dt} \\ \vdots \end{pmatrix} = \begin{pmatrix} -(\lambda_0 + \lambda_1 + \mu_1 + \xi_1) & (\mu_2 - \lambda_0) & -\lambda_0 & -\lambda_0 & \dots & \dots \\ \lambda_1 & -(\lambda_2 + \mu_2 + \xi_2) & \mu_3 & 0 & 0 & \dots \\ 0 & \lambda_2 & -(\lambda_3 + \mu_3 + \xi_3) & \mu_4 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \\ \vdots \end{pmatrix} + \begin{pmatrix} \lambda_0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}$$

или, в векторном виде,

$$\frac{d\mathbf{z}(t)}{dt} = B(t)\mathbf{z}(t) + \mathbf{f}(t). \quad (12)$$

Решение этого неоднородного уравнения можно записать в виде

$$\mathbf{z}(t) = V(t, 0)\mathbf{z}(0) + \int_0^t V(t, z)\mathbf{f}(z) dz,$$

где $V(t, z)$ — оператор Коши уравнения (12).

Рассмотрим матрицу

$$D = \begin{pmatrix} d_0 & d_0 & d_0 & \dots \\ 0 & d_1 & d_1 & \dots \\ 0 & 0 & d_2 & \dots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

и пространство последовательностей

$$\ell_{1D} = \{ \mathbf{z} = (p_1, p_2, \dots) : \|\mathbf{z}\|_{1D} = \|D\mathbf{z}\|_1 < \infty \}.$$

Имеем

$$D^{-1} = \begin{pmatrix} d_0^{-1} & -d_1^{-1} & 0 & \dots & \dots \\ 0 & d_1^{-1} & -d_2^{-1} & 0 & \ddots \\ \vdots & 0 & d_2^{-1} & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \dots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Рассмотрим теперь логарифмическую норму $\gamma(B(t))_{1D} = \gamma(DB(t)D^{-1})_1$. Имеем

$$\begin{aligned} DB(t)D^{-1} &= \\ &= \begin{pmatrix} -(\lambda_0 + \mu_1 + \xi) & \frac{d_0\mu_1}{d_1} & 0 & \dots & \dots \\ \frac{d_1\lambda_1}{d_0} & -(\lambda_1 + \mu_2 + \xi) & \frac{d_1\mu_2}{d_2} & 0 & \ddots \\ 0 & \frac{d_2\lambda_2}{d_1} & \ddots & \ddots & \ddots \\ \vdots & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & \dots & \ddots & \ddots \end{pmatrix} \end{aligned}$$

и, следовательно,

$$\gamma(B)_{1D} = \sup_{i \geq 0} \left(\frac{d_{i+1}}{d_i} \lambda_{i+1}(t) - (\lambda_i(t) + \mu_{i+1}(t) + \xi(t)) + \frac{d_{i-1}}{d_i} \mu_i(t) \right) \leq -\alpha(t)$$

в соответствии с (11). Теперь, используя доказательство теоремы 1 из [9], получаем

$$\begin{aligned} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} &\leq \\ &\leq e^{-\int_s^t \alpha(\tau) d\tau} \|\mathbf{p}^*(s) - \mathbf{p}^{**}(s)\|_{1D}. \end{aligned}$$

Сравним теперь нормы вектора $\mathbf{z} = (z_1, z_2, \dots)^T$ в пространствах \mathcal{B} и l_{1D} . Имеем

$$\begin{aligned} \|\mathbf{z}\|_{\mathcal{B}} &= \sum_{i \geq 1} d_i z_i = d_1 \left(\left| \sum_{i \geq 1} z_i + \sum_{i \geq 2} -z_i \right| \right) + \\ &+ d_2 \left(\left| \sum_{i \geq 2} z_i + \sum_{i \geq 3} -z_i \right| \right) + \dots \leq \\ &\leq d_1 \left| \sum_{i \geq 1} z_i \right| + 2d_2 \left| \sum_{i \geq 2} z_i \right| + \dots \leq 2\|\mathbf{z}\|_{1D}, \end{aligned}$$

откуда в итоге и вытекает требуемое утверждение.

Следствие 2. Пусть при выполнении условий теоремы последовательность d_i возрастает достаточно быстро так, что $\inf_{k \geq 1} d_k/k = \omega > 0$. Тогда $X(t)$ имеет предельное среднее $\phi^*(t)$ и справедлива следующая оценка:

$$|\phi^*(t) - E_k(t)| \leq \frac{2}{\omega} e^{-\int_0^t \alpha(\tau) d\tau} \|\mathbf{p}^*(0) - \mathbf{e}_k\|_{1D}.$$

Теорема 5. Пусть при выполнении условий предыдущего следствия все интенсивности 1-периодичны. Тогда существуют 1-периодический предельный режим распределения вероятностей состояний

$$\pi(t) = (\pi_0(t), \pi_1(t), \dots)^T$$

и соответствующее ему предельное 1-периодическое среднее $\phi(t)$. Кроме того, справедливы следующие оценки:

$$\|\mathbf{p}(t) - \pi(t)\|_{\mathcal{B}} \leq 2e^{-\int_0^t \alpha(\tau) d\tau} \|\mathbf{p}(0) - \pi(0)\|_{1D}; \quad (13)$$

$$|\phi(t) - E_k(t)| \leq \frac{2}{\omega} e^{-\int_0^t \alpha(\tau) d\tau} \|\pi(0) - \mathbf{e}_k\|_{1D}. \quad (14)$$

К сожалению, полученные оценки (13) и (14) имеют один существенный недостаток, связанный

с отсутствием информации о $\pi(0)$, а значит, и возможностью реального применения. Рассмотрим способ, позволяющий получить такую информацию. Пусть $X(0) = k$, тогда имеем $\|\mathbf{e}_k\|_{1D} = \sum_{i=1}^k d_i$ при $k \geq 1$ и $\|\mathbf{e}_0\|_{1D} = 0$. Далее, используя подход из [9], можно оценить $\|\pi(0)\|_{1D}$ следующим образом. Имеем

$$\sup_{|t-s| \leq 1} \int_s^t \alpha(\tau) d\tau = K < \infty,$$

далее получаем

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\pi(t)\|_{1D} &\leq \left\| \int_0^t V(t, \tau) \mathbf{f}(\tau) d\tau \right\|_{1D} \leq \\ &\leq L\nu_0 \int_0^t e^{-\int_{\tau}^t \alpha(u) du} d\tau \leq \\ &\leq Le^K \nu_0 \int_0^t e^{-\alpha^*(t-\tau)} d\tau \leq \frac{Le^K \nu_0}{\alpha^*}, \end{aligned}$$

где $\alpha^* = \int_0^1 \alpha(u) du$.

Кроме того, с учетом 1-периодичности $\pi(t)$ справедливо неравенство

$$\|\pi(0)\|_{1D} \leq \limsup_{t \rightarrow \infty} \|\pi(t)\|_{1D}.$$

А тогда получаем

$$\|\pi(0) - \mathbf{e}_k\|_{1D} \leq \limsup_{t \rightarrow \infty} \|\pi(t)\|_{1D} + \|\mathbf{e}_k\|_{1D}$$

и следующее утверждение.

Следствие 3. Пусть в условиях теоремы 5 выполняется равенство $X(0) = k$. Тогда справедливы следующие оценки скорости сходимости:

$$\|\mathbf{p}(t) - \pi(t)\|_{\mathcal{B}} \leq 2e^{-\int_0^t \alpha(\tau) d\tau} \left(\sum_{i=1}^k d_i + \frac{Le^K \nu_0}{\alpha^*} \right)$$

и

$$|\phi(t) - E_k(t)| \leq \frac{2}{\omega} e^{-\int_0^t \alpha(\tau) d\tau} \left(\sum_{i=1}^k d_i + \frac{Le^K \nu_0}{\alpha^*} \right).$$

4 Аппроксимации

Рассмотрим семейство «усеченных» процессов $X_n(t)$, $n \geq S$, с фазовыми пространствами $E_n =$

$= \{0, 1, \dots, n\}$, теми же интенсивностями при $k \leq n$ и матрицами интенсивностей $A_n(t)$.

Пусть $\{h_k\}$ — последовательность положительных чисел такая, что $1 = h_1 \leq h_2 \leq \dots$, и

$$w_n = \sup_{k \geq n} \frac{h_k}{d_k}.$$

Обозначим через $\|\mathbf{z}\|_{\mathcal{B}_d}$ и $\|\mathbf{z}\|_{\mathcal{B}_h}$ нормы, соответствующие пространствам \mathcal{B} для последовательностей $\{d_k\}$ и $\{h_k\}$ соответственно.

Теорема 6. Пусть при выполнении условий теоремы 5 дополнительно выполнено условие $\lim_{n \rightarrow \infty} w_n = 0$.

Пусть $X(0) = X_n(0) = 0$. Тогда

$$\|\mathbf{p}(t) - \mathbf{p}_n(t)\|_{\mathcal{B}_h} \leq \frac{6L^2 M w_n e^{K\nu_0 t}}{\alpha^*}$$

при всех $t \geq 0$ и любом n .

Доказательство. Будем отождествлять векторы $(x_1, \dots, x_n, 0, 0, \dots)^T$ и $(x_1, \dots, x_n)^T$. Рассмотрим прямую систему Колмогорова (2) для исходного процесса в следующей форме:

$$\frac{d\mathbf{p}}{dt} = A_n(t)\mathbf{p} + \mathbf{g}(t) + (A(t) - A_n(t))\mathbf{p},$$

а также соответствующую систему

$$\frac{d\mathbf{p}_n}{dt} = A_n(t)\mathbf{p}_n + \mathbf{g}(t) \quad (15)$$

для усеченного процесса.

Имеем

$$\mathbf{p}_n(t) = U_n(t)\mathbf{p}(0) + \int_0^t U_n(t, \tau)\mathbf{g}(\tau) d\tau$$

при $\mathbf{p}(0) = \mathbf{p}_n(0)$ и

$$\begin{aligned} \mathbf{p}(t) &= U_n(t)\mathbf{p}(0) + \int_0^t U_n(t, \tau)\mathbf{g}(\tau) d\tau + \\ &+ \int_0^t U_n(t, \tau)(A(\tau) - A_n(\tau))\mathbf{p}(\tau) d\tau. \end{aligned}$$

Тогда (в любой норме) получаем

$$\begin{aligned} \|\mathbf{p}(t) - \mathbf{p}_n(t)\| &= \\ &= \left\| \int_0^t U_n(t, \tau)(A(\tau) - A_n(\tau))\mathbf{p}(\tau) d\tau \right\|. \quad (16) \end{aligned}$$

Рассмотрим матрицу Коши

$$U_n = \begin{pmatrix} u_{00}^n & \dots & u_{0n}^n & 0 & 0 & \dots \\ u_{10}^n & \dots & u_{1n}^n & 0 & 0 & \dots \\ \dots & \dots & \dots & \vdots & \vdots & \ddots \\ u_{n0}^n & \dots & u_{nn}^n & 0 & 0 & \dots \\ 0 & \dots & 0 & 1 & 0 & \dots \\ 0 & \dots & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Тогда

$$(A - A_n)\mathbf{p} = (0, \dots, 0, -\lambda_n p_n + \mu_{n+1} p_{n+1}, \lambda_n p_n - (\lambda_{n+1} + \mu_{n+1} + \xi) p_{n+1} + \mu_{n+2} p_{n+2}, \dots)^T$$

и, следовательно,

$$\begin{aligned} U_n(A - A_n)\mathbf{p} &= \\ &= \begin{pmatrix} u_{0n}^n (-\lambda_n p_n + \mu_{n+1} p_{n+1}) \\ u_{1n}^n (-\lambda_n p_n + \mu_{n+1} p_{n+1}) \\ \vdots \\ u_{nn}^n (-\lambda_n p_n + \mu_{n+1} p_{n+1}) \\ \lambda_n p_n - (\lambda_{n+1} + \mu_{n+1} + \xi) p_{n+1} + \mu_{n+2} p_{n+2} \\ \vdots \end{pmatrix}. \end{aligned}$$

С учетом неравенств $u_{ij}^n(t, \tau) \geq 0$ (при всех i, j, t, τ) и равенств $\sum_i u_{ij}^n(t, \tau) = 1$ (при всех j, t, τ) получаем оценку

$$\begin{aligned} \|U_n(A - A_n)\mathbf{p}\|_{\mathcal{B}_h} &= \\ &= |-\lambda_n p_n + \mu_{n+1} p_{n+1}| \sum_{k \leq n} h_k u_{kn}^n + \\ &+ \sum_{k > n} h_{k+1} |\lambda_k p_k - (\lambda_{k+1} + \mu_{k+1} + \xi) p_{k+1} + \\ &+ \mu_{k+2} p_{k+2}| \leq 3LM \sum_{k \geq n} h_k p_k = 3LM \sum_{k \geq n} \frac{h_k}{d_k} d_k p_k \leq \\ &\leq 3LM w_n \limsup_{t \rightarrow \infty} \|\pi(t)\|_{\mathcal{B}_d} \leq \frac{6L^2 M w_n e^{K\nu_0}}{\alpha^*}. \quad (17) \end{aligned}$$

Теперь из (16) и (17) получаем

$$\|\mathbf{p}(t) - \mathbf{p}_n(t)\|_{\mathcal{B}_h} \leq 3LM w_n \int_0^t \limsup_{\tau \rightarrow \infty} \|\pi(\tau)\|_{\mathcal{B}_d} d\tau$$

и требуемое утверждение.

Замечание 3. По-видимому, наиболее интересные оценки получаются, если выбрать $h_k = 1$ или $h_k = k$

(при всех k). Первый случай дает возможность построить 1-периодические предельные вероятности, а второй — предельное среднее.

Следствие 4. Пусть выполнены условия теоремы 6, а $X(0) = X_n(0) = 0$. Тогда при всех $t \geq 0$ и любом n справедливы оценки

$$\|\pi(t) - \mathbf{P}_n(t)\|_1 \leq 2e^{-\int_0^t \alpha(\tau) d\tau} \frac{Le^K \nu_0}{\alpha^*} + \frac{6L^2 M w_n^1 e^K \nu_0 t}{\alpha^*} \quad (18)$$

и

$$|\phi(t) - E_{0,n}(t)| \leq \frac{2}{\omega} e^{-\int_0^t \alpha(\tau) d\tau} \frac{Le^K \nu_0}{\alpha^*} + \frac{6L^2 M w_n^2 e^K \nu_0 t}{\alpha^*},$$

где $w_n^1 = \sup_{k \geq n} 1/d_k$, $w_n^2 = \sup_{k \geq n} k/d_k$ и $E_{0,n} = E_k(t) = E\{X_n(t) | X_n(0) = 0\}$.

Пусть $J_k(t) = \Pr\{X(t) \leq k\}$ — вероятность того, что число требований в системе в момент t не превышает k . Полученные оценки позволяют приближенно вычислить предельные 1-периодические $J_k(t)$ и предельное 1-периодическое среднее следующим образом.

Пусть ε — произвольное положительное число.

1. Выбираем целое m так, чтобы первое слагаемое в правой части (18) было меньше $\varepsilon/3$ при всех $t \geq m$.
2. Находим n так, чтобы второе слагаемое в правой части (18) было меньше $\varepsilon/3$ при всех $t \leq m + 1$.

3. Тогда решение задачи Коши для усеченной прямой системы Колмогорова (15) с начальным условием \mathbf{e}_0 на отрезке $[m; m + 1]$ (вычисленное с погрешностью $\varepsilon/3$) дает предельный 1-периодический режим $\pi(t) = (\pi_0(t), \pi_1(t), \dots)^T$ с погрешностью меньше ε .

4. Наконец, предельное 1-периодическое выражение для $J_k(t) = \Pr\{X(t) \leq k\}$ вычисляется как $\sum_{i=0}^k \pi_i(t)$ с той же погрешностью ε .

Предельное среднее находится аналогично.

5 Пример

Рассмотрим систему обслуживания $M(t)/M(t)/100$ с катастрофами, считая, что интенсивности задаются следующим образом:

$$\begin{aligned} \lambda(t) &= 3 + \sin 2\pi t; \\ \mu(t) &= 2 + \cos 2\pi t; \\ \xi(t) &= 2 + \sin 4\pi t, \end{aligned}$$

и вычислим предельное среднее $\phi(t)$, а также предельные величины $J_k(t)$ при некоторых значениях k . В частности, $J_0(t)$ — вероятность того, что в момент t очередь пуста, т. е. в системе обслуживания нет ни одного требования.

Используя подход, описанный в предыдущих параграфах, выберем $d = 1,3$ и $d_k = d^k$. Тогда будет: $L = 10$, $M = 100$, $\nu_0 = 1$, $w_n^1 = 1,3^{-n}$, $w_n^2 = n/1,3^n$, далее $\alpha(t) = \mu(t) - (d - 1)\lambda(t) = 1,1 + \cos 2\pi t - 0,3 \sin 2\pi t$, $K \leq 4$ и $\alpha^* = 1,1$.

Пусть $\varepsilon = 10^{-6}$. Тогда, как следует из полученных оценок, можно выбрать $m = 30$ и $n = 155$. При этом предельное среднее $\phi(t)$ и все предельные величины $J_k(t)$ строятся с точностью до $\varepsilon = 10^{-6}$ как соответствующие характеристики решения с

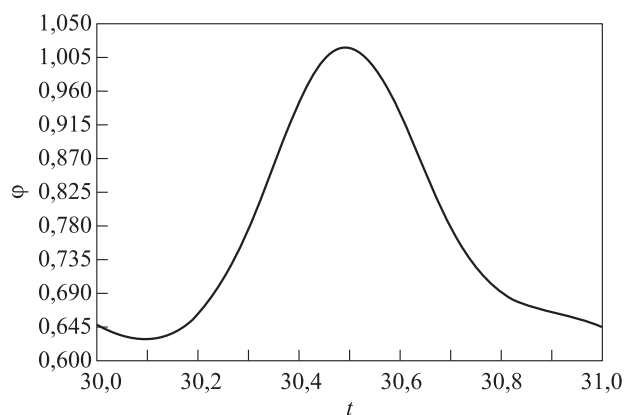


Рис. 1 Предельное среднее

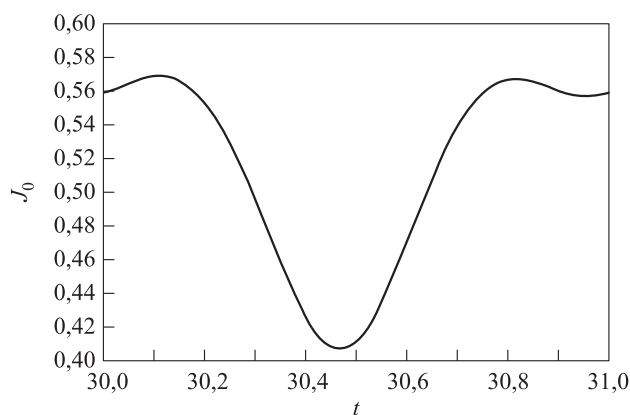


Рис. 2 Вероятность $\Pr\{X(t) = 0\}$

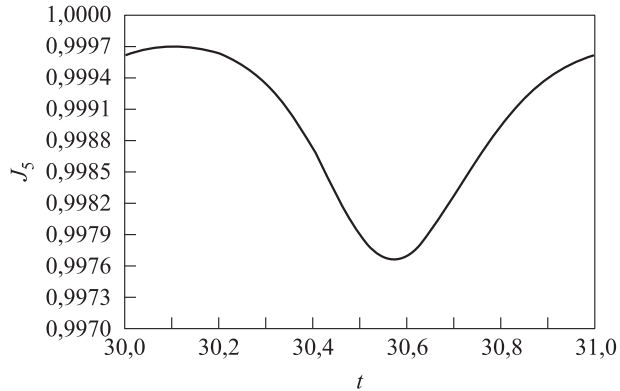


Рис. 3 Вероятность $Pr \{X(t) \leq 5\}$

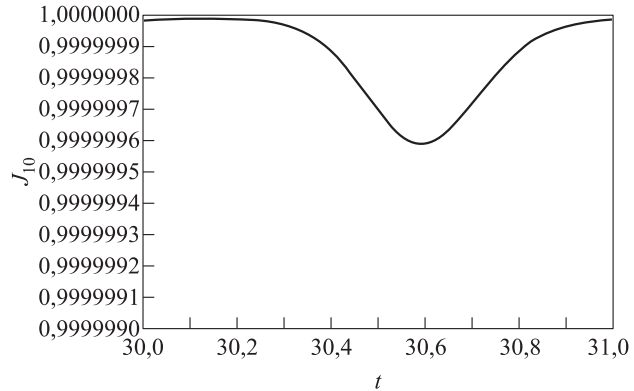


Рис. 4 Вероятность $Pr \{X(t) \leq 10\}$

начальным условием e_0 задачи Коши для соответствующей усеченной прямой системы Колмогорова на отрезке $[m, m + 1]$. Приведенные рис. 1–4 дают приближенные с описанной погрешностью предельные характеристики $\phi(t)$ и $J_0(t)$, $J_5(t)$, $J_{10}(t)$ соответственно. Отметим, что при $k \geq 11$ получается уже $J_k(t) \approx 1$.

Замечание 4. Отметим в заключение, что аналогично можно рассмотреть несколько более общую ситуацию, когда интенсивности катастроф зависят от состояния процесса. Основные изменения в этом случае будут касаться результатов параграфа 2.

Литература

1. Krishna Kumar B., Arivudainambi D. Transient solution of an $M|M|1$ queue with catastrophes // *Comput. Math. Appl.*, 2000. Vol. 40. P. 1233–1240.
2. Di Crescenzo A., Giorno V., Nobile A. G., Ricciardi L. M. On the $M|M|1$ queue with catastrophes and its continuous approximation // *Queueing Syst.*, 2003. Vol. 43. P. 329–347.
3. Van Doorn E. A., Zeifman A. Extinction probability in a birth–death process with killing // *J. Appl. Probab.*, 2005. Vol. 42. P. 185–198.
4. Di Crescenzo A., Giorno V., Nobile A. G., Ricciardi L. M. A note on birth–death processes with catastrophes // *Statist. Probab. Lett.*, 2008. Vol. 78. P. 2248–2257.
5. Zeifman A., Satin Ya., Chegodaev A., Bening V., Shorgin V. Some bounds for $M(t)/M(t)/S$ queue with catastrophes // *SMCTools08 Proceedings*. Athens, Greece, 2008.
6. Гнеденко Б. В., Макаров И. П. Свойства решений задачи с потерями в случае периодических интенсивностей // *Дифф. уравнения*, 1971. Т. 7. С. 1696–1698.
7. Zeifman A. I. Stability for continuous-time nonhomogeneous Markov chains // *Lect. Notes Mathem.*, 1985. Vol. 1155. P. 401–414.
8. Zeifman A. I. Upper and lower bounds on the rate of convergence for nonhomogeneous birth and death processes // *Stoch. Proc. Appl.*, 1995. Vol. 59. P. 157–173.
9. Zeifman A., Leorato S., Orsingher E., Satin Ya., Shilova G. Some universal limits for nonhomogeneous birth and death processes // *Queueing Syst.*, 2006. Vol. 52. P. 139–151.
10. Granovsky B., Zeifman A. Nonstationary queues: Estimation of the rate of convergence // *Queueing Syst.*, 2004. Vol. 46. P. 363–388.

БАЙЕСОВСКИЕ МОДЕЛИ МАССОВОГО ОБСЛУЖИВАНИЯ И НАДЕЖНОСТИ: ЭКСПОНЕНЦИАЛЬНО-ЭРЛАНГОВСКИЙ СЛУЧАЙ*

А. А. Кудрявцев¹, С. Я. Шоргин²

Аннотация: В данной работе продолжается изучение байесовских моделей массового обслуживания и надежности. Соответствующий метод предусматривает рандомизацию характеристик систем относительно некоторых априорных распределений параметров систем. Такой подход может использоваться, в частности, для вычисления моментных характеристик для вероятностно-временных и надежностных характеристик больших групп систем или устройств. В работе представлены новые результаты для случая, когда в качестве пары априорных распределений рассматриваются экспоненциальное распределение и распределение Эрланга.

Ключевые слова: байесовский подход; системы массового обслуживания; надежность; смешанные распределения; моделирование; эрланговское распределение; экспоненциальное распределение

1 Введение и основные предположения

Подробное изложение основ байесовского подхода к моделированию систем массового обслуживания (СМО) и ненадежных восстанавливаемых систем содержится в [1, 2]. Здесь коснемся этого вопроса лишь вкратце.

В реальной практике нередки ситуации, когда исследуемая система задана в определенном смысле «неточно»; скажем, если даже говорить о простейших системах типа $M|G|1$, исследователю могут быть априори неизвестны параметр входящего потока λ и параметры обслуживания μ и σ^2 . Такие ситуации возникают, скажем, в случае, когда рассматривается целый класс устройств, описываемых однотипными СМО, относительно которых известны только типы входящего потока и распределения обслуживания, а также дисциплина обслуживания, но конкретные параметры этих потоков и распределений, вообще говоря, различны для различных СМО данного класса. В этом случае, поскольку неизвестными являются именно «исходные» параметры потоков и времени обслуживания, естественным является рандомизационный подход, при котором элементами вероятностного пространства становятся (если рассматривать приведенный выше пример) значения λ , μ и σ^2 (а в общем случае можно говорить о вероятностном пространстве, элементами которого являются сами однотипные СМО).

При этом подлежащие вычислению характеристики такой «рандомизированной» СМО, естественно, являются рандомизацией аналогичных характеристик «обычной» СМО аналогичного типа — с учетом того априорного распределения входных параметров СМО, которое взято исследователем за основу.

Таким образом, в том же примере с системой типа $M|G|1$ возникают задачи рандомизации «обычных» характеристик таких систем с учетом априорных распределений входных параметров. Скажем, может приниматься предположение о показательном, равномерном или каком-то другом распределении одной или нескольких из величин λ , μ и σ^2 (которые при таком подходе становятся случайными величинами), об их независимости или зависимости и т. п. Полученные результаты могут применяться, например, для вычисления средних значений, построения доверительных интервалов для тех или иных характеристик рассматриваемого класса СМО «в целом». Такой подход к построению моделей массового обслуживания естественно называть *байесовским*.

Другим направлением применения байесовского подхода является оценка надежности. Как известно (см. [3]), коэффициент готовности восстанавливаемого устройства в стационарном режиме может быть вычислен по формуле

$$k = \frac{\lambda^{-1}}{\lambda^{-1} + \mu^{-1}} = \frac{\mu}{\lambda + \mu},$$

* Работа выполнена при поддержке РФФИ, проекты 08–07–00152 и 08–01–00567.

¹Московский государственный университет им. М. В. Ломоносова, факультет ВМиК, nubigena@hotmail.com

²Институт проблем информатики Российской академии наук, sshorgin@ipiran.ru

где λ^{-1} — среднее время безотказной работы, μ^{-1} — среднее время восстановления. Если принять сформулированное выше предположение, в соответствии с которым любое изучаемое устройство выбирается случайным образом из некоторого множества сходных устройств, различающихся средними величинами показателей надежности, то, согласно приведенным выше рассуждениям, значения λ и μ могут рассматриваться в качестве случайных. Следовательно, при таких предположениях коэффициент готовности k также является случайной величиной и его распределение зависит от распределений величин λ и μ . Результаты, получаемые в рамках этой постановки, могут использоваться, в частности, для вычисления средних значений и построения доверительных интервалов для надежностных характеристик всей изучаемой группы устройств.

Основным объектом рассмотрения на настоящем этапе является СМО $M|M|1$, в которой интенсивность входящего потока λ и интенсивность обслуживания μ независимы и имеют некоторые априори известные распределения. При этом нагрузка рассматриваемой системы имеет вид $\rho = \lambda/\mu$. Как известно, от значения ρ зависит наличие стационарного режима у рассматриваемой системы; величина ρ входит во многие формулы, описывающие характеристики разнообразных СМО. В связи с этим рассмотрение величины ρ избрано одной из первоочередных задач, которые следует рассмотреть в рамках байесовской теории СМО. Кроме того, рассматриваются распределения такой рандомизированной характеристики, как вероятность потерь $1 - \pi$ (здесь π — вероятность того, что входящий в СМО вызов не будет потерян).

Отметим, что значение коэффициента готовности k при среднем времени безотказной работы λ^{-1} и среднем времени восстановления μ^{-1} совпадает с величиной

$$\pi = \frac{1}{1 + \rho} = \frac{\mu}{\lambda + \mu}$$

для системы $M|M|1$, в которой λ — интенсивность входящего потока, а μ — интенсивность обслуживания. Поэтому вычисление вероятностных характеристик величины π означает одновременное вычисление вероятностных характеристик величины k при соответствующих распределениях среднего времени безотказной работы и среднего времени восстановления.

В настоящей работе всюду говорится об определении распределений величин, относящихся к байесовской модели СМО, включая π . При этом подразумевается, что распределение величины k в соответствующей «надежностной» постановке спе-

циально вычислять не нужно, поскольку оно совпадает с распределением величины π .

Данная работа является логическим продолжением статей [1, 2], в которых авторы рассматривали вероятностные характеристики коэффициента загрузки ρ , вероятности потерь $1 - \pi$ в системе $M|M|1|0$ в предположении, что пару «априорных распределений» (т.е. пару «распределение параметра входящего потока — распределение параметра обслуживания») составляют: «равномерное—равномерное», «экспоненциальное—экспоненциальное», «вырожденное — распределение Эрланга» (естественно, одновременно вычислялись характеристики коэффициента готовности k в соответствующей «надежностной» постановке).

В настоящей статье рассматриваются пары «экспоненциальное — распределение Эрланга» и «распределение Эрланга — экспоненциальное».

В дальнейшем авторы предполагают продолжить расширение множества пар априорных распределений, по которым производится рандомизация параметров λ и μ . В табл. 1 отображены этапы рассмотрения предложенной задачи. Буквы D, M, R, E и P обозначают вырожденное, экспоненциальное, равномерное, Эрланга и Парето распределения соответственно; символ «*» относится к классической постановке задачи, символ «+» соответствует уже рассмотренным ранее распределениям, символ «⊕» — распределениям, о которых пойдет речь в данной работе, символом «-» обозначаются распределения, для которых авторы планируют получить аналогичные результаты в дальнейшем.

Таблица 1 Этапы рассмотрения задачи

$\frac{\lambda}{\mu}$	D	M	R	E	P
D	*	+	+	+	-
M	-	+	-	⊕	-
R	+	-	+	-	-
E	-	⊕	-	-	-
P	-	-	-	-	-

2 Основные результаты

Итак, рассмотрим систему $M|M|1|0$. Пусть интенсивность входящего потока λ имеет экспоненциальное распределение с параметром $\theta > 0$, а интенсивность обслуживания μ имеет распределение Эрланга с параметрами $n \geq 1$ и $\alpha > 0$, причем λ и μ независимы. Целью исследования является нахождение функции распределения, плотности и первых моментов случайных величин $\rho = \lambda/\mu$ и $\pi = 1/(1 + \rho)$.

Найдем функцию распределения $F_\rho(x)$ случайной величины ρ . Имеем

$$\begin{aligned} F_\rho(x) &= P(\lambda < \mu x) = \\ &= \int_0^\infty (1 - e^{-\theta xy}) \frac{y^{n-1} \alpha^n e^{-\alpha y}}{(n-1)!} dy = \\ &= 1 - \left(\frac{\alpha}{\alpha + \theta x} \right)^n, \quad x > 0. \end{aligned}$$

Продифференцировав последнее выражение по x , найдем плотность величины ρ :

$$f_\rho(x) = \frac{n\theta\alpha^n}{(\alpha + \theta x)^{n+1}}, \quad x > 0.$$

Найдем первые два момента случайной величины ρ . Очевидно, что $E\rho = \infty$ при $n = 1$ и $E\rho^2 = \infty$ при $n \leq 2$. Для остальных натуральных n по формуле 856.12 из [4] имеем

$$\begin{aligned} E\rho &= \int_0^\infty \frac{n\theta\alpha^n x}{(\alpha + \theta x)^{n+1}} dx = \int_0^\infty \frac{n\alpha^n dx}{(\alpha + \theta x)^n} - \\ &- \int_0^\infty \frac{n\alpha^{n+1} dx}{(\alpha + \theta x)^{n+1}} = \frac{n\alpha}{\theta} \frac{\Gamma(1)\Gamma(n-1)}{\Gamma(n)} - \\ &- \frac{n\alpha}{\theta} \frac{\Gamma(1)\Gamma(n)}{\Gamma(n+1)} = \frac{\alpha}{\theta(n-1)}, \quad n \geq 2. \end{aligned}$$

Аналогично получаем

$$\begin{aligned} E\rho^2 &= \int_0^\infty \frac{n\theta\alpha^n x^2}{(\alpha + \theta x)^{n+1}} dx = \int_0^\infty \frac{n\alpha^n dx}{\theta(\alpha + \theta x)^{n-1}} - \\ &- \int_0^\infty \frac{2n\alpha^{n+1} dx}{\theta(\alpha + \theta x)^n} + \int_0^\infty \frac{n\alpha^{n+2} dx}{\theta(\alpha + \theta x)^{n+1}} = \\ &= \frac{n\alpha^2(n-3)!}{\theta^2(n-2)!} - \frac{2n\alpha^2(n-2)!}{\theta^2(n-1)!} + \frac{n\alpha^2(n-1)!}{\theta^2 n!} = \\ &= \frac{2\alpha^2}{\theta^2(n-1)(n-2)}, \quad n \geq 3. \end{aligned}$$

Рассмотрим характеристики вероятности «непотери» вызова π . Для функции распределения имеем

$$\begin{aligned} F_\pi(x) &= 1 - P\left(\rho < \frac{1-x}{x}\right) = \\ &= \left(\frac{\alpha x}{\theta + (\alpha - \theta)x}\right)^n, \quad x \in (0, 1). \end{aligned}$$

В этом случае плотность π , очевидно, имеет вид

$$f_\pi(x) = \frac{n\theta\alpha^n x^{n-1}}{(\theta + (\alpha - \theta)x)^{n+1}}, \quad x \in (0, 1).$$

Для вычисления моментов случайной величины π потребуется следующая формула, являющаяся следствием формулы 2.111 из [5]:

$$\int \frac{x^n dx}{(a + bx)^{n+1}} = - \sum_{k=0}^{n-1} \frac{x^{n-k}}{(n-k)b^{k+1}(a + bx)^{n-k}} + \frac{\ln(a + bx)}{b^{n+1}} + C, \quad (1)$$

где C — некоторая константа.

При $\alpha = \theta$, очевидно, $E\pi = n/(n+1)$, $E\pi^2 = n/(n+2)$. При $\alpha \neq \theta$, воспользовавшись (1), имеем

$$\begin{aligned} E\pi &= \int_0^1 \frac{n\theta\alpha^n x^n dx}{(\theta + (\alpha - \theta)x)^{n+1}} = -n\theta\alpha^n \times \\ &\times \left[\sum_{k=0}^{n-1} \frac{x^{n-k}}{(n-k)(\alpha - \theta)^{k+1}(\theta + (\alpha - \theta)x)^{n-k}} - \frac{\ln(\theta + (\alpha - \theta)x)}{(\alpha - \theta)^{n+1}} \right] \Bigg|_0^1 = \\ &= -\frac{n\theta}{\alpha} \sum_{k=0}^{n-1} \frac{1}{n-k} \left(\frac{\alpha}{\alpha - \theta}\right)^{k+1} + \\ &+ \frac{n\theta}{\alpha - \theta} \left(\frac{\alpha}{\alpha - \theta}\right)^n \ln \frac{\alpha}{\theta}. \end{aligned}$$

Для вычисления второго момента π при $\alpha \neq \theta$ потребуется следующая формула, аналогичная (1):

$$\begin{aligned} \int \frac{x^n dx}{(a + bx)^n} &= \frac{x^n}{b(a + bx)^{n-1}} + \\ &+ \frac{an}{b} \sum_{k=0}^{n-2} \frac{x^{n-k-1}}{(n-k-1)b^{k+1}(a + bx)^{n-k-1}} - \\ &- \frac{an}{b^{n+1}} \ln(a + bx) + C. \quad (2) \end{aligned}$$

Имеем

$$\begin{aligned} E\pi^2 &= \int_0^1 \frac{n\theta\alpha^n x^{n+1} dx}{(\theta + (\alpha - \theta)x)^{n+1}} = \\ &= n\theta\alpha^n \left[\frac{x^{n+1}}{(\alpha - \theta)(\theta + (\alpha - \theta)x)^n} + \frac{(n+1)\theta}{\alpha - \theta} \times \right. \\ &\times \left. \sum_{k=0}^{n-1} \frac{x^{n-k}}{(n-k)(\alpha - \theta)^{k+1}(\theta + (\alpha - \theta)x)^{n-k}} - \frac{(n+1)\theta \ln(\theta + (\alpha - \theta)x)}{(\alpha - \theta)^{n+2}} \right] \Bigg|_0^1 = \end{aligned}$$

$$= \frac{n\theta}{\alpha - \theta} + \frac{n(n+1)\theta^2}{\alpha(\alpha - \theta)} \sum_{k=0}^{n-1} \frac{1}{n-k} \left(\frac{\alpha}{\alpha - \theta}\right)^{k+1} - \frac{(n+1)\theta}{(\alpha - \theta)^{n+2}} \ln \frac{\alpha}{\theta}.$$

$$= 1 - \frac{k! \alpha \theta^k}{(\alpha + \theta x)^k} \sum_{l=0}^{k-1} \frac{\alpha^l x^{k-l-1}}{(k-l-1)!(l+1)! \theta^{l+1}} = \left(\frac{\theta x}{\alpha + \theta x}\right)^k, \quad x > 0.$$

Зная первые два момента случайной величины π , можно легко вычислить дисперсию рассматриваемой характеристики.

Теперь в рамках системы $M|M|1|0$ рассмотрим задачу, в которой интенсивность входящего потока λ имеет распределение Эрланга с параметрами $k \geq 1$ и $\theta > 0$, а интенсивность обслуживания μ имеет экспоненциальное распределение с параметром $\alpha > 0$, причем λ и μ независимы. Как и ранее, найдем функции распределения, плотности и первые моменты случайных величин $\rho = \lambda/\mu$ и $\pi = 1/(1 + \rho)$.

Для плотности случайной величины ρ имеем следующие соотношения:

$$f_\rho(x) = \int_0^\infty \frac{\alpha \theta^k x^{k-1} y^k e^{-\alpha y} e^{-\theta xy}}{(k-1)!} dy = \frac{\alpha \theta^k x^{k-1}}{(k-1)!(\alpha + \theta x)^{k+1}} \int_0^\infty z^k e^{-z} dz = \frac{k \alpha \theta^k x^{k-1}}{(\alpha + \theta x)^{k+1}}, \quad x > 0.$$

Откуда, очевидно, $E\rho = \infty$. Найдем функцию распределения случайной величины ρ . Будем использовать следующее следствие формулы 2.111 из [5] для $m \geq 2$:

$$\int \frac{x^n dx}{(a + bx)^{n+m}} = -\frac{x^n}{(m-1)b C_{n+m-1}^n (a + bx)^{n+m-1}} \times \sum_{l=0}^n C_{n+m-1}^{n-l} \left(\frac{a}{bx}\right)^l + C. \quad (3)$$

По формуле (3) имеем

$$F_\rho(x) = \int_0^x \frac{k \alpha \theta^k t^{k-1}}{(\alpha + \theta t)^{k+1}} dt = -k \alpha \theta^k \left[\frac{(k-1)!}{(\alpha + \theta t)^k} \sum_{l=0}^{k-2} \frac{\alpha^l t^{k-l-1}}{(k-l-1)!(l+1)! \theta^{l+1}} + \frac{(k-1)! \alpha^{k-1}}{k! \theta^k (\alpha + \theta t)^k} \right] \Big|_0^x =$$

Теперь найдем основные характеристики распределения случайной величины π .

$$F_\pi(x) = 1 - F_\rho\left(\frac{1-x}{x}\right) = 1 - \frac{\theta^k (1-x)^k}{(\theta + (\alpha - \theta)x)^k}, \quad x \in [0, 1].$$

Для соответствующей плотности имеем следующие равенства:

$$f_\pi(x) = \frac{1}{x^2} f_\rho\left(\frac{1-x}{x}\right) = \frac{k \alpha \theta^k (1-x)^{k-1}}{(\theta + (\alpha - \theta)x)^{k+1}}, \quad x \in [0, 1].$$

Очевидно, что при $\alpha = \theta$ для первых двух моментов случайной величины π выполнены соотношения

$$E\pi = kB(2, k) \quad \text{и} \quad E\pi^2 = kB(3, k),$$

где $B(n, k)$ — β -функция:

$$B(n, k) = \int_0^1 x^{n-1} (1-x)^{k-1} dx.$$

При $\alpha \neq \theta$ получаем

$$E\pi = k \alpha \theta^k \int_0^1 \frac{x(1-x)^{k-1} dx}{(\theta + (\alpha - \theta)x)^{k+1}} = k \alpha \theta^k \int_0^1 \frac{x \sum_{m=0}^{k-1} C_{k-1}^m (-1)^m x^m dx}{(\theta + (\alpha - \theta)x)^{k+1}} = k \alpha \theta^k \sum_{m=0}^{k-1} (-1)^m C_{k-1}^m \int_0^1 \frac{x^{m+1} dx}{(\theta + (\alpha - \theta)x)^{k+1}}.$$

Введем для фиксированных $\alpha \neq \theta$ обозначение

$$A(p, q) = \int_0^1 \frac{x^p dx}{(\theta + (\alpha - \theta)x)^q}.$$

Вычислим $A(k, k + 1)$, используя (1). Имеем

$$A(k, k + 1) = - \sum_{l=0}^{k-1} \frac{1}{(k-l)\alpha^{k-l}(\alpha-\theta)^{l+1}} + \frac{\ln(\alpha/\theta)}{(\alpha-\theta)^{k+1}}. \quad (4)$$

Для $m = 0, \dots, k - 2$, используя (3), получаем

$$A(m + 1, k + 1) = - \frac{\alpha^{-k}}{(k-m-1)(\alpha-\theta)C_k^{m+1}} \sum_{l=0}^{m+1} \frac{C_k^{m-l+1}\theta^l}{(\alpha-\theta)^l} + \frac{\theta^{m-k+1}}{(k-m-1)C_k^{m+1}(\alpha-\theta)^{m+2}}. \quad (5)$$

Таким образом, при $\alpha \neq \theta$

$$E\pi = k\alpha\theta^k \sum_{m=0}^{k-2} (-1)^m C_{k-1}^m A(m + 1, k + 1) + k\alpha\theta^k (-1)^{k-1} A(k, k + 1),$$

где $A(p, k + 1)$, $p = 1, \dots, k$, вычисляются по формулам (4) и (5).

Аналогично для $\alpha \neq \theta$ находим второй момент случайной величины π :

$$E\pi^2 = k\alpha\theta^k \sum_{m=0}^{k-1} (-1)^m C_{k-1}^m \int_0^1 \frac{x^{m+2} dx}{(\theta + (\alpha - \theta)x)^{k+1}}.$$

Воспользовавшись формулой (2), получаем представление для интеграла $A(k + 1, k + 1)$:

$$A(k + 1, k + 1) = \frac{1}{\alpha^k(\alpha - \theta)} + \frac{(k + 1)\theta}{\alpha - \theta} \sum_{l=0}^{k-1} \frac{1}{(k-l)\alpha^{k-l}(\alpha-\theta)^{l+1}} - \frac{(k + 1)\theta \ln(\alpha/\theta)}{(\alpha - \theta)^{k+2}}. \quad (6)$$

Следовательно,

$$E\pi^2 = k\alpha\theta^k \left((-1)^{k-1} A(k + 1, k + 1) + (-1)^k (k - 1)A(k, k + 1) + \sum_{m=0}^{k-3} (-1)^m C_{k-1}^m A(m + 2, k + 1) \right),$$

где $A(p, k + 1)$, $p = 2, \dots, k + 1$, вычисляются по формулам (4)–(6).

Представленные в статье результаты не являются полными в рамках проблематики байесовских СМО и байесовских моделей надежности даже в рамках рассмотрения байесовских моделей систем $M|M|1|0$. Очевидно, что дальнейшее продвижение в рамках данной проблематики требует рассмотрения и других априорных распределений величин λ , μ и других традиционных входных параметров для СМО и восстанавливаемых устройств, которые могут представлять интерес для практики, и вычисление соответствующих распределений показателей функционирования и надежности различных типов систем (в том числе систем вида $M|G|1$, $M|M|n|0$ и др.) после их рандомизации с учетом наиболее важных для практики априорных распределений параметров. В частности, объектом для рандомизации может послужить стандартное отклонение обслуживания σ в системе $M|G|1$, число каналов n в системе $M|M|n$ и т. п.

Литература

1. *D'Apice C., Manzo R., Shorgin S.* Some Bayesian queuing and reliability models // *Electronic J. "Reliability: Theory & Applications"*, 2006. Vol. 1. No. 4.
2. *Кудрявцев А. А., Шоргин С. Я.* Байесовский подход к анализу систем массового обслуживания и показателей надежности // *Информатика и её применения*, 2007. Т. 1. Вып. 2. С. 76–82.
3. *Kozlov B. A., Ushakov I. A.* Reliability handbook. — Holt, Rinehart & Winston, 1970.
4. *Двайт Г.* Таблицы интегралов и другие математические формулы / Пер. с англ. — М.: Наука, 1966. 228 с.
5. *Градштейн И. С., Рыжик И. М.* Таблицы интегралов, сумм, рядов и произведений. — М.: Наука, 1971. 1108 с.

ОБ ОДНОМ ПОДХОДЕ К ФОРМИРОВАНИЮ ИЗОБРАЖЕНИЙ БЕЗ ИСПОЛЬЗОВАНИЯ ЭКРАНА

А. В. Торчигин¹

Аннотация: Рассматриваются свойства изображений, возникающих при наблюдении в колеблющемся зеркале линейки модулируемых по яркости светодиодов. Проанализированы возможные области применения этого подхода.

Ключевые слова: формирование изображений; виртуальное окружение; виртуальная реальность; объемное изображение

1 Введение

Привычные изображения обычно формируются на некотором основании. Это может быть скала, на которую первобытный человек наносил свои примитивные рисунки. Это может быть холст, на котором гениальные художники эпохи Возрождения создавали свои шедевры. Это может быть экран телевизора или монитора. В любом из подобных случаев изображение представляет собой двумерную картину. С недавнего времени стали известны стереоизображения, которые невозможно совместить с плоским экраном. Однако и такие изображения создаются с помощью экрана, на который проецируют два изображения соответственно для левого и правого глаза. Зритель снабжается очками, пропускающими свет от изображения, приготовленного специально для левого и правого глаза. Этот способ требует больших помещений, двух мощных проекторов и специального дорогостоящего экрана.

Примером изображения, которое существует без экрана, является изображение картины реального мира в обыкновенном плоском зеркале. Разглядывая такое изображение в зеркале, зритель видит объемную картину, которая находится в глубине зеркала. Размеры этой картины могут существенно превосходить размеры зеркала. Например, в зеркале над умывальником зритель может видеть себя в полный рост. В предлагаемой статье рассматриваются возможности существующих технических средств для формирования подобных изображений без использования экранов и соответствующих проекторов.

2 Принцип действия

Разглядывая в зеркале отражение светящихся светодиодов, расположенных вдоль некоторой

вертикальной линии на одинаковом минимальном расстоянии друг от друга, зритель видит светящуюся вертикальную линейку светодиодов. При повороте зеркала вокруг вертикальной оси эта линейка сдвигается в зеркале в горизонтальном направлении. В зеркале, совершающем вращательные колебания вокруг вертикальной оси, зритель увидит светящийся прямоугольник, ширина которого зависит от амплитуды колебаний зеркала. Если при этом светодиоды модулируются во времени соответствующим образом, то вместо однородно светящегося прямоугольника зритель может видеть соответствующую картину. Эта картина расположена не в плоскости зеркала (плоскость колеблется с частотой около 25 Гц), а в глубине зеркала на удалении, равном расстоянию от зеркала до линейки светодиодов.

Если амплитуда колебаний зеркала достаточно велика, то зритель будет видеть в зеркале только часть картины. Однако если повернуть колеблющееся зеркало вокруг вертикальной оси на некоторый угол, то зритель увидит другой фрагмент формируемой картины. Подобно человеку, который разглядывает в карманное зеркало различные фрагменты отражения реального мира путем изменения ориентации этого зеркала, зритель, изменяя ориентацию колеблющегося зеркала, может разглядывать различные фрагменты формируемого изображения. Более того, он может не изменять ориентацию зеркала, а смотреть на него под разными углами. В этом случае зритель также будет видеть различные фрагменты формируемой картины. Приближая зеркало к глазам, он будет видеть картину под большим углом зрения, т. е. будет видеть больший фрагмент картины. При этом сам размер изображения остается неизменным. В качестве иллюстрации на рис. 1 показаны фотографии одного и того же изображения, сделанные фото-

¹Институт проблем информатики Российской академии наук, torchigin_a@mail.ru



Рис. 1 Фотографии одного и того же изображения, формируемого в прикрепленном к стене колеблющемся зеркале линейкой модулируемых по яркости светодиодов. Правая фотография сделана фотоаппаратом, находящимся на меньшем расстоянии от зеркала, чем левая

аппаратом, находящимся на разном расстоянии от колеблющегося зеркала. При приближении фотоаппарата к зеркалу увеличивается фиксируемый аппаратом фрагмент надписи.

Зеркало можно рассматривать как окно в виртуальный мир. Находясь в реальном мире и подходя к колеблющемуся зеркалу, зритель видит не зеркало (зеркало он видеть не может, так как оно колеблется и подобно спицам вращающегося колеса становится невидимым), а нечто, нарушающее привычную картину реального мира. Заглядывая в это нечто, зритель видит картину виртуального мира, созданного модулируемыми по яркости светодиодами. Для зрителя это нечто представляется в виде некоторого отверстия в реальном мире, через которое он может разглядывать виртуальный мир. Чем ближе его глаза находятся к этому отверстию, тем больше телесный угол, в котором виден виртуальный мир. Подобно тому как, рассматривая через обычное отверстие реальный мир, зритель может видеть при изменении позиции относительно отверстия разные фрагменты реального мира, так и разглядывая через это нечто виртуальный мир, зритель может видеть при изменении позиции относительно этого нечто разные фрагменты виртуального мира. Отличие лишь в том, что зритель может изменять ориентацию этого нечто и видеть при этом различные фрагменты виртуального мира. Приведенный способ формирования изображений описан в патенте автора [1].

3 Теоретические предпосылки

Построим изображение светодиода при различных угловых положениях вращающегося зеркала, показанного жирной линией на рис. 2. Имея в виду,

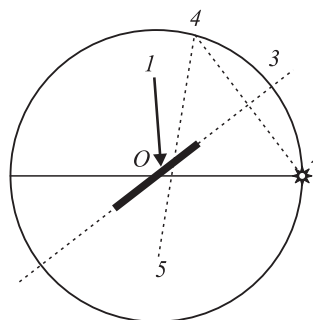


Рис. 2 Траектория 3 перемещения мнимого изображения 4 источника 2 при вращении зеркала 1

что изображение светодиода в зеркале расположено в точке, симметричной светодиоду относительно плоскости зеркала, получим, что изображение светодиода 2 при положении зеркала 1 окажется в точке 4. При вращении зеркала 1 изображение светодиода 2 описывает окружность 3 с центром O на оси вращающегося зеркала 1. Радиус этой окружности равен расстоянию от оси вращающегося зеркала 1 до светодиода 2. Однако зритель, расположенный в некоторой точке 5, может не увидеть всю окружность по нескольким причинам.

Во-первых, угол обзора ограничен строением глаз человека, который, не поворачивая головы, может видеть в секторе около 90° .

Во-вторых, угол обзора изображения ограничен размерами зеркала и расстоянием от оси вращения зеркала до зрителя. Для того чтобы изображение светодиода было видно зрителю, необходимо, чтобы прямая линия, соединяющая изображение светодиода в точке 4 с точкой 5, где находится зритель, пересекала зеркало 1.

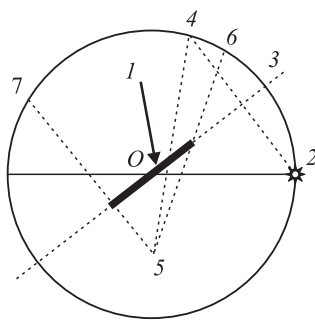


Рис. 3 Дуга, которую видит зритель 5 во вращающемся зеркале 1 ограниченных размеров

На рис. 3 показан в качестве примера для конкретного положения зрителя и размера зеркала фрагмент окружности между точками 6 и 7, по которому перемещается изображение светодиода. Только этот фрагмент виден зрителю, расположенному в точке 5. Чем ближе зритель расположен к вращающемуся зеркалу и чем больше горизонтальный размер $2R$ вращающегося зеркала, тем больший фрагмент окружности может видеть зритель.

На рисунках показана лишь одна окружность, создаваемая одним светодиодом. В том случае, если сразу несколько светодиодов расположено на прямой, параллельной оси вращения, изображение каждого из них при вращении зеркала будет описывать идентичную окружность, расположенную в плоскости, параллельной плоскости рассмотренной окружности. Набор таких окружностей формирует изображение, подобно тому как набор прямых линий в виде строк формирует изображение на экране телевизора. Формируемое изображение состоит из набора вертикальных линеек светящихся светодиодов, сдвинутых относительно друг друга в горизонтальном направлении на некоторое расстояние. Сдвиг образуется из-за того, что линейка светодиодов отражается в колеблющемся зеркале.

Угол, в котором зритель может видеть формируемое изображение, определяется отношением s/r , где s — размер зеркала в горизонтальном направлении, r — расстояние от центра зеркала до зрачков зрителя. Это означает, что при приближении зеркала к глазам зрителя поперечные размеры зеркала могут быть уменьшены. В случае если колеблющееся зеркало расположено в непосредственной близости от зрачков зрителя, его размеры могут быть сделаны сравнимыми с размером стекол в очках. При этом угол формируемого изображения может даже превосходить угол, в котором глаза зрителя способны видеть изображение.

Использование зеркал, совершающих вращательные колебания вокруг оси вращения (колеб-

лющихся зеркал), имеет несколько преимуществ перед использованием вращающихся зеркал. Во-первых, при использовании вращающихся зеркал картина формируется только в течение небольшого времени, когда положение зеркал таково, что лучи света от модулируемых светодиодов попадают в глаза зрителя. Обычно это время составляет менее 10% от периода вращения. Это приводит к тому, что большую часть времени свет от модулируемых светодиодов не может попадать в глаза зрителя, что снижает максимальную яркость формируемого изображения. Во-вторых, если в помещении находятся другие источники света, например окна, лампы освещения, то при определенном положении вращающихся зеркал свет от них попадает в глаза зрителя. Это приводит к тому, что изображение формируется на неконтролируемом светлом фоне. В-третьих, колеблющиеся зеркала можно располагать ближе к зрачкам зрителя, чем вращающиеся. Это позволяет увеличить максимальный телесный угол, в котором может формироваться изображение.

Если ось колеблющегося зеркала закрепить на платформе, жестко связанной с головой зрителя, например на шлеме или очках, то при повороте головы зритель будет видеть другой фрагмент окружности на рис. 3. У него создается впечатление, что он со всех сторон окружен виртуальным миром. При повороте глаз зритель начинает видеть другой фрагмент виртуального мира, точно так же как в реальном мире при повороте глаз он видит другой фрагмент реального мира. Как известно, движение зрачков происходит произвольно и достаточно быстро. Это движение не регистрируется известными системами создания виртуальной реальности, которые обычно реагируют только на поворот головы.

До настоящего времени речь шла об одном зрителе. Однако все вышесказанное справедливо для любого другого зрителя, который разглядывает линейку светодиодов через свою систему зеркал. Таким образом, в виртуальный мир может быть одновременно погружено сколь угодно много зрителей, каждый из которых будет иметь перед глазами свой фрагмент виртуального мира, точно так же как многие зрители в реальном мире видят каждый свой фрагмент, который им доступен из их местоположения. Однако, так же как и в реальном мире, иногда зрители могут мешать друг другу, закрывая собой некоторые фрагменты виртуальной картины.

Рассмотренный подход можно легко приспособить для формирования стереоизображений. Схема получения стереоизображения показана на рис. 4. Здесь светодиоды 1 формируют изображение для левого глаза 3 в тот интервал времени, когда свет

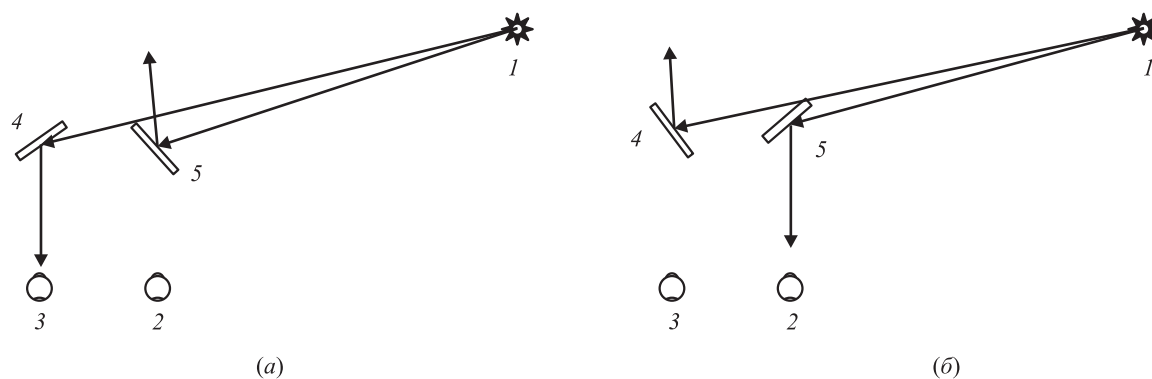


Рис. 4 Схема формирования различных изображений для левого и правого глаза от одного комплекта светодиодов. При первом положении зеркал (а) светодиоды видны левым глазом и не видны правым. При втором положении зеркал (б) светодиоды видны правым глазом и не видны левым

от источников попадает в левый глаз, отражаясь от зеркала 4, совершающего вращательные колебания. При этом свет от тех же светодиодов не попадает в правый глаз, так как от зеркала 5 он отражается в другую сторону (рис. 4, а). В следующий интервал времени — через полпериода колебаний зеркал — последние принимают положение, показанное на рис. 4, б. В этом случае ситуация изменяется на противоположную. Свет от светодиодов, отражаясь от колеблющегося зеркала 5, попадает в правый глаз 2 и не попадает в левый после отражения от зеркала 4. Таким образом, каждый глаз видит формируемую для него картину, а у зрителя создается впечатление, что он видит стереоизображение.

Отметим, что в рассматриваемом подходе при желании могут быть использованы все те средства, которые используются в существующих системах. Действительно, можно использовать два комплекта светодиодов, предназначенных соответственно для левого и правого глаза. Один комплект покрыт прозрачной пленкой, пропускающей поляризацию, предназначенную для левого глаза, а другой — для правого. Дополнительно к очкам с колеблющимися зеркалами добавляются обычные поляризационно-чувствительные очки. В этом случае интенсивность формируемой картины увеличивается вдвое, так как оба глаза воспринимают одновременно предназначенные для них изображения. В этом случае отпадает необходимость в дорогостоящем экране, сохраняющем поляризацию отраженного света, и мощных двух проекторов, обеспечивающих формирование на экране двух изображений. Возможны и другие варианты получения необходимых изображений для левого и правого глаза.

Следует обратить внимание, что при изменении местоположения зрителя картина виртуального мира, которую он видит, изменяется. Таким образом, если несколько зрителей снабжены очками с ко-

леблющимися зеркалами, то каждый зритель оказывается погруженным в виртуальное окружение. Это обстоятельство позволяет погружать в виртуальное окружение большие зрительские аудитории. При этом, в отличие от существующих систем, требующих больших помещений, экранов большой площади, мощных дорогостоящих проекторов, в рассматриваемом подходе виртуальное окружение может быть создано в любом пространстве. Для этого достаточно лишь поместить в любую комнату вертикальную линейку светодиодов. Из любого места, откуда видна эта линейка, зритель, надев очки, погружается в виртуальное окружение.

Весьма интересен подход, который позволяет создавать стереоизображение для многих зрителей, не требуя при этом, чтобы зрители надевали какие-либо очки. Рассмотрим простейший пример. На рис. 5 показаны две линейки светодиодов 1 и 2, расположенные на разных расстояниях от вращающегося или колеблющегося зеркала. Если, например, с помощью линейки 1 формировать изображение одного человека, а с помощью линейки 2 — изображение другого человека, то зритель будет видеть

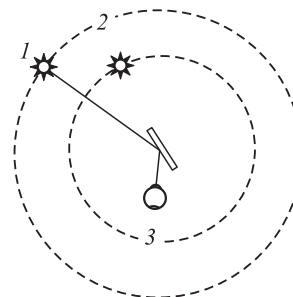


Рис. 5 Зритель 3 видит две картины на разных расстояниях от зеркала. Без каких-либо очков получается объемное изображение

в зеркале двух людей, расположенных от него на разных расстояниях. Действительно, в обычном зеркале зритель прекрасно отличает человека, стоящего вблизи зеркала, от человека, расположенного от зеркала на большом расстоянии.

Аналогичным образом в колеблющемся зеркале зритель отличает изображение, созданное линейкой светодиодов, расположенной вблизи зеркала, от изображения, сформированного линейкой, расположенной на значительном удалении от зеркала.

Обобщая этот прием, можно расположить N линеек светодиодов на расстояниях $R, R + \Delta R, \dots, R + (N - 1)\Delta R$ от оси вращения зеркала, а отображаемую объемную картину разбить на систему колец толщиной ΔR и внутренними диаметрами $R, R + \Delta R, \dots, R + (N - 1)\Delta R$. При этом фрагмент изображения, попадающий в i -е кольцо, следует показывать с помощью i -й линейки светодиодов. Так как i -я линейка формирует изображение, отстоящее от оси вращения на расстояние $R + (i - 1)\Delta R$, то всеми линейками светодиодов будут сформированы фрагменты изображения, отстоящие от оси вращения на разные расстояния, т. е. будет сформировано изображение, фрагменты которого находятся на разном удалении от оси вращения. Иными словами, будет сформировано объемное изображение. Как следует из проведенного выше рассмотрения, это объемное изображение могут наблюдать без каких-либо очков зрители, находящиеся в любом месте, откуда видны в колеблющемся зеркале линейки светодиодов. Имеются ограничения на формируемые таким образом объемные изображения, однако подобные ограничения существуют и для известных систем формирования стереоизображений. В простейшем частном случае двумя линейками светодиодов можно создавать передний и задний планы, которые воспринимаются зрителем в виде объемной картины.

4 Особенности реализации

Любая из рассмотренных систем включает в себя три основных компонента:

- (1) линейку светодиодов с системой, обеспечивающей модуляцию их яркости во времени;
- (2) зеркало, совершающее вращательно-колебательные движения;
- (3) средства, обеспечивающие синхронизацию колебаний зеркала с модуляцией светодиодов.

Казалось бы, наиболее сложным является первый компонент. Однако при тщательном анализе оказалось, что для его реализации с успехом могут быть использованы аппаратные средства, разработанные для жидкокристаллических (ЖК) мониторов, структурная схема которых приведена на рис. 6.

Здесь подготовленная в персональном компьютере (ПК) информация для показа на экране ЖК-монитора передается по стандартному интерфейсу на вход контроллера монитора, который формирует соответствующие сигналы, поступающие на горизонтальные и вертикальные шины ЖК-панели. С целью сокращения числа проводов между контроллером и панелью (ЖК-панель имеет более 1000 вертикальных шин) сигналы от контроллера передаются последовательно по меньшему числу проводов (52 провода при представлении сигналов в виде стандартных сигналов транзисторно-транзисторной логики (TTL — transistor-transistor logic) либо около 20 проводов при представлении сигналов в виде паразитных низковольтных дифференциальных сигналов (LVDS — low voltage differential signal)). На ЖК-панели в качестве ее неразрывной части изготавливается специальный драйвер, преобразующий последовательность входных сигналов в выходные, поступающие на вертикальные и

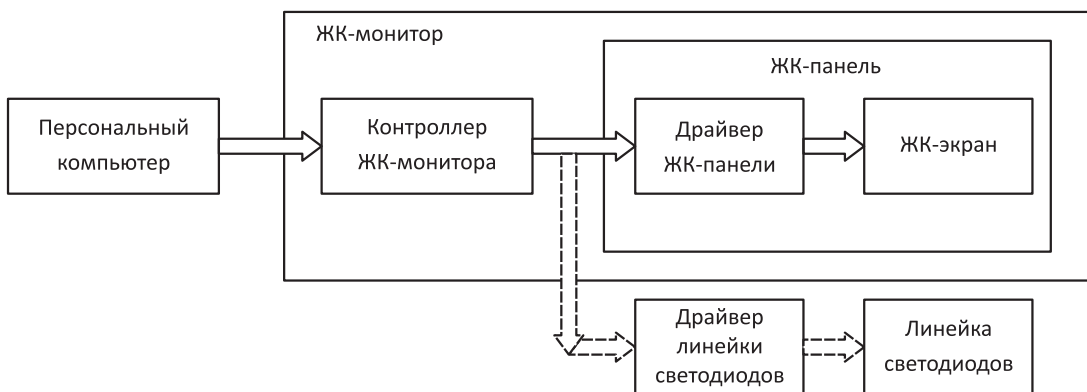


Рис. 6 Аппаратные средства для управления светодиодами

горизонтальные шины ЖК-матрицы. Вообще говоря, для управления яркостью светодиодов в линейке можно было бы использовать эти сигналы. Однако три обстоятельства, относящиеся к технической реализации, препятствуют этому.

Во-первых, электрические параметры этих сигналов (мощность и напряжение) не соответствуют требованиям светодиодов. Необходимо ставить дополнительный преобразователь для каждого такого сигнала.

Во-вторых, на выходе драйвера шаг проводников, по которым сигналы поступают на шины ЖК-матрицы, составляет около 0,05 мм. Подсоединиться к таким проводникам без специального технологического оборудования не представляется возможным.

В-третьих, различаются способы получения полутонов для жидкокристаллических ячеек и светодиодов.

Таким образом, только чисто технические проблемы не позволяют использовать уже имеющееся оборудование для модуляции светодиодов. Что касается алгоритмов работы аппаратуры монитора, то они такие же, как и для линейки светодиодов. Действительно, существующая аппаратура обеспечивает подачу сигналов на вертикальные шины ЖК-матрицы и одну из горизонтальных шин. При этом высвечивается одна строка пикселей ЖК-матрицы. Затем в следующий временной интервал высвечивается следующая строка пикселей.

При получении изображения с помощью светодиодов линейка светодиодов соответствует одной строке матрицы пикселей. Следующая строка изображения формируется той же линейкой светодиодов за счет того, что в следующем временном интервале изображение линейки светодиодов оказывается сдвинутым из-за колебаний зеркала. Таким образом, требования к аппаратуре со стороны линейки светодиодов даже слабее, чем требования к аппаратуре ЖК-монитора, так как отсутствует необходимость подачи сигналов на горизонтальные шины (такие шины просто отсутствуют). Число светодиодов в линейке приблизительно в 1000 раз (по числу строк в мониторе) меньше числа пикселей в мониторе.

С учетом приведенных соображений наиболее целесообразно использовать контроллер ЖК-монитора, а драйвер для светодиодов разработать с учетом требований со стороны последних и имеющегося стандартного оборудования, выпускаемого в настоящее время многими фирмами для управления яркостью светодиодов. На альтернативное оборудование, не входящее в состав монитора, указывают на рис. 6 пунктирные стрелки. В настоящее время развитие техники ЖК-мониторов достигло

такого уровня, что стали коммерчески доступны контроллеры для ЖК-мониторов. Примером может служить появившийся в последнее время контроллер типа SVT-1920. До этого каждая фирма, производящая ЖК-мониторы, разрабатывала собственные контроллеры, правда на основе стандартных, предназначенных для этих целей сверхбольших интегральных схем (СБИС).

Аппаратные средства светодиодных табло типа бегущая строка также могут быть использованы для модуляции линейки светодиодов. В этом случае табло также представляется в виде матрицы, число строк в которой ограничено несколькими десятками. Формирование изображения происходит точно так же, как и в ЖК-матрице. Однако в бегущих строках отсутствует реализация полутонов. Использование техники широтно-импульсной модуляции (ШИМ) позволяет решить эту проблему. Таким образом, уже существующие в настоящее время аппаратные средства можно приспособить для формирования изображений от тех же источников, что и на экране монитора ПК (телевизионный приемник, видеокамера, CD, Интернет и т. д.).

Что касается второго компонента рассматриваемого устройства, т. е. зеркала, совершающего вращательно-колебательные движения, то такое зеркало может быть самых разных размеров в зависимости от области применения. В работе [2] приведены фотографии нескольких конструкций таких зеркал. Различные зеркала могут использоваться с одной и той же линейкой светодиодов. Необходимо лишь обеспечить, чтобы период и фаза колебаний различных зеркал были одинаковы.

Для этих целей предназначен третий компонент, состоящий из радиопередатчика, связанного с линейкой светодиодов и радиоприемниками, по одному около каждого колеблющегося зеркала. При начале показа очередного кадра на линейке светодиодов передатчик выдает радиосигнал, который принимается каждым приемником и поступает на систему автоподстройки частоты, обеспечивающей определенную фазу колебаний зеркала в момент прихода этого сигнала. Любая из существующих в настоящее время систем беспроводной связи может быть использована для решения этой простейшей задачи.

Некоторые сомнения у потенциальных потребителей вызывает то обстоятельство, что в устройстве используется механическое движение в виде вращательных колебаний. Такой тип механического движения рассматривается исключительно в целях наглядности. Сплошное плоское зеркало может быть заменено уже существующими и широко используемыми в DLP (digital light processing) проекторах DMD (digital micromirror device) матри-

цами [3]. Эти матрицы представляют собой множество плоских микрзеркал, угол поворота каждого из которых задается соответствующим электрическим сигналом. Такие зеркала в состоянии совершать колебательно-вращательные движения с частотой в десятки кГц. В рассматриваемом применении условия работы таких зеркал гораздо проще. Все они должны синхронно совершать одинаковые вращательно-колебательные движения с частотой около 100 Гц.

Кроме того, необходимое изменение угла отражения лучей от зеркала может быть получено и без использования механических перемещений. С этой целью колеблющиеся зеркала могут быть заменены неподвижными дифракционными решетками, период которых управляется электрическими сигналами. Угол отклонения проходящих через дифракционную решетку лучей зависит от ее периода. При изменении периода изменяется угол отклонения лучей и у зрителя создается впечатление, что лучи попадают в его глаза по разным направлениям. Один из вариантов такой дифракционной решетки описан в патенте [4]. Известно также несколько более ранних патентов с описанием дифракционных решеток, в которых угол отклонения светового луча управляется электрическими сигналами [5].

5 Области применения

Изложенный подход к формированию изображений может быть использован в различных применениях. Рассмотрим некоторые из них.

5.1 Просмотр телевизионных передач

Как известно, в настоящее время громоздкие телевизоры с электронно-лучевыми трубками постепенно вытесняются еще более громоздкими и дорогостоящими телевизорами с использованием жидкокристаллических или плазменных плоских панелей. Габариты телевизора определяются размером его экрана, который уже может достигать двух метров по диагонали. Такой телевизор весит более 100 кг, и для него требуется достаточно большое помещение. Вместе с тем рассмотренный подход позволяет формировать телевизионные изображения, размер которых значительно превышает этот предел и превосходит даже размеры комнаты.

На рис. 7 показана схема такого формирования. Для этого требуется линейка светодиодов и колеблющееся зеркало. В частном случае линейка светодиодов может быть прикреплена к потолку, а колеблющееся зеркало размещено на столе. Размеры этого зеркала сравнимы с тарелкой, которая

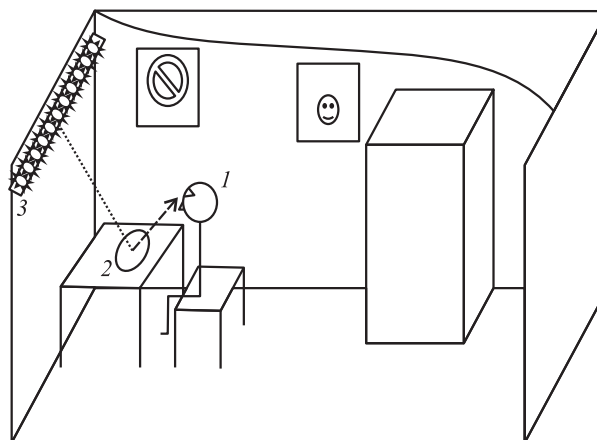


Рис. 7 Схема просмотра телевизионной передачи зрителем 1 в колеблющемся зеркале 2, в котором отражается линейка светодиодов 3

представляется зрителю в виде волшебного блюда, в котором он может наблюдать изображение, превосходящее размеры комнаты. Действительно, если зеркало колеблется вокруг горизонтальной оси, параллельной линейке светодиодов, то горизонтальный размер изображения равен длине этой линейки, а вертикальный размер определяется амплитудой колебаний зеркала и может быть значительно больше. Весьма ценным свойством такой установки является то обстоятельство, что зритель при просмотре телевизионных программ не мешает другим обитателям помещения. Ему предоставлено отверстие в виртуальный мир, через которое он может наблюдать крупноформатное изображение. При желании другие зрители могут наблюдать виртуальный мир через то же или другое колеблющееся зеркало. Следует отметить, что колеблющееся зеркало представляет собой чрезвычайно простое и дешевое устройство, состоящее из обычного плоского зеркала и электромотора от компьютерного вентилятора, обеспечивающего изменение ориентации зеркала. Что касается линейки светодиодов, то она может использоваться не только для формирования изображений, но и в качестве осветительного прибора. По существующим прогнозам, светодиодные источники света придут на смену традиционным. В этом случае линейка светодиодов может использоваться в качестве источника света, который может регулироваться в широких пределах по интенсивности, спектральному составу и местоположению.

5.2 Полиэкранный монитор

В некоторых применениях одного экрана обычного монитора оказывается недостаточно.

Например, рабочее место брокера на бирже окружено многочисленными экранами, которые позволяют ему оперативно отслеживать текущую ситуацию. При применении рассмотренного устройства брокер получит возможность наблюдать фрагмент крупноформатного изображения, состоящего из изображений многих экранов. Чтобы перевести взгляд с одного экрана на другой, ему достаточно изменить угол зрения или ориентацию зеркала. Отметим, что сложность обычных мониторов пропорциональна числу пикселей, т. е. пропорциональна квадрату линейного размера. Сложность рассматриваемого устройства пропорциональна линейному размеру. При этом увеличение числа пикселей по направлению движения изображений светодиодов не требует дополнительного оборудования, при условии что используемая аппаратура имеет запас по быстродействию.

5.3 Экран мобильного телефона

Рассмотренный подход позволяет преодолеть основное противоречие современных мобильных телефонов между необходимостью иметь малогабаритный аппарат, умещающийся в кармане, и желанием рассматривать с помощью телефона крупноформатные изображения. Если на экране мобильного телефона 1 (рис. 8) реализовать отверстие в виртуальный мир, то через это отверстие можно рассматривать крупноформатные изображения 22, размеры которых превосходят длину линейки светодиодов 2. Такая линейка размером в несколько десятков сантиметров может быть размещена на голове зрителя. В настоящее время многие слушатели аудиоплееров размещают на голове наушники. Зрители видеопрограмм или читатели книг, газет и журналов могут размещать на голове линейку светодиодов. Так они смогут просматривать электронные тексты не на экране размером со спичечную коробку, а на площади развернутой кни-

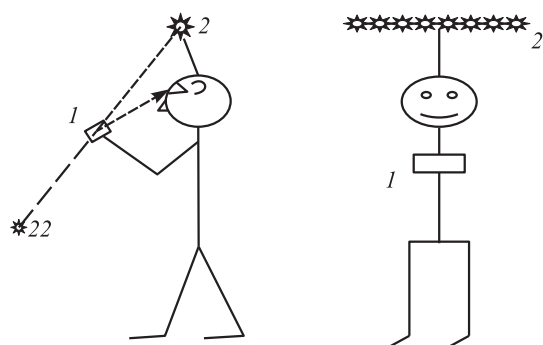


Рис. 8 Схема получения крупноформатного изображения на экране мобильного телефона

ги. При этом места в пространстве для размещения этой площади не требуется.

5.4 Рекламный щит в виде окна в виртуальный мир

В этом приложении размеры колеблющегося зеркала желательно иметь максимально возможными, чтобы получить «окно» в виртуальный мир максимального размера. Однако при этом возникают трудности, связанные с обеспечением необходимого размаха колебаний зеркала большого размера. Эти трудности могут быть преодолены несколькими способами.

Во-первых, плоское зеркало может быть составлено из нескольких плоских зеркал меньшего размера, как показано на рис. 9, а. Зеркала вращаются вдоль оси, перпендикулярной плоскости рисунка. Нетрудно убедиться, что при отклонении зеркал от горизонтального положения изображение светодиода будет сначала видно в одном зеркале и при его вращении будет перемещаться от одного края к другому. После этого изображение будет видно в соседнем зеркале, где оно будет также перемещаться в том же направлении. В результате такая система зеркал обеспечивает перемещение изображения светодиода в горизонтальном направлении и, следовательно, точно так же может быть использована для формирования изображения.

Во-вторых, может быть использована система зеркал, составленная из правильных треугольных призм, как показано на рис. 9, б. Призмы вращаются вокруг осей, перпендикулярных их основаниям и проходящих через центр основания. Боковые грани призм образованы плоскими зеркалами. В этом случае, так же как и в предыдущем варианте, изображение светодиода перемещается в горизонтальном направлении и переходит последовательно от одной призмы к соседней. За один оборот призмы

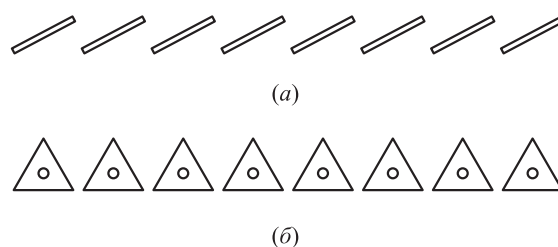


Рис. 9 Варианты замены колеблющегося сплошного зеркала: (а) множество одинаковых синхронно колеблющихся зеркал; (б) множество одинаковых правильных призм с зеркальными боковыми поверхностями, синхронно вращающихся в одном направлении

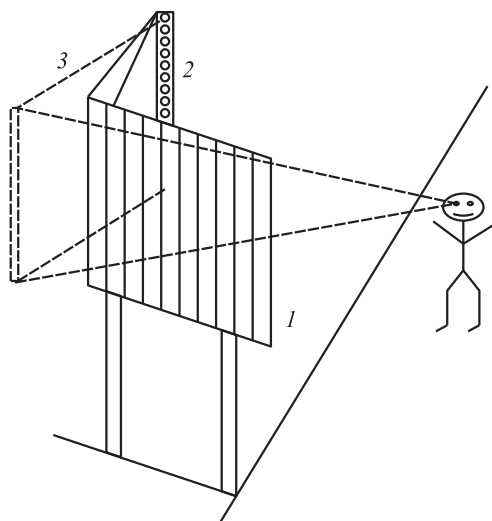


Рис. 10 Взаимное расположение вращающихся призм, линейки светодиодов и зрителя, при котором для него создается «окно» в виртуальный мир

изображение светодиода перемещается в горизонтальном направлении 3 раза. Это обстоятельство позволяет втрое уменьшить угловую скорость вращения призм по сравнению со скоростью вращения плоских зеркал на рис. 9, а. Заметим, что устройства, аналогичные представленным на рис. 9, б, широко используются в рекламе для поочередного показа трех изображений, фрагменты которых прикреплены к соответствующим боковым граням призм. При этом формируются картины стандартного для рекламных щитов размера 6×3 м.

На рис. 10 показана схема формирования динамических изображений на стандартном рекламном щите с использованием рассматриваемого подхода. Рекламный щит состоит из вращающихся призм с зеркальными боковыми стенками, показанными на рис. 9, б. В зеркалах отражается линейка светодиодов 2. Эту линейку могут наблюдать зрители, один из которых показан на рис. 10. Высота получаемого при этом изображения может несколько превосходить высоту зеркальных призм. Заметим, что формируемое изображение может быть видно практически из любого места полупространства, удаленного от рекламного щита на расстояние, превосходящее некоторый предел R . При этом чем вы-

ше расположены вращающиеся призмы, тем больше R .

Например, если расстояние между светодиодами 3 мм, то линейка из 1000 таких светодиодов имеет длину 3 м. В этом случае формируется изображение в 3 м высотой с разрешением 1000 пикселей по вертикали. Размеры изображения и разрешение по горизонтали могут быть в несколько раз больше.

У зрителя создается впечатление, что он наблюдает изображение в виртуальном мире через отверстие, площадь которого совпадает с площадью рекламного щита. Это впечатление усиливается, когда он движется относительно щита. В этом случае для него постепенно открываются новые фрагменты изображения и пропадают некоторые из уже открытых. Аналогичная ситуация имеет место при наблюдении с разных позиций в обычном неподвижном зеркале отраженных предметов реального мира.

Представлено лишь несколько областей применения рассматриваемого подхода с целью демонстрации его перспектив. В настоящее время создаются опытные образцы, чтобы исследовать его особенности. Результаты этих исследований предполагается опубликовать в отдельном сообщении. Однако уже в настоящее время на основе проведенного теоретического анализа и первых испытаний экспериментальных образцов можно утверждать, что рассматриваемый подход обладает многими весьма привлекательными свойствами, которые могут быть использованы в различных применениях.

Литература

1. Торчигин А. В. Способ формирования изображений и устройство для его осуществления. Патент РФ RU № 2328024 С2 от 29.12.2003.
2. Торчигин А. В. Формирование изображений движущимися источниками света. — М.: ИПИ РАН, 2008.
3. Бурняшев А. Современные мощные светодиоды и их оптика // Современная электроника, 2006. № 1. С. 24–27.
4. Торчигин В. П. Способ формирования стереоизображений. Патент РФ RU № 2337386 С2 от 02.11.2006.
5. Патенты США № 7420737 В2 (2008), № 7286292 В2 (2007), № 6903872 В2 (2005), № 5151814 (1992).

ОЦЕНКИ СКОРОСТИ СХОДИМОСТИ РАСПРЕДЕЛЕНИЙ ЭКСТРЕМУМОВ ОБОБЩЕННЫХ ПРОЦЕССОВ КОКСА С НЕНУЛЕВЫМ СРЕДНИМ К СДВИГОВЫМ СМЕСЯМ НОРМАЛЬНЫХ ЗАКОНОВ

С. В. Артюхов¹

Аннотация: В статье рассматриваются математические модели катастрофически накапливающихся эффектов, связанных с неоднородными хаотическими потоками экстремальных событий, — экстремумы обобщенных дважды стохастических пуассоновских процессов (обобщенных процессов Кокса) с ненулевым средним. Получены оценки скорости сходимости в предельных теоремах для экстремумов обобщенных процессов Кокса. Приведен пример существования нетривиального предела одномерных распределений экстремумов таких процессов с бесконечной дисперсией при нормировке, традиционной для сумм с конечной дисперсией.

Ключевые слова: экстремум; обобщенный дважды стохастический пуассоновский процесс; обобщенный процесс Кокса; сдвиговая смесь нормальных законов; оценка скорости сходимости

В данной статье рассматриваются математические модели катастрофически накапливающихся эффектов, связанных с неоднородными хаотическими потоками экстремальных событий. В качестве таких моделей берутся экстремумы обобщенных дважды стохастических пуассоновских процессов. Как показано, например, в книге [1], такие модели, во-первых, могут быть вполне адекватны при оценивании рисков некоторых природных катастроф и, во-вторых, дают более реалистичные оценки по сравнению с классическими моделями, недооценивающими риски в случае существенно непостоянной интенсивности потока экстремальных событий.

Пусть X_1, X_2, \dots — одинаково распределенные случайные величины, а $N(t)$ — дважды стохастический пуассоновский процесс (процесс Кокса), управляемый процессом $\Lambda(t)$, т. е.

$$N(t) = N_1(\Lambda(t)), \quad t \geq 0,$$

где $\Lambda(t)$ — случайный процесс с неубывающими непрерывными справа траекториями, выходящими из нуля, $N_1(t)$ — стандартный пуассоновский процесс (однородный пуассоновский процесс с единичной интенсивностью), причем процессы $N_1(t)$ и $\Lambda(t)$ независимы.

Предположим, что при каждом $t \geq 0$ случайные величины $N(t), X_1, X_2, \dots$ независимы. Процесс

$$S(t) = \sum_{j=1}^{N(t)} X_j, \quad t \geq 0, \quad (1)$$

назовем *обобщенным процессом Кокса* (при этом для определенности будем считать, что $\sum_{j=1}^0 = 0$). Процессы вида (1) играют чрезвычайно важную роль во многих прикладных задачах. Достаточно сказать, что при $\Lambda(t) \equiv \ell t$ с $\ell > 0$ процесс $S(t)$ превращается в классический обобщенный пуассоновский процесс, широко используемый при моделировании многих явлений в физике, теории надежности, финансовой и актуарной деятельности, биологии и т. д. Большое число разнообразных прикладных задач, приводящих к обобщенным пуассоновским процессам, описано в книгах [1–3]. Обобщенные процессы Кокса играют важную роль при моделировании характеристик неоднородных хаотических стохастических потоков случайных событий. При этом особую важность при использовании обобщенных процессов Кокса вида (1), скажем в страховой математике при моделировании неоднородных потоков страховых выплат или в теории управления запасами при моделировании неоднородных потоков заявок на поставку некоторого продукта, имеет тот случай, когда математическое ожидание слагаемых X_j в сумме (1) отлично от нуля (в частности, в условиях приведенных примеров — положительно).

В данной статье изучается асимптотическое поведение экстремумов обобщенного процесса Кок-

¹Московский государственный университет им. М. В. Ломоносова, факультет ВМиК, ArtyuhovSV@yandex.ru

са. Более точно, объектом изучения является процесс

$$\bar{S}(t) = \max_{0 \leq \tau \leq t} S(\tau),$$

где процесс $S(t)$ определен в (1). Всюду далее предполагается, что $EX_1 = a \neq 0$, $DX_1 = \sigma^2 < \infty$.

В качестве модели для катастрофически накапливающихся неблагоприятных воздействий в неоднородных хаотических потоках событий рассматриваются экстремумы обобщенных дважды стохастических пуассоновских процессов. Целесообразность рассмотрения подобных моделей при прогнозировании рисков катастроф диктуется следующими примерами.

Пример 1. Эволюция финансовых индексов хорошо описывается обобщенным дважды стохастическим пуассоновским процессом, накопленная интенсивность потока событий в котором определяется потоком новостей (см., например, [4]). При этом, как известно, если изменение этого индекса в течение биржевого дня оказывается слишком большим, то во избежание слишком больших потерь (т. е. финансовых катастроф) торги автоматически прекращаются. Другими словами, когда экстремум процесса, описывающего динамику финансового индекса, достигает критического значения, торги прекращаются.

Пример 2. Для снабжения некоторой отрасли некоторого региона в течение фиксированного периода времени (например, квартала или зимнего периода) на склад (хранилище) выделяется определенное количество некоторого ресурса (скажем, топлива). Если суммарный (накопленный) расход этого ресурса в течение указанного времени превысит выделенный лимит, то в данной отрасли в данном регионе наступит катастрофический коллапс. При этом естественно предположить, что ресурс отпускается потребителям партиями, вообще говоря, случайного объема согласно запросам, возникающим, вообще говоря, в случайные моменты времени. Поскольку интенсивность потока требований обусловлена не поддающимися абсолютно точному прогнозированию факторами (например, погодными условиями или политическими событиями), наиболее разумной моделью потока требований является обобщенный процесс Кокса.

Пример 3. Количество воды в некотором резервуаре (водохранилище, бассейне реки, озере и т. п.) изменяется случайным образом: оно увеличивается за счет выпадения осадков (в случайные моменты времени) и уменьшается за счет испарения. Если накопленное количество воды выходит за критический уровень, то происходят события катастрофи-

ческого характера. Избыток воды вызывает наводнения, ее недостаток — засуху. При этом испарение или забор воды из водохранилища происходит со случайной интенсивностью, обусловленной, к примеру, погодными условиями.

С формальной точки зрения целью данной статьи является изучение скорости сходимости в предельной теореме для экстремумов случайных сумм независимых одинаково распределенных случайных величин с ненулевым средним, в которых число слагаемых также случайно и изменяется во времени в соответствии с некоторым процессом Кокса.

В книге [1] доказан следующий результат.

Теорема 1. Пусть $a \neq 0$. Предположим, что $E\Lambda(t) \equiv t$ и $\Lambda(t) \xrightarrow{P} \infty$ при $t \rightarrow \infty$. Тогда одномерные распределения неслучайно центрированных и нормированных обобщенных процессов Кокса слабо сходятся к распределению некоторой случайной величины Z при $t \rightarrow \infty$, т. е.

$$\frac{\bar{S}(t) - at}{\sqrt{t(a^2 + \sigma^2)}} \Rightarrow Z \quad (t \rightarrow \infty),$$

тогда и только тогда, когда существует случайная величина V такая, что

$$P(Z < x) = E\Phi\left(x - \frac{aV}{\sqrt{\sigma^2 + a^2}}\right), \quad x \in \mathbb{R}, \quad (2)$$

и

$$\frac{\Lambda(t) - t}{\sqrt{t}} \Rightarrow V \quad (t \rightarrow \infty).$$

Здесь и далее $\Phi(x)$ — стандартная нормальная функция распределения, $x \in \mathbb{R}$. Соотношение (2) означает, что

$$Z \stackrel{d}{=} X + \frac{a}{\sqrt{a^2 + \sigma^2}} V,$$

где X — случайная величина со стандартным нормальным распределением, независимая от случайной величины V (символ $\stackrel{d}{=}$ обозначает совпадение распределений).

В книге [1] также приведены некоторые оценки скорости сходимости в теореме 1, но они справедливы для довольно узкого класса распределений предельной случайной величины V , довольно громоздки, содержат трудно вычисляемые характеристики и неудобны для анализа и применения. Здесь будет приведена легко интерпретируемая оценка скорости сходимости в теореме 1 в традиционных терминах.

Прежде чем сформулировать соответствующие результаты, введем дополнительные обозначения:

$$\begin{aligned} \mu^3 &= E|X_1 - a|^3; \\ L_3 &= \frac{\mu^3}{\sigma^3}; \\ L_3^* &= L_3 + \frac{\sigma^2}{a^2}; \\ F_t(x) &= P\left(\frac{\bar{S}(t) - at}{\sigma\sqrt{t}} < x\right); \\ \rho_t &= \sup_x \left| F_t(x) - E\Phi\left(\frac{\sigma x - aV}{\sqrt{a^2 + \sigma^2}}\right) \right|; \\ \Delta_t &= \sup_v \left| P\left(\frac{\Lambda(t) - t}{\sqrt{t}} < v\right) - P(V < v) \right|. \end{aligned}$$

Как видно из введенных обозначений, впредь будет использоваться несколько иная нормировка, нежели в теореме 1. Это сделано для удобства вычислений. Более того, в теореме 1 величина $\bar{S}(t)$ нормируется не ее дисперсией. Как будет показано ниже, для существования нетривиального предела экстремум обобщенного процесса Кокса можно нормировать просто величиной \sqrt{t} , т.е. приводимые результаты справедливы и для случая, когда дисперсия процесса $\bar{S}(t)$ бесконечна (за счет бесконечности дисперсии управляющего процесса).

Для получения оценки скорости сходимости в теореме 1 используется следующее утверждение.

Лемма 1. Пусть $\mu^3 < \infty$. Пусть N_λ — случайная величина, имеющая пуассоновское распределение с параметром $\lambda > 0$ и независимая от последовательности $\{X_j\}_{j \geq 1}$ независимых одинаково распределенных случайных величин. Тогда существует конечная положительная абсолютная константа C такая, что

$$\sup_x \left| P\left(\frac{\bar{S}_{N_\lambda}(t) - a\lambda}{\sigma\sqrt{\lambda}} < x\right) - \Phi\left(\frac{\sigma x}{\sqrt{a^2 + \sigma^2}}\right) \right| \leq \frac{C}{\lambda} L_3^*.$$

Доказательство см. в книге [1], с. 103.

Главным результатом данной статьи является следующее утверждение.

Теорема 2. Пусть $\mu^3 < \infty$, $E|V| < \infty$. Тогда справедлива оценка

$$\rho_t \leq \Delta_t + \frac{1}{\sqrt{t}} \inf_{\epsilon \in (0,1)} \left\{ \frac{CL_3^*}{\sqrt{1-\epsilon}} + \frac{E|V|}{\epsilon} + Q(\epsilon) E \left| \frac{\Lambda(t) - t}{\sqrt{t}} \right| \right\},$$

где

$$Q(\epsilon) = \max \left\{ \frac{1}{\epsilon}, \frac{\sqrt{1+\epsilon}}{\sqrt{2\pi e(1-\epsilon)}(1+\sqrt{1-\epsilon})} \right\}.$$

Доказательство. Имеем

$$\begin{aligned} \rho_t &= \sup_x \left| P\left(\frac{\bar{S}(t) - at}{\sigma\sqrt{t}} < x\right) - E\Phi\left(\frac{\sigma x - aV}{\sqrt{a^2 + \sigma^2}}\right) \right| = \\ &= \sup_x \left| \int_0^\infty P\left(\frac{\bar{S}_{N_\lambda}(t) - at}{\sigma\sqrt{t}} < x\right) dP(\Lambda(t) < \lambda) - E\Phi\left(\frac{\sigma x - aV}{\sqrt{a^2 + \sigma^2}}\right) \right|. \end{aligned}$$

Полученное представление величины ρ_t позволяет провести дальнейшее доказательство практически по тому же пути, которым доказана теорема 2 в работе [5] (устанавливающая аналогичные оценки для скорости сходимости распределений самих обобщенных процессов Кокса). Доказательства различаются лишь переобозначениями и тем, что вместо неравенства Берри–Эссеена для обобщенных пуассоновских распределений (см., например, [6]), использованного в работе [5], здесь используется лемма 1.

Отметим, что оценка, полученная в теореме 2, имеет компактную запись, легко вычислима и удобна для дальнейшего анализа. Однако, к сожалению, до настоящего момента конкретное значение абсолютной константы C не найдено.

Если дополнительно предположить, что семейство случайных величин $\{ |(\Lambda(t) - t)/\sqrt{t}| \}_{t>0}$ равномерно интегрируемо, то с помощью неравенства Ляпунова можно получить неравенство

$$\begin{aligned} E|V| &= \lim_{t \rightarrow \infty} E \left| \frac{\Lambda(t) - t}{\sqrt{t}} \right| \leq \\ &\leq \lim_{n \rightarrow \infty} \sqrt{D\left(\frac{\Lambda(t) - t}{\sqrt{t}}\right)} = 1. \quad (3) \end{aligned}$$

При этом из теоремы 2 очевидным образом вытекает следующий результат.

Следствие 1. Пусть в дополнение к условиям теоремы 1 при каждом $t > 0$ выполнено (3). Тогда справедлива оценка

$$\rho_t \leq \Delta_t + \frac{1}{\sqrt{t}} \inf_{\epsilon \in (0,1)} \left\{ \frac{CL_3^*}{\sqrt{1-\epsilon}} + \frac{1}{\epsilon} + Q(\epsilon) \right\},$$

где

$$Q(\epsilon) = \max \left\{ \frac{1}{\epsilon}, \frac{\sqrt{1+\epsilon}}{(1+\sqrt{1-\epsilon})\sqrt{2\pi e(1-\epsilon)}} \right\}.$$

Замечание 1. Пусть, к примеру, процесс $\Lambda(t)$ является пуассоновским с единичной интенсивностью. В таком случае распределение случайной величины $N(t)$ (см. (1)) при каждом $t > 0$ является так называемым пуассон-пуассоновским или (инфекционным) распределением Неймана типа А. Класс таких распределений введен Ю. Нейманом в работе [7] в связи с некоторыми задачами из области бактериологии и энтомологии. В этом случае величина Δ_t также имеет порядок $O(t^{-1/2})$, поскольку в условиях примера справедлива оценка

$$\Delta_t \leq \frac{C_0}{\sqrt{t}},$$

где C_0 — абсолютная постоянная, $C_0 \leq 0,7005$, причем случайная величина V имеет стандартное нормальное распределение (см., например, [6, 8, 9]), а следовательно, распределение случайной величины Z также является стандартным нормальным. В таком случае имеет место оценка

$$p_t \leq \frac{1}{\sqrt{t}} \inf_{\epsilon \in (0,1)} \left\{ C_0 + \frac{CL_3^*}{\sqrt{1-\epsilon}} + \frac{1}{\epsilon} + Q(\epsilon) \right\}.$$

Замечание 2. Другим примером управляющего процесса $\Lambda(t)$, при котором величина Δ_t также имеет порядок $O(t^{-1/2})$, служит ситуация, когда $\Lambda(t)$ является гамма-процессом Леви, приращение которого на единичном интервале имеет гамма-распределение с параметром масштаба $\lambda = \ell$ и параметром формы $\alpha = \ell^2$. В таком случае распределение случайной величины $N(t)$ (см. (1)) при каждом $t > 0$ является отрицательным биномиальным. При этом случайная величина V имеет стандартное нормальное распределение (см., например, [6]), а следовательно, распределение случайной величины Z опять-таки является стандартным нормальным.

Замечание 3. Если случайный процесс $N(t)$ является однородным пуассоновским, то логично считать, что распределение случайной величины V является вырожденным в нуле и $\Delta_t = 0$. Поэтому в таком случае оценка, приведенная в теореме 2, естественно переходит в оценку скорости сходимости распределений пуассоновских случайных сумм к нормальному закону, приведенную, например, в работе [8] с константой C_0 , уточненной в работе [9].

Замечание 4. Теорема 2 верна и в некоторых ситуациях, в которых дисперсия управляющего процесса $\Lambda(t)$ бесконечна (очевидно, в таком случае бесконечна и дисперсия процесса $\bar{S}(t)$). В качестве примера ситуации, на которую распространяется

действие теоремы 2, можно привести такую, в которой

$$\Lambda(t) = \max\{0, \sqrt{t} \cdot V + t\} + \frac{1}{2t^{\alpha/2}} \left(\frac{2\alpha + 1}{\alpha} \sqrt{t} - 1 \right),$$

где $2 < \alpha < 3$, а V — случайная величина с плотностью

$$p(x) = \frac{\alpha + 1}{2(|x| + 1)^\alpha}, \quad x \in \mathbb{R}.$$

Несложно проверить, что $E\Lambda(t) = t$ при любом $t > 0$, но второй момент случайной величины $\Lambda(t)$ бесконечен вследствие бесконечности второго момента случайной величины V (и, следовательно, бесконечен и второй момент экстремума обобщенного процесса Кокса $S(t)$, управляемого таким процессом $\Lambda(t)$). Однако при этом, как легко видеть,

$$\frac{\Lambda(t) - t}{\sqrt{t}} = \max\{-\sqrt{t}, V\} + \frac{1}{2t^{(\alpha+1)/2}} \left(\frac{2\alpha + 1}{\alpha} \sqrt{t} - 1 \right) \Rightarrow V$$

при $t \rightarrow \infty$. Этот случай является наглядной иллюстрацией очень интересного и нетривиального факта: в отличие от классической теории суммирования, для экстремумов сумм со случайным числом слагаемых (в частности, для обобщенных процессов Кокса) с бесконечной дисперсией существование нетривиальных слабых пределов возможно и при нормировке порядка $t^{1/2}$, являющейся «стандартной» в классической теории лишь для сумм с конечной дисперсией.

Литература

1. Королёв В. Ю., Соколов И. А. Математические модели неоднородных потоков экстремальных событий. — М.: ТОРУС-ПРЕСС, 2008.
2. Gnedenko B. V., Korolev V. Yu. Random summation: Limit theorems and applications. — Boca Raton: CRC Press, 1996.
3. Bening V. E., Korolev V. Yu. Generalized Poisson models and their applications in insurance and finance. — Utrecht: VSP, 2002.
4. Королёв В. Ю. Вероятностно-статистический анализ хаотических процессов с помощью смешанных гауссовских моделей. Декомпозиция волатильности финансовых индексов и турбулентной плазмы. — М.: ИПИ РАН, 2007.
5. Артюхов С. В., Королёв В. Ю. Оценки скорости сходимости распределений обобщенных дважды стохастических пуассоновских процессов с ненулевым средним к сдвиговым смесям нормальных законов //

- Обозрение промышленной и прикладной математики, 2008. Т. 15. Вып. 6. С. 988–998.
6. *Бенинг В. Е., Королёв В. Ю., Шоргин С. Я.* Математические основы теории риска. — М.: Физматлит, 2007.
7. *Neuman J.* On a new class of “contagious” distributions, applicable in entomology and bacteriology // *Ann. Math. Statist.*, 1939. Vol. 10. P. 35–57.
8. *Шевицова И. Г.* Уточнение структуры оценок скорости сходимости в центральной предельной теореме для сумм независимых случайных величин. Дис. . . . канд. физ.-мат. наук. — М.: МГУ, 2006.
9. *Шевицова И. Г.* Об абсолютной постоянной в неравенстве Берри–Эссеена // Сб. статей молодых ученых факультета ВМиК МГУ. Вып. 5. — М.: Изд-во факультета ВМиК МГУ, 2008. С. 101–110.

METHODS FOR INFORMATION MODEL BUILDING FOR THE EARTH TIDAL HEREDITARY IRREGULAR ROTATION

I. N. Sinitsyn

IPI RAN, sinitsin@dol.ru

Methods for information hereditary model building for the Earth tidal rotation fluctuations based on *a priori* and *a posteriori* data are considered. Linear and quasi-linear methods are developed. Equivalence of different hereditary disturbances is discussed. Experimental software is the part of the informational resources “Statistical dynamics of the Earth rotation.”

Keywords: *a priori* and *a posteriori* data; informational model; informational resources; quasi-linear methods; spectral-correlational characteristics; hereditary fluctuations of the Earth rotation; hereditary kernel

PARALLEL COMPUTING IN LARGE-SCALE MULTIMODAL BIOMETRIC SYSTEMS

O. S. Ushmaev

IPI RAN, oushmaev@ipiran.ru

The main topic is parallel computing in large-scale biometric identification systems. Approach of a biometric cluster throughput estimation, a cluster configuration estimation is proposed. Methods of organization of parallel computing for multimodal biometrics are developed.

Keywords: biometric identification; multimodal biometrics; parallel computing

DEVELOPMENT OF SUPERRESOLUTION-BASED FACE VIDEO ENHANCEMENT

A. V. Nasonov¹, A. S. Krylov², and O. S. Ushmaev³

¹M. V. Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, nasonov@cs.msu.ru

²M. V. Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, kryl@cs.msu.ru

³IPI RAN, oushmaev@ipiran.ru

General superresolution-based method of face image enhancement for video data has been suggested. The superresolution is modeled as inverse problem to image downsampling, i.e., it finds an image that gives the minimal value of the quadratic discrepancy with initial low-resolution images after the motion dependent downsampling. High-quality superresolution method and fast superresolution methods are considered. Special deringing method for fast superresolution is proposed. New multiscale motion estimation method has been developed.

Keywords: superresolution; deringing; facial video sequence; multiscale motion estimation; fast superresolution

RECONSTRUCTION OF PROBABILISTIC CHARACTERISTICS OF RANDOM FUNCTIONS IN SPECT PROBLEMS

V. G. Ushakov¹ and O. V. Shestakov²

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University; IPI RAN, vgushakov@mail.ru

²Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, oshestakov@cs.msu.su

The problem of reconstruction of probabilistic characteristics of an object which structure changes in a random manner during the process of projection data acquisition is considered. Within the frames of proposed tomography experiment model, a method to reconstruct distributions of a random function from distributions of projections in the case when random function has at most denumerable number of states is developed.

Keywords: single-photon emission computer tomography (SPECT); stochastic tomography; exponential Radon transform; random functions; projection data

A DESIGN CONCEPT OF DOMESTIC INTEGRATED COMMUNICATION MICROCONTROLLERS FOR PACKET SWITCHING

V. B. Egorov

¹IPI RAN, vegorov@ipiran.ru

A concept of simplified integrated communication microcontrollers, which could be applied in various domestic packet switching and routing devices, with leveraging their functionality and facilitating development is suggested.

Keywords: integrated communication microcontroller; PowerQUICC; decentralized switching; routing switch

ON NONSTATIONARY QUEUEING SYSTEMS WITH CATASTROPHES

A. I. Zeifman¹, Ya. A. Satin², and A. V. Chegodaev³

¹Vologda State Pedagogical University; IPI RAN; VSCC CEMI RAN, a_zeifman@mail.ru

²Vologda State Pedagogical University, yacovi@mail.ru

³Vologda State Pedagogical University, cheg_al@mail.ru

Nonstationary birth and death processes with catastrophes are considered. The bounds of the rate of convergence to the limit regime and the estimates of the limit probabilities are obtained. Also, the bounds for the mean of the process are studied and a queueing example is considered.

Keywords: nonstationary queues; Markovian models with catastrophes; weak ergodicity; bounds; limiting characteristics; approximations

BAYESIAN QUEUING AND RELIABILITY MODELS: AN EXPONENTIAL-ERLANG CASE

A. A. Kudriavtsev¹ and S. Ya. Shorgin²

¹Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University, nubigena@hotmail.com

²IPI RAN, sshorgin@ipiran.ru

The investigation of Bayesian queuing and reliability models is continued in the paper. The method provides the randomization of system characteristics with regard to *a priori* distributions of input parameters. This approach could be used to calculate average values of performance and reliability characteristics for the large groups of systems or devices. The new results are presented for a case of exponential and Erlang *a priori* distributions.

Keywords: Bayesian approach; queuing systems; reliability; mixed distributions; modeling; Erlang distribution; exponential distribution

ON ONE APPROACH TO IMAGE PRODUCTION WITHOUT SCREENS

A. V. Torchigin

IPI RAN, torchigin.a@mail.ru

Properties of images observed in an oscillating mirror, where LEDs modulated by brightness are reflected, are considered. Possible areas of application of this approach are analyzed.

Keywords: image production; virtual environment; virtual reality; stereo images

CONVERGENCE RATE ESTIMATES OF DISTRIBUTIONS OF EXTREMA OF COMPOUND COX PROCESSES WITH NONZERO MEANS TO LOCATION MIXTURES OF NORMAL LAWS

S. V. Artyukhov

Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University, ArtyuhovSV@yandex.ru

Mathematical models of catastrophically accumulating effects related to nonhomogeneous chaotic flows of extremal events are considered, namely, extrema of compound doubly stochastic Poisson processes (compound Cox processes) with nonzero expectation. Convergence rate estimates are obtained in limit theorems for extrema of compound Cox processes. An example is given of existence of nontrivial limit of one-dimensional distributions of extrema of such processes with infinite variance under normalization which is traditional for sums with finite variance.

Keywords: extremum; compound doubly stochastic Poisson process; compound Cox process; location mixture of normal laws; convergence rate estimates

Об авторах

Артюхов Сергей Владимирович (р. 1983) — аспирант факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Егоров Владимир Борисович (р. 1948) — кандидат технических наук, и.о. ведущего научного сотрудника ИПИ РАН

Зейфман Александр Израилевич (р. 1954) — доктор физико-математических наук, профессор, декан факультета прикладной математики и компьютерных технологий Вологодского государственного педагогического университета; старший научный сотрудник ИПИ РАН; ведущий научный сотрудник ВНКЦ ЦЭМИ РАН

Крылов Андрей Серджевич (р. 1956) — кандидат физико-математических наук, заведующий лабораторией математических методов обработки изображений факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Кудрявцев Алексей Андреевич (р. 1978) — кандидат физико-математических наук, старший научный сотрудник кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Насонов Андрей Владимирович (р. 1985) — младший научный сотрудник лаборатории математических методов обработки изображений факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Сатин Яков Александрович (р. 1978) — кандидат физико-математических наук, старший преподава-

тель кафедры математического анализа Вологодского государственного педагогического университета

Синицин Игорь Николаевич (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, заведующий отделом ИПИ РАН.

Торчигин Александр Владимирович (р. 1974) — научный сотрудник ИПИ РАН

Ушаков Владимир Георгиевич (р. 1952) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Ушмаев Олег Станиславович (р. 1981) — кандидат технических наук, старший научный сотрудник ИПИ РАН

Чегодаев Александр Вячеславович (р. 1983) — аспирант факультета прикладной математики и компьютерных технологий Вологодского государственного педагогического университета

Шестаков Олег Владимирович (р. 1976) — кандидат физико-математических наук, ассистент кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, заместитель директора ИПИ РАН

About Authors

Artuykhov Sergei V. (b. 1983) — PhD student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Chegodayev Alexander V. (b. 1983) — PhD student, Faculty of Applied Mathematics and Computer Technologies, Vologda State Pedagogical University

Egorov Vladimir B. (b. 1948) — Candidate of Technical Sciences (PhD), leading scientist, Institute of Informatics Problems, Russian Academy of Sciences

Krylov Andrey S. (b. 1956) — Candidate of Science (PhD) in physics and mathematics; associate professor; Head of the Laboratory of Mathematical Methods of Image Processing, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Kudriavtsev Alexei A. (b. 1978) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Nasonov Andrey V. (b. 1985) — junior researcher, Laboratory of Mathematical Methods of Image Processing, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Satin Yakov A. (b. 1978) — Candidate of Science in physics and mathematics, senior lecturer, Vologda State Pedagogical University

Shestakov Oleg V. (b. 1976) — Candidate of Science (PhD) in physics and mathematics, assistant profes-

sor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Shorgin Sergey Ya. (b. 1952) — Doctor of Science in physics and mathematics, professor, Deputy Director of Institute of Informatics Problems, Russian Academy of Sciences

Sinitsyn Igor N. (b. 1940) — Doctor of Science in technology, professor, honored scientist of RF, Head of the Department in Institute of Informatics Problems, Russian Academy of Sciences

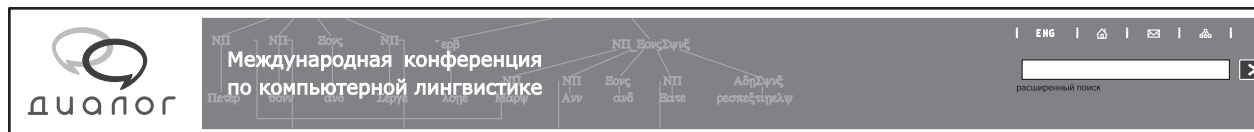
Torchigin Alexander V. (b. 1974) — scientist, Institute of Informatics Problems, Russian Academy of Sciences

Ushakov Vladimir G. (b. 1952) — Doctor of Science in physics and mathematics, professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Ushmaev Oleg S. (b. 1981) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Russian Academy of Sciences

Zeifman Alexander I. (b. 1954) — Doctor of Science in physics and mathematics; professor, Dean of the Faculty of Applied Mathematics and Computer Technologies, Vologda State Pedagogical University; senior researcher, Institute of Informatics Problems, Russian Academy of Sciences; leading researcher, VSCC CEMI, Russian Academy of Sciences

ОБЪЯВЛЕНИЯ О КОНФЕРЕНЦИЯХ



<http://www.dialog-21.ru/>

XV Международная конференция «ДИАЛОГ 2009» «Компьютерная лингвистика и интеллектуальные технологии»

27–31 мая 2009 г., Московская область, Наро-Фоминский район, Россия

Конференция проходит под патронажем Российского фонда фундаментальных исследований и при организационной поддержке компании АBBYY.

В подготовке и проведении «Диалога 2009» участвуют:

- Филологический факультет МГУ
- Институт лингвистики РГГУ
- Институт проблем информатики РАН
- Институт проблем передачи информации РАН
- Российский НИИ искусственного интеллекта
- Яндекс (Москва)

Конференция проводится по следующим направлениям, сочетающим теоретические исследования и приложения:

- Лингвистическая семантика и семантический анализ
- Формальные модели языка и их применение
- Теоретическая и компьютерная лексикография
- Создание и применение компьютерных лексических ресурсов
- Корпусная лингвистика. Создание, применение, оценка корпусов
- Интернет как лингвистический ресурс. Лингвистические технологии в интернете
- Извлечение знаний из текстов
- Модели общения. Коммуникация, диалог и речевой акт
- Анализ и синтез речи
- Компьютерный анализ документов: реферирование, классификация, поиск
- Машинный перевод
- Вопросно-ответные системы

ОБЪЯВЛЕНИЯ О КОНФЕРЕНЦИЯХ



**ЭЛЕКТРОННЫЕ БИБЛИОТЕКИ:
ПЕРСПЕКТИВНЫЕ МЕТОДЫ И ТЕХНОЛОГИИ,
ЭЛЕКТРОННЫЕ КОЛЛЕКЦИИ**
XI Всероссийская научная конференция

➤ English version

Петрозаводск,
17 - 21 сентября, 2009

<http://rcdl2009.krc.karelia.ru/>

XI Всероссийская научная конференция RCDL 2009

Электронные библиотеки: перспективные методы и технологии, электронные коллекции

17–21 сентября 2009 г., Петрозаводск, Россия

Электронные библиотеки — область исследований и разработок, направленных на развитие теории и практики обработки, распространения, хранения, поиска и анализа цифровых данных различной природы.

Основная цель серии конференций RCDL заключается в том, чтобы способствовать формированию сообщества специалистов России, ведущих исследования и разработки в области электронных библиотек. Конференция также способствует изучению зарубежного опыта, развитию международного сотрудничества в области электронных библиотек.

Значительное внимание в тематике RCDL уделяется практическим проектам, в которых решаются сложные задачи. RCDL придает большое значение исследованиям в области создания крупномасштабных электронных библиотек (Very Large Digital Libraries — VLDL), включая использование сервисных архитектур, архитектур, основанных на грид, и обеспечение их качества, развитие техники интероперабельности и устойчивости VLDL, а также разработке организационных моделей крупных электронных библиотек. Особый интерес представляет применение современных научных подходов в контексте высоких нагрузок: сотни тысяч пользователей, десятки гигабайт данных, терабайты трафика.

За 10 лет проведения RCDL в работе конференции приняло участие несколько сотен ученых из ведущих российских и зарубежных научных центров Австрии, Германии, Греции, Италии, Новой Зеландии, США, Украины и других стран.

Традиционно совместно с RCDL проводятся Всероссийские научные семинары по оценке методов текстового поиска РОМИП. В 2009 году планируется совмещение с RCDL Семинара РОМИП и Третьей Российской летней школы по информационному поиску RuSSIR'2009, во время которой ведущие российские и зарубежные ученые прочитают обзорные лекции по актуальным проблемам развития поиска цифровых данных для решения фундаментальных и прикладных задач.

Организаторы конференции RCDL'2009:

- Российская академия наук
- Российский фонд фундаментальных исследований
- Карельский научный центр РАН
- Институт прикладных математических исследований
- Петрозаводский государственный университет
- Институт проблем информатики РАН
- Московская секция ACM SIGMOD

Правила подготовки рукописей статей для публикации в журнале «Информатика и её применения»

Журнал «Информатика и её применения» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информатики и ее приложений. Журнал издается на русском языке. Тематика журнала охватывает следующие направления:

- теоретические основы информатики;
- математические методы исследования сложных систем и процессов;
- информационные системы и сети;
- информационные технологии;
- архитектура и программное обеспечение вычислительных комплексов и сетей.

1. В журнале печатаются результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях. Публикация не должна нарушать закон об авторских правах. Направляя свою рукопись в редакцию, авторы автоматически передают учредителям и редколлегии неисключительные права на издание данной статьи на русском языке и на ее распространение в России и за рубежом. При этом за авторами сохраняются все права как собственников данной рукописи. В связи с этим авторами должно быть представлено в редакцию письмо в следующей форме: Соглашение о передаче права на публикацию:

«Мы, нижеподписавшиеся, авторы рукописи « _____ », передаем учредителям и редколлегии журнала «Информатика и её применения» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторского права других лиц или организаций. Подписи авторов: (ф. и. о., дата, адрес)».

Редколлегия вправе запросить у авторов экспертное заключение о возможности опубликования представленной статьи в открытой печати.

2. Статья подписывается всеми авторами. На отдельном листе представляются данные автора (или всех авторов): фамилия, полное имя и отчество, телефон, факс, e-mail, почтовый адрес. Если работа выполнена несколькими авторами, указывается фамилия одного из них, ответственного за переписку с редакцией.

3. Редакция журнала осуществляет самостоятельную экспертизу присланных статей. Возвращение рукописи на доработку не означает, что статья уже принята к печати. Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редакционной коллегии о принятии статьи к печати или ее отклонении сообщается авторам. Редколлегия не обязуется направлять рецензию авторам отклоненной статьи.

5. Корректурa статей высылается авторам для просмотра. Редакция просит авторов присылать свои замечания в кратчайшие сроки.

6. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки. Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху — 2, снизу — 2, от края до нижнего колонтитула — 1,3. Основной текст: стиль — «Обычный»; шрифт Times New Roman, размер 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине. Рекомендуемый объем рукописи — не свыше 25 страниц указанного формата. Ознакомиться с шаблонами, содержащими примеры оформления, можно по адресу в Интернете: <http://www.ipiran.ru/journal/template.doc>.

7. К рукописи, предоставляемой в 2-х экземплярах, обязательно прилагается электронная версия статьи (как правило, в форматах MS WORD (.doc) или LaTeX (.tex), а также — дополнительно — в формате .pdf) на дискете, лазерном диске или по электронной почте. Сокращения слов, кроме стандартных, не применяются. Все страницы рукописи должны быть пронумерованы.

8. Статья должна содержать следующую информацию на русском и английском языках: название, Ф.И.О. авторов, места работы авторов и их электронные адреса, аннотация (не более 100 слов), ключевые слова. Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в порядке их первого упоминания. Все фамилии авторов, заглавия статей, названия книг, конференций и т. п. даются на языке оригинала, если этот язык использует кириллический или латинский алфавит.

9. Присланные в редакцию материалы авторам не возвращаются.

10. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- использовать attach (присоединение);
- в случае больших объемов информации возможно использование общеизвестных архиваторов (ZIP, RAR);
- в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.

11. Журнал «Информатика и её применения» является некоммерческим изданием, и гонорар авторам не выплачивается.

Адрес редакции: Москва 119333, ул. Вавилова, д. 44, корп. 2, ИПИ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05 E-mail: rust@ipiran.ru