

Информатика и её применения

Том 18 Выпуск 1 Год 2024

СОДЕРЖАНИЕ

Субоптимальная фильтрация в стохастических системах, не разрешенных относительно производных, со случайными параметрами И. Н. Сеницын	2
Алгебраическая спецификация распределенных систем с изменяющейся архитектурой С. П. Ковалёв	11
Управление линейным выходом автономной дифференциальной системы по квадратичному критерию на бесконечном горизонте А. В. Босов	18
Рынок с марковской скачкообразной волатильностью IV: алгоритм мониторинга рыночной цены риска по потоку высокочастотных наблюдений базовых активов и деривативов А. В. Борисов	26
Равномерные оценки скорости сходимости для интегрального индекса баланса А. А. Кудрявцев, О. В. Шестаков	33
Сравнительный анализ узловых мультиточков в многопользовательской сетевой системе Ю. Е. Малашенко, И. А. Назарова	40
Локальная древовидность в конфигурационных графах со степенным распределением М. М. Лери, Ю. Л. Павлов	46
Статистический критерий стабильности системы массового обслуживания, основанный на входном и выходном потоках М. П. Кривенко	54
Модели синтеза компоновочной схемы в задаче формирования облика самолета Л. Л. Вышинский, Ю. А. Флёров	61
О порождении синтетических признаков на основе опорных цепей и произвольных метрик в рамках топологического подхода к анализу данных. Часть 1. Включение в формализм эмпирических функций расстояния И. Ю. Торшин	71
Логика обмана в машинном обучении А. А. Грушо, Н. А. Грушо, М. И. Забейло, В. О. Писковский, Е. Е. Тимонина, С. Я. Шоргин	78
Жизненный цикл методологии построения рефлексивно-активных систем искусственных гетерогенных интеллектуальных агентов С. В. Листопад	84
Анализ подходов к определению нечеткой резольвенты Т. М. Леденева, М. В. Лещинская	92
Об авторах	100
Правила подготовки рукописей	102
Requirements for manuscripts	105

СУБОПТИМАЛЬНАЯ ФИЛЬТРАЦИЯ В СТОХАСТИЧЕСКИХ СИСТЕМАХ, НЕ РАЗРЕШЕННЫХ ОТНОСИТЕЛЬНО ПРОИЗВОДНЫХ, СО СЛУЧАЙНЫМИ ПАРАМЕТРАМИ

И. Н. Синицын¹

Аннотация: Для наблюдаемых гауссовских дифференциальных стохастических систем (СтС), не разрешенных относительно производных (НРОП), со случайными параметрами в виде интегральных канонических представлений (ИКП), приводимых к дифференциальным СтС, разработано методическое обеспечение анализа точности субоптимальной фильтрации (СОФ). Представлен обзор результатов в области аналитического моделирования и СОФ, экстраполяции и идентификации. Приведены необходимые сведения из корреляционной теории скалярных и векторных многокомпонентных (МК) ИКП. Особое внимание уделено среднеквадратичной оптимальной регрессионной линеаризации, в том числе на основе МК ИКП. Представлено методическое обеспечение СОФ гауссовской дифференциальной СтС НРОП, приведенной к дифференциальной, на основе метода нормальной аппроксимации (МНА) для условных и безусловных относительно переменных случайных параметров, заданных МК ИКП. Особое внимание уделено фильтрационным уравнениям. Иллюстративный пример для одномерной системы, нелинейной относительно старшей производной и линейной относительно параметров и возмущений, иллюстрирует методику синтеза нормальной (гауссовской) СОФ (НСОФ) на основе МК ИКП.

Ключевые слова: метод нормальной аппроксимации (МНА); регрессионная линеаризация 1-го и 2-го рода; стохастический процесс; стохастическая система, не разрешенная относительно производной (СтС НРОП); субоптимальная фильтрация (СОФ)

DOI: 10.14357/19922264240101

EDN: KUWMKJ

1 Введение

В [1–4] рассмотрены вопросы аналитического моделирования процессов в СтС НРОП. Особое внимание в них уделено нормальным (гауссовским) стохастическим процессам (СтП). В [5] предложены методы нормализации систем, стохастически НРОП. Теория распределений с инвариантной мерой в СтС НРОП развита в [6].

Статья [7] посвящена нелинейным корреляционным методам аналитического моделирования (МAM) процессов в дифференциальных СтС НРОП. Дан обзор работ в области аналитического моделирования СтС НРОП. Приводятся необходимые сведения из теории ИКП случайных процессов и их линейных и нелинейных преобразований. Отдельный раздел посвящен СтС НРОП, приводимым к дифференциальным системам. Представлены основные алгоритмы МAM качества на основе ИКП для типовых СтС НРОП. Приведен пример, иллюстрирующий особенности СтС НРОП при нестационарных случайных параметрах, заданных ИКП.

В [8] рассмотрены вопросы синтеза НСОФ для дифференциальных СтС НРОП. Представлены

уравнения состояния и наблюдения нелинейных дифференциальных СтС НРОП. Синтез НСОФ выполнен при следующих условиях:

- (1) отсутствуют пуассоновские шумы в наблюдениях;
- (2) коэффициент при гауссовском шуме не зависит от состояния.

Подробно рассмотрен синтез НСОФ при аддитивных шумах в уравнениях состояния и наблюдения.

В [9] для нелинейных интегро-дифференциальных (ИД) СтС НРОП, приводимых к дифференциальным методом сингулярных ядер, разработаны алгоритмы аналитического моделирования нормальных СтП, при этом нелинейность под интегралом может быть разрывной, а также синтеза НСОФ для онлайн-обработки информации в ИД СтС. Предложены алгоритмы оценки качества НСОФ на основе теории чувствительности.

В [3] разработано методическое обеспечение для негладких правых частей уравнений СтС НРОП. Рассмотрены вопросы аналитического моделирования нормальных СтП на основе нелинейных регрессионных моделей. Особое внимание уделено методам гауссовской фильтрации и экстраполяции.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Московский авиационный институт, sinitsin@dol.ru

Изучены вопросы условно-оптимальной фильтрации и экстраполяции для СтС НРОП с параметрическими шумами.

В [10] разработано методическое и алгоритмическое обеспечение аналитического моделирования оценивания и идентификации для существенно нестационарных процессов (например, ударных) в СтС НРОП. Дан обзор профильных публикаций и изучены основные классы регрессионных уравнений СтС НРОП. Основные результаты:

- (1) для общего вида нелинейных СтС НРОП приведены оптимальные алгоритмы совместной фильтрации и распознавания;
- (2) для линейных гауссовских СтС НРОП получены простые алгоритмы;
- (3) для СтС НРОП, линейных относительно состояния X_t и нелинейных относительно Y_t наблюдений, получены соответствующие алгоритмы;
- (4) в случае 3 методом нормальной аппроксимации получен простой алгоритм. Приводится иллюстративный пример скалярной нелинейной гауссовской СтС НРОП.

Ставится задача разработки методического обеспечения среднеквадратичной (с. к.) СОФ процессов в СтС НРОП со случайными переменными параметрами, описываемыми ИКП. В разд. 2 приводятся сведения из теории ИКП, их линейных и нелинейных преобразований, допускающих регрессионную линеаризацию. Представлены результаты в области с. к. СОФ для СтС НРОП, приводимых к дифференциальным СтС. Приводится иллюстративный пример.

2 Интегральные канонические представления

Как известно [11], для скалярного СтП $X = X(t)$ ИКП называется его и его ковариационной функции интегральное представление через непрерывный скалярный белый шум $V = V(\lambda)$ параметра $\lambda \in \Lambda$:

$$X(t) = m_x(t) + \int_{\Lambda} V(\lambda)x(t, \lambda) d\lambda; \quad (1)$$

$$K_x(t, t') = M [X(t)\overline{X(t')}] = \int_{\Lambda} G(\lambda)x(t, \lambda)\overline{x(t', \lambda)} d\lambda.$$

Здесь $x(t, \lambda)$ — координатная функция; $G(\lambda)$ — интенсивность $V(\lambda)$, причем $K_v(\lambda, \lambda') = G(\lambda)\delta(\lambda - \lambda')$.

Формула для координатной функции ИКП имеет вид:

$$x(t, \lambda) = \frac{1}{G(\lambda)} K_{xv}(t, \lambda) = \frac{1}{G(\lambda)} M [X^0(t)\overline{V(\lambda)}].$$

Для построения ИКП используются следующие необходимые и достаточные условия:

$$V(\lambda) = \int_T \overline{a(t, \lambda)} X^0(t) dt;$$

$$x(t, \lambda) = \frac{1}{G(\lambda)} \int_T a(s, \lambda) K_x(t, s) ds; \quad (2)$$

$$\int_T \overline{a(t, \lambda)} x(t, \lambda') dt = \delta(\lambda - \lambda'); \quad (3)$$

$$\int_{\Lambda} \overline{a(t', \lambda)} x(t, \lambda) d\lambda = \delta(t - t'). \quad (4)$$

Теорема 2.1. Для того чтобы скалярный СтП $X = X(t)$ в области T допускал ИКП (1), условия (2) и (3) необходимы, а условия (2)–(4) — достаточны.

Аналогично для векторного СтП $X(t) = [X_1(t) \cdots X_n(t)]^T$ имеем:

$$\left. \begin{aligned} X(t) &= m^x(t) + \int_{\Lambda} V(\lambda)x(t, \lambda) d\lambda; \\ X_h(t) &= m_h^x(t) \int_{\Lambda} V(\lambda)x_h(t, \lambda) d\lambda \quad (h = \overline{1, n}). \end{aligned} \right\} \quad (5)$$

Здесь

$$m^x(t) = [m_1^x(t) \cdots m_n^x(t)]^T.$$

При этом ИКП матрицы ковариационных функций $K^x(t, t') = [K_{hi}^x(t, t')]$ имеет вид:

$$K_{hi}^x(t, t') = \int_{\Lambda} G(\lambda)x_h(t, \lambda)\overline{x_i(t', \lambda)} d\lambda \quad (h, i = \overline{1, n}), \quad (6)$$

где белый шум $V(\lambda)$ определяется формулой

$$V(\lambda) = \sum_{h=1}^n \int_T \overline{a_h(t, \lambda)} X_h^0(t) dt,$$

а его интенсивность $G(\lambda)$ равна

$$G(\lambda) = \sum_{h, l=1}^n \int_T \int_T \overline{a_h(t, \lambda)} a_l(t', \lambda') K_{hl}^x(t, t') dt dt' d\lambda'.$$

Для вычисления координатных функций $x_h(t, \lambda)$ и функций $a_h(t, \lambda)$ получим вместо (2), (3) и (4) уравнения:

$$x_h(t, \lambda) = \frac{1}{G(\lambda)} \sum_{l=1}^n \int_T a_l(t', \lambda) K_{hl}^x(t, t') dt' \quad (h = \overline{1, n}); \quad (7)$$

$$\sum_{h=1}^n \int_T \overline{a_h(t, \lambda)} x_h(t, \lambda') dt = \delta(\lambda - \lambda'); \quad (8)$$

$$\int_{\Lambda} \overline{a_l(t', \lambda)} x_h(t, \lambda) d\lambda = \delta_{hl}(t - t').$$

Если параметр λ принимает все возможные значения, принадлежащие нескольким областям $\Lambda_1, \dots, \Lambda_r$, то такое МК ИКП векторного СтП $X(t) = [X_1(t) \dots X_n(t)]^T$ и матрицы ковариационных функций $K^x(t, t') = [K_{hl}^x(t, t')]$ будут иметь вид:

$$\left. \begin{aligned} X(t) &= m^x(t) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_{\rho p}(t, \lambda) d\lambda; \\ X_h(t) &= m_h^x(t) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_{\rho h}(t, \lambda) d\lambda \end{aligned} \right\} \quad (9)$$

$$K_{hl}(t, t') = \sum_{\rho=1}^r \int_{\Lambda_\rho} G_\rho(\lambda) x_{\rho h}(t, \lambda) \overline{x_{\rho l}(t', \lambda)} d\lambda \quad (h, l = \overline{1, n}), \quad (10)$$

где $V_1(\lambda), \dots, V_r(\lambda)$ — некоррелированные белые шумы, определяемые формулой

$$V_\rho(\lambda) = \sum_{h=1}^n \int_T \overline{a_{\rho h}(t, \lambda)} X_h^0(t) dt \quad (\lambda \in \Lambda_\rho; \quad \rho = \overline{1, r}).$$

Интенсивности белых шумов $V_\rho(\lambda)$ определяются формулой:

$$G_\rho(\lambda) = \sum_{h, l=1}^n \int_T \int_T \overline{a_{\rho h}(t, \lambda)} a_{\rho l}(t', \lambda') K_{hl}^x(t, t') dt dt' d\lambda' \quad (\lambda \in \Lambda_\rho; \quad \rho = \overline{1, r}).$$

Для вычисления координатных функций $x_{\rho h}(t', \lambda')$ и функций $a_{\rho h}(t, \lambda)$ используются формулы:

$$x_{\rho h}(t, \lambda) = \frac{1}{G_\rho(\lambda)} \sum_{l=1}^n \int_T a_{\rho l}(t', \lambda) K_{hl}^x(t, t') dt' \quad (\lambda \in \Lambda_\rho; \quad \rho = \overline{1, r}; \quad h = \overline{1, n}); \quad (11)$$

$$\sum_{h=1}^n \int_T \overline{a_{\mu h}(t, \lambda)} a_{\rho h}(t, \lambda') dt = \delta_{\rho\mu} \delta(\lambda - \lambda') \quad (\lambda \in \Lambda_\mu; \quad \lambda' \in \Lambda_\rho; \quad \rho, \mu = \overline{1, r}); \quad (12)$$

$$\sum_{\rho=1}^r \int_{\Lambda_\rho} \overline{a_{\rho l}(t', \lambda)} x_{\rho h}(t, \lambda) d\lambda = \delta_{hl} \delta(t - t') \quad (h, l = \overline{1, n}),$$

выражающие необходимые и достаточные условия представления векторного СтП посредством МК ИКП.

Теорема 2.2. В условиях (7), (8) или (11), (12) из ИКП (5) или (9) вытекает МК ИКП матрицы его ковариационных функций (6) или (10).

Теорема 2.3. Если известно ИКП векторного СтП (5), то вектор математического ожидания и матрица ковариационных функций линейного преобразования $Y_t(t) = A_t X_t$ допускают МК ИКП, определяемые формулами [12]:

$$X_p(t) = m_p^x(t) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_{\rho p}(t, \lambda) d\lambda \quad (p = \overline{1, n});$$

$$Y_p(s) = m_p^y(s) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) y_{\rho p}(s, \lambda) d\lambda,$$

$$y_{\rho p}(s, \lambda) = \sum_{h=1}^n A_{\rho h} x_{\rho h}(t, \lambda) \quad (p = \overline{1, m});$$

$$K^y(s, s') = [K_{pq}^y(s, s')],$$

$$K_{pq}^y(s, s') = \sum_{\rho=1}^r \int_{\Lambda_\rho} G_\rho(\lambda) y_{\rho p}(s, \lambda) \overline{y_{\rho q}(s', \lambda)} d\lambda \quad (p, q = \overline{1, m}).$$

В задачах нелинейной корреляционной теории невырожденные безынерционные скалярные и векторные существенно нелинейные преобразования $Y_t = \varphi_t(X_t)$ заменяют оптимальными (в с. к. смысле) линейными регрессионными преобразованиями [12, 13]. Задача эквивалентной регрессионной линеаризации детерминированной векторной нелинейной функции $Y = \varphi(X)$ при использовании критерия минимума с. к. ошибки совпадает с классической задачей линейного регрессионного анализа. В этом случае оптимальная линейная с. к. регрессия вектора Y на вектор X определяется формулой:

$$m^y(X) = gX + a, \quad (13)$$

где

$$g = K^{yx} (K^x)^{-1}; \quad (14)$$

$$a = m^y - gm^x.$$

Пусть $f(y, x)$ — совместная плотность случайных векторов Y и X ; m^x и K^x — математическое ожидание и ковариационная матрица вектора x ,

$\det |K^x| \neq 0$. Формула (14) при этом принимает вид:

$$g = K^{yx}(K^x)^{-1} = \int_{-\infty}^{\infty} [m^y(x) - m^y](x - m^x)^T (K^x)^{-1} f_1(x) dx, \quad (15)$$

где $f_1(x)$ — плотность случайного вектора X . Эта формула вместе с приближенной формулой

$$m^y(X) \approx m^y + g(X - m^x)$$

дает статистическую линеаризацию регрессии $m^y(X)$ по Казакову.

Теорема 2.4. Если существуют конечные моменты первого и второго порядка векторного СтП $X = X(t)$, то векторное нелинейное преобразование $Y = \varphi(X)$ допускает линейную с. к. регрессию $Y = Y(t)$ на $X = X(t)$, определяемую по Казакову формулами (13) и (15).

Первый подход к линеаризации основан на использовании формул для ИКП X_t в теореме 2.2 в случае одной области Λ , а второй подход — для нескольких областей Λ_r . При этом имеют место следующие утверждения.

Теорема 2.5. Оптимальная с. к. линеаризация посредством ИКП (теорема 2.2) первого рода определяется следующими формулами:

$$\varphi_t(X_t) \approx m_t^{(1)y}(X_t) = \varphi_{0t}^{(1)}(m_t^x, K_t^x) + g_t^{(1)}(m_t^x, K_t^x) \left[\sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_t(\lambda) d\lambda - m_t^x \right],$$

где

$$\varphi_{0t}^{(1)}(m_t^x, K_t^x) = M_N[\varphi(X_t)];$$

$$g_t^{(1)}(m_t^x, K_t^x) = K_t^{yx}(K_t^x)^{-1}.$$

Здесь M_N — символ математического ожидания для нормального распределения.

Теорема 2.6. Оптимальная с. к. линеаризация посредством МК ИКП (теорема 2.2) второго рода определяется следующими формулами:

$$\varphi_t(X_t) \approx m_t^{(2)y}(X_t) = \varphi_{0t}^{(2)}(m_t^x, K_{\rho t}^x) + \sum_{\rho=1}^r \int_{\Lambda_\rho} g_{0t}^{(2)}(m_t^x, K_{\rho t}^x) V_\rho(\lambda) x_{\rho t} d\lambda - g_t^{(2)}(m_t^x, K_t^x) m_t^x,$$

где

$$\varphi_{0t}^{(2)}(m_t^x, K_{\rho t}^x) = M_N[\varphi(X_t)];$$

$$g_{0t}^{(2)}(m_t^x, K_{\rho t}^x) = K_{\rho t}^{yx}(K_{\rho t}^x)^{-1}.$$

3 Основные результаты

В качестве исходной СтС НРОП рассмотрим детерминированную векторную систему уравнений

$$\dot{\Phi} = \Phi(t, X_t, \bar{X}_t, Y_t, \Theta_t, U_t) = 0. \quad (16)$$

Здесь X_t — вектор состояния; $\bar{X}_t = [\dot{X}_t^T \dots (X^{(l)})^T]^T$ — расширенный вектор состояния, состоящий из l -го порядка производных по времени; Y_t — вектор наблюдений; Θ_t — вектор случайных параметров, описываемый МК ИКП (см. разд. 2); Φ — нелинейная функция переменных $\bar{X}_t = [X_t^T \bar{X}_t^T U_t^T]^T$ и Θ_t , допускающая при фиксированном Θ_t с. к. регрессионную линеаризацию по \bar{X}_t и U_t вида

$$\Phi \approx \Phi_0 + \sum_{j=1} k_{\bar{X},j}^\Phi X_t^j + k_U^\Phi U_t^0. \quad (17)$$

Вектор возмущений U_t связан с гауссовским белым шумом V_0 линейным уравнением формирующего фильтра:

$$\dot{U}_t = a_t^U U_t + a_{0t}^U + b_t^U V_0, \quad (18)$$

где

$$M V_0 = 0; \quad M[\nu_0(t)\nu_0(t)^T] = \nu_0 \delta(t - \tau).$$

Тогда при фиксированном $\Theta = \Theta_t$ и при условиях $\det k_{\bar{X}l}^\Phi \neq 0$ и $\det K_U^\Phi \neq 0$ дифференциальная СтС НРОП (16) приводится к дифференциальной СтС следующего вида (**теорема 3.1**):

$$\dot{\bar{X}}_{1t} = \bar{X}_{2t}, \dots, \dot{\bar{X}}_{(l-1)t} = \bar{X}_{lt};$$

$$\dot{\bar{X}}_{lt} = - (k_{\bar{X}l}^\Phi)^{-1} \bar{X}_t^{(l)} - (k_{\bar{X}l}^\Phi)^{-1} (k_U^\Phi)^{-1} U_t$$

и (18). Матрицы коэффициентов $k_{\bar{X}l}^\Phi$ и k_U^Φ неявно зависят от первых двух вероятностных моментов $m_t^{\bar{X}_t}$ и $K_t^{\bar{X}_t}$.

Пусть объектовая СтС НРОП допускает приведение к дифференциальной (теорема 3.1), измерительная система вполне наблюдаема, наблюдения влияют на объект, а уравнение наблюдения разрешено относительно Y_t . Тогда в качестве исходных приведенных дифференциальных уравнений для объекта и измерительной системы можно принять следующую:

$$\dot{X}_t = A^\Pi(X_t, Y_t, m_t^{\bar{X}_t}, K_t^{\bar{X}_t}, \Theta_t, t) =$$

$$= a^\Pi(X_t, Y_t, m_t^{\bar{X}_t}, K_t^{\bar{X}_t}, \Theta_t, t) +$$

$$+ b^\Pi(X_t, Y_t, m_t^{\bar{X}_t}, K_t^{\bar{X}_t}, \Theta_t, t) V_0(\Theta_t);$$

$$Z_t = \dot{Y}_t = B(X_t, Y_t, \Theta_t, t) = a_1(X_t, Y_t, \Theta_t, t) + b_1(X_t, Y_t, \Theta_t, t) V_0(\Theta_t).$$

Здесь a^Π , a_1 , b^Π и b_1 — известные векторно-матричные функции; V_0 — векторный нормальный (гауссовский) белый шум интенсивности $\nu_0 = \nu_0(\Theta_t)$.

В рамках МНА апостериорной плотности вероятности, учитывая, что гауссовское (нормальное) распределение, аппроксимирующее апостериорное распределение вектора \hat{X}_t , полностью определяется апостериорными математическим ожиданием \hat{X}_t и ковариационной матрицей R_t вектора \hat{X}_t , при аппроксимации апостериорного распределения вектора \hat{X}_t нормальным распределением все математические ожидания будут представлять собой стохастические дифференциальные уравнения, определяющие \hat{X}_t и R_t (**теорема 3.2**):

$$\begin{aligned} \dot{\hat{X}}_t &= f\left(\hat{X}_t, Y_t, R_t, t\right) + h\left(\hat{X}_t, Y_t, R_t, \Theta_t, t\right) \times \\ &\quad \times \left[\dot{Y}_t - f^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta_t, t\right) dt \right]; \quad (19) \\ \dot{R}_t &= \left\{ f^{(2)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) - h\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \times \right. \\ &\quad \times \left. b_1 \nu_0 b_1^T(Y_t, \Theta, t) h\left(\hat{X}_t, Y_t, R_t, \Theta, t\right)^T \right\} + \\ &\quad + \sum_{r=1}^{n_y} \rho_r\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \times \\ &\quad \times \left[\dot{Y}_r - f_r^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \right]. \quad (20) \end{aligned}$$

Здесь

$$\begin{aligned} f\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) &= \\ &= [(2\pi)^n |R_t|]^{-1/2} \int_{-\infty}^{\infty} a(Y_t, x, \Theta, t) \times \\ &\quad \times \exp\left\{ -\frac{\left(x^T - \hat{X}_t^T\right) R_t^{-1} \left(x - \hat{X}_t\right)}{2} \right\} dx; \\ f^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) &= \left\{ f_r^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \right\} = \\ &= [(2\pi)^{n_x} |R_t|]^{-1/2} \int_{-\infty}^{\infty} a_1(Y_t, x, \Theta, t) \times \\ &\quad \times \exp\left\{ -\frac{\left(x^T - \hat{X}_t^T\right) R_t^{-1} \left(x - \hat{X}_t\right)}{2} \right\} dx; \end{aligned}$$

$$\begin{aligned} h\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) &= \left\{ [(2\pi)^{n_x} |R_t|]^{-1/2} \times \right. \\ &\quad \times \int_{-\infty}^{\infty} \left[x a_1(Y_t, x, \Theta, t)^T + b \nu_0 b_1^T(Y_t, x, \Theta, t) \right] \times \\ &\quad \times \exp\left\{ -\frac{\left(x^T - \hat{X}_t^T\right) R_t^{-1} \left(x - \hat{X}_t\right)}{2} \right\} dx - \\ &\quad \left. - \hat{X}_t f^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right)^T \right\} \left(b_1 \nu_0 b_1^T \right)^{-1}(Y_t, \Theta, t); \end{aligned}$$

$$\begin{aligned} f^{(2)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) &= [(2\pi)^{n_x} |R_t|]^{-1/2} \times \\ &\quad \times \int_{-\infty}^{\infty} \left\{ \left(x - \hat{X}_t\right) a(Y_t, x, \Theta, t)^T + \right. \\ &\quad \left. + a(Y_t, x, \Theta, t) \left(x^T - \hat{X}_t^T\right) + b \nu_0 b_1^T(Y_t, x, \Theta, t) \right\} \times \\ &\quad \times \exp\left\{ -\frac{\left(x^T - \hat{X}_t^T\right) R_t^{-1} \left(x - \hat{X}_t\right)}{2} \right\} dx; \end{aligned}$$

$$\begin{aligned} \rho_r\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) &= [(2\pi)^{n_x} |R_t|]^{-1/2} \times \\ &\quad \times \int_{-\infty}^{\infty} \left\{ \left(x - \hat{X}_t\right) \left(x^T - \hat{X}_t^T\right) a_r(Y_t, x, \Theta, t) + \right. \\ &\quad \left. + \left(x - \hat{X}_t\right) b_r(Y_t, x, \Theta, t)^T \left(x^T - \hat{X}_t^T\right) + \right. \\ &\quad \left. + b_r(Y_t, x, \Theta, t) \left(x^T - \hat{X}_t^T\right) \right\} \times \\ &\quad \times \exp\left\{ -\frac{\left(x^T - \hat{X}_t^T\right) R_t^{-1} \left(x - \hat{X}_t\right)}{2} \right\} dx \\ &\quad \quad \quad (r = \overline{1, n_y}), \quad (21) \end{aligned}$$

где a_r — r -й элемент матрицы-строки $(a_1^T - \hat{a}_1^T)(b_1 \nu_0 b_1^T)^{-1}$; b_{kr} — элемент k -й строки и r -го столбца матрицы $(b_1 \nu_0 b_1^T)^{-1}$. Тогда, обозначив через b_r r -й столбец матрицы $b \nu_0 b_1^T (b_1 \nu_0 b_1^T)^{-1}$, имеем $b_r = [b_{1r} \dots b_{pr}]^T$ ($r = \overline{1, n_1}$).

Число уравнений МНА одномерного апостериорного распределения определяется по формуле:

$$Q_{\text{МНА}} = n_x + \frac{n_x(n_x + 1)}{2} = \frac{n_x(n_x + 3)}{2}.$$

За начальные значения \hat{X}_t и R_t при интегрировании уравнений (19) и (20), естественно, следует принять условные математическое ожидание и ковариационную матрицу величины \hat{X}_0 относительно Y_0 :

$$\hat{X}_0 = M_N \left[\hat{X}_0 \mid Y_0 \right];$$

$$R_0 = M \left[(\hat{X}_0 - \hat{X}_0) \left(\hat{X}_0^T - \hat{X}_0^T \right) \mid Y_0 \right].$$

Если нет информации об условном распределении \hat{X}_0 относительно Y_0 , то начальные условия можно взять в виде:

$$\hat{X}_0 = M_N \hat{X}_0;$$

$$R_0 = M_N (\hat{X}_0 - M_N \hat{X}_0) \left(\hat{X}_0^T - M_N \hat{X}_0^T \right).$$

Если же и об этих величинах нет никакой информации, то начальные значения \hat{X}_t и R_t придется задавать произвольно.

Из формулы (21) видно, что если функция a_1 линейна относительно \hat{X}_t , а функция b не зависит от \hat{X}_t , то при нормальной аппроксимации апостериорного распределения все матрицы ρ_r равны нулю, вследствие чего уравнение (20) не содержит \dot{Y}_t .

Наконец, применяя регрессионную линеаризацию к уравнениям (19) и (20) посредством МК ИКП, согласно теоремам 2.5 и 2.6, для условных характеристик, придем к следующим уравнениям аналитического моделирования параметров СОФ (теорема 3.3):

$$\begin{aligned} \hat{X}_t &= \tilde{X}_t + \delta \hat{X}_t; \quad \tilde{R}_t = \tilde{R}_t + \delta \tilde{R}_t; \\ \frac{d}{dt} \hat{X}_t &= f_0 + h_0 \dot{Y}_t - (h f^{(1)})_0; \end{aligned} \quad (22)$$

$$\begin{aligned} \frac{d}{dt} \tilde{R}_t &= f_0^{(2)} - h_0 \alpha_0 h_0^T + \sum_{i=1}^{n_y} \left[\rho_{i0} \dot{Y}_t - (\rho_i f_i^{(1)})_0 \right] \\ &(\alpha = b_1 \nu_0 b_1^T); \end{aligned} \quad (23)$$

$$\begin{aligned} \frac{d}{dt} \delta \hat{X}_t &= f_{1,\delta X} \delta \hat{X}_t + f_{1,\delta R} \delta \tilde{R}_t + f_{1,\Theta} \Theta_t + \\ &+ \left(h_{1,\delta X} \delta \hat{X}_t + h_{1,\delta R} \delta \tilde{R}_t + h_{1,\Theta} \Theta_t \right) \dot{Y}_t - \\ &- \left[\left(h f^{(1)} \right)_{1,\delta X} \delta \hat{X}_t + \left(h f^{(1)} \right)_{1,\delta R} \delta \tilde{R}_t + \right. \\ &\quad \left. + \left(h f^{(1)} \right)_{1,\Theta} \Theta_t \right]; \end{aligned} \quad (24)$$

$$\begin{aligned} \frac{d}{dt} \delta \tilde{R}_t &= f_{1,\delta X}^{(2)} \delta \hat{X}_t + f_{1,\delta R}^{(2)} \delta \tilde{R}_t + f_{1,\Theta}^{(2)} \Theta_t - \\ &- \left[(h \alpha h)_{1,\delta X} \delta \hat{X}_t + (h \alpha h)_{1,\delta R} \delta \tilde{R}_t + (h \alpha h)_{1,\Theta} \Theta_t \right] + \end{aligned}$$

$$\begin{aligned} &+ \sum_{i=1}^{n_y} \left\{ \left(\rho_{i,1,\delta X} \delta \hat{X}_t + \rho_{i,1,\delta R} \delta \tilde{R}_t + \rho_{i,1,\Theta} \Theta_t \right) \dot{Y}_t - \right. \\ &- \left[\left(\rho_i f_i^{(1)} \right)_{1,\delta X} \delta \hat{X}_t + \left(\rho_i f_i^{(1)} \right)_{1,\delta R} \delta \tilde{R}_t + \right. \\ &\quad \left. \left. + \left(\rho_i f_i^{(1)} \right)_{1,\Theta} \Theta_t \right] \right\}. \end{aligned} \quad (25)$$

Здесь при регрессионной линеаризации первого рода аргументами функций f , $f^{(1)}$, $f^{(2)}$, h и $f_i^{(1)}$ служат переменные \tilde{X}_t , Y_t , \tilde{R}_t , m_t^Θ и K_t^Θ , а для линеаризации второго рода — \hat{X}_t , Y_t , \tilde{R}_t , $m_t^{\Theta_r}$ и $K_t^{\Theta_r}$.

Анализ уравнений теоремы 3.3 позволяет сделать следующие выводы:

- (1) фильтрационные уравнения (22) и (23) нелинейны относительно безусловных характеристик \hat{X}_t и \tilde{R}_t , определяемых уравнением Риккати, а уравнения точности НСОФ (24) и (25) линейны относительно $\delta \hat{X}_t$, $\delta \tilde{R}_t$;
- (2) если приравнять правые части уравнений (22) и (23) нулю, найдем стационарные решения, а уравнения (24) и (25) позволят найти как условия существования и устойчивости стационарных решений, так и оценить точность НСОФ. Для определения весовых функций физически реализуемых линейных уравнений (23) и (24) следует образовать соответствующие коэффициенты при переменных количествах;
- (3) в прикладных задачах при оценке точности НСОФ в условиях экстремальных значений переменных случайных параметров суждение о точности НСОФ делается на основе анализа уравнений Риккати (23) и (24) путем численного решения системы взаимосвязанных детерминированных уравнений.

4 Пример

Рассмотрим наблюдаемую дифференциальную систему вида

$$\dot{\varphi} = \varphi_1(\dot{X}_{1t}) + \gamma X_{1t} + X_{2t} = 0; \quad (26)$$

$$\dot{X}_{2t} = a^U X_{2t} + V_1; \quad (27)$$

$$Z_t = \dot{Y}_t = c_1 X_{1t} + V_2. \quad (28)$$

Здесь $[X_{1t} X_{2t}]^T = X_t$ — вектор состояния; V_1 и V_2 — независимые скалярные белые шумы с интенсивностями ν_1 и ν_2 ; $\varphi_1 = \varphi_1(\dot{X}_{1t})$ — функция, допускающая регрессионную линеаризацию

$$\varphi_1 \approx \varphi_{10} + k_1 \dot{X}_{1t}^0, \quad (29)$$

где $\varphi_{10} = \varphi_{10}(m_{1t}^{\dot{X}}, D_{1t}^{\dot{X}})$; $k_1 = k_1(m_{1t}^{\dot{X}}, D_{1t}^{\dot{X}})$. При $k_1 \neq 0$ уравнение (26) с учетом (29) приводится к виду:

$$\dot{X}_{1t} = -(\varphi_{10} + \gamma X_{1t} - X_{2t}) k_1^{-1}. \quad (30)$$

Обозначая

$$a_{0t} = \begin{bmatrix} -\varphi_{10} k_1^{-1} \\ 0 \end{bmatrix}; \quad a_{1t} = \begin{bmatrix} \gamma k_1^{-1} \\ c_1 \quad 0 \end{bmatrix};$$

$$b_{1t} = [c_1 \quad 0]; \quad \tilde{R} = \begin{bmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{bmatrix},$$

запишем уравнения (30), (27) и (28) в форме

$$\left. \begin{aligned} \dot{X}_t &= a_{1t} X_t + a_{0t} + V_1; \\ Z_t &= \dot{Y}_t = b_{1t} X_t + V_2. \end{aligned} \right\} \quad (31)$$

Для (31) уравнения НСОФ [13] имеют следующий вид:

$$\begin{aligned} \dot{\hat{X}}_t &= a_{1t} \hat{X}_t + a_{0t} + \tilde{\beta}_t (Z_t - b_{1t} \hat{X}_t); \\ \tilde{\beta}_t &= \tilde{R}_t b_{1t}^T \nu_{2t}^{-1}; \\ \dot{\tilde{R}}_t &= a_{1t} \tilde{R}_t + \tilde{R}_t a_{1t}^T - \tilde{R}_t b_{1t}^T \nu_{2t}^{-1} b_{1t} \tilde{R}_t + \\ &\quad + (\nu_{10} + \nu_{11} \Theta_t). \end{aligned} \quad (32)$$

Здесь

$$\Theta_t = m_t^\Theta + \int_{\Lambda} V(\lambda) x^\Theta(\lambda, t) d\lambda;$$

$$K_t^\Theta(t_1, t_2) = \int_{\Lambda} G(\lambda) x^\Theta(\lambda, t_1) x^\Theta(\lambda, t_2) d\lambda,$$

где

$$x(\lambda, t) = q_1(t) q_2^{-1}(\lambda) \mathbf{1}(t - \lambda),$$

$$q_1(t) = e^{-\alpha t}, \quad q_2(t) = q_1(t) \int_{t_0}^t q_1^{-2}(\lambda) d\lambda;$$

$$K^\Theta(t_1, t_2) = \begin{cases} q_1(t_2) q_2(t_1) & \text{при } t_1 < t_2; \\ q_1(t_1) q_2(t_2) & \text{при } t_1 > t_2. \end{cases}$$

При $\nu_{11}^* \ll \nu_{10}^*$, $a_{0t} = a_0^*$, $a_{1t} = a_1^*$, $b_{1t} = b_1^*$, $\nu_{10} = \nu_{10}^*$, $m_t^\Theta = m_*^\Theta$ и $\nu_2 = \nu_2^*$ уравнение (32) имеет стационарное решение:

$$a_1^* R^* + R^* a_1^{*\Gamma} - R^* b_1^{*\Gamma} (\nu_2^*)^{-1} R^* + \nu_{10}^* = 0.$$

При этом для оценки точности СОФ после линеаризации относительно $\delta R_t = \tilde{R}_t - R^*$ получаем линейное уравнение с постоянными коэффициентами:

$$\begin{aligned} \delta \dot{R}_t &= [a_1^* + R^* b_1^{*\Gamma} (\nu_2^*)^{-1} b_1^*] \delta R_t + \\ &\quad + \delta R_t [b_1^{*\Gamma} (\nu_2^*)^{-1} b_1^* R^*] + \nu_{11}^* \Theta_t. \end{aligned}$$

Отсюда сразу вытекают уравнения для $m^{\delta R} = M \delta R_t$, а также ковариационной матрицы $K_t^{\delta R}$.

5 Заключение

Для наблюдаемых гауссовских дифференциальных СтС НРОП со случайными параметрами в виде ИКП, приводимых к дифференциальным СтС, разработано методическое обеспечение анализа точности СОФ, основанное на МНА при фиксированном векторе параметров для условных вероятностных характеристик и регрессионной линеаризации посредством ИКП безусловных характеристик.

Результаты могут быть использованы для наблюдаемых негауссовских СтС НРОП, приводимых к негауссовским дифференциальным СтС. Дальнейшее развитие методического обеспечения для негауссовских СтС НРОП связано с использованием методов моментов, квазимоментов и ортогональных разложений одно- и многомерных плотностей для условных вероятностных характеристик.

Литература

1. Синецын И. Н. Аналитическое моделирование широкополосных процессов в стохастических системах, не разрешенных относительно производных // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 3–10. doi: 10.14357/19922264170101. EDN: YOCMVL.
2. Синецын И. Н. Параметрическое аналитическое моделирование процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2017. Т. 27. № 1. С. 21–45. doi: 10.14357/08696527170102. EDN: YODCZL.
3. Sinityn I. N. Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives // J. Mathematics Statistics Research, 2021. Vol. 3. Iss. 1. Art. 139. 7 p. doi: 10.36266/JMSR/139.
4. Синецын И. Н. Аналитическое моделирование и оценивание нестационарных нормальных процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2022. Т. 32. № 2. С. 58–71. doi: 10.14357/08696527220206. EDN: YMGERJ.
5. Синецын И. Н. Нормализация систем, стохастически не разрешенных относительно производных // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 32–38. doi: 10.14357/19922264220105. EDN: LDFJJB.
6. Синецын И. Н. Аналитическое моделирование распределений с инвариантной мерой в стохастических системах, не разрешенных относительно производных // Информатика и её применения, 2023. Т. 17. Вып. 1. С. 2–10. doi: 10.14357/19922264230101. EDN: QWXVXS.
7. Синецын И. Н. Аналитическое моделирование стохастических систем, не разрешенных относительно

- производных, со случайными параметрами // Системы и средства информатики, 2024 (в печати). Т. 34. № 1.
8. Синицын И. Н. Нормальные субоптимальные фильтры для дифференциальных стохастических систем, не разрешенных относительно производных // Информатика и её применения, 2021. Т. 15. Вып. 1. С. 3–10. doi:10.14357/19922264210101. EDN: UPEHRI.
 9. Синицын И. Н. Аналитическое моделирование и фильтрация нормальных процессов в интегродифференциальных стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2021. Т. 31. № 1. С. 37–56. doi: 10.14357/08696527210104. EDN: PLYOSF.
 10. Синицын И. Н. Совместная фильтрация и распознавание нормальных процессов в стохастических системах, не разрешенных относительно производных // Информатика и её применения, 2022. Т. 16. Вып. 2. С. 85–93. doi: 10.14357/19922264220211. EDN: SMJCBB.
 11. Синицын И. Н. Канонические представления случайных функций. Теория и применения. — М.: ТОРУС ПРЕСС, 2023. 816 с.
 12. Синицын И. Н. Фильтры Калмана и Пугачёва. — 2-е изд. — М.: Логос, 2007. 776 с.
 13. Пугачёв В. С. Теория вероятностей и математическая статистика. — 2-е изд. — М.: Физматлит, 2002. 496 с.

Поступила в редакцию 26.09.23

SUBOPTIMAL FILTERING IN STOCHASTIC SYSTEMS WITH RANDOM PARAMETERS AND UNSOLVED DERIVATIVES

I. N. Sinitsyn^{1,2}

¹Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

²Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation

Abstract: For observable Gaussian differential stochastic systems (StS) with random parameters in the form of integral canonical expansions (ICE) and StS with unsolved derivatives (USD), methodological support for synthesis of suboptimal filters is presented. Survey in fields of analytical modeling and suboptimal filtering (SOF), extrapolation, and identification is presented. Necessary information concerning multicomponent (MC) ICE is given. Special attention is paid to mean square regressive linearization including MC ICE. Basic results in normal SOF (NSOF) are presented for StS USD reducible to differential StS. Stationary and nonstationary NSOF are considered. An illustrative example for scalar StS USD reducible to differential is given. For future, SOF generalization methods of moments, quasi-moments, and one- and multidimensional densities of orthogonal expansions are recommended.

Keywords: normal approximation method (NAM); regression linearization; stochastic process; stochastic systems with unsolved derivatives (StS USD); suboptimal filtering (SOF)

DOI: 10.14357/19922264240101

EDN: KUWMKJ

References

1. Sinitsyn, I. N. 2017. Analiticheskoe modelirovanie shirokopolosnykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling of wide band processes in stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):3–10. doi: 10.14357/19922264170101. EDN: YOVMVL.
2. Sinitsyn, I. N. 2017. Parametricheskoe analiticheskoe modelirovanie protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Parametric analytical modeling of wide band processes in stochastic systems with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(1):21–45. doi: 10.14357/08696527170102. EDN: YODCZL.
3. Sinitsyn, I. N. 2021. Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives. *J. Mathematics Statistics Research* 3(1):139. 7 p. doi: 10.36266/JMSR/139.
4. Sinitsyn, I. N. 2022. Analiticheskoe modelirovanie i otsenivanie nestatsionarnykh normal'nykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling and estimation of nonstationary normal processors with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(2):58–71. doi: 10.14357/08696527220206. EDN: YMGERJ.
5. Sinitsyn, I. N. 2022. Normalizatsiya sistem, stokhasticheski ne razreshennykh otnositel'no proizvodnykh [Normalization of systems with stochastically unsolved derivatives].

- Informatika i ee Primeneniya — Inform. Appl.* 16(1):32–38. doi: 10.14357/19922264220105. EDN: LDFJJB.
6. Sinitsyn, I. N. 2023. Analiticheskoe modelirovanie raspredeleniy s invariantnoy meroy v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling of distributions with invariant measure in stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 17(1):2–10. doi: 10.14357/19922264230101. EDN: QWXVXC.
 7. Sinitsyn, I. N. 2024 (in press). Analiticheskoe modelirovanie stokhasticheskikh sistem, ne razreshennykh otnositel'no proizvodnykh, so sluchaynymi parametrami [Analytical modeling of stochastic systems with random parameters and unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 34(1).
 8. Sinitsyn, I. N. 2021. Normal'nye suboptimal'nye fil'try dlya differentsial'nykh stokhasticheskikh sistem, ne razreshennykh otnositel'no proizvodnykh [Normal suboptimal filtering for differential stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 15(1):3–10. doi: 10.14357/19922264210101. EDN: UPEHRI.
 9. Sinitsyn, I. N. 2021. Analiticheskoe modelirovanie i fil'tratsiya normal'nykh protsessov v integrodifferentsial'nykh stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling and filtering for integrodifferential systems with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 31(1):37–56. doi: 10.14357/08696527210104. EDN: PLYOSF.
 10. Sinitsyn, I. N. 2022. Sovmestnaya fil'tratsiya i raspoznavanie normal'nykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Joint filtration and recognition of normal processes in stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 16(2):85–93. doi: 10.14357/19922264220211. EDN: SMJCBB.
 11. Sinitsyn, I. N. 2023. *Kanonicheskie predstavleniya sluchaynykh funktsiy. Teoriya i primeneniya* [Canonical expansion of random functions. Theory and application]. Moscow: TORUS PRESS. 816 p.
 12. Sinitsyn, I. N. 2007. *Fil'try Kalmana i Pugacheva* [Kalman and Pugachev filters]. 2nd ed. Moscow: Logos. 776 p.
 13. Pugachev, V. S. 2002. *Teoriya veroyatnostey i matematicheskaya statistika* [Probability theory and mathematical statistics]. 2nd ed. Moscow: Fizmatlit. 496 p.

Received September 26, 2023

Contributor

Sinitsyn Igor N. (b. 1940) — Doctor of Science in technology, professor, Honored scientist of RF, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; professor, Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; sinitsin@dol.ru

АЛГЕБРАИЧЕСКАЯ СПЕЦИФИКАЦИЯ РАСПРЕДЕЛЕННЫХ СИСТЕМ С ИЗМЕНЯЮЩЕЙСЯ АРХИТЕКТУРОЙ

С. П. Ковалёв¹

Аннотация: Развивается предложенный ранее метод обобщенной алгебраической спецификации распределенных систем на базе новой теоретико-категорной конструкции графалгебры. Основой графалгебраической спецификации служит ориентированный мультиграф, ребра которого представляют вычислительные операции, выполняемые в узлах системы, а вершины задают порты обмена данными между компонентами. Изменение архитектуры системы в ходе жизненного цикла приводит к изменению формы графа, алгоритмов вычислений и содержания обмена данными. Для формального описания таких изменений предложена техника типа трансформации графов для графалгебр. Введена новая теоретико-категорная конструкция гибких графалгебр, тесно связанная с известной монадой диаграмм. Описан функтор, представляющий конструирование категории гибких графалгебр по сигнатуре. Теоретические результаты проиллюстрированы примерами из области автоматического синтеза архитектуры нейронных сетей путем пошаговых трансформаций.

Ключевые слова: алгебраическая спецификация; распределенная система; эволюция архитектуры; теория категорий; графалгебра; монада диаграмм

DOI: 10.14357/19922264240102

EDN: MZAHBS

1 Введение

Для специфицирования программ с математическим уровнем строгости нередко применяются алгебраические методы (см., например, [1]). Они позволяют описывать операции, выполняемые программами, термами языка первого порядка (или его расширений) и доказывать утверждения об операциях, в том числе индукцией по синтаксической структуре термов. Для описания программ, изменяющихся во времени (при фиксированной сигнатуре операций), используется аппарат эволюционирующих алгебр [2].

В то же время классической алгебраической спецификации присущ недостаток, состоящий в невозможности явно описать структуру многокомпонентной распределенной программной системы. Все операции описываются так, будто выполняются одним универсальным абстрактным вычислительным устройством. Для преодоления этого недостатка была предложена конструкция графалгебры [3] — структуры алгебраического типа над произвольным ориентированным мультиграфом. Ребра этого графа представляют вычислительные операции, выполняемые в узлах системы, а вершины — порты обмена данными между компонентами. Сигнатура графалгебры состоит из графа и правил представления данных в портах.

Однако фиксированная форма сигнатуры ограничивает применение категорий графалгебр в при-

ложениях, поскольку современным распределенным системам присуща гибкость — способность изменять граф структуры, в том числе в темпе, близком к реальному времени. Например, архитектура многослойной нейронной сети может быть оптимизирована под класс решаемых задач (по числу слоев, мощности каждого слоя, топологии и весам связей) путем целенаправленных многошаговых трансформаций некоторой небольшой начальной сети [4], а распределенные программно-аппаратные системы всегда изменяются в ходе создания и модернизации.

Для формального описания и верификации процедур эволюции систем применяется техника типа трансформации графов [5]. Для графалгебр такая техника впервые предложена в настоящей статье. Введена новая теоретико-категорная конструкция гибких графалгебр, тесно связанная с известной монадой диаграмм.

2 Графалгебры

С точки зрения теории категорий алгебра над некоторой категорией \mathcal{C} — это морфизм $a : FA \rightarrow A$, где F — эндифунктор в \mathcal{C} , представляющий сигнатуру; A — носитель. У алгебры первопорядковой сигнатуры Σ функтор F имеет «полиномиальный» вид:

$$FA = \coprod_{\varphi \in \Sigma} A^{n(\varphi)},$$

¹Институт проблем управления им. В. А. Трапезникова Российской академии наук, kovalyov@sipnet.ru

где $n(\varphi)$ — местность (число аргументных мест) функционального символа $\varphi \in \Sigma$. Алгебру можно рассматривать как C -диаграмму в форме **2** (так обозначается граф, состоящий из двух вершин 0 и 1 и одного ребра $0 \rightarrow 1$), а гомоморфизм алгебр — как морфизм носителей, индуцирующий естественное преобразование таких диаграмм. Путем обобщения этой точки зрения получается графалгебра — диаграмма произвольной формы, вершины которой помечены результатами действия заданных сигнатурных функторов на носитель [3]. Все сигнатурные функторы имеют одну и ту же область и одну и ту же кообласть (в общем случае отличную от области). Поэтому сигнатурой графалгебры служит любая пара $\langle I, F : D \rightarrow C^{\text{Ob} I} \rangle$, где I — малая категория, называемая формой графалгебры; категория D — пространство данных; C — пространство вычислений. Графалгебра сигнатуры $\langle I, F \rangle$ — это произвольная пара (A, Δ) , состоящая из носителя $A \in \text{Ob } D$ и диаграммы структуры $\Delta : I \rightarrow C$, такая что $\Delta i = \text{Pr}_i F A$, где $\text{Pr}_i : C^{\text{Ob} I} \rightarrow C$ — проекция степени на i -й экземпляр основания, $i \in \text{Ob } I$. Гомоморфизмом графалгебры (A, Δ) в (A', Δ') называется морфизм носителей $f : A \rightarrow A'$ такой, что семейство морфизмов $\text{Pr}_i F f : \Delta i \rightarrow \Delta' i$, $i \in \text{Ob } I$, образует естественное преобразование Δ в Δ' . Все графалгебры некоторой сигнатуры и все их гомоморфизмы образуют категорию, обозначаемую через $A_I F$, на которой действуют канонические забывающие функторы: функтор носителя $U_I^F : A_I F \rightarrow D : (A, \Delta) \mapsto A$ и функтор структуры $S_I^F : A_I F \rightarrow C^I : (A, \Delta) \mapsto \Delta$.

Графалгебры выступают спецификациями распределенных систем: стрелки диаграммы структуры помечаются вычислительными операциями в узлах системы, а вершины описывают передаваемые между узлами данные. С этой точки зрения обычная алгебра действительно представляет автономный вычислительный узел, а, например, многослойная нейросеть прямого распространения имеет форму нетривиальной конечной цепочки: это граф множеств и отображений вида

$$A^{k_0} \rightarrow A^{k_1} \rightarrow \dots \rightarrow A^{k_l},$$

где A — множество значений распознаваемого сигнала (например, $A = \mathbb{R}$); $l > 1$ — глубина сети; $k_i > 0$ — число нейронов в i -м слое, $i = \overline{0, l-1}$. Обработка сигнала в каждом слое $f_i : A^{k_i} \rightarrow A^{k_{i+1}}$ состоит из отображений-компонентов (например:

$$\begin{aligned} f_{is} : \mathbb{R}^{k_i} &\rightarrow \mathbb{R} : (x_1, \dots, x_{k_i}) \mapsto \\ &\mapsto \sum_{j=1}^{k_i} w_{ijs} \sigma_{ij}(x_j + b_{ij}), \quad s = \overline{1, k_{i+1}}, \end{aligned}$$

где $\sigma_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ — функция активации j -го нейрона i -го слоя; w_{ijs} и b_{ij} — веса и смещения, подбираемые при обучении нейросети). Сигнатура графалгебраической спецификации такой нейросети имеет форму цепочки из l стрелок, каждой вершине которой соответствует эндифунктор возведения в конечную степень в **Set**.

Распределенные системы общего вида, в том числе сложно устроенные нейросети, могут иметь графалгебраическую спецификацию практически любой конечной формы (с учетом ресурсных ограничений). В дополнение к конечным цепочкам, представляющим последовательную конвейерную обработку данных, форма может содержать подграфы, представляющие параллельную обработку (например, $\bullet \rightrightarrows \bullet$), обратную связь (например, $\bullet \rightleftarrows \bullet$), разделяемую память (например, $\subset, \bullet \supset$), экземпляры других шаблонов проектирования. Сигнатурные функторы полиномиальны, коль скоро функционирование системы сводится к традиционным вычислениям и детерминированному взаимодействию.

3 Гибкие графалгебры

Гомоморфизмы графалгебр фиксированной формы I индуцируют естественные преобразования диаграмм структуры — морфизмы в категории-экспоненте C^I . Чтобы связать морфизмами графалгебры произвольных форм, необходимо объединить и «протянуть» между собой все экспоненты с фиксированным основанием C и малыми категориями в качестве показателя. Для этого существует канонический способ, а именно: уплощающая конструкция Гротендика (Grothendieck flattening construction [6, п. 12.2.10]). Произвольному функтору вида $F : Q \rightarrow \mathbf{CAT}$ эта конструкция сопоставляет категорию $\int F$ типа зависимой суммы (dependent sum), объектом которой служит любая пара (A, X) , $X \in \text{Ob } Q$, $A \in \text{Ob } F X$, а морфизмом пары (A, X) в (A', X') служит любая пара $(f : (Fg)A \rightarrow A', g : X \rightarrow X')$ (с законом композиции вида $(f, g) \circ (h, q) = (f \circ (Fg)h, g \circ q)$). Имеется забывающий функтор $U_{(F)} : \int F \rightarrow Q : (A, X) \mapsto X$, $(f, g) \mapsto g$. И обратно: каждый $X \in \text{Ob } Q$ порождает вложение $i_X : F X \hookrightarrow \int F : A \mapsto (A, X)$, $f \mapsto (f, 1_X)$, причем в сумме из таких вложений получается вложение

$$[i_X]_{X \in \text{Ob } D} : \coprod_{X \in \text{Ob } D} F X \hookrightarrow \int F,$$

биективное на объектах.

Для функтора $C^{(-)} : \mathbf{Cat}^{\text{op}} \rightarrow \mathbf{CAT} : X \mapsto C^X$ конструкция Гротендика дает категорию диа-

грамм \mathbf{DC} , класс объектов которой состоит из всех C -диаграмм, а морфизмом диаграммы $\Delta : X \rightarrow C$ в $\Sigma : Y \rightarrow C$ служит любая пара вида $\langle \gamma, f \rangle$, состоящая из функтора $f : X \rightarrow Y$ и естественного преобразования $\gamma : \Delta \rightarrow \Sigma f$; закон композиции морфизмов диаграмм имеет вид:

$$\langle \gamma, f \rangle \circ \langle \varphi, g \rangle = \langle \gamma_{g(-)} \circ \varphi, fg \rangle.$$

С прикладной точки зрения такие морфизмы описывают структурные преобразования моделей систем в алгебраическом представлении [7]. Имеется забывающий функтор формы $\mathbf{D}!_C : \mathbf{DC} \rightarrow \mathbf{Cat} : \Delta \mapsto \text{dom } \Delta, \langle \gamma, f \rangle \mapsto f$. Такое обозначение функтора связано с тем, что соответствие $C \mapsto \mathbf{DC}$ функториально: оно расширяется до эндоморфизма $\mathbf{D} : \mathbf{CAT} \rightarrow \mathbf{CAT}$ (а функтор формы получается путем его действия на функтор $!_C : C \rightarrow \mathbf{1}$, где $\mathbf{1}$ — сингулярная категория). Эндоморфизм \mathbf{D} индуцирует 2-монаду в \mathbf{CAT} , известную как метамодель системного моделирования, поскольку в ней «закодированы» основные применяемые в нем категорные конструкции [8].

Форма графалгебры фигурирует не только в диаграмме структуры, но и в кообласти сигнатурного функтора, в части множества вершин, которое выступает показателем экспоненты. Каноническое объединение всех экспонент с фиксированным основанием и множеством в качестве показателя также получается путем конструкции Гротендика, только для функтора $C^{(-)} : \mathbf{Set}^{\text{op}} \rightarrow \mathbf{CAT}$. Получается полная подкатегория в \mathbf{DC} , состоящая из всех дискретных C -диаграмм, которую будем обозначать через $\mathbf{D}_{\text{Set}}C$. Она корефлексивна в \mathbf{DC} : корефлектор убирает из произвольной диаграммы $\Delta : X \rightarrow C$ все нетождественные морфизмы, превращая ее в дискретную диаграмму $]\Delta[: \text{Ob } X \rightarrow C : x \mapsto \Delta x$. Соответствие $C \mapsto \mathbf{D}_{\text{Set}}C$ расширяется до эндоморфизма в \mathbf{CAT} , индуцирующего подмонаду в \mathbf{D} , причем указанные корефлекторы образуют ретрацию монады \mathbf{D} на нее. Тем самым из «материала» монады диаграмм получается следующее обобщение конструкции графалгебры.

Определение 1. Гибкой графалгеброй сигнатуры $F : D \rightarrow \mathbf{D}_{\text{Set}}C$ называется любая пара $(A, \Delta : I \rightarrow C), A \in \text{Ob } D, I \in \mathbf{Cat}$, такая что $]\Delta[= FA$. Морфизмом такой гибкой графалгебры в $(A', \Delta' : I' \rightarrow C)$ называется любая пара $(f : A \rightarrow A', g : I \rightarrow I')$, такая что существует (обязательно единственное) семейство морфизмов $\theta_i : \Delta_i \rightarrow \Delta'_i g_i, i \in \text{Ob } I$, такое что пара $\langle \theta,]g[: \text{Ob } I \rightarrow \text{Ob } I' \rangle$ совпадает с $Ff :]\Delta[\rightarrow]\Delta'[$, а пара $\langle \theta, g \rangle$ представляет собой морфизм диаграммы Δ в Δ' в \mathbf{DC} . \square

Теорема 1. Все гибкие графалгебры сигнатуры $F : D \rightarrow \mathbf{D}_{\text{Set}}C$ и все их морфизмы образуют кате-

горию, изоморфную вершине следующего декартова квадрата в \mathbf{CAT} :

$$\begin{array}{ccc} AF & \overset{\text{---}}{\longrightarrow} & \mathbf{DC} \\ \downarrow & & \downarrow]-[\\ D & \xrightarrow{F} & \mathbf{D}_{\text{Set}}C \end{array}$$

Доказательство. Проверяется непосредственно по правилам вычисления пределов и экспонент в \mathbf{CAT} . \square

Ребра декартова квадрата из теоремы 1 задают канонические забывающие функторы, которые, аналогично графалгебрам фиксированной формы, выделяют из гибких графалгебр сигнатуры F носитель (левая вертикальная стрелка $U^F : AF \rightarrow D : (A, \Delta) \mapsto A$) и структуру (верхняя горизонтальная стрелка $S^F : AF \rightarrow \mathbf{DC} : (A, \Delta) \mapsto \Delta$). Имеет значение также функтор выделения формы

$$H^F = (\mathbf{D}!_C)S^F : AF \rightarrow \mathbf{Cat} : (A, \Delta) \mapsto \text{dom } \Delta, (f, g) \mapsto g.$$

В частности, функтор $\langle U^F, H^F \rangle : AF \rightarrow D \times \mathbf{Cat} : (A, \Delta) \mapsto (A, \text{dom } \Delta)$ унивалентен, а функтор носителя, как правило, нет. Зато есть возможность строить свободные гибкие графалгебры.

Следствие 1.1. Функтор носителя гибких графалгебр любой сигнатуры представляет собой корефлектор.

Доказательство. Непосредственно проверяется, что класс всех корефлекторов устойчив относительно декартовых квадратов в \mathbf{CAT} . В явной форме для произвольной сигнатуры гибких графалгебр $F : D \rightarrow \mathbf{D}_{\text{Set}}C$ левым сопряженным к U^F служит функтор

$$G^F : D \rightleftarrows AF : A \mapsto (A, FA), f \mapsto (f, (\mathbf{D}_{\text{Set}}!_C)Ff).$$

Единица этого сопряжения тождественна, а коединица состоит из морфизмов гибких графалгебр $(1_A, \text{Ob } \text{dom } \Delta \rightleftarrows \text{dom } \Delta) : (A, FA) \rightarrow (A, \Delta), (A, \Delta) \in \text{Ob } AF$. \square

Следствие 1.2. Для любой сигнатуры гибких графалгебр $F : D \rightarrow \mathbf{D}_{\text{Set}}C$ следующие утверждения эквивалентны:

- (i) функтор носителя представляет собой изоморфизм;
- (ii) функтор носителя унивалентен;
- (iii) F сопоставляет любому носителю пустую диаграмму.

Доказательство.

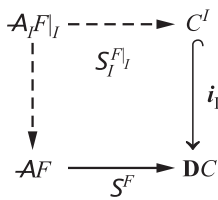
(iii) \Rightarrow (i). Если Δ — пустая диаграмма, то и Δ пуста, поэтому при выполнении условия (iii) коединица сопряжения $G^F \dashv U^F$ тождественна, т. е. $G^F U^F = 1_{AF}$.

(i) \Rightarrow (ii). Тривиально.

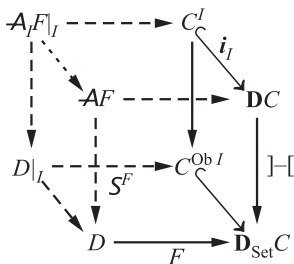
(ii) \Rightarrow (iii). Предположим противное: пусть диаграмма FA при некотором $A \in \text{Ob } D$ содержит точку i . Обозначим через Δ C -диаграмму, полученную из FA путем добавления эндострелки $e : i \rightarrow i$ с законом композиции $e \circ e = 1_i$, помеченной морфизмом $1_{(FA)i}$. Гибкая графалгебра (A, Δ) имеет два разных эндоморфизма с тождественным носителем: один переводит e в e , а другой — в 1_i . \square

Чтобы выделить в AF графалгебры произвольной фиксированной формы I , нужно найти в D все носители, которым сигнатурный функтор сопоставляет объекты категории $C^{\text{Ob } I}$. Для этого используется конструкция полного прообраза сигнатурного функтора — декартова квадрата с соответствующим вложением. Любая гибкая графалгебра сигнатуры F содержится в одной и только одной из категорий графалгебр фиксированной формы, полученных таким способом.

Теорема 2. Пусть $F : D \rightarrow \mathbf{DSet}C$ — сигнатура гибких графалгебр, I — произвольная форма, $F|_I : D|_I \rightarrow C^{\text{Ob } I}$ — полный прообраз функтора F относительно канонического вложения $i_{\text{Ob } I} : C^{\text{Ob } I} \hookrightarrow \mathbf{DSet}C$. Полный прообраз функтора структуры $S^F : AF \rightarrow \mathbf{DC}$ относительно канонического вложения $i_I : C^I \hookrightarrow \mathbf{DC}$ совпадает (с точностью до изоморфизма) с функтором структуры категории $\langle I, F|_I \rangle$ -алгебр, которая тем самым канонически вкладывается в AF :



Доказательство. Рассмотрим следующую диаграмму в форме куба:



Здесь стрелки с закругленным началом обозначают канонические вложения (так что, в частности,

правая грань является коммутативным квадратом); нижняя грань представляет построение функтора $F|_I$ как полного прообраза (т. е. является декартовым квадратом); передняя грань — это декартов квадрат из теоремы 1; задняя грань (с правой вертикальной стрелкой $C^{(\text{Ob } I \hookrightarrow I)} : C^I \rightarrow C^{\text{Ob } I}$) — это декартов квадрат, представляющий построение категории графалгебр $A_I F|_I$ [3, теорема 1]; стрелка из $A_I F|_I$ в AF изображает единственный функтор, делающий всю диаграмму коммутативной. Ввиду композиционных свойств декартовых квадратов [9, предложение 11.10(1)], прямоугольник $A_I F|_I - C^I - \mathbf{DSet}C - D$ является декартовым, а значит, и верхняя грань куба — декартов квадрат. \square

Чтобы описать динамику изменения формы структуры распределенной системы, введем пространство данных D вида $\dot{D} \times T$, где \dot{D} — инвариантное (не зависящее от времени) представление данных; T — вполне упорядоченное множество, представляющее время. (Напомним, что любое предупорядоченное множество стандартным образом представляется в виде категории [10, § 1.2].) Тогда графалгебраические спецификации систем, которые могут существовать в некоторый момент времени $t \in T$, определяются как объекты подкатегории $AF|_t \subseteq AF$ — полного прообраза функтора носителя относительно сечения

$$(1_{\dot{D}}, [t]!_{\dot{D}}) : \dot{D} \hookrightarrow D : X \mapsto \langle X, t \rangle.$$

Сценарий эволюции некоторой системы $(A, \Delta) \in \text{Ob } AF$ может быть задан функтором вида $S : T \hookrightarrow AF$, обратным справа к функтору проекции носителя на T и принимающим значение (A, Δ) в начальный момент времени. Появляется возможность строго специфицировать и верифицировать структурные требования к трансформациям архитектуры системы в ходе эволюции, такие как переход от реализации некоторого шаблона проектирования к реализации другого. Требования такого рода задаются как свойства входящих в область значений функтора S морфизмов гибких графалгебр.

Например, автоматический синтез архитектуры нейросетей путем пошаговых трансформаций описывается гибкими графалгебрами следующим образом. Выберем в качестве \dot{D} подкатегорию в \mathbf{Set} , состоящую из единственного объекта \mathbb{R} и подходящих вещественных функций («перекодировок» в составе допустимых трансформаций), в качестве T некоторый начальный отрезок натуральных чисел (счетчик шагов), а в качестве пространства вычислений C категорию \mathbf{Set} . Сигнатурный функтор F сопоставляет паре (\mathbb{R}, i) , $i \in T$, конечную дискретную диаграмму, размеченную конечными степенями множества \mathbb{R} . Тем самым сигнатура фиксирует

число и мощность слоев нейросети на каждом шаге, но не накладывает ограничений на связи между слоями, что позволяет искать среди объектов подкатегории $AF|_i$ наиболее подходящую конфигурацию (топологию и веса) связей. Можно порождать и параллельно анализировать разные варианты сигнатуры, например с помощью генетического алгоритма. В итоге любой сценарий синтеза сети $S : T \hookrightarrow AF$ включает морфизм результата i -го шага в результат $(i + 1)$ -го, гарантирующий сохранение функциональности при трансформации — необходимое условие вычислительной эффективности алгоритма [4]. Действительно, требование наличия такого морфизма может существенно сузить пространство поиска конфигурации связей, ограничивая его кообластями морфизмов результата предыдущего шага. Теоретико-категорное описание открывает путь к дальнейшему повышению эффективности путем привлечения средств компьютерной алгебры для навигации вдоль морфизмов графалгебр в ходе поиска.

В еще более общем случае эволюции распределенной системы от времени зависит не только ее форма, но и представление данных. Чтобы обеспечивать преемственность данных по ходу эволюции такой системы, указанная зависимость задается функтором $D^\circ : T \rightarrow \mathbf{CAT}$ таким, что $D^\circ t$ — представление данных в момент времени $t \in T$. В качестве носителя гибких графалгебр целесообразно выбрать конструкцию Гротендика $\int D^\circ$. Если и представление вычислений зависит от времени, то вводится функтор $C^\circ : T \rightarrow \mathbf{CAT}$ и сигнатура приобретает вид:

$$F : \int D^\circ \rightarrow \mathbf{D}_{\text{Set}} \int C^\circ.$$

Чтобы она не нарушала причинность, в любой момент времени t она должна удовлетворять соотношению

$$\begin{aligned} (\mathbf{D}_{\text{Set}}([t]!_{\int C^\circ})) Fi_t &= \\ &= (\mathbf{D}_{\text{Set}}U_{(C^\circ)}) Fi_t : D^\circ t \rightarrow \mathbf{D}_{\text{Set}}T. \end{aligned}$$

Сценарий эволюции системы с такой сильно изменчивой архитектурой — это функтор $S : T \hookrightarrow AF$ такой, что

$$U_{(D^\circ)}U^F S = 1_T.$$

Например, в этой форме можно будет описать технологии развития компьютерного представления объектов управления от грубой расчетной/информационной модели до полнофункционального цифрового двойника. Такие технологии активно разрабатываются в настоящее время [11].

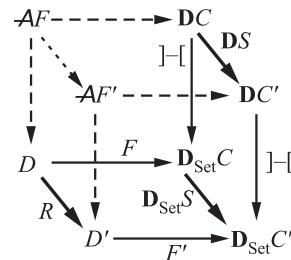
4 Функтор гибких графалгебр

Примечательно, что сигнатуры гибких графалгебр также строятся из «материала» монады диаграмм: в совокупности все они образуют класс морфизмов категории Клейсли монады \mathbf{D}_{Set} . Однако, чтобы выявить функториальную природу конструкции гибкой графалгебры, требуется категория, в которой сигнатуры выступают объектами, а не морфизмами. В таких случаях применяется другая универсальная конструкция — категория запятой (comma category) [10, § II.6] (которая, в свою очередь, представляет собой частный случай графалгебры [3]).

Теорема 3. *Отображение $F \mapsto AF$ служит функцией объектов функтора*

$$A : (1_{\mathbf{CAT}} \downarrow \mathbf{D}_{\text{Set}}) \rightarrow \mathbf{CAT}.$$

Доказательство. По определению категории запятой объектом в $(1_{\mathbf{CAT}} \downarrow \mathbf{D}_{\text{Set}})$ служит в точности любой функтор $F : D \rightarrow \mathbf{D}_{\text{Set}}C$, т.е. сигнатура гибких графалгебр, а его морфизмом в $F' : D' \rightarrow \mathbf{D}_{\text{Set}}C'$ служит любая пара функторов $(R : D \rightarrow D', S : C \rightarrow C')$, такая что $F'R = (\mathbf{D}_{\text{Set}}S)F$. Эта пара индуцирует единственный функтор из AF в AF' , который делает коммутативным весь нижеприведенный куб, передняя и задняя грани которого представляют собой декартовы квадраты из теоремы 1:



5 Заключение

Конструкция графалгебры открывает новые горизонты в применении алгебраических методов для анализа и проектирования распределенных систем. В ее гибком варианте появляются возможности формально специфицировать, верифицировать и в конечном счете эффективно автоматизировать процедуры эволюционного изменения архитектуры систем. Для практической реализации этих возможностей предстоит создать удобные языки и инструменты.

С теоретической точки зрения примечательно, что гибкие графалгебры строятся из «материала»

монады диаграмм: это подчеркивает ключевую роль указанной монады в качестве математического фундамента системной инженерии. В частности, намечается «системно-инженерный» подход к конструированию классического понятия алгебры: категорию обычных алгебр можно получить как частный случай комбинации декартовых квадратов, приведенных в теоремах 1 и 2.

Литература

1. Ковалёв С. П. Формальный подход к разработке программных систем. — Новосибирск: НГУ, 2004. 180 с.
2. Gurevich Y. Evolving algebras 1993: Lipari guide // Specification and validation methods. — Oxford: Oxford University Press, 1995. P. 9–36.
3. Ковалёв С. П. Алгебраическая спецификация графовых вычислительных структур // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 2–9. doi: 10.14357/19922264220101. EDN: GXHPZI.
4. Cai H., Chen T., Zhang W., Yu Y., Wang J. Efficient architecture search by network transformation // 32th AAAI Conference “Artificial Intelligence” Proceedings / Eds. S. A. McIlraith, K. Q. Weinberger. — Palo Alto, CA, USA: AAAI Press, 2018. P. 2787–2794.
5. Mens T., Magee J., Rumpe B. Evolving software architecture descriptions of critical systems // Computer, 2010. Vol. 43. No. 5. P. 42–48. doi: 10.1109/MC.2010.136.
6. Barr M., Wells C. Category theory for computing science. — London: Prentice Hall, 1990. 538 p.
7. Ковалёв С. П. Методы теории категорий в модельно-ориентированной системной инженерии // Информатика и её применения, 2017. Т. 11. Вып. 3. С. 42–50. doi: 10.14357/19922264170305. EDN: ZGIGGD.
8. Ковалёв С. П. Теория категорий как математическая прагматика модельно-ориентированной системной инженерии // Информатика и её применения, 2018. Т. 12. Вып. 1. С. 95–104. EDN: YTTRFC.
9. Adámek J., Herrlich H., Strecker G. E. Abstract and concrete categories. — New York, NY, USA: John Wiley, 1990. 507 p.
10. Маклейн С. Категории для работающего математика / Пер. с англ. — М.: Физматлит, 2004. 352 с. (Mac Lane S. Categories for the working mathematician. — New York, NY, USA: Springer, 1978. 317 p.)
11. Семенов П. В., Семишкур Р. П., Дяченко И. А. Концептуальная модель реализации технологии «цифровых двойников» для предприятий нефтегазового комплекса // Газовая промышленность, 2019. № 7(787). С. 24–30. EDN: EZXDVI.

Поступила в редакцию 12.09.22

ALGEBRAIC SPECIFICATION OF GRAPH COMPUTATIONAL STRUCTURES

S. P. Kovalyov

V. A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, 65 Profsoyuznaya Str., Moscow 117997, Russian Federation

Abstract: The previously proposed generalized approach to algebraic specification of distributed systems is developed based on the novel category-theoretic construction called graphalgebra. The graphalgebraic specification is based upon a directed multigraph, the edges of which represent computational operations performed in the nodes of the system and the vertices denote the data exchange ports between the components. Changing the system architecture during the life cycle leads to changes in the graph shape, computation algorithms, and data exchange contents. For a formal description of such changes, a graph transformation technique for graphalgebras is proposed. A novel category-theoretic construction called flexible graphalgebra is introduced which appeared to be closely related to the well-known monad of diagrams. A functor is presented that produces all categories of flexible graphalgebras from their signatures. The theoretical results are illustrated by examples from the field of automatic synthesis of neural network architecture by step-by-step transformations.

Keywords: algebraic specification; distributed system; architecture evolution; category theory; graphalgebra; monad of diagrams

DOI: 10.14357/19922264240102

EDN: MZAHBS

References

1. Kovalyov, S. P. 2004. *Formal'nyy podkhod k razrabotke programnykh sistem* [Formal approach to software systems development]. Novosibirsk: NSU. 180 p.
2. Gurevich, Y. 1995. Evolving algebras 1993: Lipari guide. *Specification and validation methods*. Oxford: Oxford University Press. 9–36.
3. Kovalyov, S. P. 2022. Algebraicheskaya spetsifikatsiya grafovykh vychislitel'nykh struktur [Algebraic specifica-

- tion of graph computational structures]. *Informatika i ee Primeneniya — Inform. Appl.* 16(1):2–9. doi: 10.14357/19922264220101. EDN: GXHPZI.
4. Cai, H., T. Chen, W. Zhang, Y. Yu, and J. Wang. 2018. Efficient architecture search by network transformation. *32th AAAI Conference “Artificial Intelligence” Proceedings*. Eds. S. A. McIlraith and K. Q. Weinberger. Palo Alto, CA: AAAI Press. 2787–2794.
 5. Mens, T., J. Magee, and B. Rumpe. 2010. Evolving software architecture descriptions of critical systems. *Computer* 43(5):42–48. doi: 10.1109/MC.2010.136.
 6. Barr, M., and C. Wells. 1990. *Category theory for computing science*. London: Prentice Hall. 538 p.
 7. Kovalyov, S. P. 2017. Metody teorii kategoriy v model’no-orientirovannoy sistemnoy inzhenerii [Methods of category theory in model-based systems engineering]. *Informatika i ee Primeneniya — Inform. Appl.* 11(3):42–50. doi: 10.14357/19922264170305. EDN: ZGIGGD.
 8. Kovalyov, S. P. 2018. Teoriya kategoriy kak matematicheskaya pragmatika model’no-orientirovannoy sistemnoy inzhenerii [Category theory as a mathematical pragmatics of model-based systems engineering]. *Informatika i ee Primeneniya — Inform. Appl.* 12(1):95–104. EDN: YTTRFC.
 9. Adámek, J., H. Herrlich, and G. E. Strecker. 1990. *Abstract and concrete categories*. New York, NY: John Wiley. 507 p.
 10. Mac Lane, S. 1978. *Categories for the working mathematician*. New York, NY: Springer. 317 p.
 11. Semenov, P. V., R. P. Semishkur, and I. A. Dyachenko. 2019. Kontseptual’naya model’ realizatsii tekhnologii “tsifrovyykh dvoynikov” dlya predpriyatiy neftegazovogo kompleksa [Conceptual model of digital twin technology implementation for oil and gas industry]. *Gazovaya promyshlennost’* [Gas Industry] 7(787):24–30. EDN: EZXDVJ.

Received September 12, 2022

Contributor

Kovalyov Sergey P. (b. 1972) — Doctor of Science in physics and mathematics, leading scientist, V. A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, 65 Profsoyuznaya Str., Moscow 117997, Russian Federation; kovalyov@sibnet.ru

УПРАВЛЕНИЕ ЛИНЕЙНЫМ ВЫХОДОМ АВТОНОМНОЙ ДИФФЕРЕНЦИАЛЬНОЙ СИСТЕМЫ ПО КВАДРАТИЧНОМУ КРИТЕРИЮ НА БЕСКОНЕЧНОМ ГОРИЗОНТЕ*

А. В. Босов¹

Аннотация: Решена задача оптимального управления линейным выходом стохастической дифференциальной системы на бесконечном горизонте. Решение получено как предельная форма оптимального управления в соответствующей задаче с конечным горизонтом. Приведены достаточные условия существования управления, состоящие из требований стационарности нелинейной динамики, конечности квадратичного целевого функционала, стабилизируемости линейного выхода и существования предела в формуле Фейнмана–Каца, описывающей нелинейную часть управления. Условия для линейной части управления связаны с классическими результатами существования решения автономного уравнения Риккати. Существование предела в формуле Фейнмана–Каца — с решением параболического уравнения, задающего коэффициенты для нелинейной части управления. Рассмотрен частный случай линейного сноса, при котором сохраняется нелинейный характер задачи, но оптимальное управление оказывается линейным и по выходу, и по переменной состояния. Приведены результаты численного эксперимента, позволяющего проанализировать переходный процесс в задаче с конечным горизонтом и эргодическим процессом в динамике. Для коэффициентов управления проиллюстрирован предельный переход к оптимальным значениям соответствующего оптимального автономного управления.

Ключевые слова: стохастическая дифференциальная система Ито; управление по выходу; оптимальное управление; квадратичный критерий; параболическое уравнение; формула Фейнмана–Каца

DOI: 10.14357/19922264240103

EDN: UEESFO

1 Введение

Типовым вариантом классической задачи управления [1] линейной дифференциальной гауссовской системой по квадратичному критерию качества (LQG, linear-quadratic-Gaussian) стала постановка с бесконечным горизонтом управления [2, 3]. Естественно, что на систему управления в этом случае накладываются жесткие ограничения, обеспечивающие ее существование бесконечное время. Прежде всего это автономность, т. е. независимость от времени всех функций в модели динамики и в целевом функционале, устойчивость и стабилизируемость. Последние свойства обеспечивают принципиальную возможность управления и конечность квадратичного критерия. Кроме того, вполне «понятным» элементом постановки варианта LQG с бесконечным горизонтом становится поиск оптимального автономного управления в виде линейной функции выхода. Линейный класс допустимых управлений сводит исследование задачи управления к изучению свойств уравнения Риккати, так что решение задачи управления существует, если имеет решение предельное уравнения

Риккати. Постановка задачи управления, используемая в данной статье, похожа на модель LQG, т. е. используется квадратичный целевой функционал и линейный выход, дающий в оптимальном решении типовое, т. е. линейное по выходу, слагаемое. Принципиальное отличие заключается в нелинейной неуправляемой динамике, из-за которой в оптимальном управлении в варианте с конечным горизонтом появляется нелинейная часть, для вычисления которой нужно решать параболическое дифференциальное уравнение в частных производных.

Подробно эта задача исследована в [4], а цель данной статьи — представить решение аналогичной задачи в постановке с бесконечным горизонтом. Эта постановка сформулирована в следующем разделе статьи, а в разд. 3 сформулирован основной результат. Ровно так же, как управление [4], «похожее» на LQG, все-таки оказывается нелинейным, полученное оптимальное автономное управление также остается нелинейным, но вместо параболического уравнения описывается теперь обыкновенным дифференциальным уравнением, хотя и нелинейным. Чтобы посмотреть, как на практике приме-

* Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва).

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ABosov@frcscs.ru

нится автономное управление, в разд. 4 рассмотрен частный случай линейного сноса, когда управление оказывается линейным, несмотря на остающуюся нелинейной задачу. Этот вариант интересен тем, что позволяет еще и посмотреть, как численные решения параболического уравнения традиционным методом сеток [5] и методом имитационного моделирования формулы Фейнмана–Каца [6], которые «не знают» о наличии точного линейного решения, ведут себя в связке при переходе обычного решения в автономный режим.

2 Постановка задачи

На каноническом вероятностном пространстве $(\Omega, \mathcal{F}, \mathcal{P}, \mathcal{F}_t)$, $t \in [0, \infty)$, рассмотрим автономную стохастическую динамическую систему, состояние которой представляет диффузионный процесс $y_t \in \mathbb{R}^{n_y}$, описываемый системой нелинейных стохастических дифференциальных уравнений Ито:

$$dy_t = A(y_t) dt + \Sigma(y_t) dv_t, \quad y_0 = Y, \quad t \in [0, \infty), \quad (1)$$

где $v_t \in \mathbb{R}^{n_v}$ — стандартный векторный винеровский процесс; $Y \in \mathbb{R}^{n_y}$ — случайная величина с конечным вторым моментом; векторная функция $A = A(y) : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y}$ и матричная функция $\Sigma = \Sigma(y) : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_y \times n_v}$ удовлетворяют условиям Ито:

$$|A(y)| + |\Sigma(y)| \leq C(1 + |y|);$$

$$|A(y_1) - A(y_2)| + |\Sigma(y_1) - \Sigma(y_2)| \leq C|y_1 - y_2|, \\ y_1, y_2 \in \mathbb{R}^{n_y},$$

обеспечивающим существование единственного потраекторного решения уравнения (1) на любом конечном интервале $t \in [0, T]$ [7] (здесь и далее $|\cdot|$ обозначает евклидову норму вектора или матрицы). Чтобы рассматривать решения (1) на интервале $[0, \infty)$, дополнительно потребуем от процесса y_t стационарности в широком и узком смысле [8]. Это предположение избыточно, пока рассматривается задача с полной информацией, т.е. состояние y_t предполагается известным. Но в дальнейшем при переходе к постановке задачи с косвенными наблюдениями за y_t стационарность примет вполне содержательный смысл.

Состоянием y_t формируется управляемый выход, описываемый процессом $z_t \in \mathbb{R}^{n_z}$, линейно связанным с состоянием:

$$dz_t = ay_t dt + bz_t dt + cu_t dt + \sigma dw_t, \\ z_0 = Z, \quad t \in [0, \infty), \quad (2)$$

где $a \in \mathbb{R}^{n_z \times n_y}$, $b \in \mathbb{R}^{n_z \times n_z}$, $c \in \mathbb{R}^{n_z \times n_u}$ и $\sigma \in \mathbb{R}^{n_z \times n_w}$ — известные матрицы; $w_t \in \mathbb{R}^{n_w}$ — не

зависящий от v_t , Y и Z стандартный векторный винеровский процесс; $Z \in \mathbb{R}^{n_z}$ — случайная величина с конечным вторым моментом, не зависящая от v_t и Y ; $u_t \in \mathbb{R}^{n_u}$ — управление, целью которого ставится стабилизация выхода z_t около некоторой траектории, формируемой состоянием y_t . Цель управления формализуется ниже квадратичным критерием качества общего вида.

Задача управления формулируется в предположении наличия полной информации о состоянии y_t и выходе z_t (соответствующая σ -алгебра обозначается $\mathcal{F}_t^{y,z}$ и выполнено $\mathcal{F}_t^{y,z} \subseteq \mathcal{F}_t \subseteq \mathcal{F}$), т.е. управление u_t предполагается $\mathcal{F}_t^{y,z}$ -измеримым. Класс U_0^∞ допустимых управлений составляют автономные (не зависящие прямо от времени) управления с полной обратной связью, т.е. функции вида $u_t = u(y, z) \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$ и $z \in \mathbb{R}^{n_z}$, в предположении, что соответствующая реализация $u_t = u(y_t, z_t)$ обеспечивает выполнение условий существования y_t и z_t для $u_t \in U_0^\infty$. Поскольку состояние y_t не зависит от управления u_t , а выход z_t описывается линейным автономным уравнением с винеровским процессом, то данное формальное требование ограничивает допустимые управления процессами второго порядка, что обеспечивает существование решения (2) на любом конечном интервале $t \in [0, T]$. Для управления на интервале $[0, \infty)$ дополнительно потребуются типовые условия стабилизируемости [2], обсуждаемые далее.

Качество управления U_0^∞ определяется целевым функционалом следующего вида:

$$J(U_0^\infty) = \lim_{T \rightarrow \infty} J(U_0^T), \\ J(U_0^T) = \mathbb{E} \left\{ \frac{1}{T} \int_0^T \|Py_t + Qz_t + Ru_t\|_S^2 dt \right\}, \quad (3)$$

где $P \in \mathbb{R}^{n_j \times n_y}$, $Q \in \mathbb{R}^{n_j \times n_z}$, $R \in \mathbb{R}^{n_j \times k_u}$ и $S \in \mathbb{R}^{n_j \times n_j}$ ($S \geq 0$, $S = S'$) — заданные матрицы, весовая функция $\|x\|_S^2 = x'Sx$, единичной матрице $S = E$ соответствует евклидова норма $\|x\|_E^2 = |x|^2$, «'» — операция транспонирования; $\mathbb{E}\{\cdot\}$ — оператор математического ожидания (далее еще используется обозначение $\mathbb{E}\{\cdot|\mathcal{F}\}$ для условного математического ожидания относительно σ -алгебры \mathcal{F}). Кроме того, предполагается выполненным обычное условие невырожденности, в данных обозначениях принимающее вид $R'SR > 0$.

Задача состоит, таким образом, в поиске $(U^*)_0^\infty = \{u^*(y, z), y \in \mathbb{R}^{n_y}, z \in \mathbb{R}^{n_z}\}$, допустимого управления с обратной связью, реализации $u_t^* = u^*(y_t, z_t^*)$, $t \in [0, \infty)$, которого доставляют минимум квадратичному функционалу $J(U_0^\infty)$:

$$(U^*)_0^\infty = \arg \min_{u_t \in U_0^\infty} J(U_0^\infty). \quad (4)$$

Далее через z_t^* обозначается решение (2), отвечающее u_t^* , и учитывается, что y_t от u_t не зависит.

Следует обратить внимание на некоторые отличия данной постановки от классической задачи LQG с бесконечным горизонтом (хорошее описание которой дано в [2]). В классическом варианте вопрос с определением класса допустимых управлений однозначно решается его описанием линейными функциями состояния и выхода, а именно: если обозначить через $u_t^\# = u_t^\#(y_t, z_t^\#, T)$ решение (неавтономное, поэтому $u_t^\#(y, z, T)$ с нижним индексом t) LQG-задачи с конечным горизонтом $t \in [0, T]$, оптимальное на классе нелинейных допустимых управлений, то оно получится линейным:

$$u_t^\# = L_t^{y\#} y_t + L_t^{z\#} z_t^\# + l_t^\#.$$

Здесь, как и аналогично выше, через $z_t^\#$ обозначено решение (2), отвечающее $u_t^\#$. Соответственно, в описании класса допустимых управлений в задаче с бесконечным горизонтом будут фигурировать только линейные управления

$$u_t = u(y_t, z_t) = L^y y_t + L^z z_t + l,$$

что естественным образом вытекает из оптимальности в задаче с конечным горизонтом именно линейного управления. В рассматриваемой задаче этого нет, поскольку и состояние (1), и, главное, решение задачи с конечным горизонтом, полученное в [4], нелинейные. Так что, хотя искомое решение и находится в виде предельной формы оптимального управления с конечным горизонтом, оно не будет относиться к классу линейных.

3 Основной результат

Очевидно, что основу для решения задачи (4) обеспечивает решение соответствующей задачи с конечным горизонтом, которое получено в [4]. Для корректного предельного перехода от постановки с конечным горизонтом к бесконечному времени требуется выполнение ряда условий, объединенных в следующем утверждении.

Теорема. *Решение задачи (4) может быть записано в виде*

$$u^*(y, z) = -\frac{1}{2} (R'SR)^{-1} (c'(2\alpha_* z + \beta_*(y)) + 2R'S(Py + Qz)), \quad (5)$$

где симметричная неотрицательно определенная матрица $\alpha_* \in \mathbb{R}^{n_z \times n_z}$ и вектор-функция $\beta_* = \beta_*(y) = (\beta_*^{(1)}(y), \dots, \beta_*^{(n_z)}(y))' \in \mathbb{R}^{n_z}$ представляют собой решения уравнений

$$\begin{aligned} & (b' - Q'SR(R'SR)^{-1}c') \alpha_* + \\ & + \alpha_* (b - c(R'SR)^{-1}R'SQ) + \\ & + Q'(S - SR(R'SR)^{-1}R'S)Q - \\ & - \alpha_* c(R'SR)^{-1}c' \alpha_* = 0; \quad (6) \end{aligned}$$

$$\begin{aligned} & \frac{1}{2} \text{tr} \left\{ \Sigma' \frac{\partial^2 \beta_*^{(i)}}{\partial y^2} \Sigma \right\} + A' \frac{\partial \beta_*^{(i)}}{\partial y} + \\ & + \sum_{j=1}^{n_y} y^{(j)} [M_*]^{(ji)} + \sum_{j=1}^{n_z} \beta_*^{(j)} [N_*]^{(ji)} = 0, \quad (7) \end{aligned}$$

где

$$\begin{aligned} M_* &= 2((a' - P'SR(R'SR)^{-1}c') \alpha_* + \\ & + P'(S - SR(R'SR)^{-1}R'S)Q); \\ N_* &= b - c(R'SR)^{-1}R'SQ - c(R'SR)^{-1}c' \alpha_*, \\ & i = \overline{1, n_z}, \end{aligned}$$

если для параметров системы (1), (2) и целевого функционала (3) выполнены следующие условия:

- (1) матрица b устойчива;
- (2) матрицы (K_b, c) стабилизируемы, $K_b = b - c(R'SR)^{-1}R'SQ$;
- (3) матрицы (K_b', K_Q) стабилизируемы, $K_Q = Q'S^{1/2}(E - S^{1/2}R(R'SR)^{-1}R'S^{1/2})$;
- (4) для любого $y \in \mathbb{R}^{n_y}$ существует и не зависит от t конечный предел

$$\lim_{T \rightarrow \infty} \mathbb{E} \left\{ I^{-1}(t) \int_t^T I^{-1}(\tau) M_*' y_\tau d\tau | \mathcal{F}_t^y \right\}, \quad (8)$$

где $I^{-1}(\tau) = \exp\{N_*' \tau\}$; y_τ — решение уравнения (1) с переменной времени $\tau \in [t, \infty)$ и начальным условием $y_t = y$.

Уравнение (7) записано с использованием обозначений $y = (y^{(1)}, \dots, y^{(n_y)})'$ для элементов вектора y , $[A]^{(ji)}$ — для элемента j -й строки i -го столбца матрицы A , $\text{tr}\{A\}$ — для следа матрицы A .

Доказательство. Перечисленные в теореме условия обеспечивают существование предельного решения соответствующей (4) задачи с конечным горизонтом, т.е. управления $(U^\#)_0^T = \min_{u_t \in U_0^T} J(U_0^T)$, где $U_0^T = \{u_t(y, z, T), y \in \mathbb{R}^{n_y}, z \in \mathbb{R}^{n_z}, t \in [0, T]\}$. Оптимальное управление $u_t^\# = u_t^\#(y_t, z_t^\#, T)$ получено в [4] в виде:

$$\begin{aligned} u_t^\# &= u_t^\#(y, z, T) = -\frac{1}{2} (R'SR)^{-1} \times \\ & \times (c'(2\alpha_t z + \beta_t) + 2R'S(Py + Qz)), \quad (9) \end{aligned}$$

где матричный коэффициент $\alpha_t = \alpha_t(T) \in \mathbb{R}^{n_z \times n_z}$ представляет собой решение задачи Коши для уравнения Риккати

$$\begin{aligned} \frac{d\alpha_t}{dt} + (b' - Q'SR(R'SR)^{-1}c')\alpha_t + \\ + \alpha_t(b - c(R'SR)^{-1}R'SQ) + \\ + Q'(S - SR(R'SR)^{-1}R'S)Q - \\ - \alpha_t c(R'SR)^{-1}c'\alpha_t = 0, \quad \alpha_T = Q'SQ, \end{aligned} \quad (10)$$

а векторный коэффициент $\beta_t = \beta_t(y, T) = (\beta_t^{(1)}(y, T), \dots, \beta_t^{(n_z)}(y, T))' \in \mathbb{R}^{n_z}$ — задачи Коши для системы дифференциальных уравнений в частных производных

$$\begin{aligned} \frac{\partial \beta_t^{(i)}}{\partial t} + \frac{1}{2} \text{tr} \left\{ \Sigma' \frac{\partial^2 \beta_t^{(i)}}{\partial y^2} \Sigma \right\} + A' \frac{\partial \beta_t^{(i)}}{\partial y} + \\ + \sum_{j=1}^{n_y} y^{(j)} [M_t^{(j)}]^{(ji)} + \sum_{j=1}^{n_z} \beta_t^{(j)} [N_t^{(j)}]^{(ji)} = 0, \end{aligned} \quad (11)$$

где

$$\begin{aligned} M_t = M_t(T) = 2((a' - P'SR(R'SR)^{-1}c')\alpha_t + \\ + P'(S - SR(R'SR)^{-1}R'S)Q); \\ N_t = N_t(T) = b - c(R'SR)^{-1}R'SQ - \\ - c(R'SR)^{-1}c'\alpha_t; \end{aligned}$$

$$\beta_T^{(i)} = 2 \sum_{j=1}^{n_y} y^{(j)} [Q'SP]^{(ji)}, \quad i = \overline{1, n_z}.$$

Перечисленные в теореме условия вместе с исходным предположением о стационарности y_t обеспечивают, во-первых, конечность целевого функционала (3). Во-вторых, из (10) для коэффициента оптимального управления α_* получается уравнение (6) как предел $\alpha_* = \lim_{T \rightarrow \infty} \alpha_t(T)$. Предельное решение матричного уравнения Риккати (10) существует, не зависит от граничного условия и является неотрицательно определенным, поскольку сформулированные в теореме условия стабилизируемости приводят коэффициенты уравнения Риккати к формулировкам условий теоремы 12.2 [2]. Действительно, стабилизируемость пары матриц (K_b, c) — это условие стабилизируемости наблюдений, стабилизируемость пары (K'_b, K_Q) — это условие детектируемости. Последнее основано на том факте, что матрица K_Q обеспечивает представление

$$Q'(S - SR(R'SR)^{-1}R'S)Q = K_Q K'_Q,$$

так как

$$\begin{aligned} (E - S^{1/2}R(R'SR)^{-1}R'S^{1/2}) \times \\ \times (E - S^{1/2}R(R'SR)^{-1}R'S^{1/2})' = \end{aligned}$$

$$\begin{aligned} = (E - S^{1/2}R(R'SR)^{-1}R'S^{1/2})^2 = \\ = E - S^{1/2}R(R'SR)^{-1}R'S^{1/2}. \end{aligned}$$

Кроме того, эта же теорема гарантирует устойчивость матрицы N_* , требуемую для выполнения (8). И последним из (11) получается коэффициент β_* , для которого предельный переход $\beta_*(y) = \lim_{T \rightarrow \infty} \beta_t(y, T)$ обеспечивается условием 4 теоремы. Действительно, поскольку решение (11) может быть представлено с помощью формулы Фейнмана—Каца [9],

$$\begin{aligned} \beta_t(y, T) = \mathbb{E} \left\{ 2I^{-1}(t)I^{-1}(T)Q'SPy(T) + \right. \\ \left. + I^{-1}(t) \int_y^T I^{-1}(\tau)M'(\tau)y(\tau)d\tau | \mathcal{F}_t^y \right\}, \\ I^{-1}(t) = \exp\{N_t^* t\}. \end{aligned} \quad (12)$$

Поскольку матрица $N_* = \lim_{T \rightarrow \infty} N_t(T)$ (предел существует и не зависит от t , так как $N_t(T)$ выражается линейно через $\alpha_t(T)$) устойчива, первое слагаемое в (12) обращается в ноль, что вместе с условием 4 теоремы обеспечивает существование предела при $T \rightarrow \infty$ в (11), не зависящего от t и граничного условия $\beta_T^{(i)}(y, T)$, что и дает уравнения (7), завершая доказательство.

Отметим, что использованный результат в [4] включает еще соотношения, описывающие решение задачи с конечным горизонтом полностью, т. е. определяющие функцию Беллмана. Эти соотношения можно трансформировать и для рассматриваемой задачи (4), но ничего практически содержательного они не дадут, поэтому не используются.

Принципиально важно в полученном решении то, что оптимальное управление (5) остается нелинейным, как и в допредельной постановке, что и является, как уже упоминалось, главным отличием классической задачи LQG от рассматриваемой. Другой вопрос, что к условиям существования решения ничего конструктивного, кроме традиционных требований к коэффициентам уравнения Риккати, добавить не удастся. Это ожидаемо, так как даже конструктивные условия существования решения системы параболических уравнений (11) представляют проблему, тем более затруднительно изучение свойств самих решений. С другой стороны, аналитические исследования можно вполне успешно заменить практическими расчетами. В работах [5, 6] предложены два эффективных метода численного решения (11), а при наличии готовых расчетов анализ сходимости $\lim_{T \rightarrow \infty} \beta_t(y, T)$ тру-

да не составляет. Другой вопрос, что реальные возможности реализации нелинейного управления для общей модели состояния (1) представляются весьма ограниченными. Более практически интересным будет частный случай линейного сноса или сводящийся к нему случай косвенных наблюдений за состоянием дискретной цепи Маркова, которому планируется посвятить будущую работу.

4 Частный случай линейного сноса

Итак, практическая реализация управления $(U^*)_0^\infty$ сведена к решению алгебраического уравнения (6) и системы обыкновенных дифференциальных уравнений (7) вместо уравнения Риккати (10) и системы параболических уравнений (11). При этом если упрощение в части отказа от решения уравнения Риккати не представляется сколь-либо значимым, так как с эффективными численными методами для него нет проблем, то замена параболических уравнений на обыкновенные — это действительно большое упрощение. Тем не менее работа с системой (7) все еще остается довольно сложной, поскольку решать требуется хоть и обыкновенные дифференциальные уравнения второго порядка, но для сетки, покрывающей область возможных значений вектора $y \in \mathbb{R}^{n_y}$, что даже для значений $n_y = 3, 4$ (это минимальные размерности для моделей, имеющих практический смысл) становится уже крайне ресурсоемким. Вместе с тем заменять уравнение (1) линейным неинтересно, так как тогда задача сведется к классической LQG. Интерес, таким образом, представляет случай линейного сноса, т.е. линейной функции $A(y)$, а нелинейность задачи останется за счет диффузии $\Sigma(y)$.

Случай неавтономной системы с линейным сносом $A_t(y) = A_t^a y + A_t^b$, где матрица $A_t^a \in \mathbb{R}^{n_y \times n_y}$ и вектор $A_t^b \in \mathbb{R}^{n_y}$ не зависят от y , рассмотрен в [4]. В этом случае $\beta_t(y, T) = \beta_t^a(T)y + \beta_t^b(T)$ и для коэффициентов получены уравнения:

$$\left. \begin{aligned} \frac{d\beta_t^a}{dt} + \beta_t^a A^a + M_*' + N_*' \beta_t^a &= 0; \\ \frac{d\beta_t^b}{dt} + \beta_t^a A^b + N_*' \beta_t^b &= 0; \\ \beta_T^a(T) &= 2Q' SP; \quad \beta_T^b(T) = 0. \end{aligned} \right\} \quad (13)$$

В условиях сформулированной теоремы существуют пределы $\beta_*^a = \lim_{T \rightarrow \infty} \beta_t^a$ и $\beta_*^b = \lim_{T \rightarrow \infty} \beta_t^b$, не зависящие от t и от начального условия $\beta_T^a(T), \beta_T^b(T)$ в задаче Коши, описываемые уравнениями:

$$\beta_*^a A^a + M_*' + N_*' \beta_*^a = 0; \quad \beta_*^a A^b + N_*' \beta_*^b = 0. \quad (14)$$

Соответственно, коэффициент $\beta_*(y) = \beta_*^a y + \beta_*^b$, а оптимальное управление (5) принимает вид:

$$u^*(y, z) = -\frac{1}{2} (R' SR)^{-1} (2(c' \alpha_* + R' SQ)z + (c' \beta_*^a + 2R' SP)y + c' \beta_*^b). \quad (15)$$

5 Численный анализ стационарного режима

Для иллюстрации перехода неавтономного управления (9) $u_t^\# = u_t^\#(y, z, T)$ в автономный режим (5) $u^* = u^*(y, z)$, который должен иметь место при достаточно больших T , используем пример, детально рассмотренный в [1]. Простая модель эволюции процентных ставок Кокса–Ингерсолла–Росса (Cox–Ingersoll–Ross model) [10], применяемая также для описания показателя RTT (Round-Trip Time) сетевого протокола TCP (Transmission Control Protocol) [11], имеет вид:

$$dy_t = (1 - y_t) dt + 2,5 \sqrt{y_t} dv_t, \quad y_0 = Y \sim \mathbb{N}(15, 9). \quad (16)$$

Выход и целевой функционал описываются уравнениями:

$$dz_t = y_t dt - z_t dt + u_t dt + 2,5 dw_t, \quad z_0 = Z \sim \mathbb{N}(9, 9);$$

$$J(U_0^T) = \mathbb{E} \left\{ \int_0^T ((y_t - z_t)^2 + z_t^2 + u_t^2) dt + (y_T - z_T)^2 + z_T^2 \right\}, \quad T = 5.$$

Процессы y_t, z_t и u_t и возмущения v_t и w_t — скалярные, $\mathbb{N}(M, D)$ обозначает нормальное распределение со средним M и дисперсией D . Известно, что траектории y_t неотрицательны, т.е. $y_t > 0$, процесс — эргодический, известны предельное распределение и переходная вероятность. При этом начальные условия в (15) выбраны так, что M и D значительно отличаются от предельных моментов, так что динамика исследуется в переходном процессе. Соответственно, в стационарном режиме процесс удовлетворяет предположениям модели (1), а также имеет линейный снос, т.е. можно аналитически вычислить коэффициенты (13) и оптимальное автономное управление (14). Также очевидна устойчивость z_t , так что вопрос остается только к выполнению условия 4 теоремы, т.е. к характеру поверхности $\beta_t(y, T)$, определяемой уравнением (11) (в данном скалярном случае — од-

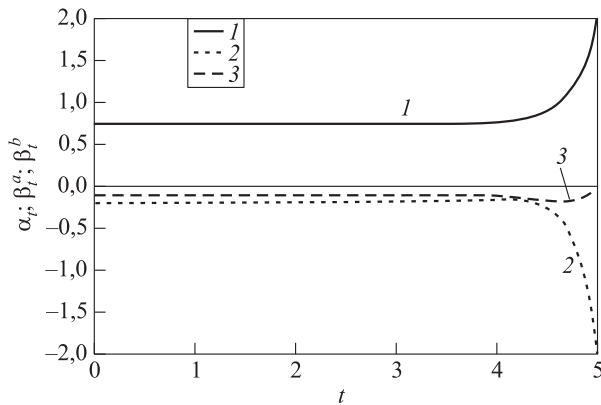


Рис. 1 Коэффициенты оптимального управления: 1 — α_t ; 2 — β_t^a ; 3 — β_t^b

ним параболическим уравнением, а не системой) или формулой Фейнмана—Каца (12).

Отметим сначала, что в рассматриваемом примере нетрудно найти точное решение уравнения Риккати (10), а именно:

$$\alpha_t = \frac{C_\alpha e^{2\sqrt{3}t}(1 + \sqrt{3}) - 1 + \sqrt{3}}{1 - C_\alpha e^{2\sqrt{3}t}},$$

$$C_\alpha = \frac{3 - \sqrt{3}}{3 + \sqrt{3}} e^{-10\sqrt{3}}.$$

Уравнение (6) дает $\alpha_* = \sqrt{3} - 1 \approx 0,73$, что, как нетрудно видеть, совпадает с $\lim_{t \rightarrow -\infty} \alpha_t$. Для решения (11) есть три варианта: не учитывать частный случай линейного сноса и решать (11) традиционным сеточным методом (как предлагается в [5]), или методом имитационного моделирования формулы Фейнмана—Каца (как предлагается в [4]), или решать обыкновенные дифференциальные уравнения (7). Все три расчета представлены в [4]. Применительно к рассматриваемой автономной задаче все три метода подтверждают сходимость $\beta_t(y, T)$ и дают $\beta_*(y) = \beta_*^a y + \beta_*^b$, где $\beta_*^a \approx -0,19$, $\beta_*^b \approx -0,11$, т. е. при любом варианте расчета $\beta_*^a y = \lim_{t \rightarrow -\infty} \beta_t^a(y, 5)$, $\beta_*^b = \lim_{t \rightarrow -\infty} \beta_t^b(5)$. Значения y определялись в результате предварительного моделирования (16) и заданы отрезком $[0, 40]$. Иллюстрируют эти сходимости рис. 1 и 2.

6 Заключение

Представленное исследование дополняет ранее решенную задачу оптимального управления линейным выходом стохастической системы по квадратичному критерию традиционным автономным вариантом с бесконечным временем. Основной

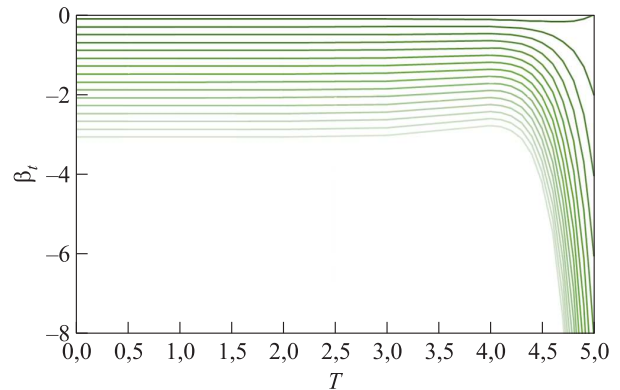


Рис. 2 Примеры сечений поверхности $\beta_t(y, T)$ для разных значений y : сверху вниз — $y = 0; 1; \dots; y = 15$

результат эксплуатирует классические методы исследования уравнения Риккати, как и в традиционной LQG-задаче. Важное отличие при этом заключается в том, что решение остается в классе нелинейных управлений, что было также важным свойством задачи с конечным горизонтом. Завершить изучение данной задачи должен случай косвенных наблюдений и практический анализ качества автономного управления в сравнении с оптимальным на конечном горизонте. Эти исследования запланированы на будущее.

Литература

1. Athans M. The role and use of the stochastic linear-quadratic-Gaussian problem in control system design // IEEE T. Automat. Contr., 1971. Vol. 16. No. 6. P. 529–552. doi: 10.1109/TAC.1971.1099818.
2. Wonham W. M. Linear multivariable control. A geometric approach. — Lecture notes in economics and mathematical systems ser. — Berlin: Springer-Verlag, 1974. Vol. 101. 347 p.
3. Девис М. Х. А. Линейное оценивание и стохастическое управление / Пер. с англ. — М.: Наука, 1984. 206 с. (Davis M. H. A. Linear estimation and stochastic control. — London: Chapman and Hall, 1977. 224 p.)
4. Босов А. В. Задача управления линейным выходом нелинейной неуправляемой стохастической дифференциальной системы по квадратичному критерию // Известия РАН. Теория и системы управления, 2021. № 5. С. 52–73. doi: 10.31857/S000233882104003X.
5. Босов А. В., Стефанович А. И. Управление выходом стохастической дифференциальной системы по квадратичному критерию. II. Численное решение уравнений динамического программирования // Информатика и её применения, 2019. Т. 13. Вып. 1. С. 9–15. doi: 10.14357/19922264190102. EDN: ZASZFR.

6. Босов А. В., Стефанович А. И. Управление выходом стохастической дифференциальной системы по квадратичному критерию. IV. Альтернативное численное решение // Информатика и её применения, 2020. Т. 14. Вып. 1. С. 24–30. doi: 10.14357/19922264200104. EDN: XNHVFT.
7. Флеминг У., Ришел Р. Оптимальное управление детерминированными и стохастическими системами / Пер. с англ. — М.: Мир, 1978. 316 с. (Fleming W. H., Rishel R. W. Deterministic and stochastic optimal control. — New York, NY, USA: Springer-Verlag, 1975. 222 p.)
8. Ширяев А. Н. Вероятность. — 2-е изд. — М.: Наука, 1989. 640 с.
9. Øksendal B. Stochastic differential equations. An introduction with applications. — New York, NY, USA: Springer-Verlag, 2003. 324 p.
10. Cox J. C., Ingersoll J. E., Ross S. A. A theory of the term structure of interest rates // Econometrica, 1985. Vol. 53. Iss. 2. P. 385–407. doi: 10.2307/1911242.
11. Bohacek S., Rozovskii B. A diffusion model of roundtrip time // Comput. Stat. Data An., 2004. Vol. 45. Iss. 1. P. 25–50. doi: 10.1016/S0167-9473(03)00114-2.

Поступила в редакцию 07.12.23

AUTONOMOUS DIFFERENTIAL SYSTEM LINEAR OUTPUT CONTROL BY SQUARE CRITERION ON AN INFINITE HORIZON

A. V. Bosov

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The problem of optimal control of the stochastic differential system linear output on an infinite horizon is solved. The solution is considered as the limit form of optimal control in the corresponding problem with a finite horizon. Sufficient conditions for the existence of control are given. They consist of the requirements of the stationarity of nonlinear dynamics, the finiteness of the quadratic target functional, the stabilizability of the linear output, and the existence of a limit in the Feynman–Katz formula describing the nonlinear part of control. The conditions for the linear part of the control are related to the classical results of the existence of a solution to the autonomous Riccati equation. The existence of a limit in the Feynman–Katz formula is associated with the solution of a parabolic equation that sets the coefficients for the nonlinear part of the control. A special case of linear drift is considered in which the nonlinear nature of the problem is preserved but optimal control turns out to be linear both in output and in the state variable. The results of a numerical experiment are presented which makes it possible to analyze the transient process in a problem with a finite horizon and an ergodic process in dynamics. For the control coefficients, the limiting transition to the optimal values of the corresponding optimal autonomous control is illustrated.

Keywords: stochastic differential Ito system; output control; optimal control; quadratic criterion; parabolic equation; Feynman–Katz formula

DOI: 10.14357/19922264240103

EDN: UEESFO

Acknowledgments

The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (СКР “Informatics”) of FRC CSC RAS (Moscow).

References

1. Athans, M. 1971. The role and use of the stochastic linear-quadratic-Gaussian problem in control system design. *IEEE T. Automat. Contr.* 16(6):529–552. doi: 10.1109/TAC.1971.1099818.
2. Wonham, W. M. 1974. *Linear multivariable control. A geometric approach*. Lecture notes in economics and mathematical systems ser. Berlin: Springer-Verlag. 347 p.
3. Davis, M. H. A. 1977. *Linear estimation and stochastic control*. London: Chapman and Hall. 224 p.
4. Bosov, A. V. 2021. The problem of controlling the linear output of a nonlinear uncontrollable stochastic differential system by the square criterion. *J. Comput. Sys. Sc. Int.* 60(5):719–739. doi: 10.1134/S1064230721040031.
5. Bosov, A. V., and A. I. Stefanovich. 2019. Управление выходом стохастической дифференциальной системы по квадратичному критерию. II. Численное решение уравнений динамической системы управления по квадратичному критерию. II. Динамическое программирование численного решения. *Информатика и её Применение — Inform. Appl.* 13(1):9–15. doi: 10.14357/19922264190102. EDN: ZASZFR.
6. Bosov, A. V., and A. I. Stefanovich. 2020. Управление выходом стохастической дифференциальной системы по квадратичному критерию. IV. Альтернативное численное

- reshenie [Stochastic differential system output control by the quadratic criterion. IV. Alternative numerical decision]. *Informatika i ee Primeneniya — Inform. Appl.* 14(1):24–30. doi: 10.14357/19922264200104. EDN: XNHVFT.
7. Fleming, W. H., and R. W. Rishel. 1975. *Deterministic and stochastic optimal control*. New York, NY: Springer-Verlag. 222 p.
 8. Shiryaev, A. N. 1996. *Probability*. New York, NY: Springer Verlag. 624 p.
 9. Øksendal, B. 2003. *Stochastic differential equations. An introduction with applications*. New York, NY: Springer-Verlag. 324 p.
 10. Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1985. A theory of the term structure of interest rates. *Econometrica* 53(2):385–407. doi: 10.2307/1911242.
 11. Bohacek, S., and B. Rozovskii. 2004. A diffusion model of roundtrip time. *Comput. Stat. Data An.* 45(1):25–50. doi: 10.1016/S0167-9473(03)00114-2.

Received December 7, 2023

Contributor

Bosov Alexey V. (b. 1969) — Doctor of Science in technology, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; avbosov@ipiran.ru

РЫНОК С МАРКОВСКОЙ СКАЧКООБРАЗНОЙ ВОЛАТИЛЬНОСТЬЮ IV: АЛГОРИТМ МОНИТОРИНГА РЫНОЧНОЙ ЦЕНЫ РИСКА ПО ПОТОКУ ВЫСОКОЧАСТОТНЫХ НАБЛЮДЕНИЙ БАЗОВЫХ АКТИВОВ И ДЕРИВАТИВОВ*

А. В. Борисов¹

Аннотация: Четвертая часть цикла посвящена субоптимальному алгоритму мониторинга рыночной цены риска по наблюдениям цен базовых активов и деривативов. Как и во всем цикле, базовые активы описываются моделью со стохастической волатильностью, представленной скрытым марковским скачкообразным процессом (МСП). В условиях безарбитражности рыночная цена риска представляет собой функцию текущего состояния этого процесса. Особенность исследуемой модели рынка как стохастической системы наблюдений заключена в структуре доступных наблюдений. Они представляют собой последовательности цен базовых активов и деривативов, фиксируемых в случайные моменты времени. Цены базовых активов наблюдаются точно, в то время как цены деривативов содержат случайные шумы. Распределения отрезков времени между поступлениями наблюдений, а также шумов в наблюдениях зависят от состояния оцениваемого процесса. Важнейшее свойство наблюдений заключено в высокой интенсивности их поступления по сравнению с интенсивностью изменения состояния оцениваемого марковского процесса. Это свойство позволяет использовать в предлагаемом фильтре центральную предельную теорему для обобщенных процессов восстановления (ЦПТ ОПВ). Качество оценок фильтрации в зависимости от состава используемых наблюдений проиллюстрировано численным примером.

Ключевые слова: рыночная цена риска; марковский скачкообразный процесс; высокочастотные наблюдения; мультивариантный точечный процесс; численный алгоритм

DOI: 10.14357/19922264240104

EDN: ZRQKIT

1 Введение

Заметка продолжает цикл работ [1–3], посвященных решению прикладных задач системного анализа для класса моделей финансового рынка со стохастической волатильностью [4], описываемой внешним ненаблюдаемым МСП. Подобный класс моделей, несмотря на свою простоту, используется достаточно активно [5–7] как для определения справедливой цены различного рода деривативов, так и для решения задач хеджирования, оптимального инвестирования и пр.

Одно из основных качеств исследуемого рынка, затрудняющих решение данных задач, — неполнота рынка, порожденная именно наличием скрытого МСП. Эта статистическая неопределенность определяет актуальность задач оценивания этого процесса по имеющейся статистической информации. В [3] наблюдения представляли собой временную дискретизацию непрерывных процессов, выполненную с некоторым неслучайным шагом. Такая

структура наблюдений ближе к реальности, однако и она остается идеализацией. В действительности биржевая информация представляет собой последовательность заявок на покупку/продажу бумаг, а также проведенных сделок. Ключевое свойство этих данных заключается в том, что все события и цены фиксируются в *случайные* моменты времени. Очевидно, что полезная информация об оцениваемом МСП содержится не только в ценах бумаг, но и в случайных отрезках времени между наблюдениями [8].

Цель работы заключается в представлении субоптимального алгоритма оценивания рыночной цены риска (Market Price of Risk, MPR), или, что эквивалентно, скрытого состояния рынка по наблюдениям цен базовых активов и деривативов, поступающих в случайные моменты времени.

Статья организована следующим образом. Раздел 2 содержит сведения об исследуемой модели рынка, а также составе и свойствах доступных наблюдений. Раздел 3 посвящен представлению ал-

* Работа выполнялась с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва).

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, aborISOV@frcsc.ru

горитма оценивания состояния рынка. Он основан на важнейшем свойстве имеющихся наблюдений — высокой интенсивности их поступления по сравнению со сменой режимов рынка. Данное свойство позволяет использовать ЦПТ ОПВ [9] применительно к исходным наблюдениям, преобразованным специальным образом. Использование этого утверждения позволяет сконструировать достаточно простой и эффективный субоптимальный фильтр. Его качество, а также влияние на точность оценивания наличия дополнительных наблюдений цен деривативов проиллюстрированы численным примером в разд. 4. Раздел 5 содержит заключительные замечания.

2 Модель рынка и постановка задачи мониторинга рыночной цены риска

На вероятностном базисе с фильтрацией $(\Omega, \mathcal{F}, \mathcal{P}, \{\mathcal{F}_t\}_{t \in [0, T]})$ рассматривается модель финансового рынка, состоящего:

- из банковского вклада с детерминированной ставкой r_t ;
- N базовых финансовых инструментов $S_t = \text{col}(S_t^1, \dots, S_t^N)$;
- M деривативов $F_t = \text{col}(F_t^1, \dots, F_t^M)$ с единым сроком погашения T , определяемых платежным требованием $H(S_T) = \text{col}(H^1(S_T), \dots, H^M(S_T))$.

Цена S_t описывается стохастической дифференциальной системой (СДС)

$$dS_t = \text{diag} S_t a(t, Z_t) dt + \text{diag} S_t \sigma(t, Z_t) dw_t, \quad t \in (0, T], \quad S_0 \sim \pi_0(s), \quad (1)$$

где $w_t \triangleq \text{col}(w_t^1, \dots, w_t^K)$ — K -мерный стандартный винеровский процесс ($K \geq N$), а случайные функции мгновенной процентной ставки a и внутренней волатильности σ имеют вид:

$$a(t, Z_t) = \sum_{\ell=1}^L Z_t^\ell a^\ell(t); \quad \sigma(t, Z_t) = \sum_{\ell=1}^L Z_t^\ell \sigma^\ell(t)$$

с набором известных детерминированных функций $\{a^\ell(t)\}_{\ell=1, \dots, L}$ и $\{\sigma^\ell(t)\}_{\ell=1, \dots, L}$. Распределение начальной цены S_0 таково, что все ее компоненты положительны с вероятностью 1.

Мгновенная процентная ставка $a(t, Z_t)$ и стохастическая волатильность $\sigma(t, Z_t)$ определяются

внешним скрытым МСП $Z_t \triangleq \text{col}(Z_t^1, \dots, Z_t^L) \in \{e_1, \dots, e_L\}$ с матрицей интенсивностей переходов $\Lambda(\cdot)$ и начальным распределением π_0^Z . Марковский скачкообразный процесс Z_t представляет собой решение СДС с мартингалом M_t в правой части

$$dZ_t = \Lambda^\top(t) Z_t dt + dM_t, \quad t \in (0, T], \quad Z_0 \sim \pi_0^Z. \quad (2)$$

Согласно [1], MPR θ_t в данной модели определяется формулой:

$$\theta_t = \sum_{\ell=1}^L Z_t^\ell (\sigma^\ell(t))^+ (a^\ell(t) - r_t \mathbf{1})$$

(здесь $\mathbf{1}$ — вектор-столбец подходящей размерности, составленный из единиц) и однозначно определяется значением МСП Z_t . Таким образом, задача оценивания MPR очевидным образом сводится к задаче фильтрации скрытого процесса Z_t .

Процессы (S_t, Z_t, F_t) — ненаблюдаемые. В качестве доступной статистической информации выступают мультивариантные точечные процессы (МТП) [10] $\{(\tau_j^S, S_j)\}_j$ и $\{(\tau_j^F, F_j)\}_j$, где (τ_j^S, S_j) — моменты наблюдений и собственно наблюдения цен базового актива; (τ_j^F, F_j) — соответствующие значения производных ценных бумаг.

Задача фильтрации заключается в построении оценки \hat{Z}_t состояния Z_{t_i} , $t_i = i\Delta$, по всем наблюдениям $\{(\tau_j^S, S_j), (\tau_j^F, F_j) : 0 \leq \tau_j^S, \tau_k^F \leq t_i\}$.

Следует отметить, что в данной задаче структура доступных наблюдений в точности совпадает с реальностью: таким же образом выглядит часть электронного журнала биржевых транзакций, посвященная регистрации успешных сделок.

Поставленная задача может быть решена оптимально в смысле среднеквадратичного критерия, подобно [11]. Однако вычисление оптимальных оценок требует достаточно большого объема вычислений. Помимо этого, оптимальные оценки чувствительны к неточному знанию параметров системы наблюдения. В следующем разделе представлен субоптимальный фильтр, обеспечивающий высокую точность оценивания и требующий умеренных вычислительных затрат.

3 Алгоритм субоптимальной фильтрации

Прежде всего уточним свойства наблюдений. Без ограничения общности для простоты изложения будем считать, что на рынке присутствуют один базовый и один производный актив и что рынок стационарен, т. е. функции r , a , σ и Λ не зависят от

времени t . Дополнительно предположим, что для модели рынка и комплекса наблюдений выполнены следующие предположения.

1. $S_j \triangleq S_{\tau_j^S}$, т. е. в случайные моменты τ_j^S наблюдаются точные цены базовой бумаги.
2. При известной траектории Z интервалы времени $\delta_j^S \triangleq \tau_j^S - \tau_{j-1}^S$ между наблюдениями цен базового актива независимы в совокупности. Распределение δ_j^S зависит от $Z_{\tau_{j-1}^S}$ и имеет известные условные моменты

$$m_\ell^S \triangleq \mathbb{E} \left\{ \delta_j^S | Z_{\tau_{j-1}^S} = e_\ell \right\},$$

$$D_\ell^S \triangleq \mathbb{E} \left\{ (\delta_j^S - m_\ell^S)^2 | Z_{\tau_{j-1}^S} = e_\ell \right\}, \quad \ell = \overline{1, L}.$$

3. $F_k \triangleq F_{\tau_k^F} v_k$, т. е. в случайные моменты времени τ_k^F цены дериватива $\{F_{\tau_k^F}\}$ наблюдаются с мультипликативными ошибками $\{v_k\}$, которые при известной траектории Z независимы в совокупности. Распределение v_k также зависит от текущего состояния $Z_{\tau_{k-1}^F}$ и имеет известные условные моменты:

$$m_\ell^v \triangleq \mathbb{E} \left\{ \ln(v_k) | Z_{\tau_k^F} = e_\ell \right\},$$

$$D_\ell^v \triangleq \mathbb{E} \left\{ (\ln(v_k) - m_\ell^v)^2 | Z_{\tau_k^F} = e_\ell \right\}, \quad \ell = \overline{1, L}.$$

4. При известной траектории Z интервалы времени $\delta_k^F \triangleq \tau_k^F - \tau_{k-1}^F$ между наблюдениями цен дериватива независимы в совокупности. Распределение δ_k^F зависит от текущего состояния $Z_{\tau_{k-1}^F}$ и

$$m_\ell^F \triangleq \mathbb{E} \left\{ \delta_k^F | Z_{\tau_{k-1}^F} = e_\ell \right\},$$

$$D_\ell^F \triangleq \mathbb{E} \left\{ (\delta_k^F - m_\ell^F)^2 | Z_{\tau_{k-1}^F} = e_\ell \right\}, \quad \ell = \overline{1, L}.$$

5. При известной траектории Z последовательности $\{\delta_j^S\}_j$, $\{\delta_k^F\}_k$ и $\{v_k\}_k$ независимы в совокупности.
6. Среднее время между скачками считающих процессов $\{\tau_j^S\}_i$ и $\{\tau_k^F\}_i$ много меньше шага дискретизации Δ , который, в свою очередь, много меньше среднего времени между сменой состояний процесса Z :

$$\max_{1 \leq \ell_1, \ell_2 \leq L} (m_{\ell_1}^S, m_{\ell_2}^L) \ll \Delta \ll \min_{1 \leq \ell \leq L} |\Lambda_{\ell\ell}|^{-1}.$$

Предположения 1–5 гарантируют, что при условии $Z_t \equiv e_\ell$, $t \in [t_{i-1}, t_i]$, наблюдаемые МТП могут быть преобразованы к некоторым обобщенным процессам восстановления; функция преобразования определена ниже. Аргументация отсутствия

на практике непрерывных наблюдений цен бумаг, а также наличия шумов в наблюдениях деривативов представлена в [3]. Предположение 6 — ключевое для обеспечения корректного функционирования предлагаемого алгоритма фильтрации. С одной стороны, оно гарантирует малую вероятность того, что на отдельном интервале времени $[t_{i-1}, t_i]$ произойдет смена состояния Z . С другой стороны, интервала длины Δ должно хватить, чтобы на нем начала действовать асимптотика, определенная ЦПТ ОПВ [9].

На каждом отрезке дискретизации $[t_{i-1}, t_i]$ преобразуем имеющиеся наблюдения следующим образом:

$$W_i = [W_i^1, W_i^2, W_i^3, W_i^4] \triangleq$$

$$\triangleq \begin{bmatrix} \sum_{\substack{j: \tau_{j-1}^S, \tau_j^S \in \\ \in [t_{i-1}, t_i]}} 1 & \sum_{\substack{j: \tau_{j-1}^S, \tau_j^S \in \\ \in [t_{i-1}, t_i]}} \ln \frac{S_j}{S_{j-1}} & \sum_{\substack{k: \tau_{k-1}^F, \tau_k^F \in \\ \in [t_{i-1}, t_i]}} 1 & \sum_{\substack{k: \tau_k^F \in \\ \in [t_{i-1}, t_i]}} \ln F_k \end{bmatrix}.$$

Найдем асимптотику распределения вектора W_i при достаточно большом значении Δ для каждого из условий $Z_t \equiv e_\ell$, $t \in [t_{i-1}, t_i]$. Прежде всего при фиксированной траектории Z подвекторы $\text{col}(W_i^1, W_i^2)$ и $\text{col}(W_i^3, W_i^4)$ независимы в совокупности. Из правила Ито следует, что

$$\ln \frac{S_j}{S_{j-1}} = \left(a^\ell - \frac{B^\ell}{2} \right) \delta_j^S + \sqrt{B^\ell} \delta_j^S u_j,$$

где последовательность $\{u_j\}$ — стандартный гауссовский дискретный белый шум; $B^\ell \triangleq \sigma^\ell (\sigma^\ell)^\top$. Каждый вектор W_i — случайная сумма слагаемых вида

$$w_{ik} \triangleq$$

$$\triangleq \left[1 \quad \left(a^\ell - \frac{B^\ell}{2} \right) \delta_i^S + \sqrt{B^\ell} \delta_i^S u_j \quad 1 \quad \ln F_{\tau_k^F} + \ln(v_k) \right].$$

Для этих слагаемых выполнено равенство:

$$\mathbb{E} \left\{ w_{ik} - \begin{bmatrix} 0 & 0 & 0 & \ln F_{\tau_k^F} \end{bmatrix} \right\} =$$

$$= \begin{bmatrix} 1 & \left(a^\ell - \frac{B^\ell}{2} \right) m_\ell^S & 1 & m_\ell^v \end{bmatrix}.$$

Векторы

$$\bar{w}_{ik} \triangleq w_{ik} - \begin{bmatrix} 0 & 0 & 0 & \ln F_{\tau_k^F} \end{bmatrix} -$$

$$- \mathbb{E} \{ w_{ik} \}^\top \text{diag} \left(\frac{\delta_i^S}{m_\ell^S}, \frac{\delta_i^S}{m_\ell^S}, \frac{\delta_k^F}{m_\ell^F}, \frac{\delta_k^F}{m_\ell^F} \right)$$

имеют нулевое среднее и ковариационную матрицу

$$K_\ell^w \triangleq \text{cov}(\bar{w}_{ik}, \bar{w}_{ik}) = \text{diag} \left(\frac{D_\ell^S}{(m_\ell^S)^2}, B^\ell m_\ell^S, \frac{D_\ell^F}{(m_\ell^F)^2}, D_\ell^v + \frac{(m_\ell^v)^2}{m_\ell^F} D_\ell^F \right).$$

Согласно ЦПТ ОПВ [9] при $\Delta \rightarrow \infty$ имеет место следующая слабая сходимость:

$$\frac{1}{\sqrt{\Delta}} \left(\mathbf{w}_i - \begin{bmatrix} 0 & 0 & 0 & \sum_{\substack{k: \tau_k^F \in \\ \in [t_{i-1}, t_i]}} \ln F_{\tau_k^F} \end{bmatrix} - \mathbf{E} \{ \bar{w}_{ik} \}^\top \text{diag} \left(\frac{\Delta}{m_\ell^S}, \frac{\Delta}{m_\ell^S}, \frac{\Delta}{m_\ell^F}, \frac{\Delta}{m_\ell^F} \right) \right) \xrightarrow{\text{Law}} \mathcal{N} \left(\mathbf{0}, \text{diag} \left(\frac{1}{m_\ell^S}, \frac{1}{m_\ell^S}, \frac{1}{m_\ell^F}, \frac{1}{m_\ell^F} \right) K_\ell^w \right).$$

Таким образом, можно считать, что распределение семплированных наблюдений $(1/\sqrt{\Delta})\mathbf{w}_i$ близко к гауссовскому $\mathcal{N}(\mathbf{m}_{i,\ell}, \mathbf{K}_i)$ со средним

$$\mathbf{m}_{i,\ell} \triangleq \begin{bmatrix} \frac{\sqrt{\Delta}}{m_\ell^S} & \left(a^\ell - \frac{b^\ell}{2} \right) \sqrt{\Delta} & \frac{\sqrt{\Delta}}{m_\ell^F} \\ \frac{1}{\sqrt{\Delta}} \sum_{\substack{k: \tau_k^F \in \\ \in [t_{i-1}, t_i]}} \ln F_{\tau_k^F} + \frac{\sqrt{\Delta} m_\ell^v}{m_\ell^F} \end{bmatrix}$$

и ковариационной матрицей

$$\mathbf{K}_i = \text{diag} \left(\frac{D_\ell^S}{(m_\ell^S)^3}, B^\ell, \frac{D_\ell^F}{(m_\ell^F)^3}, \frac{D_\ell^v}{m_\ell^F} + \frac{(m_\ell^v)^2 D_\ell^F}{(m_\ell^F)^3} \right).$$

Предлагаемый субоптимальный алгоритм фильтрации имеет следующий вид.

Шаг 1. $i := 0$.

Шаг 2. Начальное условие: $\hat{Z}_0 = \pi_0^Z$.

Шаг 3. $i := i + 1$.

Шаг 4. Прогноз: $\bar{Z}_i = \exp \{ \Delta \Lambda^\top \} \hat{Z}_{i-1}$.

Шаг 5. Коррекция:

$$\hat{Z}_i^\ell = \frac{\bar{Z}_i^\ell \mathcal{G} \left(\left((1/\sqrt{\Delta}) \mathbf{w}_i, \mathbf{m}_{i,\ell}, \mathbf{K}_i \right) \right)}{\sum_{k=1}^L \bar{Z}_i^k \mathcal{G} \left(\left((1/\sqrt{\Delta}) \mathbf{w}_i, \mathbf{m}_{i,k}, \mathbf{K}_k \right) \right)}.$$

Шаг 6. Перейти к шагу 3.

4 Численный пример

Предлагаемый пример иллюстрирует влияние наблюдений цен дериватива в дополнение к имеющимся наблюдениям цены базового актива на качество оценивания скрытого процесса состояния рынка Z .

Параметры модели исследуемого рынка совпадают с рассмотренными в [3]. На отрезке времени $[0, 1]$ (1 год, 250 торговых дней по 8 ч каждый) моделируется поведение одного базового актива, стохастическая волатильность которого представляет собой МСП с четырьмя возможными состояниями: «рост – состояние, предшествующее панике, – паника – рецессия». Ставка по депозиту $r = 2\%$ годовых, на рынке присутствует один дериватив – call-опцион с ценой исполнения 1. Для моделирования (1) и (2) выбраны следующие параметры: $S_0 \equiv 1$; $a = (0,04; 0,035; 0,015; 0,02)$; $\sigma = (0,1; 0,12; 0,25; 0,15)$;

$$\Lambda = \begin{bmatrix} -12,5 & 12,5 & 0 & 0 \\ 100 & -500 & 400 & 0 \\ 0 & 0 & -125 & 125 \\ 40 & 0 & 10 & -50 \end{bmatrix}; \quad \pi_0^Z = \begin{bmatrix} 0,7273 \\ 0,0182 \\ 0,0727 \\ 0,1818 \end{bmatrix}.$$

Интегрирование СДС выполнялось методом Эйлера–Маруямы с шагом $\tau = 10^{-6}$.

В МТП $\{(\tau_j^S, S_j)\}$ интервалы между наблюдениями базового актива при фиксированной траектории Z представляют собой независимые экспоненциально распределенные случайные величины [12]. Их параметры зависят от текущего состояния Z_t и определяются вектором $\text{col}(150\,000, 130\,000, 110\,000, 140\,000)$.

В МТП $\{(\tau_j^F, F_j)\}$ интервалы между наблюдениями цен опциона также экспоненциально распределены с параметрами $\text{col}(60\,000, 57\,000, 51\,000, 54\,000)$. Шумы в наблюдениях имеют логнормальное распределение с векторами средних $(0; 0; 0; 0)$ и среднеквадратичных отклонений $(0,5; 0,51; 0,53; 0,52)$.

Состояние Z необходимо оценить с шагом $\Delta = 0,0001$, который соответствует 12 мин операционного времени.

Рисунок 1 представляет наблюдаемые МТП $\{(\tau_j^S, S_j)\}$ и $\{(\tau_j^F, F_j)\}$, полученные на отдельном интервале Δ .

Рисунок 2 содержит результаты фильтрации: покомпонентно представлено истинное значение состояния Z_t в сравнении с оценкой фильтрации \hat{Z}_t^S , построенной только по наблюдениям базового актива, и оценкой \hat{Z}_t^{SF} , построенной по наблюдениям базового актива и опциона.

По результатам расчетов можно сделать следующие заключения.

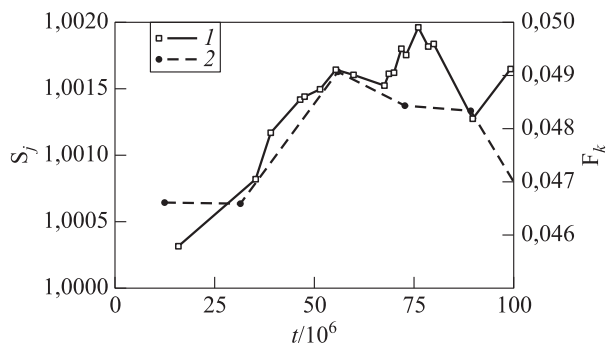


Рис. 1 Наблюдаемые МТП: 1 — $\{(\tau_j^S, S_j)\}$; 2 — $\{(\tau_j^F, F_j)\}$

1. Число наблюдений, попавших в интервал длиной Δ , достаточно для применения к преобразованным наблюдениям ЦПТ ОПВ.
2. Использование в процессе фильтрации только наблюдений базовых активов не позволяет определять второе состояние рынка («состояние, предшествующее панике») с приемлемой точностью в связи со слишком коротким временем пребывания в этом состоянии и малым

объемом статистической информации, получаемым за это время.

3. Использование в процессе фильтрации дополнительных зашумленных наблюдений цен опциона значительно повышает качество оценивания.
4. Различия в характеристиках распределений интервалов между поступлениями измерений в МТП, а также в характеристиках шумов наблюдений для различных значений рынка играют положительную роль.

5 Промежуточные выводы

В работе представлен субоптимальный алгоритм фильтрации скрытого состояния рынка по доступным наблюдениям рискованных ценных бумаг, поступающих в случайные моменты времени. Оценки состояния рынка однозначным образом позволяют оценить MPR. Численный эксперимент проиллюстрировал высокое качество полученных оценок.

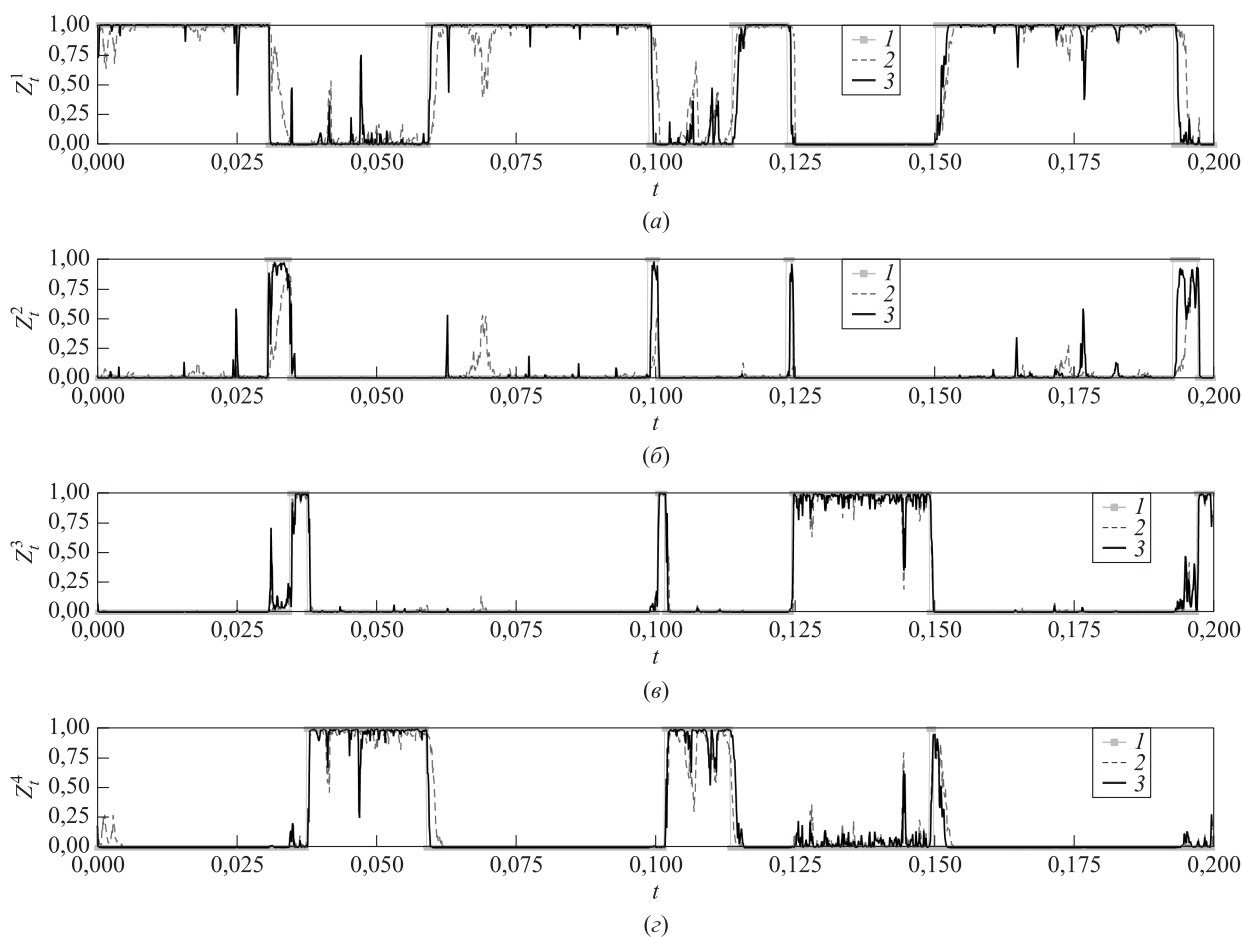


Рис. 2 Результаты фильтрации: 1 — Z_t ; 2 — \hat{Z}_t^S ; 3 — \hat{Z}_t^{SF}

В процедуре фильтрации использовались не только наблюдения цен бумаг, но и собственно поток моментов наблюдений. Ключевое свойство имеющихся наблюдений, позволяющее применить этот алгоритм, состоит в их высокой интенсивности.

Оценивание скрытого состояния рынка, как и оценивание MPR, — лишь один из возможных способов снижения статистической неопределенности данной модели рынка, связанной с его неполнотой. Другой подход заключается в пополнении рынка некоторым набором производных финансовых инструментов. Решению этой задачи посвящена финальная часть данного цикла заметок.

Литература

1. *Борисов А.* Рынок с марковской скачкообразной волатильностью I: мониторинг цены риска как задача оптимальной фильтрации // Информатика и её применения, 2023. Т. 17. Вып. 2. С. 27–33. doi: 10.14357/1992264230204. EDN: GAXCHQ.
2. *Борисов А.* Рынок с марковской скачкообразной волатильностью II: алгоритм вычисления справедливой цены деривативов // Информатика и её применения, 2023. Т. 17. Вып. 3. С. 18–24. doi: 10.14357/1992264230303. EDN: DNXXJB.
3. *Борисов А.* Рынок с марковской скачкообразной волатильностью III: алгоритм мониторинга цены риска по дискретным наблюдениям цен активов // Информатика и её применения, 2023. Т. 17. Вып. 4. С. 9–16. doi: 10.14357/1992264230402. EDN: OFYELT.
4. *Andersen T. G., Benzoni L.* Stochastic volatility // Complex systems in finance and econometrics / Ed. R. A. Meyers. — New York, NY, USA: Springer, 2009. Vol. 1. P. 694–726. doi: 10.1007/978-1-4419-7701-4_38.
5. *Mamon R. S., Rodrigo M. R.* Explicit solutions to European options in a regime-switching economy // Oper. Res. Lett., 2005. Vol. 33. P. 581–586. doi: 10.1016/j.orl.2004.12.003.
6. *Boyle P., Draviam T.* Pricing exotic options under regime switching // Insur. Math. Econ., 2007. Vol. 40. P. 267–282. doi: 10.1016/j.insmatheco.2006.05.001.
7. *Mamon R., Elliott R.* Hidden Markov models in finance: Further developments and applications. — New York, NY, USA: Springer, 2014. Vol. II. 283 p.
8. *Cvitanic J., Liptser R., Rozovskii B.* A filtering approach to tracking volatility from prices observed at random times // Ann. Appl. Probab., 2006. Vol. 16. P. 1633–1652. doi: 10.1214/105051606000000222.
9. *Smith W.* Regenerative stochastic processes // P. Roy. Soc. A — Math. Phys., 1955. Vol. 232. P. 6–31. doi: 10.1098/rspa.1955.0198.
10. *Jacod J., Shiryaev A.* Limit theorems for stochastic processes. — Berlin: Springer, 2010. 684 p.
11. *Борисов А., Миллер Б., Семенихин К.* Фильтрация марковского скачкообразного процесса по наблюдениям мультивариантного точечного процесса // Автоматика и телемеханика, 2015. Вып. 2. С. 34–60.
12. *Scalas E., Gorenflo R., Lueckock H., Mainardi F., Mantelli M., Raberto M.* Anomalous waiting times in high-frequency financial data // Quant. Financ., 2004. Vol. 4. P. 695–702. doi: 10.1080/14697680500040413.

Поступила в редакцию 05.01.24

MARKET WITH MARKOV JUMP VOLATILITY IV: PRICE OF RISK MONITORING ALGORITHM GIVEN HIGH FREQUENCY OBSERVATION FLOWS OF ASSETS PRICES

A. V. Borisov

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The fourth part of the series presents a suboptimal algorithm of the market price of risk monitoring given the observations of the underlying and derivative asset prices. As in the previous papers, the market model contains the stochastic volatility described by a hidden Markov jump process. This market has no arbitrage; so, the market price of risk is a function of the Markov process state. The key feature of the investigated market lies in the structure of the available observations. They represent the underlying and derivative prices registered at random instants. The underlying prices are observed accurately, while the derivative prices are corrupted by a random noise. The distribution of the interarrival times between the observable prices and the observation noises depends on the estimated process. The essential feature of the obtained observations is their high arrival intensity compared with the hidden process transition rate. This property allows one to use the central limit theorem for

generalized regenerative processes for the filter design. The influence of the estimation performance depending on the observation complexes is illustrated with a numerical example.

Keywords: market price of risk; Markov jump process; high frequency observations; multivariate point process; numerical algorithm

DOI: 10.14357/19922264240104

EDN: ZRQKIT

Acknowledgments

The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (CKP “Informatics”) of FRC CSC RAS (Moscow).

References

1. Borisov, A. 2023. Rynok s markovskoy skachkoobraznoy volatil'nost'yu I: monitoring tseny riska kak zadacha optimal'noy fil'tratsii [Market with Markov jump volatility I: Price of risk monitoring as an optimal filtering problem]. *Informatika i ee primeneniya — Inform. Appl.* 17(2):27–33. doi: 10.14357/19922264230204. EDN: GAXCHQ.
2. Borisov, A. 2023. Rynok s markovskoy skachkoobraznoy volatil'nost'yu II: algoritm vychisleniya spravedlivoy tseny derivativov [Market with Markov jump volatility II: Algorithm of derivative fair price calculation]. *Informatika i ee primeneniya — Inform. Appl.* 17(3):18–24. doi: 10.14357/19922264230303. EDN: DNXJGB.
3. Borisov, A. 2023. Rynok s markovskoy skachkoobraznoy volatil'nost'yu III: algoritm monitoringa tseny riska po diskretnym nablyudeniya tsen aktivov [Market with Markov jump volatility III: Price of risk monitoring algorithm given discrete-time observations of asset prices]. *Informatika i ee primeneniya — Inform. Appl.* 17(4):9–16. doi: 10.14357/19922264230402. EDN: OFYELT.
4. Andersen, T. G., and L. Benzoni. 2009. Stochastic volatility. *Complex systems in finance and econometrics*. Ed. R. A. Meyers. New York, NY: Springer. 1:694–726. doi: 10.1007/978-1-4419-7701-4_38.
5. Mamon, R. S., and M. R. Rodrigo. 2005. Explicit solutions to European options in a regime-switching economy. *Oper. Res. Lett.* 33(6):581–586. doi: 10.1016/j.orl.2004.12.003.
6. Boyle, P., and T. Draviam. 2007. Pricing exotic options under regime switching. *Insur. Math. Econ.* 40(2):267–282. doi: 10.1016/j.insmatheco.2006.05.001.
7. Mamon, R. S., and R. Elliott. 2014. *Hidden Markov models in finance: Further developments and applications*. New York, NY: Springer. Vol. II. 283 p.
8. Cvitanić, J., R. Liptser, and B. Rozovskii. 2006. A filtering approach to tracking volatility from prices observed at random times. *Ann. Appl. Probab.* 16(3):1633–1652. doi: 10.1214/105051606000000222.
9. Smith, W. 1955. Regenerative stochastic processes. *P. Roy. Soc. A — Math. Phys.* 232(1188):6–31. doi: 10.1098/rspa.1955.0198.
10. Jacod, J., and A. N. Shiryaev. 2010. *Limit theorems for stochastic processes*. Berlin: Springer. 664 p.
11. Borisov, A., B. Miller, and K. Semenikhin. 2015. Filtering of the Markov jump process given the observations of multivariate point process. *Automat. Rem. Contr.* 76(2):219–240. doi: 10.1134/S0005117915020034.
12. Scalas, E., R. Gorenflo, H. Luckock, F. Mainardi, M. Mantelli, and M. Raberto. 2004. Anomalous waiting times in high-frequency financial data. *Quant. Financ.* 4:695–702. doi: 10.1080/14697680500040413.

Received January 5, 2024

Contributor

Borisov Andrey V. (b. 1965) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; aborisov@frcsc.ru

РАВНОМЕРНЫЕ ОЦЕНКИ СКОРОСТИ СХОДИМОСТИ ДЛЯ ИНТЕГРАЛЬНОГО ИНДЕКСА БАЛАНСА*

А. А. Кудрявцев¹, О. В. Шестаков²

Аннотация: Рассматривается байесовская модель баланса, в рамках которой исследуется скорость слабой сходимости нормированного интегрального индекса баланса к дигамма-распределению в терминах равномерной метрики. Рассматриваются интегральные факторы, негативно и позитивно влияющие на функционирование системы, и их отношение — интегральный индекс баланса системы. Предполагается, что число факторов априори не известно и описывается смешанным пуассоновским распределением со структурным обобщенным гамма-распределением. Исследуется скорость слабой сходимости в описанной схеме к дигамма-распределению. В качестве вспомогательного приводится утверждение о скорости слабой сходимости нормированной случайной суммы с индексом, имеющим обобщенное отрицательное биномиальное распределение, к предельному обобщенному гамма-распределению. Результаты работы могут быть востребованы при исследовании моделей, использующих для описания процессов распределения с неограниченным неотрицательным носителем.

Ключевые слова: дигамма-распределение; смешанные распределения; индекс баланса; случайное суммирование; слабая сходимость; оценки скорости сходимости

DOI: 10.14357/19922264240105

EDN: WNUUHY

1 Введение

Большинство аспектов современной жизни усложнилось настолько, что определение критериев эффективности путем детерминированного анализа стало практически невозможным. По этой причине активно применяются разного рода показатели, индексы и рейтинги, дающие возможность оперативного принятия решений.

В основе построения рейтингов и индексов обычно лежит условное разделение факторов, влияющих на эффективность функционирования системы, на два класса. Первый класс включает параметры, способствующие функционированию целевого объекта и позитивно влияющие на исследуемый процесс (*p-факторы*); второй класс включает параметры, препятствующие и негативно влияющие (*n-факторы*).

Вполне естественно, что функционирование исследуемой системы в итоге зависит не столько от значений *n*- и *p*-факторов, сколько от их соотношения. При этом большое расхождение между величинами факторов обычно свидетельствует либо о чрезмерных затратах на борьбу с негативным влиянием, либо о недооценке негативного воздействия. Таким образом, для того чтобы система была

сбалансированной, имеет смысл стремиться уравновесить *n*- и *p*-факторы. Для этой цели было предложено использовать и исследовать свойства *индекса баланса* — отношения *n*-фактора к *p*-фактору [1]. В [1, 2] приводится ряд примеров применения моделей баланса в политологии, науке о безопасности жизнедеятельности, медицине, экономике, демографии, физике, теории массового обслуживания и надежности.

С течением времени *n*- и *p*-факторы, а следовательно, и индексы баланса претерпевают изменения. Это связано с неустойчивостью среды, в которой происходит функционирование, — изменяется экономическая ситуация, политическая система, технологии производства, пристрастия населения и т. д. Такие изменения зачастую невозможно предугадать. Измерения при помощи приборов и методик нельзя считать точными, ввиду неизбежных вносимых погрешностей, имеющих в большинстве случаев изменчивый (случайный) характер. Это дает предпосылки для рассмотрения факторов и индексов как случайных величин. При этом стоит учитывать, что глобальные изменения окружающей среды происходят достаточно редко, поэтому законы, влияющие на значения факторов, можно считать (в рамках конкретной модели) не-

* Работа выполнена при финансовой поддержке РФФИ (проект № 22–11–00212).

¹ Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Московский центр фундаментальной и прикладной математики, aakudryavtsev@cs.msu.ru

² Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Московский центр фундаментальной и прикладной математики; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, oshestakov@cs.msu.ru

изменными. Из этого следует, что распределения рассматриваемых случайных величин следует полагать заданными априори. Приведенные рассуждения обуславливают применение к моделям баланса байесовского метода, а исследование рандомизированных индексов баланса, таким образом, имеет смысл осуществлять при помощи метода анализа масштабных смесей априорных распределений негативных и позитивных факторов. Кроме того, точное число факторов, оказывающих влияние на функционирование системы, невозможно определить априори. Это приводит к естественному обобщению [3] байесовских моделей баланса, заключающемуся в использовании интегральных характеристик, основанных на случайном суммировании.

В статье обсуждается асимптотика интегрального индекса баланса и приводятся равномерные оценки скорости его сходимости к предельному распределению.

2 Математическое описание моделей баланса

Рассмотрим некоторую сложную систему, на функционирование которой влияют негативные n -факторы ξ_1, ξ_2, \dots и позитивные p -факторы η_1, η_2, \dots . Поскольку априори невозможно предугадать число влияющих на систему в данный момент времени t негативных и позитивных факторов, имеет смысл говорить о том, что их случайное число, которое будем соответственно обозначать $N_\xi(t)$ и $N_\eta(t)$. Таким образом, интегральные величины n - и p -факторов системы представляют собой случайные суммы, а интегральный индекс баланса системы — их отношение.

При моделировании наступления случайного числа событий и исследовании априори неизвестного числа наблюдений широко используются смешанные пуассоновские процессы и процессы Кокса с соответствующими непрерывными структурными распределениями, имеющими смысл соответственно рандомизированных интенсивностей и накопленных интенсивностей. Для проекции процесса Кокса $N(t) = N_1(\Lambda(t))$ в точке $t > 0$ справедливо

$$P(N(t) = n) = P(N_1(\Lambda(t)) = n) = \int_0^\infty e^{-y} \frac{y^n}{n!} dP(\Lambda(t) < y), \quad n = 0, 1, \dots$$

В случае когда проекция независимой от стандартного пуассоновского процесса $N_1(t)$ случайной

меры $\Lambda(t)$ имеет обобщенное гамма-распределение [4, 5] $GG(\nu, p, \delta(t))$ с функцией распределения

$$G_{\nu, p, \delta(t)}(x) = \int_0^x \frac{|\nu| y^{\nu p - 1} e^{-(y/\delta(t))^\nu}}{(\delta(t))^{\nu p} \Gamma(p)} dy, \\ \nu \neq 0, \quad p > 0, \quad \delta(t) > 0, \quad t > 0, \quad x > 0,$$

говорят, что проекции процесса Кокса $N(t)$ имеют обобщенное отрицательное биномиальное распределение [6]. Такие распределения широко применяются в страховании, финансовой математике, физике и других областях [7–12].

Предположим, что n -факторы ξ_1, ξ_2, \dots и p -факторы η_1, η_2, \dots суть неотрицательные и одинаково распределенные внутри каждой последовательности случайные величины, имеющие конечные положительные математические ожидания $E\xi_1 = a$ и $E\eta_1 = b$. Пусть $N_{1,1}(t)$ и $N_{1,2}(t)$ — стандартные пуассоновские процессы и $N_{1,1}(t), N_{1,2}(t), \Lambda_1(t), \Lambda_2(t), \xi_1, \xi_2, \dots, \eta_1, \eta_2, \dots$ независимы, а $\Lambda_1(t)$ и $\Lambda_2(t)$ имеют соответственно обобщенные гамма-распределения $GG(\nu_1, p, \delta_1(t))$ и $GG(\nu_2, q, \delta_2(t))$. Предположим, что число негативных факторов $N_\xi(t)$ и число позитивных факторов $N_\eta(t)$ случайны и имеют вид:

$$N_\xi(t) = N_{1,1}(\Lambda_1(t)) + 1; \quad N_\eta(t) = N_{1,2}(\Lambda_2(t)) + 1,$$

т.е. в каждой группе присутствует хотя бы один фактор.

Определим интегральные n - и p -факторы и интегральный индекс баланса:

$$\xi(t) = \sum_{k=1}^{N_\xi(t)} \xi_k; \quad \eta(t) = \sum_{k=1}^{N_\eta(t)} \eta_k; \quad \rho(t) = \frac{\xi(t)}{\eta(t)}.$$

3 Вспомогательные утверждения

Поскольку на систему влияет случайное число случайных факторов, вопрос о сбалансированности ее функционирования следует рассматривать при помощи исследования распределения интегрального индекса баланса. Классический подход теории вероятностей при исследовании такого рода объектов основан на предельных теоремах и фактической замене анализа исходного распределения на анализ предельного.

Основные результаты работы базируются на приведенном в данном разделе аналоге результата, доказанного в [3], для формулирования которого понадобится следующее распределение [13].

Определение 1. Будем говорить, что случайная величина $\hat{\rho}$ имеет дигамма-распределение

DiG (r, ν, p, q, δ) с характеристическим показателем $r \in \mathbb{R}$ и параметрами формы $\nu \neq 0$, концентрации $p, q > 0$ и масштаба $\delta > 0$, если ее преобразование Меллина задается соотношением:

$$\mathcal{M}_{\hat{\rho}}(z) = \frac{\delta^z \Gamma(p + z/\nu) \Gamma(q - rz/\nu)}{\Gamma(p)\Gamma(q)},$$

$$p + \frac{\operatorname{Re}(z)}{\nu} > 0, \quad q - \frac{r\operatorname{Re}(z)}{\nu} > 0. \quad (1)$$

Замечание 1. Случайная величина $\hat{\rho}$, имеющая распределение (1), представима в виде масштабной смеси обобщенных гамма-законов GG (ν, p, δ) и GG ($\nu/r, q, 1$):

$$\hat{\rho} \stackrel{d}{=} \delta \left(\frac{\lambda}{\mu^r} \right)^{1/\nu},$$

где независимые случайные величины λ и μ соответственно имеют гамма-распределения GG ($1, p, 1$) и GG ($1, q, 1$).

Аналогично работе [3] можно показать справедливость следующего утверждения.

Лемма 1. Пусть $\delta_i(t) \rightarrow \infty$ при $t \rightarrow \infty, i = 1, 2$. При введенных в разд. 2 предположениях

$$\frac{\rho(t)}{\delta_1(t)/\delta_2(t)} \Rightarrow \hat{\rho}, \quad t \rightarrow \infty,$$

где $\hat{\rho} \sim \text{DiG}(\nu_1/\nu_2, \nu_1, p, q, a/b)$.

Поскольку интегральный индекс баланса, асимптотическое поведение которого описано в лемме 1, представляет собой отношение двух независимых случайных сумм, вопрос об оценивании скорости его сходимости к предельному закону можно свести к вопросу о скорости сходимости числителя и знаменателя к соответствующим предельным распределениям. Для получения необходимых оценок понадобится следующее утверждение из [14].

Лемма 2. Если $\nu \in (0, 1], p \in (0, 1)$, то обобщенное гамма-распределение — это смешанное показательное распределение:

$$\mathcal{G}_{\nu, p, \delta}(x) = \int_0^x \int_0^1 \frac{y}{1-y} e^{-yz/(1-y)} h_{\nu, p, \delta}(y) dy dz, \quad x > 0,$$

где

$$h_{\nu, p, \delta}(y) = \frac{\delta^2}{\Gamma(1-p)\Gamma(p)} \frac{1}{(1-y)^2} \times \int_1^\infty \frac{f_{\nu, 1}(\delta^\nu y(1-y)^{-1} x^{-1/\nu}) dx}{(x-1)^p x^{1+2/\nu}}, \quad 0 < y < 1; \quad (2)$$

$f_{\nu, 1}(x), 0 < \nu < 1$, — плотность одностороннего строго устойчивого закона, сосредоточенного на положительной полуоси, с характеристической функцией

$$\phi_{\nu, 1}(t) = \exp \left\{ -|t|^\nu \exp \left\{ -\frac{1}{2} i \pi \nu \operatorname{sgn} t \right\} \right\}.$$

Воспользуемся леммой 2 для доказательства аналога теоремы 1 из [14].

Лемма 3. Предположим, что $\nu_1 \in (0, 1], p \in (0, 1)$. Тогда для интегрального n -фактора $\xi(t)$ имеет место оценка

$$\Delta_\xi \equiv \sup_{x>0} \left| \mathbb{P} \left(\frac{\xi(t)}{\delta_1(t)} < x \right) - \mathcal{G}_{\nu_1, p, a}(x) \right| \leq \int_0^1 \min \left\{ 1, \frac{y}{1-y} \frac{\mathbb{E} \xi_1^2}{a^2} \max \left\{ 1-y, \frac{1}{2} \right\} \right\} \times h_{\nu_1, p, \delta_1(t)}(y) dy,$$

где плотность $h_{\nu, p, \delta}(y)$ определена в (2).

Доказательство. Имеем

$$\begin{aligned} \Delta_\xi &\equiv \sup_{x>0} \left| \mathbb{P} \left(\frac{\xi(t)}{\delta_1(t)} < x \right) - \mathcal{G}_{\nu_1, p, a}(x) \right| = \\ &= \sup_{x>0} \left| \mathbb{P} \left(\frac{\xi(t)}{\delta_1(t)} < ax \right) - \mathcal{G}_{\nu_1, p, a}(ax) \right| = \\ &= \sup_{x>0} \left| \mathbb{P}(\xi(t) < x) - \mathcal{G}_{\nu_1, p, \delta_1(t)} \left(\frac{x}{a} \right) \right|. \end{aligned}$$

Рассмотрим геометрическую сумму

$$S_{p, 0} = \sum_{k=1}^{N_{p, 0}} \xi_k,$$

где

$$\mathbb{P}(N_{p, 0} = n) = p(1-p)^n, \quad n = 0, 1, \dots, \quad p \in (0, 1).$$

В работе [15] было показано, что

$$\mathbb{P}(\xi(t) < x) = \int_0^1 \mathbb{P}(S_{y, 0} < x) h_{\nu_1, p, \delta_1(t)}(y) dy,$$

поэтому по лемме 2

$$\begin{aligned} \Delta_\xi &= \sup_{x>0} \left| \int_0^1 \mathbb{P}(S_{y, 0} < x) h_{\nu_1, p, \delta_1(t)}(y) dy - \right. \\ &\quad \left. - \int_0^{x/a} \int_0^1 \frac{y}{1-y} e^{-yz/(1-y)} h_{\nu_1, p, \delta_1(t)}(y) dy dz \right| \leq \end{aligned}$$

$$\begin{aligned} &\leq \int_0^1 \sup_{x>0} \left| \mathbb{P}(S_{y,0} < x) - \mathcal{G}_{1,1,(1-y)/y} \left(\frac{x}{a} \right) \right| \times \\ &\quad \times h_{\nu_1,p,\delta_1(t)}(y) dy = \\ &= \int_0^1 \sup_{x>0} \left| \mathbb{P} \left(\frac{yS_{y,0}}{(1-y)a} < x \right) - \right. \\ &\quad \left. - \mathcal{G}_{1,1,(1-y)/y} \left(\frac{(1-y)x}{y} \right) \right| h_{\nu_1,p,\delta_1(t)}(y) dy = \\ &= \int_0^1 \sup_{x>0} \left| \mathbb{P} \left(\frac{yS_{y,0}}{(1-y)a} < x \right) - \mathcal{G}_{1,1,1}(x) \right| \times \\ &\quad \times h_{\nu_1,p,\delta_1(t)}(y) dy. \end{aligned}$$

Используя оценку из [14], получаем

$$\begin{aligned} \Delta_\xi &\leq \int_0^1 \min \left\{ 1, y \frac{EX_1^2}{a^2} \max \left\{ 1, \frac{1}{2(1-y)} \right\} \right\} \times \\ &\quad \times h_{\nu_1,p,\delta_1(t)}(y) dy = \\ &= \int_0^1 \min \left\{ 1, \frac{y}{1-y} \frac{EX_1^2}{a^2} \max \left\{ 1-y, \frac{1}{2} \right\} \right\} \times \\ &\quad \times h_{\nu_1,p,\delta_1(t)}(y) dy. \end{aligned}$$

Лемма доказана.

Замечание 2. Поскольку интегральный р-фактор $\eta(t)$ имеет ту же структуру, что и $\xi(t)$, справедлива аналогичная оценка

$$\begin{aligned} \Delta_\eta &\equiv \sup_{x>0} \left| \mathbb{P} \left(\frac{\eta(t)}{\delta_2(t)} < x \right) - \mathcal{G}_{\nu_2,q,b}(x) \right| \leq \\ &\leq \int_0^1 \min \left\{ 1, \frac{y}{1-y} \frac{E\eta_1^2}{b^2} \max \left\{ 1-y, \frac{1}{2} \right\} \right\} \times \\ &\quad \times h_{\nu_2,q,\delta_2(t)}(y) dy. \end{aligned}$$

4 Оценивание скорости сходимости нормированного интегрального индекса баланса к дигамма-распределению

Рассмотрим оценки скорости сходимости в лемме 1 в терминах равномерной метрики. Справедливо следующее утверждение.

Теорема 1. *Предположим, что $\nu_1, \nu_2 \in (0, 1]$, $p, q \in (0, 1)$. Пусть интегральные n - и p -факторы удовлетворяют условиям разд. 2. Тогда*

$$\begin{aligned} \Delta &\equiv \sup_{x>0} \left| \mathbb{P} \left(\frac{\rho(t)}{\delta_1(t)/\delta_2(t)} < x \right) - \mathbb{P}(\hat{\rho} < x) \right| \leq \\ &\leq \int_0^1 \min \left\{ 1, \frac{y}{1-y} \frac{E\xi_1^2}{a^2} \max \left\{ 1-y, \frac{1}{2} \right\} \right\} \times \\ &\quad \times h_{\nu_1,p,\delta_1(t)}(y) dy + \\ &\quad + \int_0^1 \min \left\{ 1, \frac{y}{1-y} \frac{E\eta_1^2}{b^2} \max \left\{ 1-y, \frac{1}{2} \right\} \right\} \times \\ &\quad \times h_{\nu_2,q,\delta_2(t)}(y) dy. \quad (3) \end{aligned}$$

Доказательство. Согласно замечанию 1 предельная случайная величина $\hat{\rho}$ в лемме 1 может быть представлена в виде

$$\hat{\rho} \stackrel{d}{=} \frac{\hat{\xi}}{\hat{\eta}},$$

где случайные величины $\hat{\xi}$ и $\hat{\eta}$ соответственно имеют распределения GG(ν_1, p, a) и GG(ν_2, q, b).

Рассмотрим равномерное расстояние между допредельными и предельными величинами. Имеем

$$\begin{aligned} \Delta &= \sup_{x>0} \left| \mathbb{P} \left(\frac{\rho(t)}{\delta_1(t)/\delta_2(t)} < x \right) - \mathbb{P} \left(\frac{\hat{\xi}}{\hat{\eta}} < x \right) \right| = \\ &= \sup_{x>0} \left| \mathbb{P} \left(\ln \frac{\rho(t)}{\delta_1(t)/\delta_2(t)} < \ln x \right) - \right. \\ &\quad \left. - \mathbb{P} \left(\ln \frac{\hat{\xi}}{\hat{\eta}} < \ln x \right) \right| = \\ &= \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\ln \frac{\xi(t)}{\delta_1(t)} - \ln \frac{\eta(t)}{\delta_2(t)} < x \right) - \right. \\ &\quad \left. - \mathbb{P}(\ln \hat{\xi} - \ln \hat{\eta} < x) \right|. \end{aligned}$$

Поскольку равномерная метрика обладает свойством регулярности [16], имеем

$$\begin{aligned} \Delta &\leq \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\ln \frac{\xi(t)}{\delta_1(t)} < x \right) - \mathbb{P}(\ln \hat{\xi} < x) \right| + \\ &\quad + \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\ln \frac{\eta(t)}{\delta_2(t)} < x \right) - \mathbb{P}(\ln \hat{\eta} < x) \right| = \\ &= \sup_{x>0} \left| \mathbb{P} \left(\frac{\xi(t)}{\delta_1(t)} < x \right) - \mathbb{P}(\hat{\xi} < x) \right| + \\ &\quad + \sup_{x>0} \left| \mathbb{P} \left(\frac{\eta(t)}{\delta_2(t)} < x \right) - \mathbb{P}(\hat{\eta} < x) \right| = \Delta_\xi + \Delta_\eta. \end{aligned}$$

Таким образом, утверждение теоремы следует из леммы 3 и замечания 2. Теорема доказана.

В качестве примера применения оценки (3) рассмотрим случай, когда $a = b = 1$ и проекции управляющих процессов Кокса, отвечающих за число негативных и позитивных факторов, имеют гамма-распределения с параметрами масштаба

$$\delta_i(t) = \frac{1 - \theta_i(t)}{\theta_i(t)},$$

где $\theta_i(t) \rightarrow 0$ при $t \rightarrow \infty$, $i = 1, 2$. Для этого воспользуемся соответствующим результатом из работы [14]:

$$\begin{aligned} \sup_{x>0} \left| P(\xi(t) < x) - \mathcal{G}_{1,p,(1-\theta_1(t))/\theta_1(t)}(x) \right| &\leq \\ &\leq \frac{(\theta_1(t))^p (2E\xi_1^2 + p^{-1} + 1)}{\Gamma(p)\Gamma(2-p)}. \end{aligned}$$

Из последнего неравенства получаем

$$\begin{aligned} \sup_{x>0} \left| P\left(\frac{\xi(t)}{\delta_1(t)} < x\right) - \mathcal{G}_{1,p,1}(x) \right| &\leq \\ &\leq \frac{(\theta_1(t))^p (2E\xi_1^2 + p^{-1} + 1)}{\Gamma(p)\Gamma(2-p)}. \end{aligned}$$

Поскольку интегральный p -фактор $\eta(t)$ имеет структуру, аналогичную $\xi(t)$, окончательно для интегрального индекса баланса получаем оценку:

$$\begin{aligned} \sup_{x>0} \left| P\left(\frac{\rho(t)}{\delta_1(t)/\delta_2(t)} < x\right) - P(\hat{\rho} < x) \right| &\leq \\ &\leq \frac{(\theta_1(t))^p (2E\xi_1^2 + p^{-1} + 1)}{\Gamma(p)\Gamma(2-p)} + \\ &+ \frac{(\theta_2(t))^q (2E\eta_1^2 + q^{-1} + 1)}{\Gamma(q)\Gamma(2-q)}. \end{aligned}$$

5 Заключение

В работе доказана теорема о скорости слабой сходимости нормированного интегрального индекса баланса к дигамма-распределению. Получена верхняя оценка равномерного расстояния между функциями распределения допредельного и предельного законов. Предполагается, что интегральный индекс баланса представляет собой отношение двух случайных сумм, индексы которых описываются обобщенными отрицательными биномиальными распределениями. Результаты работы могут найти применение при исследовании масштабных смесей в схеме случайного суммирования, в частности при анализе байесовских моделей баланса.

Литература

1. Кудрявцев А. А. Байесовские модели баланса // Информатика и её применения, 2018. Т. 12. Вып. 3. С. 18–27. doi: 10.14357/19922264180303. EDN: YAMDFH.
2. Kudryavtsev A. A., Shestakov O. V. Asymptotically normal estimators for the parameters of the gamma-exponential distribution // Mathematics, 2021. Vol. 9. Iss. 3. Art. 273. 13 p. doi: 10.3390/math9030273.
3. Кудрявцев А. А., Шестаков О. В. Дигамма-распределение как предельное для интегрального индекса баланса // Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн., 2022. № 3. С. 26–32.
4. Крицкий С. Н., Менкель М. Ф. О приемах исследования случайных колебаний речного стока // Труды НИУ ГУГМС. Сер. IV, 1946. Вып. 29. С. 3–32.
5. Крицкий С. Н., Менкель М. Ф. Выбор кривых распределения вероятностей для расчетов речного стока // Известия АН СССР. Отд. техн. наук, 1948. № 6. С. 15–21.
6. Кудрявцев А. А. О представлениях гамма-экспоненциального и обобщенного отрицательного биномиального распределений // Информатика и её применения, 2019. Т. 13. Вып. 4. С. 76–80. doi: 10.14357/19922264190412. EDN: YTZLNY.
7. Korolev V. Yu., Zeifman A. I. Generalized negative binomial distributions as mixed geometric laws and related limit theorems // Lith. Math. J., 2019. Vol. 59. Iss. 3. P. 366–388. 10.1007/s10986-019-09452-x.
8. Wang X., Zhao X., Sun J. A compound negative binomial distribution with mutative termination conditions based on a change point // J. Comput. Appl. Math., 2019. Vol. 351. P. 237–249. doi: 10.1016/j.cam.2018.11.009.
9. Korolev V. Yu., Gorshenin A. K. Probability models and statistical tests for extreme precipitation based on generalized negative binomial distributions // Mathematics, 2020. Vol. 8. Iss. 4. Art. 604. 30 p. doi: 10.3390/math8040604.
10. Zhang J., Wang D., Yang K. A study of RCINAR(1) process with generalized negative binomial marginals // Commun. Stat. B — Simul., 2020. Vol. 49. Iss. 6. P. 1487–1510. doi: 10.1080/03610918.2018.1498891.
11. Bhati D., Ahmed I. S. On uniform-negative binomial distribution including Gauss hypergeometric function and its application in count regression modeling // Commun. Stat. — Theor. M., 2021. Vol. 50. Iss. 13. P. 3106–3122. doi: 10.1080/03610926.2019.1682163.
12. Mangiola S., Thomas E. A., Modrák M., Vehtari A., Papenfuss A. T. Probabilistic outlier identification for RNA sequencing generalized linear models // NAR Genomics Bioinformatics, 2021. Vol. 3. Iss. 1. Art. lqab005. 9 p. doi: 10.1093/nargab/lqab005.
13. Кудрявцев А. А., Недоливко Ю. Н., Шестаков О. В. Оновые вероятностные характеристики дигамма-распределения и метод оценивания его параметров // Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн., 2022. № 3. С. 22–29.
14. Shevtsova I., Tselishchev M. On the accuracy of the generalized gamma approximation to generalized negative binomial random sums // Mathematics, 2021. Vol. 9. Iss. 13. Art. 1571. 8 p. doi: 10.3390/math9131571.
15. Kudryavtsev A. A., Shestakov O. V. Estimates of the convergence rate in the generalized Renyi theorem with a structural digamma distribution using zeta metrics // Mathematics, 2023. Vol. 11. Iss. 21. Art. 4477. 10 p. doi: 10.3390/math11214477.
16. Senatov V. V. Normal approximation: New results, methods, and problems. — Utrecht, The Netherland: VSP, 1998. 363 p.

Поступила в редакцию 15.01.24

UNIFORM CONVERGENCE RATE ESTIMATES FOR THE INTEGRAL BALANCE INDEX

A. A. Kudryavtsev^{1,2} and O. V. Shestakov^{1,2,3}

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

²Moscow Center for Fundamental and Applied Mathematics, M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

³Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The paper considers the Bayesian balance model in which the rate of weak convergence of the normalized integral balance index to the digamma distribution is studied in terms of a uniform metric. The integral factors negatively and positively influencing the functioning of the system and their ratio, the integral balance index of the system, are considered. It is assumed that the number of factors is not known *a priori* and is described by a mixed Poisson distribution with a structural generalized gamma distribution. The rate of weak convergence to the digamma distribution in the described scheme is studied. As an auxiliary statement, the rate of weak convergence of a normalized random sum with an index having a generalized negative binomial distribution to a limiting generalized gamma distribution is estimated. The results of the work may be in demand in the study of models used to describe processes with distributions having an unlimited nonnegative support.

Keywords: digamma distribution; mixed distributions; balance index; random summation; weak convergence; estimates of convergence rate

DOI: 10.14357/19922264240105

EDN: WNUUHY

Acknowledgments

The research was supported by the Russian Science Foundation, project No. 22-11-00212.

References

1. Kudryavtsev, A. A. 2018. Bayesovskie modeli balansa [Bayesian balance models]. *Informatika i ee Primeneniya — Inform. Appl.* 12(3):18–27. doi: 10.14357/19922264180303. EDN: YAMDFH.
2. Kudryavtsev, A. A., and O. V. Shestakov. 2021. Asymptotically normal estimators for the parameters of the gamma-exponential distribution. *Mathematics* 9(3):273. 13 p. doi: 10.3390/math9030273.
3. Kudryavtsev, A. A., and O. V. Shestakov. 2022. Digamma distribution as a limit for the integral balance index. *Moscow University Computational Mathematics Cybernetics* 46:133–139. doi: 10.3103/S0278641922030062.
4. Kritskiy, S. N., and M. F. Menkel'. 1946. O priemakh issledovaniya sluchaynykh kolebaniy rechnogo stoka [Methods of investigation of random fluctuations of river flow]. *Trudy NIU GUGMS Ser. IV* [Proceedings of GUGMS Research Institutions Ser. IV] 29:3–32.
5. Kritsky, S. N., and M. F. Menkel'. 1948. Vybor krivyykh raspredeleniya veroyatnostey dlya raschetov rechnogo stoka [Selection of probability distribution curves for river flow calculations]. *Izvestiya AN SSSR. Otd. tekhn. nauk* [Herald of the Russian Academy of Sciences. Technical Sciences] 6:15–21.
6. Kudryavtsev, A. A. 2019. O predstavlenii gamma-eksponentsial'nogo i obobshchennogo otritsatel'nogo binomial'nogo raspredeleniy [On the representation of gamma-exponential and generalized negative binomial distributions]. *Informatika i ee Primeneniya — Inform. Appl.* 13(4):76–80. doi: 10.14357/19922264190412. EDN: YTZLNY.
7. Korolev, V. Yu., and A. I. Zeifman. 2019. Generalized negative binomial distributions as mixed geometric laws and related limit theorems. *Lith. Math. J.* 59(3):366–388. doi: 10.1007/s10986-019-09452-x.
8. Wang, X., X. Zhao, and J. Sun. 2019. A compound negative binomial distribution with mutative termination conditions based on a change point. *J. Comput. Appl. Math.* 351:237–249. doi: 10.1016/j.cam.2018.11.009.
9. Korolev, V. Yu., and A. K. Gorshenin. 2020. Probability models and statistical tests for extreme precipitation based on generalized negative binomial distributions. *Mathematics* 8(4):604. 30 p. doi: 10.3390/math8040604.
10. Zhang, J., D. Wang, and K. Yang. 2020. A study of RCINAR(1) process with generalized negative binomial marginals. *Commun. Stat. B — Simul.* 49(6):1487–1510. doi: 10.1080/03610918.2018.1498891.
11. Bhati, D., and I. S. Ahmed. 2021. On uniform-negative binomial distribution including Gauss hypergeometric function and its application in count regression mod-

- eling. *Commun. Stat. — Theor. M.* 50(13):3106–3122. doi: 10.1080/03610926.2019.1682163.
12. Mangiola, S., E. A. Thomas, M. Modrák, A. Vehtari, and A. T. Papenfuss. 2021. Probabilistic outlier identification for RNA sequencing generalized linear models *NAR Genomics Bioinformatics* 3(1):lqab005. 9 p. doi: 10.1093/nar-gab/lqab005.
 13. Kudryavtsev, A. A., Yu. N. Nedolivko, and O. V. Shestakov. 2022. Main probabilistic characteristics of the digamma distribution and the method of estimating its parameters. *Moscow University Computational Mathematics Cybernetics* 46:81–88. doi: 10.3103/S0278641922020054.
 14. Shevtsova, I., and M. Tselishchev. 2021. On the accuracy of the generalized gamma approximation to generalized negative binomial random sums. *Mathematics* 9(13):1571. 8 p. doi: 10.3390/math9131571.
 15. Kudryavtsev, A. A., and O. V. Shestakov. 2023. Estimates of the convergence rate in the generalized Renyi theorem with a structural digamma distribution using zeta metrics. *Mathematics* 11(21):4477. 10 p. doi: 10.3390/math11214477.
 16. Senatov, V. V. 1998. *Normal approximation: New results, methods, and problems*. Utrecht, The Netherland: VSP. 363 p.

Received January 15, 2024

Contributors

Kudryavtsev Alexey A. (b. 1978) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Moscow Center for Fundamental and Applied Mathematics, M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; aakudryavtsev@cs.msu.ru

Shestakov Oleg V. (b. 1976) — Doctor of Science in physics and mathematics, professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; leading scientist, Moscow Center for Fundamental and Applied Mathematics, M. V. Lomonosov Moscow State University, 1 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; oshestakov@cs.msu.ru

СРАВНИТЕЛЬНЫЙ АНАЛИЗ УЗЛОВЫХ МУЛЬТИПОТОКОВ В МНОГОПОЛЬЗОВАТЕЛЬСКОЙ СЕТЕВОЙ СИСТЕМЕ

Ю. Е. Малашенко¹, И. А. Назарова²

Аннотация: В рамках вычислительных экспериментов изучаются монопольные и предельные режимы передачи исходящих узловых мультипоток в многопользовательской сетевой системе. Анализируются межузловые потоки разных видов, которые передаются всеми корреспондентами из всех узлов по всем кратчайшим маршрутам. В рамках алгоритмической схемы вычисляются равные межузловые потоки. Найденные значения выступают в качестве компонент вектора мультипотока, исходящего из каждого узла-источника во все узлы-адресаты, и интерпретируются как гарантированные многокритериальные оценки показателей функционирования многопользовательской системы. Для оценки минимальных удельных затрат для каждого узла используется монопольный режим передачи исходящего мультипотока без учета всех остальных. Для поиска максимально достижимого межузлового потока рассматривается предельный режим передачи, при котором полностью загружены все ребра сети. Исследуются сети с различными структурными особенностями и одинаковой суммарной пропускной способностью.

Ключевые слова: многопродуктовая потоковая модель; распределение узловых исходящих мультипоток; удельные затраты при передаче потока

DOI: 10.14357/19922264240106

EDN: AKCMCQ

1 Введение

Работа продолжает исследование функциональных характеристик сетевых телекоммуникационных систем при различных способах передачи потока, распределения ресурсов и управления [1–6]. В [4–6] в качестве многопараметрической оценки функциональных возможностей многопользовательской сети используются векторы исходящих узловых мультипоток. Компонентами вектора выступают все межузловые потоки с источником в выбранной вершине и стоками в остальных узлах сети. В ходе вычислительных экспериментов вначале определяются распределения межузловых потоков, которые передаются корреспондентами по кратчайшим маршрутам. Затем на основе найденных значений формируются допустимые векторы исходящих мультипоток, передаваемых в монопольном режиме без учета потоков для остальных узлов сети. Анализируются величины узловых исходящих потоков в ситуациях, когда внешняя нагрузка превышает нормативные ограничения.

В работе изучаются пиковые распределения при полной загрузке сети и максимальном суммарном межузловом потоке. В рамках модели пропускная способность ребер сети рассматривается как вектор ресурсов. Ресурсы распределяются для обеспечения передачи потоков соответствующих видов. Для каждого узла оцениваются требуемые затраты ресурсов при передаче всех исходящих межузловых

поток, для которых данный узел служит источником. Предложенный способ моделирования дает возможность сравнивать и выявлять различия в распределении ресурсов при использовании различных диспетчерских правил стратегий. Полученные оценки можно трактовать как ограничения на величину исходящих узловых мультипоток, которые можно будет передать в сети одновременно с потоками из других узлов, если условия совместной передачи на остальные виды потоков не будут нарушены.

Результаты, полученные в ходе экспериментов, можно также рассматривать как описание функциональных возможностей сети при предельных внешних входных потоках. Появляется возможность на этапе предпроектного анализа проследить взаимосвязь узловых исходящих мультипоток и требуемых затрат ресурсов — пропускной способности ребер сети — при использовании монопольных и уравнивающих диспетчерских правил.

2 Математическая модель

Для описания многопользовательской сетевой системы связи воспользуемся следующей математической записью модели передачи многопродуктового потока. Сеть G задается множествами $\langle V, R, U, P \rangle$: узлов (вершин) сети $V = \{v_1, v_2, \dots, v_n, \dots, v_N\}$; неориентированных ребер $R = \{r_1, r_2, \dots, r_k, \dots, r_E\}$; ориентированных дуг

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, malash09@ccas.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, irina-nazar@yandex.ru

$U = \{u_1, u_2, \dots, u_k, \dots, u_{2E}\}$ и пар узлов-корреспондентов $P = \{p_1, p_2, \dots, p_M\}$. Предполагается, что в сети отсутствуют петли и двоянные ребра.

Ребро $r_k \in R$ соединяет смежные вершины v_{n_k} и v_{j_k} . Каждому ребру r_k ставятся в соответствие две ориентированные дуги u_k и u_{k+E} из множества U . Дуги $\{u_k, u_{k+E}\}$ определяют прямое и обратное направление передачи потока по ребру r_k между концевыми вершинами v_{n_k} и v_{j_k} . Для каждой вершины v_n формируется список $K(n)$ номеров инцидентных ей ребер: $K(n) = \{k_1, k_2, \dots, k_a\}$, где a — число инцидентных ребер для v_n .

В многопользовательской сети G рассматривается $M = N(N - 1)$ независимых, невзаимозаменяемых и равноправных межузловых потоков различных видов. Каждой паре узлов-корреспондентов p_m из множества P соответствуют: вершина-источник с номером s_m , из s_m входной поток m -го вида поступает в сеть; вершина-приемник с номером t_m , из t_m поток m -го вида покидает сеть. Для каждой вершины $v_n \in V$, $n = \overline{1, N}$, определяется подмножество $P(v_n)$ всех пар-корреспондентов, для которых вершина v_n служит узлом-источником:

$$P(v_n) = \{p_m | s_m = n, t_m \neq n, t_m = \overline{1, N}\},$$

а для каждого $P(v_n)$ — список номеров $M(n)$ пар p_m , входящих в подмножество $P(v_n)$:

$$M(n) = \{m_1(n), m_2(n), \dots, m_{N-1}(n)\}.$$

Обозначим через z_m величину межузлового потока m -го вида, поступающего в сеть через узел с номером s_m и покидающего сеть из узла с номером t_m ; x_{mk} и $x_{m(k+E)}$ — поток m -го вида, который передается по дугам u_k и u_{k+E} согласно направлению передачи, $x_{mk} \geq 0$, $x_{m(k+E)} \geq 0$, $m = \overline{1, M}$, $k = \overline{1, E}$; $S(v_n)$ — множество номеров исходящих дуг, по ним поток покидает узел v_n ; $T(v_n)$ — множество номеров входящих дуг, по ним поток поступает в узел v_n . Состав множеств $S(v_n)$ и $T(v_n)$ однозначно формируется в ходе выполнения следующей процедуры. Пусть некоторое ребро $r_k \in R$ соединяет вершины с номерами n и j , такими что $n < j$. Тогда ориентированная дуга $u_k = (v_n, v_j)$, направленная из вершины v_n в v_j , считается *исходящей* из вершины v_n , и ее номер k заносится в множество $S(v_n)$, а дуга u_{k+E} , направленная из v_j в v_n , — *входящей* для v_n , и ее номер $k + E$ помещается в список $T(v_n)$. Дуга u_k является *входящей* для v_j , и ее номер k попадает в $T(v_j)$, а дуга u_{k+E} — *исходящей*, и номер $k + E$ вносится в список исходящих дуг $S(v_j)$.

Во всех узлах сети $v_n \in V$, $n = \overline{1, N}$, для каждого вида потока должны выполняться условия сохранения потоков:

$$\begin{aligned} \sum_{i \in S(v_n)} x_{mi} - \sum_{i \in T(v_n)} x_{mi} &= \\ &= \begin{cases} z_m, & \text{если } v_n = v_{s_m}, \\ -z_m, & \text{если } v_n = v_{t_m}, \\ 0 & \text{в остальных случаях,} \end{cases} \\ n = \overline{1, N}, m = \overline{1, M}, x_{mi} \geq 0, z_m \geq 0. \end{aligned} \quad (1)$$

Величина z_m равна входному межузловому потоку m -го вида, проходящему от источника s_m к приемнику t_m пары p_m при распределении потоков x_{mi} по дугам сети.

Каждому ребру $r_k \in R$ приписывается неотрицательное число d_k — суммарный предельно допустимый поток, который можно передать по ребру r_k в обоих направлениях. В исходной сети компоненты вектора пропускных способностей $\mathbf{d} = (d_1, d_2, \dots, d_k, \dots, d_E)$ — положительные числа $d_k > 0$. Вектор \mathbf{d} определяет следующие ограничения на сумму потоков всех видов, передаваемых по ребру r_k одновременно:

$$\begin{aligned} \sum_{m=1}^M (x_{mk} + x_{m(k+E)}) &\leq d_k, \\ x_{mk} \geq 0, x_{m(k+E)} &\geq 0, k = \overline{1, E}. \end{aligned} \quad (2)$$

Ограничения (1) и (2) задают множество допустимых значений компонент вектора межузловых потоков $\mathbf{z} = (z_1, z_2, \dots, z_m, \dots, z_M)$:

$$\mathcal{Z}(\mathbf{d}) = \{\mathbf{z} \geq 0 \mid \exists \mathbf{x} \geq 0 : (\mathbf{z}, \mathbf{x}) \text{ удовлетворяют (1), (2)}\}.$$

В рамках формализма модели для каждой вершины $v_n \in V$, $n = \overline{1, N}$, формируется VMF-вектор (от *англ.* vertex multiflow) исходящего мультипоточка $\vec{\mathbf{z}}(n)$, $\vec{\mathbf{z}}(n) \in \mathcal{Z}(\mathbf{d})$, компонентами которого выступают $z_j(n) = 0$, $j \notin M(n)$, $z_j(n) = z_m$, $j = m$, $m \in M(n)$. Таким образом, ненулевыми компонентами $\vec{\mathbf{z}}(n)$ становятся межузловые потоки $z_m > 0$ для всех пар $p_m \in P(v_n)$ с узлами-источниками s_m , расположенными в вершине v_n .

3 Уравнительное распределение в монопольном и предельном режимах

При проведении вычислительных экспериментов векторы $\vec{\mathbf{z}}(n)$ формируются на основе найденных распределений совместно допустимых межузловых потоков z_m , $p_m \in P(v_n)$, передаваемых одновременно. Для каждой пары узлов-корреспон-

дентов $p_m \in P$, для некоторого заданного допустимого межзвонного потока \tilde{z}_m и соответствующих значений дуговых потоков \tilde{x}_{mk} , $k = \overline{1, 2\overline{E}}$, величина

$$\tilde{y}_m = \sum_{i=1}^{2\overline{E}} \tilde{x}_{mi}, \quad m = \overline{1, \overline{M}},$$

характеризует результирующую межзвонную нагрузку (далее — RI-нагрузку, от *англ.* resulting internodal load) на ребра сети при передаче межзвонного потока \tilde{z}_m из узла-источника s_m в узел-приемник t_m . Величина \tilde{y}_m показывает, какая суммарная пропускная способность сети потребуется для передачи дуговых потоков \tilde{x}_{mk} . В рамках модели отношение RI-нагрузки и межзвонного потока

$$\tilde{w}_m = \frac{\tilde{y}_m}{\tilde{z}_m}, \quad m = \overline{1, \overline{M}},$$

можно трактовать как удельные затраты ресурсов сети при передаче единичного потока m -го вида между узлами s_m и t_m при дуговых потоках \tilde{x}_{mi} .

Для получения гарантированных оценок VMF-векторов предполагается, что межзвонные потоки передаются по маршрутам, содержащим минимальное число ребер (далее — MER-маршруты, от *англ.* minimum edge route). При проведении вычислительных экспериментов для оценки величин «расщепленного» потока на первом этапе для каждой пары узлов $p_m = (s_m, t_m)$ в сети $G(1)$ формируется набор $H_m(1)$ всех путей, которые далее используются как MER-маршруты передачи m -го вида потока:

$$H_m(1) = \left\{ h_m^1(1), h_m^2(1), \dots, h_m^j(1), \dots, h_m^{J_m(1)}(1) \right\},$$

где $h_m^j(1)$ — список номеров дуг в j -м пути в сети $G(1)$ между узлами s_m и t_m ; $\mu_m(1)$ — минимальное число ребер в MER-маршруте $h_m^j(1)$; $J_m(1)$ — число MER-маршрутов для m -й пары.

Для каждой пары $p_m \in P$ по всем MER-маршрутам из $H_m(1)$ передается единичный межзвонный поток z_m и подсчитываются значения индикаторной функции:

$$\eta_k^j(m) = \begin{cases} 1, & \text{если } k \in h_m^j(1); \\ 0 & \text{в остальных случаях.} \end{cases}$$

Определяются дуговые потоки для пары p_m :

$$x_{mk}^0(1) = \sum_{j=1}^{J_m(1)} \eta_k^j(m), \quad m = \overline{1, \overline{M}}, \quad k = \overline{1, 2\overline{E}}. \quad (3)$$

Межзвонный поток по MER-маршрутам (далее — MER-поток) $z_m^0(1)$ между узлами s_m и t_m вычисляется по формулам (1) и (3). Рассчитываются нормирующий коэффициент

$$\omega_m^0(1) = \frac{1}{z_m^0(1)}, \quad z_m^0(1) \neq 0, \quad m = \overline{1, \overline{M}},$$

и дуговые потоки

$$x_{mk}^0 = \omega_m^0 x_{mk}^0(1), \quad m = \overline{1, \overline{M}}, \quad k = \overline{1, 2\overline{E}}. \quad (4)$$

При передаче всех потоков x_{mk}^0 по ребрам сети межзвонный поток из узла s_m в узел t_m равен единице для всех $p_m \in P$.

Распределение совместно допустимых межзвонных потоков последовательно определяется на основе MER-маршрутов и дуговых потоков x_{mk}^0 . Для каждого VMF-вектора вычисляются равные значения исходящих мультипотоков, которые могут передаваться одновременно в сети $G(1)$ без учета остальных корреспондентов. Указанный способ называется далее *монопольным* режимом передачи VMF-потока.

Задача 1. Для заданной вершины v_n найти

$$\beta^*(n) = \max_{\beta} \beta$$

при условиях:

$$\beta \sum_{m \in M(n)} \left[x_{mk}^0 + x_{m(k+E)}^0 \right] \leq d_k(1), \quad \beta \geq 0, \quad k = \overline{1, \overline{E}}.$$

С помощью решения задачи 1 для всех $p_m \in P$, $m \in M(n)$, вычисляются VMF-векторы $\vec{z}^*(n)$, все компоненты которых равны $\beta^*(n)$, т.е. $z_m^*(n) = \beta^*(n)$, $m \in M(n)$, $n = \overline{1, \overline{N}}$. Для каждого ребра r_k , $k = \overline{1, \overline{E}}$, определяется

$$\Delta_k^*(n) = \beta^*(n) \sum_{m \in M(n)} \left[x_{mk}^0 + x_{m(k+E)}^0 \right], \quad k = \overline{1, \overline{E}}.$$

Векторы ресурсов $\vec{\Delta}^*(n) = \{\Delta_1^*(n), \Delta_2^*(n), \dots, \Delta_E^*(n)\}$ показывают, какие затраты пропускной способности требуются при передаче исходящих межзвонных потоков из вершины v_n в монопольном режиме без учета всех остальных корреспондентов. Для каждого VMF-вектора подсчитываются

$$\zeta(n) = \sum_{m \in M(n)} z_m^*(n) = \beta^*(n)(N-1); \quad \delta(n) = \sum_{k=1}^E \Delta_k^*(n);$$

$$\|\vec{z}^*(n)\| = \beta^*(n)(N-1)^{1/2};$$

$$\|\vec{\Delta}^*(n)\| = \left[\sum_{k=1}^E (\Delta_k^*(n))^2 \right]^{1/2}, \quad n = \overline{1, \overline{N}}.$$

Для оценки VMF-векторов при полной загрузке всех ребер сети сформируем вектор $\vec{z}^{**}(n)$, компоненты которого определяются по следующему правилу: для каждого узла v_n и всех пар $p_m \in P(v_n)$, которые не являются смежными узлами и кратчайший маршрут соединения для них состоит из двух и более ребер, поток $z_m^{**}(n) = 0$. Для всех пар $p_a \in P(v_n)$, для которых кратчайшим маршрутом соединения служит некоторое ребро r_k , поток $z_a^{**}(n) = d_k(1)/2$, где $d_k(1)$ — пропускная способность ребра r_k в сети $G(1)$.

При одновременной передаче всех VMF-поток $\vec{z}^{**}(n)$ из всех узлов $v_n, n = \overline{1, N}$, достигается полная загрузка всех ребер сети $G(1)$, и максимальное суммарное значение допустимых межузловых потоков z^{**} равно

$$z^{**} = \sum_{n=1}^N \sum_{m=1}^{M(n)} z_m^{**}(n) = \sum_{k=1}^E d_k(1) = D(0).$$

Вычисляется норма векторов

$$\|\vec{z}^{**}(n)\| = \left[\sum_{k \in K(n)} \left(\frac{d_k(1)}{2} \right)^2 \right]^{1/2} = \|\vec{\Delta}^{**}(n)\|, \quad n = \overline{1, N},$$

где $\vec{\Delta}^{**}(n)$ — вектор ресурсов, необходимых при передаче мультипотока $\vec{z}^{**}(n)$. Значения x_{mk}^0 из (4) позволяют формировать VMF-вектор равных межузловых потоков $\vec{z}^{(=)}(n)$, которые могут передаваться одновременно из всех узлов сети.

Задача 2. Найти

$$\alpha^*(n) = \max_{\alpha} \alpha$$

при условиях:

$$\alpha \sum_{m=1}^M [x_{mk}^0 + x_{m(k+E)}^0] \leq d_k(1), \quad \alpha \geq 0, \quad k = \overline{1, E}.$$

Из решения задачи 2 определяется $z_m^{(=)} = \alpha^*$ для всех $p_m \in P$ и, соответственно, для всех VMF-векторов $\vec{z}^{(=)}(n) = \alpha^*, n = \overline{1, N}$:

$$\|\vec{z}^{(=)}(n)\| = \alpha^*(N-1)^{1/2}.$$

4 Вычислительный эксперимент

Вычислительный эксперимент проводился на моделях сетевых систем, представленных на рис. 1. В каждой сети 69 узлов. В ходе вычислительного эксперимента проводилась нормировка, и суммарная пропускная способность в обеих сетях была одинакова: $\sum_{k=1}^E d_k(1) = D(0) = 68\,256$.

На диаграммах, представленных на рис. 2, для VMF-векторов $\vec{z}(n)$ в базовой и кольцевой сетях по горизонтальной оси откладываются суммарные значения мультипотока $\zeta(n)$, а по вертикальной $\delta(n)$ — соответствующие значения требуемой

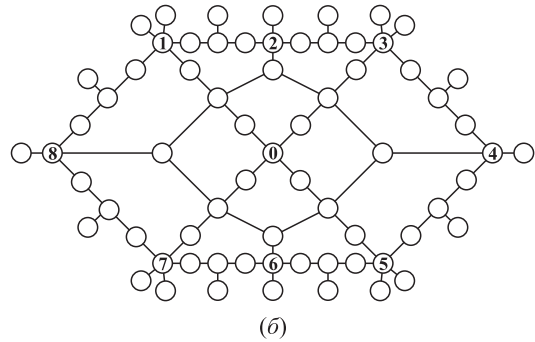
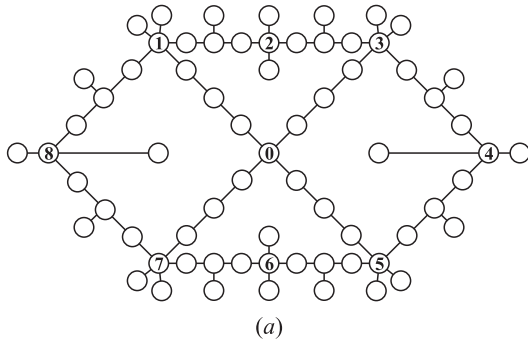


Рис. 1 Базовая (а) и кольцевая (б) сети

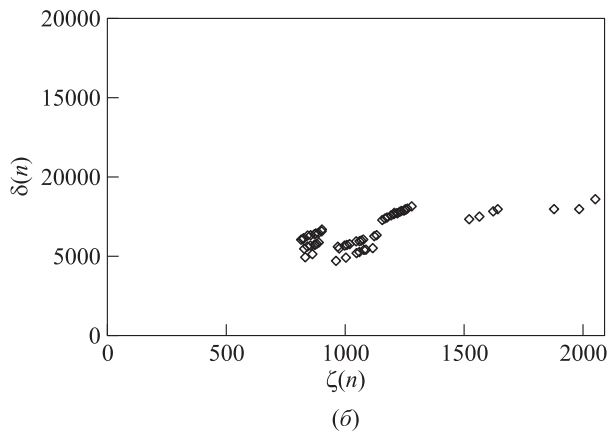
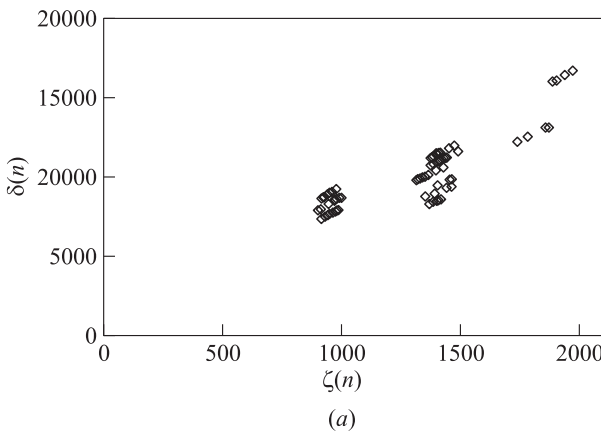


Рис. 2 Узловые мультипоток и ресурсы в базовой (а) и в кольцевой (б) сетях

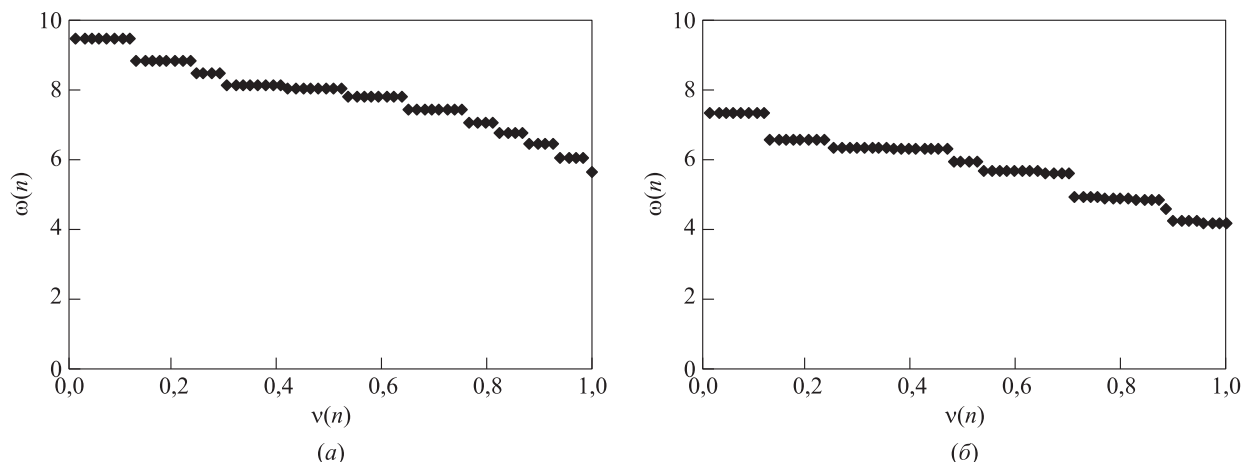


Рис. 3 Удельные затраты в базовой (а) и в кольцевой (б) сетях

суммарной пропускной способности при передаче в монопольном режиме. На рис. 2, а группа точек, для которых $\zeta(n) \leq 1000$, относится к висячим вершинам и суммарный исходящий поток ограничен величиной пропускной способности инцидентного ребра. Для большой группы узлов суммарный исходящий межузловой поток находится в диапазоне $1300 \leq \zeta(n) \leq 1500$, поскольку в монопольном режиме используется уравнивательная стратегия для всех исходящих потоков. В кольцевой сети «средняя» длина кратчайшего пути меньше, чем в базовой. В результате, как следует из рис. 2, б, затраты ресурсов — пропускной способности — значительно меньше практически для всех VMF-потоков и $\delta(n) \leq 10\,000$.

На рис. 3 представлены диаграммы распределения удельных затрат $w(n) = \delta(n)/\zeta(n)$ передачи единичного потока из узла $v(n)$. Значения $w(n)$ откладываются по вертикальной оси, а относительные номера узлов $v(n) = n/N$, $n = \overline{1, N}$, указываются по горизонтальной оси. Удельные затраты в кольцевой сети меньше, чем в базовой, поскольку добавление ребер приводит к уменьшению длины кратчайших путей для большого числа пар-корреспондентов. Для висячих узлов, соединенных единственным ребром с сетью, удельные затраты выше, чем для вершин, расположенных в центре. Самая правая точка на рис. 3, а со значением $w^*(n) < 6$ соответствует центральному узлу в базовой сети.

5 Заключение

Построенные априорные оценки VMF-векторов позволяют анализировать ситуации, когда внешняя (входная) нагрузка на сеть превышает допустимые нормативные значения. Исходящие взаимно допустимые мультитокеты $\vec{z}(n)$ указывают

на возможные ограничения, отказы в обслуживании и/или возникновение очередей на входе в систему. Оценив удельные затраты, полученные на основе VMF-векторов, можно выделить группы узлов-корреспондентов, требующих больших ресурсов для передачи соответствующих потоков. Анализ структурных потоковых характеристик, их взаимосвязи и взаимовлияния может быть использован на предпроектном этапе создания сетей различного назначения [7–9].

Литература

1. Kung H. T., Wu C. H. Content networks: Taxonomy and new approaches // The Internet as a large-scale complex system. — Oxford: Oxford University Press, 2005. P. 203–225. doi: 10.1093/oso/9780195157208.003.0007.
2. Yang R., van der Mei R. D., Roubos D., et al. Resource optimization in distributed real-time multimedia applications // Multimed. Tools Appl., 2012. Vol. 59. P. 941–971. doi: 10.1007/s11042-011-0782-5.
3. Beben A., Batalla J. M., Chai W. K., Sliwinski J. Multi-criteria decision algorithms for efficient content delivery in content networks // Ann. Telecommun., 2013. Vol. 68. P. 153–165. doi: 10.1007/s12243-012-0321-z.
4. Малащенко Ю. Е., Назарова И. А. Оценки распределения ресурсов в многопользовательской сети при равных межузловых нагрузках // Информатика и её применения, 2023. Т. 17. Вып. 1. С. 83–88. doi: 10.14357/19922264230111. EDN: BUKVGV.
5. Малащенко Ю. Е., Назарова И. А. Анализ загрузки многопользовательской сети при расщеплении потоков по кратчайшим маршрутам // Информатика и её применения, 2023. Т. 17. Вып. 3. С. 33–38. doi: 10.14357/19922264230305. EDN: NLUSQJ.
6. Малащенко Ю. Е., Назарова И. А. Анализ узловых мультитокетов в многопользовательской системе при уравнивательных стратегиях управления // Известия РАН. Теория и системы управления, 2023. № 6. С. 137–149.

7. Качанов С. А., Медведев Н. В. Алгоритм выравнивания загрузки узлов мобильной информационно-коммуникационной сети // Технологии гражданской безопасности, 2011. Т. 8. № 1. С. 26–29. EDN: NQUSAR.
8. Zhang H. P., Yin B. Q., Lu X. N. Modeling and analysis for streaming service systems // Int. J. Automation Computing, 2014. Vol. 11. P. 449–458. doi: 10.1007/s11633-014-0812-7.
9. Симаков Д. В. Управление трафиком в сети с высокой динамикой метрик сетевых маршрутов // Научное ведомство, 2016. Т. 8. № 1. Ст. 60TVN116.

Поступила в редакцию 04.12.23

ANALYSIS OF NODE MULTIFLOWS IN A MULTIUSER NETWORK SYSTEM

Yu. E. Malashenko and I. A. Nazarova

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: Within the framework of computational experiments, exclusive and limit modes of transmission of outgoing node multiflows in a multiuser network system are studied. Internodal flows of different types that are transmitted from all nodes along all shortest routes are analyzed. Within the framework of the algorithmic scheme, the transmission of all internodal flows is calculated. The found values act as components of the multiflow vector originating from each source node to all destination nodes and are interpreted as guaranteed multicriteria estimates of the functioning of a multiuser system. To estimate the minimum unit cost for each node, the monopole mode of transmitting the outgoing multiflow is used without taking into account all the others. To search for the maximum achievable internodal flow, the limiting transmission mode is considered in which all edges of the network are completely loaded. The networks with different structural features and the same total capacity are studied.

Keywords: multicommodity flow model; distribution of node multiflows; unit cost of flow transmission

DOI: 10.14357/19922264240106

EDN: AKCMCQ

References

1. Kung, H. T., and C. H. Wu. 2005. Content networks: Taxonomy and new approaches. *The Internet as a large-scale complex system*. Oxford: Oxford University Press. 203–225. doi: 10.1093/oso/9780195157208.003.0007.
2. Yang, R., R. D. van der Mei, D. Roubos, et al. 2012. Resource optimization in distributed real-time multimedia applications. *Multimed. Tools Appl.* 59(3):941–971. doi: 10.1007/s11042-011-0782-5.
3. Beben, A., J. M. Batalla, W. K. Chai, and J. Sliwinski. 2013. Multi-criteria decision algorithms for efficient content delivery in content networks. *Ann. Telecommun.* 68:153–165. doi: 10.1007/s12243-012-0321-z.
4. Malashenko, Yu. E., and I. A. Nazarova. 2023. Otsenki raspredeleniya resursov v mnogopol'zovatel'skoy seti pri ravnykh mezhuzlovykh nagruzkakh [Estimates of the resource distribution in the multiuser network with equal internodal loads]. *Informatika i ee Primeneniya — Inform. Appl.* 17(1):83–88. doi: 10.14357/19922264230111. EDN: BUKVGV.
5. Malashenko, Yu. E., and I. A. Nazarova. 2023. Analiz zagruzki mnogopol'zovatel'skoy seti pri rasshcheplenii potokov po kratchayshim marshrutam [Multiuser network load analysis by splitting flows along the shortest routes]. *Informatika i ee Primeneniya — Inform. Appl.* 17(3):33–38. doi: 10.14357/19922264230305. EDN: NLUSQJ.
6. Malashenko, Yu. E., and I. A. Nazarova. 2023. Analysis of nodal multi-flows in a multiuser system with equalizing management strategies. *J. Comput. Sys. Sc. Int.* 62(6):1022–1033.
7. Kachanov, S. A., and N. V. Medvedev. 2011. Algoritm vyravnivaniya zagruzki uzlov mobil'noy informatsionno-kommunikatsionnoy seti [Algorithm of leveling loading of mobile information and communication centers net]. *Tekhnologii grazhdanskoy bezopasnosti* [Civil Security Technologies] 8(1):26–29. EDN: NQUSAR.
8. Zhang, H. P., B. Q. Yin, and X. N. Lu. 2014. Modeling and analysis for streaming service systems. *Int. J. Automation Computing* 11:449–458. doi: 10.1007/s11633-014-0812-7.
9. Simakov, D. V. 2016. Upravlenie trafikom v seti s vysokoy dinamikoy metrik setevykh marshrutov [Traffic engineering for networks with high dynamics of routing metrics]. *Naukovedenie* [Science Studies] 8(1):60TVN116.

Received December 4, 2023

Contributors

Malashenko Yuri E. (b. 1946) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; malash09@ccas.ru

Nazarova Irina A. (b. 1966) — Candidate of Science (PhD) in physics and mathematics, scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; irina-nazar@yandex.ru

ЛОКАЛЬНАЯ ДРЕВОВИДНОСТЬ В КОНФИГУРАЦИОННЫХ ГРАФАХ СО СТЕПЕННЫМ РАСПРЕДЕЛЕНИЕМ*

М. М. Лери¹, Ю. Л. Павлов²

Аннотация: Исследуется локальная древовидность конфигурационных графов, предназначенных для моделирования сложных сетей коммуникаций. В таких графах степени вершин независимы и одинаково распределены по степенному закону. В случае ограниченного числа вершин графа найдены зависимости максимального объема подграфа, имеющего вид дерева, от числа вершин графа и параметра распределения степеней вершин. Такая же задача решалась для числа деревьев заданного объема. Даны также оценки среднего объема дерева в графе. Показано, что при ограниченном числе вершин конфигурационного графа найденные зависимости статистически значимо улучшают описание структуры сетей по сравнению с известными ранее асимптотическими моделями.

Ключевые слова: конфигурационные графы; степенное распределение; локальная древовидность; объем дерева; имитационное моделирование

DOI: 10.14357/19922264240107

EDN: ITTZEO

1 Введение

Случайные графы активно используются для моделирования топологий сложных сетей различных видов, таких как транспортные, компьютерные, биологические, социальные и др. [1–4]. При этом в качестве вершин могут рассматриваться как отдельные объекты (транспортные средства, компьютеры, люди и пр.), так и некоторые комплексные системы (транспортные узлы, компьютерные автономные системы, социальные сообщества и др.), что существенно влияет на определение степеней вершин соответствующих случайных графов. Подобное разнообразие касается и принципов установления связей между этими вершинами. Так, конфигурационные графы, модель которых впервые была предложена Б. Боллобашем [5], с независимыми одинаково распределенными степенями вершин с общим дискретным законом распределения часто используются при моделировании телекоммуникационных сетей, среди которых и сеть Интернет [1, 6]. Исследованию таких моделей посвящена обширная литература, число опубликованных монографий и статей по этой тематике исчисляется сотнями. Наиболее полные обзоры современного состояния вопроса, полученных результатов и используемых методов можно найти в фундаментальных трудах [3, 4]. Понятно, что структура и динамика конфигурационных графов

существенно зависят от распределений степеней вершин, и многие работы посвящены исследованию таких распределений. Но в последние годы стало ясно, что для построения адекватных моделей недостаточно знания законов распределения степеней, важно также иметь представление о различных структурных особенностях изучаемых объектов. Одна из таких особенностей конфигурационных графов — локальная древовидность их структуры [3, 7, 8].

Для изучения структуры конфигурационных графов успешно используются методы теории ветвящихся процессов (см., например, [3, 4]). Каждой вершине графа можно поставить в соответствие начинающийся с одной частицы ветвящийся процесс Гальтона–Ватсона. Связь между такими ветвящимися процессами и конфигурационными графами подробно изучена в [4]. С помощью этой связи при стремлении к бесконечности числе вершин графа найдены условия возникновения гигантской компоненты связности, получены оценки объемов компонент связности, расстояний между вершинами и диаметра. Была выявлена локальная древовидность топологии графа, состоящая в том, что связный подграф, содержащий произвольную вершину, асимптотически достоверно сближается с начинающейся с соответствующей этой вершине частицы частью траектории ветвящегося процесса при некотором ограничении общего числа частиц

* Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания Карельского научного центра Российской академии наук (Институт прикладных математических исследований КарНЦ РАН).

¹ Институт прикладных математических исследований Карельского научного центра Российской академии наук, leri@krc.karelia.ru

² Институт прикладных математических исследований Карельского научного центра Российской академии наук, pavlov@krc.karelia.ru

такой части траектории. Понятно, что эта часть траектории представляет собой дерево, в котором начальная частица служит корнем. Таким образом, начиная с любой вершины графа можно построить дерево, содержащее максимально возможное число вершин этого графа. Естественно, возникает вопрос об оценке максимального объема дерева среди всех таких деревьев. Другими словами, возникает задача поиска максимального по объему дерева среди подграфов конфигурационного графа. До сих пор каждое дерево в конфигурационном графе как часть траектории ветвящегося процесса изучалось отдельно. Но если рассматривать деревья в совокупности, то получится случайный лес. Теория случайных лесов в значительной степени была создана и изложена в книге [9]. В [7, 8] впервые было предложено использовать теорию случайных лесов для изучения структуры конфигурационных графов. Однако в [9] предполагалось, что леса генерируются ветвящимися процессами с конечной дисперсией числа потомков каждой частицы, а в современных моделях сетей это условие нарушено. Поэтому в [7, 8] были получены недостающие результаты о лесах и с их помощью изучалось асимптотическое поведение некоторых числовых характеристик. Были доказаны теоремы о предельных распределениях числа деревьев заданного объема [8], а также предельные теоремы для максимального объема дерева [7]. Под объемом дерева понимается число его вершин. Целью исследования настоящей работы было изучение этих же характеристик конфигурационных графов, а также среднего объема дерева в доасимптотической области и построение моделей их зависимостей от объема графа и параметра распределения степеней вершин.

2 Описание модели и алгоритма

В работе рассматриваются конфигурационные графы, состоящие из N вершин. Построение графа происходит следующим образом. Каждой вершине присваивается степень в соответствии со следующим распределением [6]:

$$\mathbf{P}\{\xi \geq k\} = k^{-\tau}, \quad k = 1, 2, \dots, \quad (1)$$

где ξ — случайная величина, равная степени любой вершины графа, а параметр $\tau > 0$ принимает фиксированные значения. Заметим, что в известных моделях сетей коммуникаций [1–4, 6] значения параметра τ не выходят за пределы интервала [2, 4]. Степень вершины равна числу полурёбер (см., например, [6]) вершины графа. Все полурёбра конфигурационного графа занумерованы в произвольном

порядке. Далее все полурёбра попарно и равномерно соединяются друг с другом, образуя рёбра графа. Сумма степеней вершин (число полурёбер конфигурационного графа) в данной работе считается случайной величиной. Легко видеть, что для построения графа эта величина должна быть четной, поэтому в противном случае степень равномерно выбранной вершины увеличивается на 1, тем самым добавляя к ней полурёбро. Очевидно, что конструкция конфигурационного графа допускает наличие циклов, петель и кратных рёбер. Помимо этого, известно [2, 3, 6], что такие графы асимптотически почти наверное состоят из более чем одной компоненты связности, а при значениях параметра τ , лежащих в интервале (1, 2), такие графы имеют так называемую гигантскую компоненту связности, число вершин которой пропорционально N при $N \rightarrow \infty$, тогда как объемы других компонент бесконечно малы по сравнению с объемом гигантской компоненты.

Пусть степени вершин графа — независимые одинаково распределенные случайные величины, имеющие распределение (1). Тогда

$$p_k = \mathbf{P}\{\xi = k\} = \frac{1}{k^\tau} - \frac{1}{(k+1)^\tau}, \quad k = 1, 2, \dots \quad (2)$$

Каждой вершине графа поставим в соответствие начинающийся с одной частицы ветвящийся процесс Гальтона–Ватсона, в котором распределение числа прямых потомков начальной частицы совпадает с (2). Обозначим η случайную величину, равную числу прямых потомков всех остальных частиц ветвящегося процесса. В [4] показано, что в этом случае

$$\mathbf{P}\{\eta = k\} = \frac{(k+1)p_{k+1}}{\mathbf{E}\xi}, \quad k = 0, 1, 2, \dots$$

Авторами разработан алгоритм нахождения дерева максимального объема с корнем в произвольно выбранной вершине графа. При построении дерева каждая вершина рассматривается вместе с инцидентными ей рёбрами. Вершины, смежные с корнем, считаются вершинами 1-го поколения. Вершины, смежные с вершинами 1-го поколения и отличные от корня, образуют 2-е поколение. Вершины 3-го поколения смежны с вершинами 2-го поколения и, естественно, не принадлежат первым двум поколениям. Аналогично находятся и вершины следующих поколений. Построение дерева по алгоритму происходит начиная с корня от поколения к поколению. Вершины очередного поколения присоединяются к дереву только в том случае, если их включение не приводит к появлению кратных рёбер и циклов, включая петли.

Приведенный ниже алгоритм находит максимальное по объему дерево с корнем в i -й вершине и вычисляет его объем v_i . Обозначим символом \square конец работы алгоритма, его результатом будет значение v_i . В описании алгоритма l означает номер поколения, значение $l = 0$ соответствует корню, а k_l равно числу вершин поколения l .

Алгоритм

1. i — произвольно выбранная корневая вершина, $l = 0$.
2. Проверяем условие: корневая вершина имеет петлю или соединена кратными ребрами с любой другой вершиной графа:
 - если условие 2 выполнено, то $k_0 = 0, v_i = 0 \square$;
 - если условие 2 не выполнено, то $k_0 = 1$ и переходим к шагу 3.
3. Полагаем $l = l + 1$.
4. Находим все вершины графа, смежные с вершинами поколения $l - 1, k_l$ — число вершин поколения l :
 - если $k_l = 0$, то $v_i = k_{l-1} \square$;
 - если $k_l > 0$, переходим к шагу 5.
5. Для всех вершин поколения l проверяем условия: (1) вершина имеет петлю; (2) вершина соединена ребром с другой вершиной поколения l ; (3) вершина соединена кратными ребрами с любой вершиной графа, не являющейся вершиной поколения l ; (4) две вершины поколения l соединены ребрами с одной любой вершиной графа, не являющейся вершиной поколения l за исключением вершин предыдущего поколения $l - 1$:

- если для любой из вершин поколения l выполнено хотя бы одно из условий (1)–(4), то $v_i = \sum_{j=0}^{l-1} k_j \square$;
- если для всех вершин поколения l все условия (1)–(4) не выполнены, переходим к шагу 3.

В качестве примера рассмотрим конфигурационный граф, состоящий из 20 вершин (рис. 1) и с помощью приведенного выше алгоритма на этом графе найдем максимальное дерево с корнем в вершине 1 (рис. 2). В данном примере первое поколение состоит из вершины с номером 6. Вершины второго поколения — 5, 14 и 15, третьего — 7, 8, 9, 10, 16, 18 и 20. Таким образом, объем максимального дерева с корнем в вершине 1 равен 12 вершинам.

Далее под деревом i -й вершины будем понимать дерево максимального объема, найденное с по-

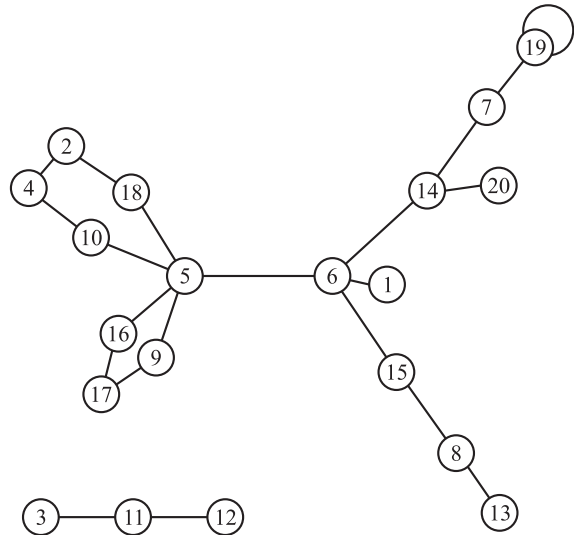


Рис. 1 Пример конфигурационного графа, состоящего из 20 вершин

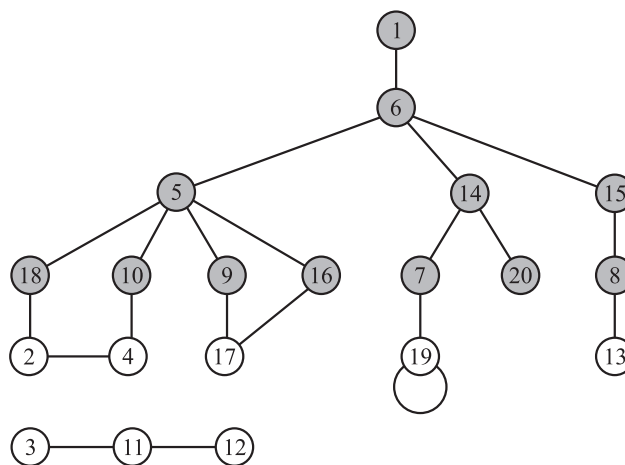


Рис. 2 Пример нахождения максимального дерева с корнем в вершине 1 для графа на рис. 1

мощью приведенного выше алгоритма. Рассмотрим каждую i -ю вершину графа ($i = \overline{1, N}$) в качестве корня и построив соответствующие корневые деревья, общее число которых равно числу вершин графа N , можно оценить следующие структурные характеристики: объем максимального дерева, средний объем дерева и число деревьев заданного объема и найти зависимости этих характеристик от числа вершин графа N и параметра распределения степеней вершин τ . Исследование проводилось посредством методов имитационного моделирования с последующей статистической обработкой данных с помощью программного обеспечения Wolfram Mathematica 9.0.

3 Результаты

Рассматривались конфигурационные графы следующих объемов: $100 \leq N \leq 1000$ с шагом 100 и $1500 \leq N \leq 10\,000$ с шагом 500. Значения параметра τ изменялись от 1,01 до 3,0 с шагом 0,01 для $1,01 \leq \tau \leq 1,1$ и 0,1 для $1,2 \leq \tau \leq 3,0$. Для каждой пары значений (N, τ) генерировалось по 100 графов (всего 81 200).

Распределение (1) имеет конечное математическое ожидание и бесконечную дисперсию при $\tau \in (1, 2]$, а при $\tau > 2$ дисперсия конечна. Поэтому при исследовании зависимостей рассматриваемых структурных характеристик от N и τ будем строить эти зависимости на двух интервалах изменения параметра распределения: $\tau \in (1, 2]$ и $\tau \in (2, 3]$.

3.1 Объем максимального дерева

Объем максимального дерева для каждого графа вычислялся следующим образом:

$$\nu_{\max} = \max_i v_i, \quad i = \overline{1, N}.$$

Таблица 1 Значения коэффициентов a и b зависимостей вида (3) для некоторых значений τ и коэффициенты детерминации R^2 этих уравнений

τ	a	b	R^2
1,01	0,336	1,646	0,09
1,05	0,466	1,097	0,13
1,1	0,593	0,883	0,13
1,3	1,792	-3,677	0,26
1,5	4,611	-16,507	0,35
1,8	13,614	-61,531	0,48
2,0	24,890	-122,947	0,55
2,3	53,950	-291,503	0,54
2,5	65,988	-370,800	0,50
3,0	17,177	-86,828	0,41

Предельное поведение максимального объема дерева конфигурационного графа при $N \rightarrow \infty$ рассматривалось в [7]. Теорема 2 этой работы позволяет предположить, что при достаточно больших N и $\tau \in (1, 2)$ максимальный объем пропорционален $\ln N$. Проверим, что и в доасимптотической области эта характеристика ведет себя так же, как и при $N \rightarrow \infty$. Для этого построим следующие зависимости объема максимального дерева ν_{\max} от числа вершин N при фиксированных значениях параметра τ :

$$\nu_{\max} = a \ln N + b. \quad (3)$$

В табл. 1 приведены значения коэффициентов a и b регрессионных уравнений вида (3) для некоторых значений параметра τ и соответствующие коэффициенты детерминации R^2 этих уравнений. Для всех уравнений гипотезы о равенстве коэффициентов детерминации нулю ($H_0 : R^2 = 0$) отвергаются, и все коэффициенты a и b в табл. 1 значимы на уровне значимости 5%.

Значения коэффициентов детерминации полученных зависимостей выше всего при значениях τ в центре интервала $(1, 3]$ и ниже на его концах, причем $R^2 > 0,5$ при $1,9 \leq \tau \leq 2,5$.

Поиск регрессионной зависимости объема максимального дерева ν_{\max} от числа вершин N при фиксированных значениях параметра τ с более высокими коэффициентами детерминации привел к зависимости другого вида:

$$\ln \nu_{\max} = a_l \ln N + b_l. \quad (4)$$

Для полученных зависимостей вида (4) наблюдаем рост значений коэффициентов детерминации, и здесь уже $R^2 > 0,5$ при $1,6 \leq \tau \leq 3,0$ (табл. 2).

Оценив значимость различия между коэффициентами множественной корреляции регрессионных

Таблица 2 Значения коэффициентов a_l и b_l зависимостей вида (4) для некоторых значений τ и коэффициенты детерминации R^2 этих уравнений

τ	a_l	b_l	R^2
1,01	0,095	0,653	0,12
1,05	0,117	0,580	0,17
1,1	0,128	0,634	0,18
1,3	0,211	0,580	0,32
1,5	0,298	0,496	0,43
1,8	0,418	0,343	0,59
2,0	0,498	0,139	0,67
2,3	0,655	-0,603	0,75
2,5	0,732	1,181	0,78
3,0	0,499	-0,310	0,67

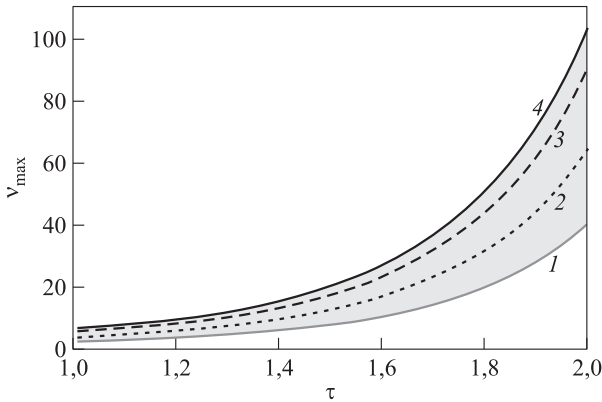


Рис. 3 Регрессионная зависимость (5) максимального объема дерева ν_{\max} от $\tau \in (1, 2]$ при фиксированных значениях N : 1 — $N = 100$; 2 — 1000; 3 — 5000; 4 — $N = 10\,000$

моделей (3) и (4) попарно при каждом фиксированном τ на 5%-ном уровне значимости, получаем, что нулевые гипотезы $H_0 : r_{(3)} = r_{(4)}$ ($r_{(3)}$ и $r_{(4)}$ — коэффициенты множественной корреляции зависимостей (3) и (4) соответственно) для каждой пары моделей отвергаются; следовательно, различие между коэффициентами корреляции значимо. И так как значения коэффициентов детерминации моделей вида (4) выше, то можно сделать вывод, что эти модели «лучше» подходят для описания зависимости максимального объема дерева ν_{\max} от N в доасимптотической области. Таким образом, зависимость объема максимального дерева от объема графа N в конфигурационных графах, число вершин в которых меньше 10 000, отличается от таковой при $N \rightarrow \infty$.

Далее были найдены зависимости максимального объема дерева ν_{\max} от размера графа N и параметра распределения степеней вершин τ для двух интервалов изменения τ :

— для $\tau \in (1, 2]$ (рис. 3):

$$\ln \nu_{\max} = 0,205 \ln N + 0,941\tau^2 - 1,013 \quad (R^2 = 0,95); \quad (5)$$

— для $\tau \in (2, 3]$ (рис. 4):

$$\ln \nu_{\max} = (-0,397\tau^2 + 1,921\tau - 1,728) \ln N \quad (R^2 = 0,97). \quad (6)$$

Здесь и далее все коэффициенты полученных регрессионных уравнений значимы на уровне значимости 5% и для всех приведенных ниже коэффициентов детерминации R^2 полученных моделей гипотезы о равенстве этих коэффициентов детерминации нулю ($H_0 : R^2 = 0$) отвергаются.

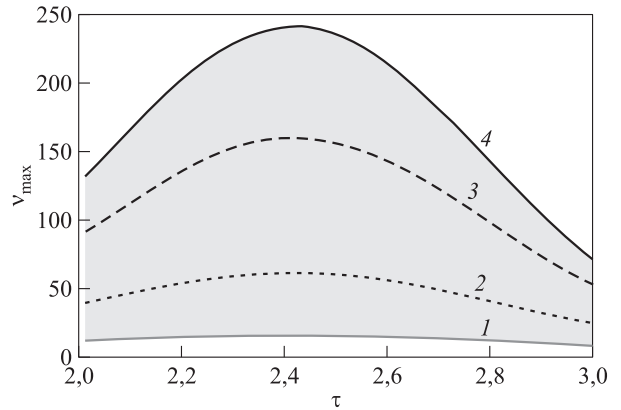


Рис. 4 Регрессионная зависимость (6) максимального объема дерева ν_{\max} от $\tau \in (2, 3]$ при фиксированных значениях N : 1 — $N = 100$; 2 — 1000; 3 — 5000; 4 — $N = 10\,000$

На рис. 3 и 4 все кривые зависимостей (5) и (6) при фиксированных значениях N лежат внутри затененных областей и расположены одна над другой по мере роста числа вершин графа в пределах граничных значений $100 \leq N \leq 10\,000$.

Таким образом, согласно полученным результатам, объем максимального дерева растет с увеличением размерности графа и с ростом значения параметра на интервале $\tau \in (1, 2]$, достигая максимальных значений при $\tau = 2$. На интервале $\tau \in (2, 3]$ вначале рост ν_{\max} продолжается, достигая своего максимума при $\tau \approx 2,4$, не превышая, однако, 250 вершин при $N = 10\,000$, и далее уменьшается с приближением к правой границе интервала.

3.2 Средний объем дерева

Следующей рассматривалась такая характеристика, как средний объем дерева, который для каждого графа вычислялся по формуле

$$\bar{\nu} = \frac{1}{N} \sum_{i=1}^N \nu_i.$$

Были построены регрессионные зависимости $\bar{\nu}$ от N и τ при $\tau \in (1, 2]$ и $\tau \in (2, 3]$:

$$\ln \bar{\nu} = 0,039 \ln N + 0,733\tau^2 - 1,079 \quad (R^2 = 0,98);$$

$$\ln \bar{\nu} = (-0,369\tau^2 + 1,745\tau - 1,745) \ln N \quad (R^2 = 0,95).$$

Эти модели показывают, что средний объем дерева ведет себя примерно так же, как и максимальный объем с ростом N и τ , достигая своих максимальных значений также при $\tau \approx 2,4$, однако сами значения среднего объема дерева значительно меньше и не превышают 20 вершин при $N = 10\,000$.

3.3 Число деревьев заданного объема

Пусть $r = 0, 1, 2, \dots$ — натуральное число, соответствующее объему дерева при произвольной вершине графа $i = \overline{1, N}$. Число деревьев объема r обозначим n_r . Объем дерева $r = 0$ соответствует случаю, когда корневая вершина либо имеет петлю, либо связана с любой другой вершиной графа кратными ребрами. Дерево объема $r = 1$ состоит только из корневой вершины.

В случае $\tau \in (1, 2]$ получена следующая регрессионная зависимость n_0 от объема графа N и параметра распределения степеней вершин $1 < \tau \leq 2$:

$$n_0 = 14,067 + \frac{0,186N}{\tau^{5,371}} \quad (R^2 = 0,99).$$

При $\tau \in (2, 3]$ для случаев $r = 0$ и 1 получены следующие регрессионные зависимости n_0 и n_1 от объема графа N и параметра τ :

$$n_0 = \frac{184,515 \ln N}{\tau^{7,279}};$$

$$\ln n_1 = 0,490 \ln N - 8,821 \ln \tau + 7,393$$

с коэффициентами детерминации 0,93 и 0,96 соответственно.

Далее были получены регрессионные зависимости n_r от r , объема графа N и параметра распределения степеней вершин τ для двух интервалов изменения параметра:

$$n_r = 9,754 + \frac{0,641N^{1,021}}{\tau^{1,866}r^{2,060}}, \quad r \geq 1, \quad \tau \in (1, 2]; \quad (7)$$

$$n_r = 2,959 + \frac{1,604 N^{0,994} \tau^{1,184}}{r^{3,148}}, \quad r \geq 2, \quad \tau \in (2, 3], \quad (8)$$

с коэффициентами детерминации 0,87 и 0,99 соответственно. На рис. 5 и 6 представлены графики зависимостей (7) и (8) при $N = 10\,000$ и некоторых фиксированных значениях r .

Из полученных результатов следует, что число деревьев заданного объема n_r увеличивается с ростом числа вершин графа и уменьшается с увеличением значения r . Что касается зависимости n_r от параметра распределения степеней вершин τ , то число деревьев n_r с ростом значения параметра сначала уменьшается на интервале $\tau \in (1, 2]$, а затем на интервале $\tau \in (2, 3]$ увеличивается.

В конфигурационных графах с параметром распределения степеней вершин $\tau \in (1, 2]$ больше всего деревьев объема $r = 1$, однако их число при $N = 10\,000$ не превышает 8000 вершин. С ростом значения r число деревьев n_r резко уменьшается.

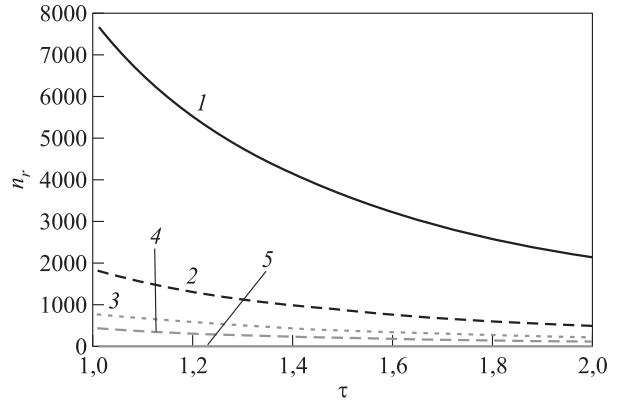


Рис. 5 Регрессионная зависимость (7) n_r — числа деревьев объема $r \geq 1$ от $\tau \in (1, 2]$ при $N = 10\,000$ при фиксированных значениях r : 1 — $r = 1$; 2 — 2; 3 — 3; 4 — 4; 5 — $r = 10$

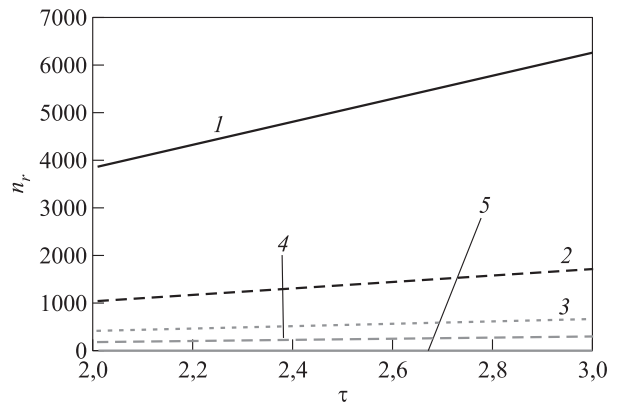


Рис. 6 Регрессионная зависимость (8) n_r — числа деревьев объема $r \geq 2$ от $\tau \in (2, 3]$ при $N = 10\,000$ при фиксированных значениях r : 1 — $r = 2$; 2 — 3; 3 — 4; 4 — 5; 5 — $r = 10$

Так, для графов объема $N = 10\,000$ при $r = 2$ оно не будет превышать 2000, при $r = 3$ — 1000, при $r = 5$ — 300, а при $r > 10$ число деревьев n_r становится меньше 100. При $\tau \in (2, 3]$ больше всего деревьев объема $r = 2$ и их число при $N = 10\,000$ не превышает 7000. На этом интервале n_r также уменьшается с ростом значения r . Для графов объема $N = 10\,000$ при $r = 3$ оно не будет превышать 2000, при $r = 4$ — 1000, при $r = 5$ — 400, а при $r > 10$ число деревьев n_r становится меньше 50.

4 Заключение

В результате проведенных исследований локальной древовидности конфигурационных графов впервые были построены модели зависимостей максимального и среднего объемов деревьев от числа вершин графа N и параметра τ распре-

ления (1). Также были построены модели зависимости числа деревьев заданного объема r от N , τ и r . Проведение экспериментов, необходимых для формирования таких моделей, требует значительных вычислительных ресурсов, поэтому рассматривались графы с ограничением $N \leq 10\,000$. Для максимального объема дерева проведено сравнение построенных моделей с асимптотическими результатами работы [7]. Оказалось, что асимптотическая модель (3) применима и в доасимптотической области, но предложенная модель (5) статистически значимо превосходит (3). Естественно поставить вопрос о нахождении достаточно большого объема графа, начиная с которого асимптотическая модель получит преимущество. Возможно, что ответ на этот вопрос можно будет получить в будущем, если использовать разработанную авторами имитационную модель и гораздо более мощную вычислительную технику.

Литература

1. *Faloutsos C., Faloutsos P., Faloutsos M.* On power-law relationships of the internet topology // *Comput. Commun. Rev.*, 1999. Vol. 29. P. 251–262. doi: 10.1145/316194.316229.
2. *Durrett R.* Random graph dynamics. — Cambridge: Cambridge University Press, 2007. 221 p. doi: 10.1017/CBO9780511546594.
3. *Hofstad R.* Random graphs and complex networks. — Cambridge: Cambridge University Press, 2017. Vol. 1. 337 p. doi: 10.1017/9781316779422.
4. *Hofstad R.* Random graphs and complex networks // *Notes RGCNII*, 2023. Vol. 2. 314 p. <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf>.
5. *Bollobas B.* A probabilistic proof of an asymptotic formula for the number of labelled regular graphs // *Eur. J. Combin.*, 1980. Vol. 1. Iss. 4. P. 311–316. doi: 10.1016/S0195-6698(80)80030-8.
6. *Reittu H., Norros I.* On the power-law random graph model of massive data networks // *Perform. Evaluation*, 2004. Vol. 55. Iss. 1-2. P. 3–23. doi: 10.1016/S0166-5316(03)00097-X.
7. *Павлов Ю. Л.* Максимальное дерево случайного леса в конфигурационном графе // *Математический сборник*, 2021. Т. 212. Вып. 9. С. 146–163. doi: 10.4213/sm9481.
8. *Павлов Ю. Л., Чеплюкова И. А.* Объемы деревьев случайного леса и конфигурационные графы // *Труды Математического института им. В. А. Стеклова*, 2022. Т. 216. С. 298–315. doi: 10.4213/tm4216.
9. *Pavlov Yu. L.* Random forests. — Utrecht: VSP, 2000. 122 p.

Поступила в редакцию 19.06.23

LOCAL TREELIKE STRUCTURE IN THE POWER-LAW CONFIGURATION GRAPHS

M. M. Leri and Yu. L. Pavlov

Institute of Applied Mathematical Research, Karelian Research Center of the Russian Academy of Sciences,
11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation

Abstract: The local treelike structure of configuration graphs intended for modeling complex communication networks is studied. In such graphs, the vertex degrees are independent and identically distributed according to the power law. In the case of a limited number of graph vertices, the dependences of the maximum volume of a treelike subgraph on the number of graph vertices and the vertex degree distribution parameter are found. The same problem was solved for the number of trees of a given size. Estimates are also given for the average size of a tree in the graph. It is shown that with a limited number of vertices of the configuration graph, the found dependences statistically significantly improve the description of the network structure in comparison with the previously known asymptotic models.

Keywords: configuration graph; power-law distribution; local treelike structure; tree size; simulations

DOI: 10.14357/19922264240107

EDN: ITTZE0

Acknowledgments

The study was carried out under state order to the Karelian Research Center of the Russian Academy of Sciences (Institute of Applied Mathematical Research KarRC RAS).

References

1. Faloutsos, C., P. Faloutsos, and M. Faloutsos. 1999. On power-law relationships of the internet topology. *Comput. Commun. Rev.* 29:251–262. doi: 10.1145/316194.316229.
2. Durrett, R. 2007. *Random graph dynamics*. Cambridge: Cambridge University Press. 221 p. doi: 10.1017/CBO9780511546594.
3. Hofstad, R. 2017. *Random graphs and complex networks*. Cambridge: Cambridge University Press. Vol. 1. 337 p. doi: 10.1017/9781316779422.
4. Hofstad, R. 2023. Random graphs and complex networks. *Notes RGCNII*. Vol. 2. 314 p. Available at: <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf> (accessed April 17, 2023).
5. Bollobas, B. 1980. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Combin.* 1(4):311–316. doi: 10.1016/S0195-6698(80)80030-8.
6. Reittu, H., and I. Norros. 2004. On the power-law random graph model of massive data networks. *Perform. Evaluation.* 55(1-2):3–23. doi: 10.1016/S0166-5316(03)00097-X.
7. Pavlov, Yu. L. 2021. The maximum tree of a random forest in the configuration graph. *Sb. Math.* 212(9):1329–1346. doi: 10.1070/SM9481.
8. Pavlov, Yu. L., and I. A. Cheplyukova. 2022. Sizes of trees in a random forest and configuration graphs. *P. Steklov Inst. Math.* 316:280–297. doi: 10.1134/S0081543822010205.
9. Pavlov, Yu. L. 2000. *Random forests*. Utrecht: VSP. 122 p.

Received June 19, 2023

Contributors

Leri Marina M. (b. 1969) — Candidate of Science (PhD) in technology, scientist, Institute of Applied Mathematical Research of the Karelian Research Center of the Russian Academy of Sciences, 11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation; leri@krc.karelia.ru

Pavlov Yuri L. (b. 1949) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Applied Mathematical Research of the Karelian Research Center of the Russian Academy of Sciences, 11 Pushkinskaya Str., Petrozavodsk 185910, Russian Federation; pavlov@krc.karelia.ru

СТАТИСТИЧЕСКИЙ КРИТЕРИЙ СТАБИЛЬНОСТИ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ, ОСНОВАННЫЙ НА ВХОДНОМ И ВЫХОДНОМ ПОТОКАХ

М. П. Кривенко¹

Аннотация: Одним из базовых свойств системы массового обслуживания (СМО) считается стабильность — способность системы функционировать, сохраняя во времени неизменной свою структуру и характеристики. Рассматривается проблема статистической проверки стабильности СМО на основе характеристик входного $A(t)$ и выходного $D(t)$ потоков заявок. Подтверждение стабильности основывается на установлении равенства интенсивностей этих потоков. Таким образом на языке статистического анализа данных получаем классическую задачу сравнения интенсивностей потоков. Для её решения период наблюдений делится на фрагменты, которые дают отдельные оценки. Вместе они составляют выборку, которая участвует в сравнении. При анализе стабильности необходимо учитывать возможную зависимость $A(t)$ и $D(t)$, поэтому приходится обращаться к методам обработки так называемых согласованных пар (matched pairs) наблюдений. Контроль стабильности вынуждает решать ряд вспомогательных задач: подбор объемных параметров оценивания интенсивностей, проверку нормальности распределения, анализ корреляций. В ходе экспериментов с реальной системой выявлен ряд особенностей: эффект подмены допредельного распределения реальным при фрагментировании, наличие зависимости оценок интенсивностей анализируемых потоков, сходящая на нет для нестабильных систем.

Ключевые слова: СМО (система массового обслуживания); стабильность выборочной траектории; критерии для согласованных пар; тесты мультиформальности

DOI: 10.14357/19922264240108

EDN: JNJMU

1 Введение

Одним из базовых свойств СМО считается стабильность — способность системы функционировать, сохраняя во времени неизменной свою структуру и характеристики.

Статья посвящена постановке задачи статистического контроля стабильности функционирования систем обслуживания как задачи анализа данных при предположении общего характера, когда исчерпываются возможности аналитических методов исследования системы, а также выбору соответствующих методов вместе с выявлением возможных проблем их применения.

Термин «стабильность» имеет много значений в различных контекстах. Подобная расплывчатость позволяет охватить ряд сходных, но далеко не идентичных поведений исследуемых процессов и рассматриваемых моделей, а также решать общие вопросы структуры системы и устойчивости ее функционирования в удивительно широких условиях. С практической точки зрения система, которая стабильна в соответствии с одним определением, скорее всего, будет стабильной в соответствии с другим определением, но не надо забывать,

что нечеткость определений стабильности, нюансы описания принятых моделей, различия в способах получения результатов могут приводить к несхожим условиям стабильности (для примера можно сравнить [1] и [2]).

Далее обсуждается проблема статистического контроля стабильности СМО на основе характеристик входных и выходных потоков заявок. При этом содержание и организация процесса обслуживания либо неизвестны, либо известны и допускают возможность воспроизведения в виде статистического моделирования, но не доступны с точки зрения теоретических исследований. Для каждого неотрицательного времени t пусть $A(t)$ обозначает число поступивших заявок (входной процесс), а $D(t)$ — покинувших систему (выходной процесс). Тогда общее число заявок в системе $N(t)$ равно

$$N(t) = N(0) + A(t) - D(t).$$

Предполагается доступность реализации перечисленных процессов для заданного интервала времени $[0, T]$.

В последние годы со стороны исследователей и специалистов, работающих в области стохастических процессов и их приложений, наблюдается

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, mkcrivenko@ipiran.ru

значительный интерес к ориентированному на выборку анализу. Изучение свойств одной выборочной траектории (sample path — SP) стохастического процесса часто приводит к лучшему и более глубокому пониманию свойств исследуемой системы. Оно также предоставляет специалистам-практикам мощный инструмент для определения того, какие свойства данной системы не зависят от обычно навязываемых вероятностных предположений.

2 Определение стабильности

В литературе были предложены различные определения стабильности, в основном в контексте стохастических моделей (см., например, обзор этих определений в [3, приложение С]). Особое место занимают определения, которые не зависят от конкретных стохастических предположений и формируются в рамках SP-анализа. Следуя разд. 5 работы [3], дадим

Определение. Входной/выходной процесс $N(t)$, $t \geq 0$, называется R-стабильным (Rate-стабильность), если

$$t^{-1}N(t) \rightarrow 0 \text{ при } t \rightarrow \infty.$$

Из определения процесса $N(t)$ следует: при условии, что $t^{-1}A(t) \rightarrow \lambda_A < \infty$, $t \rightarrow \infty$, он R-стабилен тогда и только тогда, когда $t^{-1}D(t) \rightarrow \lambda_D < \infty$ и $\lambda_D = \lambda_A$.

Таким образом, стабильность присутствует, если долгосрочные скорости поступления и убытия заявок существуют, конечны и равны, т. е. находятся в равновесии. Можно сформулировать ряд эквивалентных постановок задачи подтверждения стабильности некоторой СМО, из которых как основной далее рассматривается вариант анализа $A(t)$ и $D(t)$. После придания стохастического характера задаче установления стабильности появляется возможность оценить обоснованность принимаемых решений. При этом на языке статистического анализа данных получаем классическую задачу сравнения интенсивностей потоков.

Предполагается, что реализации процессов $A(t)$ и $D(t)$ доступны в любом требуемом объеме. Но если ограничения на процесс поступления заявок обычно входят в условия задачи контроля стабильности и их можно считать заданными, то относительно существования скорости убытия заявок и ее конечности приходится рассматривать два случая: характеристики выходного процесса известны или нет. Априорные данные появляются, когда исследуемая СМО изучена теоретически достаточно полно и факт существования и конечности предела $t^{-1}D(t)$ при $t \rightarrow \infty$ установлен. Здесь речь

идет о классических результатах теоремы Берка [4], о примерах из обзора типа [5], а также о множестве отдельных публикаций. Более интересным с прикладной точки зрения представляется случай, когда полной информации об интенсивностях выходного потока нет. В этом случае анализ стабильности системы необходимо предварять исследованием поведения $A(t)$ и $D(t)$ при растущих значениях t .

Здесь сразу же надо пояснить: если есть возможность проводить наблюдения сколь угодно долго, то сделать вывод о стабильности можно на основании того, что обычно асимптотически распределение оценки интенсивности $\tilde{\lambda}_B$ для некоторого процесса B оказывается нормальным и сходимость относительно числа событий n имеет вид:

$$\sqrt{n} \left(\tilde{\lambda}_B - \lambda_B \right) \xrightarrow{d} N(0, \sigma_B^2).$$

Тогда за счет выбора n всегда можно однозначно определиться относительно значимого отличия интенсивностей входного и выходного потоков. Данный результат остается в силе по крайней мере для стационарных и эргодических процессов [6]. Поэтому сравнение интенсивностей потоков представляет интерес с учетом сокращения объема использованного статистического материала.

Пожалуй, единственное систематическое рассмотрение соответствующих методов проведено в уже давнишней работе [7]. Однако использовать их непосредственно не получается. Во-первых, достаточно подробный разбор сравнения пуассоновских процессов малоинтересен: если подобная ситуация и возникает при выявлении стабильности, то для тривиальных СМО; поэтому ее предназначение — верификация разрабатываемого программного обеспечения сопоставления выборок. Непуассоновский случай проанализирован кратко, причем для процессов восстановления, а также, что очень важно в случае анализа входных и выходных потоков, в русле задачи проверки однородности двух независимых выборок, что не обязательно встречается реально. Хотя сами по себе методы, опирающиеся на большие выборки, в совокупности с понятием R-стабильности, носящим асимптотический характер, могут стать плодотворными. В этой связи важны результаты разд. 4.5 и 4.6 работы [7], справедливые для многих стационарных точечных процессов при $t \rightarrow \infty$. Если некоторый стационарный целочисленный процесс обозначить как $B(t)$, а его характеристики второго порядка через $M_B(t)$ и $V_B(t)$, то для индекса дисперсии $I_B(t) = V_B(t)/M_B(t)$ получим асимптотическое равенство $I_B(t) \simeq I_B$. В результате приходим к практически важным следствиям:

- (1) величину I_B можно оценить по асимптотическому значению угла наклона дисперсионной кривой, которую саму по себе можно использовать как средство визуального контроля правильности принятых предположений;
- (2) если асимптотическое равенство справедливо с хорошим приближением для всех t , больших некоторого t' , то численности событий в неперекрывающихся интервалах длины t' практически независимы;
- (3) если за продолжительный период наблюдения длины t произошли n событий, то оценкой интенсивности будет $\tilde{\lambda}_B = n/t$, а $\text{var}(\tilde{\lambda}_B)$ — величина $I_B \tilde{\lambda}_B / t \simeq I_B n / t^2$.

Если значение I_B известно, например из результатов предварительного теоретического или эмпирического исследования СМО с помощью следствия (1), то согласно следствию (3) для конкретной последовательности данных интенсивность наступления событий можно оценить обычным образом, используя далее для статистического вывода нормальную аппроксимацию с соответствующей дисперсией.

Если значение I_B неизвестно, то приходится обращаться к следствиям (2) и (3) в предположении, что тем или иным путем установлено значение t' . Далее каждая последовательность разбивается на смежные интервалы длины t' , для которых подсчитываются численности событий в каждом из них. В результате приходим к задаче сравнения средних значений из двух выборок с возможностью оценивания их точности $V(t')$.

Обе рассмотренные ситуации со значением I_B фактически отличаются друг от друга тем, что выделяется или нет в виде отдельной задачи оценивание точности $\tilde{\lambda}_B$. В остальном определенный произвол приемов формирования распределений при сравнении интенсивностей остается, что не придает уверенности исследователю. Не менее важно и другое: при анализе стабильности необходимо учитывать возможную зависимость входящего и выходящего потоков. По этой причине приходится обращаться к методам обработки так называемых согласованных пар (matched pairs) наблюдений, естественно возникающих при исследованиях входного потока заявок до и выходящего потока после обработки в СМО. Для анализа различий d_i в каждой из M пар можно использовать ряд критериев [8, с. 4579–4583, 7726–7730, 9150–9153], но ни один из них не кажется безусловно лучшим. Принимая во внимание, что объектами служат оценки средних, т. е. данные распределены, скорее всего, нормально, предпочтение из перечисленных следует отдать t -критерию.

Пусть (x_{i1}, x_{i2}) — i -е наблюдаемое значение сопоставленной пары $(\tilde{\lambda}_A, \tilde{\lambda}_D)$, $i = \overline{1, M}$. Тогда $d_i = x_{i2} - x_{i1}$ дают оценку математического ожидания Δ отличия интенсивности выходящего потока от интенсивности входного. Нулевая гипотеза о стабильности СМО заключается в том, что $\Delta = 0$, конкурирующая — $\Delta < 0$.

Статистика t -критерия с $M - 1$ степенями свободы имеет вид:

$$t_{MP} = (S_d)^{-1} \sqrt{M} \bar{d},$$

где

$$\bar{d} = M^{-1} \sum_{i=1}^M d_i; \quad S_d^2 = (M - 1)^{-1} \sum_{i=1}^M (d_i - \bar{d})^2.$$

Вводя обычным образом выборочный коэффициент корреляции r между x_1 и x_2 , получаем тождество

$$S_d^2 = S_{x_1}^2 - 2r S_{x_1} S_{x_2} + S_{x_2}^2$$

и иной эквивалентный вид статистики t_{MP} , т. е. описанный критерий включает коэффициент корреляции, характеризующий связь выходящего и входного потоков. Поэтому интерес вызывает его поведение в зависимости от стабильного или нестабильного типа СМО, а также становится актуальным описание выборочных свойств этой статистики. Если за основу принять обычную оценку коэффициента корреляции, то для нормализации ее дисперсии применяется преобразование Фишера [8, с. 1375–1385] вида

$$z = \frac{1}{2} \ln \frac{1+r}{1-r}.$$

Тогда для больших значений M при теоретическом значении корреляции ρ имеем:

$$Z \sim N \left(\frac{1}{2} \ln \frac{1+\rho}{1-\rho}, \frac{1}{M-3} \right).$$

Если Q — квантиль стандартного нормального распределения уровня $0,5(1 - \alpha)$, то сразу же можно получить доверительный интервал уровня α для среднего значения Z .

Традиционно используемое z -преобразование практично, поскольку доверительный интервал может быть вычислен на основе стандартного нормального распределения и он демонстрирует достаточную точность уже при десятке наблюдений. Но не надо забывать о наличии иных результатов [9], включающих точные.

Надо обратить внимание, что ниоткуда не следует, что совместное распределение оценок интенсивностей асимптотически также будет нормальным. На важность основополагающих предположений

при анализе корреляции нормально распределенных наблюдений обращено внимание в [10], а в случае статистического контроля функционирования СМО общего характера вряд ли можно без колебаний принять предположение о совместном нормальном распределении оценок интенсивностей входного и выходного потоков, особенно если система представляет собой «черный ящик». Поэтому так важны процедуры тестирования многомерной нормальности результатов наблюдений входного и выходного потоков СМО одновременно. В этой связи к фундаментальным работам следует отнести [11], примером сравнительного анализа подходов служит [12], специфика двухмерного случая отражена в [13].

Множественность методов вызвана сложностью постановки задачи (содержательная асимметричность альтернативы, многомерность данных, ограничение на маргинальные распределения, опора на выборочные характеристики), которая приводит к их перманентным модификациям и к необходимости соблюдать осторожность при применении.

Так, тестирование мультинормальности [8, с. 5047–5051] на основе мер асимметрии и эксцесса обладает отягощающим фактором: асимптотические распределения соответствующих статистик при нулевой гипотезе можно использовать корректно только при достаточно больших объемах данных, причем асимптотика специфична для разных показателей формы распределения статистик.

Подобные проблемы обычно решаются методом статистических испытаний, что неприемлемо при автоматизированном контроле функционирования СМО в меняющейся обстановке. Более успешной оказывается попытка точной аппроксимации предельных распределений вкуче с объединением частных критериев мультинормальности в виде сводного теста на основе коэффициентов асимметрии и эксцесса. Но здесь может проявиться так называемый эффект усложнения текста [14], затрудняющий понимание и реализацию предлагаемых процедур. В качестве соответствующего примера приведем один из вариантов критерия мультинормальности. Пусть $\tilde{\mathbf{X}}^T$ — $(p \times N)$ -матрица N центрированных наблюдаемых значений p -вектора с выборочной ковариационной матрицей

$$\mathbf{C} = \frac{1}{N} \tilde{\mathbf{X}}^T \tilde{\mathbf{X}}$$

и корреляционной матрицей

$$\mathbf{R} = \mathbf{V}^{-1/2} \mathbf{C} \mathbf{V}^{-1/2},$$

где $\mathbf{V} = \text{diag}(\hat{\sigma}_1^2, \dots, \hat{\sigma}_p^2)$. Согласно [15] определим $(p \times N)$ -матрицу \mathbf{Y}^T преобразованных наблюдений

$$\mathbf{Y}^T = \mathbf{H} \mathbf{\Lambda}^{-1/2} \mathbf{H}^T \mathbf{V}^{-1/2} \tilde{\mathbf{X}}^T,$$

где фигурируют элементы спектрального разложения корреляционной матрицы $\mathbf{R} = \mathbf{H} \mathbf{\Lambda} \mathbf{H}^T$, $\mathbf{H}^T \mathbf{H} = \mathbf{I}$; $\mathbf{\Lambda}$ — диагональная матрица с соответствующими собственными значениями на главной диагонали. Тогда в условиях нулевой гипотезы данные из многомерного нормального распределения преобразуются в совокупность независимых величин со стандартным нормальным распределением. Для каждой из получившихся переменных можно подсчитать коэффициенты асимметрии и эксцесса, которые в совокупности дадут статистику сводного критерия. Несмотря на то что в случае сравнения потоков речь идет о $p = 2$, актуальным остается уточнение этого подхода для случая, когда ранг матрицы \mathbf{R} может оказаться меньше предполагаемого значения p .

Искусственность усложнения в [15] заключается в том, что нет необходимости перехода к стандартным распределениям. Достаточно обратиться к первым главным компонентам, а затем применить набор составных критериев значимости. При этом в ходе перехода совершенно естественно решается задача снижения размерности.

3 Эксперименты

Рассмотрим задачу контроля стабильной работы двухпроцессорной системы обработки заданий со случайным выбором числа требуемых процессоров [2]. Функционирование подобной системы $M/M/2$ определяется параметрами: λ — интенсивность поступления заданий на обработку; p_1 — вероятность того, что для выполнения задания требуется один процессор; μ — среднее время обработки задания процессором/процессорами. Согласно ссылке в разд. 3 работы [2], условие стабильности функционирования этой системы принимает вид

$$\frac{\lambda}{\mu} < \frac{2}{2 - p_1^2},$$

с помощью которого будет осуществляться выбор моделируемых систем: параметры p_1 и μ принимаются базовыми со значениями $p_1 = 0,5$ и $\mu = 1$, а λ выбирается как $C_\lambda(2\mu/(2 - p_1^2))$, где C_λ — индикатор стабильности работы системы (при $C_\lambda < 1$ система стабильна).

В качестве объекта статистического анализа выступают входной и выходной процессы, время наблюдения за которыми равно T . Этот период для учета особенностей функционирования СМО делится на N_s последовательных (идущих одна за другой) фаз одинаковой длительности T_s . Каждая

Результаты проверки нулевых гипотез об адекватности модели (строки помечены буквой «М») и о стабильности (строки помечены буквой «S»): «+» — принятие; «±» — принятие с отдельными исключениями в многокритериальном случае; «-» — отвержение

Вариант СМО		Номер фазы									
		1	2	3	4	5	6	7	8	9	10
$C_\lambda = 0,95$, стабильна	M	+	+	+	±	±	+	+	+	±	+
	S	+	+	+	+	+	+	+	+	+	+
$C_\lambda = 1,05$, нестабильна	M	±	+	+	+	±	+	+	±	±	+
	S	-	-	-	-	-	-	-	-	-	-

фаза включает N_f фрагментов одинаковой продолжительности T_f (то же самое, что и t' в [7]), которые служат для оценивания интенсивности, их насыщенность данными должна обеспечивать возможность корректной аппроксимации распределения оценок. По совокупности фрагментов в фазе строится статистический вывод о свойствах наблюдаемых процессов, рост числа фрагментов повышает мощность критерия, применяемого к данным фазы.

Центральным становится задание объемных параметров T_f и N_f , при выборе которых встречается неожиданная на первый взгляд ситуация: при одной и той же длине фрагмента (например, для простейшего потока событий, когда во фрагменте в среднем 50 событий) оценка вероятности отклонения нулевой гипотезы с помощью составного критерия при увеличении числа фрагментов ($N_f = 20, 50, 100$ и 200) растет, начиная с предполагаемого значения ошибки первого рода в 5%, а не остается постоянной ($\alpha^* = 5,1\%, 5,8\%, 7,0\%$ и $10,0\%$ соответственно). Здесь проявляется эффект подмены (замещения, подстановки), когда проверка нормальности ориентирована на предельное нормальное распределение и не зависит от объема выборки, а применяется к данным с фактически иным распределением. Поэтому пока объемы T_f и N_f согласуются, используемый критерий ведет себя ожидаемо, но с ростом числа наблюдений N_f проявляется отклонение фактического распределения от предельного. Это означает, что подбор T_f и N_f надо осуществлять только совместно.

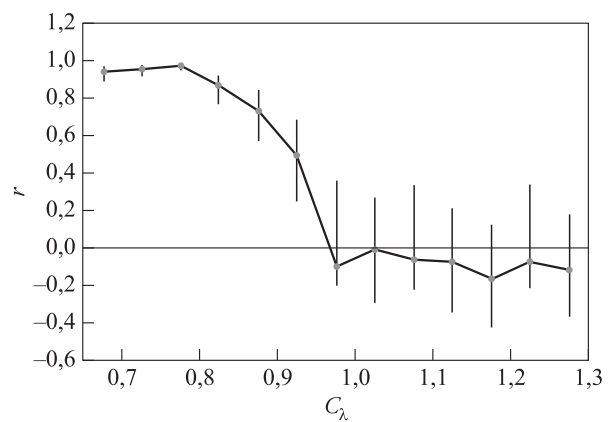
Для каждой фазы проводились:

- контроль основных характеристик $A(t)$ и $D(t)$ (интенсивность, среднее число запросов на обработку во фрагменте);
- проверка адекватности модели оценивания интенсивностей $A(t)$ и $D(t)$, включающая корреляционный анализ (наличие взаимной связи и автокорреляция) и тестирование мультинормальности и нормальности в отдельности по каждому процессу;
- решение о стабильности анализируемой СМО.

Результаты моделирования, представленные в таблице, позволяют сделать следующие выводы:

- принятые предположения (строки таблицы, помеченные буквой «М») выполняются за редкими исключениями, число которых не противоречит статистическим принципам принятия решений; все отклонения от ожидаемых результатов соответствовали критериям мультинормальности;
- критерий стабильности (строки «S») дает ожидаемые результаты для двух типов СМО;
- обнаруживается одна особенность оценок интенсивности, когда для входного и выходного процессов корреляционные связи проявляются в случае стабильных систем и исчезают для нестабильных.

Для детализации последнего вывода была оценена зависимость коэффициента корреляции r между фрагментарными оценками интенсивности процессов $A(t)$ и $D(t)$ от индикатора стабильности



Зависимость коэффициента корреляции r между фрагментарными оценками интенсивности процессов $A(t)$ и $D(t)$ от индикатора стабильности работы C_λ в виде линии, соединяющей отдельные оценки r , и соответствующих доверительных интервалов в виде вертикальных отрезков

работы C_λ , которая в виде графика отдельных значений r и соответствующих 95% доверительных интервалов дана на рисунке (взята 4-я фаза). Из него видно, что для стабильных систем (значения $C_\lambda < 1$) корреляционная связь присутствует, но нет веских оснований предполагать ее наличие для нестабильных систем (принадлежность нуля соответствующим доверительным интервалам подтверждается результатами проверки нулевой гипотезы о нулевой корреляции). Получается, что для нестабильных систем, когда нарушается равновесие их функционирования, выход «теряет связь» со входом и начинает зависеть только от характеристик обработки заданий.

4 Заключение

Критерий стабильности СМО наиболее востребован при построении и исследовании малоизученных систем, для которых известны частично или не описаны вообще следующие свойства: теоретические и выборочные характеристики входного и выходного потоков, наличие и характер статистических связей между ними, свойства и возможности применения процедур анализа данных, интерпретация получаемых результатов. Решения приходится искать среди известных моделей и методов, что усиливает роль задач проверки адекватности принятых предположений, приводит к смещению акцентов при исследованиях на технологии обработки информации со всеми обычными проблемами: соблюдение корректности; преодоление усложнений из-за комплексного характера возникающих задач; обеспечение эффективности обработки. Поэтому такими важными представляются дальнейшие исследования по порядку применения сформированных методов анализа данных и пополнению набора базовых процедур обработки информации.

Литература

1. *Green L.* A queueing system in which customers require a random number of servers // *Oper. Res.*, 1980. Vol. 28. No. 6. P. 1335–1346. doi: 10.1287/opre.28.6.1335.
2. *Brill P. H., Green L.* Queues in which customers receive simultaneous service from a random number of servers: A system point approach // *Manage. Sci.*, 1984. Vol. 30. No. 1. P. 51–68. doi: 10.1287/mnsc.30.1.51.
3. *El-Taha M., Stidham S., Jr.* Sample-path analysis of queueing systems. — New York, NY, USA: Springer Science + Business Media, 1999. 302 p. doi: 10.1007/978-1-4615-5721-0.
4. *Burke P. J.* The output of a queueing system // *Oper. Res.*, 1956. Vol. 4. No. 6. P. 699–704. doi: 10.1287/opre.4.6.699.
5. *Daley D. J.* Queueing output processes // *Adv. Appl. Probab.*, 1976. Vol. 8. No. 2. P. 395–415. doi: 10.2307/1425911.
6. *Karr A. F.* Point processes and their statistical inference. — 2nd ed. — New York, NY, USA: Marcel Dekker, 1991. 512 p.
7. *Кокс Д., Льюис П.* Статистический анализ последовательностей событий / Пер. с англ. — М.: Мир, 1969. 312 с. (*Cox D. R., Lewis P. A. W.* The statistical analysis of series of events. — New York, NY, USA: John Wiley, 1966. 285 p.)
8. *Kotz S., Read C. B., Balakrishnan N., Vidakovic B.* Encyclopedia of statistical sciences: 16 vol. set. — 2nd ed. — Hoboken, NJ, USA: Wiley, 2006. 9686 p.
9. *Taraldsen G.* The confidence density for correlation // *Sankhya Ser. A*, 2023. Vol. 85. Part 1. P. 600–616. doi: 10.1007/s13171-021-00267-y.
10. *Levy K. J.* Non-normality and testing that a correlation equals zero // *Educ. Psychol. Meas.*, 1977. Vol. 37. Iss. 3. P. 691–694.
11. *Mardia K. V.* Tests of univariate and multivariate normality // *Handbook of statistics* / Ed. R. Krishnaiah. — Amsterdam: North-Holland Publishing Co., 1980. Vol. 1. P. 279–320. doi: 10.1016/S0169-7161(80)01011-5.
12. *Romeu J. L., Ozturk A.* A comparative study of goodness-of-fit tests for multivariate normality // *J. Multivariate Anal.*, 1993. Vol. 46. Iss. 2. P. 309–334. doi: 10.1006/jmva.1993.1063.
13. *Versluis C.* Comparison of tests for bivariate normality with unknown parameters by transformation to an univariate statistic // *Commun. Stat. — Theor. M.*, 1996. Vol. 25. Iss. 3. P. 647–665. doi: 10.1080/03610929608831719.
14. *Дьякова А. А.* Усложнение текста: структурно-семантический аспект // *Известия ВГПУ. Филологические науки*, 2016. № 1(105). С. 158–165. EDN: VRWUIV.
15. *Doornik J., Hansen H.* An omnibus test for univariate and multivariate normality // *Oxford B. Econ. Stat.*, 2008. Vol. 70. Iss. S1. P. 927–939. doi: 10.1111/j.1468-0084.2008.00537.x.

Поступила в редакцию 29.11.23

STATISTICAL CRITERION FOR QUEUING SYSTEM STABILITY BASED ON INPUT AND OUTPUT FLOWS

M. P. Krivenko

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: One of the basic properties of a queuing system is stability — the ability of the system to function, maintaining its structure and characteristics unchanged over time. The problem of statistical verification of the stability of the queuing system based on the characteristics of the input $A(t)$ and output $D(t)$ order flows is considered. The confirmation of stability is based on establishing the equality of the rates of these flows. Thus, in the language of statistical data analysis, one obtains the classic problem of comparing rates of occurrence. To solve it, the observation period is divided into fragments that give separate estimates. Together, they make up the sample that participates in the comparison. When analyzing stability, it is necessary to take into account possible dependence of $A(t)$ and $D(t)$; so, it is necessary to turn to methods for processing the so-called matched pairs of observations. Stability control makes it necessary to solve a number of auxiliary tasks: selection of volumetric parameters for rate estimation, checking the normality of the distribution, and analysis of correlations. In the course of experiments with the real system, a number of features were revealed: the effect of substituting the prelimit distribution with the real one during fragmentation as well as the presence of dependence of the rate estimates of analyzed flows which comes to naught for unstable systems.

Keywords: queueing system; sample-path stability; matched pairs tests; tests of multinormality

DOI: 10.14357/19922264240108

EDN: JNJJMU

References

- Green, L. 1980. A queueing system in which customers require a random number of servers. *Oper. Res.* 28(6):1335–1346. doi: 10.1287/opre.28.6.1335.
- Brill, P. H., and L. Green. 1984. Queues in which customers receive simultaneous service from a random number of servers: A system point approach. *Manage. Sci.* 30(1):51–68. doi: 10.1287/mnsc.30.1.51.
- El-Taha, M., and S. Stidham, Jr. 1999. *Sample-path analysis of queueing systems*. New York, NY: Springer. 302 p. doi: 10.1007/978-1-4615-5721-0.
- Burke, P. J. 1956. The output of a queueing system. *Oper. Res.* 4(6):699–704. doi: 10.1287/opre.4.6.699.
- Daley, D. J. 1976. Queueing output processes. *Adv. Appl. Probab.* 8(2):395–415. doi: 10.2307/1425911.
- Karr, A. F. 1991. *Point processes and their statistical inference*. 2nd ed. New York, NY: Marcel Dekker. 512 p.
- Cox, D. R., and P. A. W. Lewis. 1966. *The statistical analysis of series of events*. New York, NY: John Wiley. 285 p.
- Kotz, S., C. B. Read, N. Balakrishnan, and B. Vidakovic. 2006. *Encyclopedia of statistical sciences*. 16 vol. set. 2nd ed. Hoboken, NJ: Wiley. 9686 p.
- Taraldsen, G. 2023. The confidence density for correlation. *Sankhya Ser. A* 85(1):600–616. doi: 10.1007/s13171-021-00267-y.
- Levy, K. J. 1977. Non-normality and testing that a correlation equals zero. *Educ. Psychol. Meas.* 37(3):691–694.
- Mardia, K. V. 1980. Tests of univariate and multivariate normality. *Handbook of statistics*. Ed. R. Krishnaiah. Amsterdam: North-Holland Publishing Co. 1:279–320. doi: 10.1016/S0169-7161(80)01011-5.
- Romeu, J. L., and A. Ozturk. 1993. A comparative study of goodness-of-fit tests for multivariate normality. *J. Multivariate Anal.* 46(2):309–334. doi: 10.1006/jmva.1993.1063.
- Versluis, C. 1996. Comparison of tests for bivariate normality with unknown parameters by transformation to an univariate statistic. *Commun. Stat. — Theor. M.* 25(3):647–665. doi: 10.1080/03610929608831719.
- D'yakova, A. A. 2016. Uslozhenie teksta: strukturno-semanticheskiy aspekt [Complication of a text: Structural and semantic aspect]. *Izvestiya VGPU. Filologicheskie nauki* [Izvestia of the Volgograd State Pedagogical University. Philological Sciences] 1(105):158–165. EDN: VRWUIV.
- Doornik, J., and H. Hansen. 2008. An omnibus test for univariate and multivariate normality. *Oxford B. Econ. Stat.* 70(S1):927–939. doi: 10.1111/j.1468-0084.2008.00537.x.

Received November 29, 2023

Contributor

Krivenko Michail P. (b. 1946) — Doctor of Science in technology, professor, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; mkrivenko@ipiran.ru

МОДЕЛИ СИНТЕЗА КОМПОНОВОЧНОЙ СХЕМЫ В ЗАДАЧЕ ФОРМИРОВАНИЯ ОБЛИКА САМОЛЕТА

Л. Л. Вышинский¹, Ю. А. Флёров²

Аннотация: Рассматриваются задачи синтеза математических моделей самолета на начальной стадии проектирования. Центральная задача начальной стадии проектирования состоит в разработке структурно-параметрической модели (СПМ) изделия и построении его компоновочной схемы. Структурно-параметрическая модель изделия описывает пространство основных конструктивных параметров самолета, а компоновочная схема представляет его структуру, множество основных агрегатов и систем летательного аппарата (ЛА), их взаимное расположение и конструктивные связи между агрегатами. Компоновочная схема самолета служит основой для дальнейшей детализации проекта. На начальной стадии проектирования определяются допустимые диапазоны значений конструктивных параметров агрегатов компоновочной схемы, в которых осуществляется поиск оптимальных конструктивных решений. Представлено описание СПМ и компоновочных схем широкого класса самолетов, которые могут быть реализованы средствами автоматизированной системы весового проектирования.

Ключевые слова: математическое моделирование; автоматизация проектирования; самолет; весовое проектирование; генератор проектов

DOI: 10.14357/19922264240109

EDN: DSPGKV

1 Введение

В работе [1] была представлена автоматизированная система весового проектирования самолетов (АСВП). Задачи весового проектирования встают на всех этапах жизненного цикла (ЖЦ) самолетов, начиная с формирования облика, эскизного и рабочего проектирования вплоть до изготовления серийных образцов и их эксплуатации [2]. Естественно, что на разных этапах необходимы разные модели и разные средства их синтеза и анализа. В настоящей статье более подробно описаны задачи, связанные с синтезом формирования облика, поскольку именно на этом этапе принимаются основные технические и проектные решения.

Ситуация на начальных стадиях проектирования самолетов может быть совершенно различной в зависимости от характера разработки — либо это модификация некоторого серийного или экспериментального самолета, либо новая разработка и создание самолета принципиально нового поколения. Поэтому каждый раз перед проектировщиком встает задача выбора моделей, которые бы позволили дать формальное описание конструкции проектируемого изделия. Широкое применение автоматизированных систем проектирования, САД-систем, систем геометрического моделирования позволяет создавать уже на ранних этапах проектирования сложные геометрические 3D-модели современных ЛА. Однако построению 3D-моде-

лей обычно предшествует большой объем расчетов, связанных с выбором структуры проектируемого изделия, состава его основных агрегатов, оптимизацией значений их конструктивных параметров и характеристик. Такие расчеты должны опираться на адекватное описание проектируемого изделия в виде СПМ, разработка которой и есть основная задача начального этапа проектирования самолета. Классическая схема формирования облика самолета предполагает решение следующих задач [3, 4]:

- (1) анализ тактико-технического задания (ТТЗ) к проектируемому изделию, поиск возможных прототипов и выбор значений основных конструктивных параметров на основании анализа опыта разработки самолетов, близких по назначению и уровню требований ТТЗ, и, если это возможно, выбор реальных прототипов для будущего проекта;
- (2) на основе проведенного анализа и выбранных прототипов синтез компоновочной схемы самолета;
- (3) анализ компоновочной схемы, оценка выполнения требований ТТЗ и оптимизация параметров компоновочной схемы.

Анализ ТТЗ и поиск прототипов проектируемого самолета должен быть обеспечен наличием представительной информации о серийных и опытных самолетах текущего и прошлых поколений. В АСВП

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, wyshinsky@mail.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, fler@ccas.ru

информация для поиска прототипов может накапливаться и храниться в базе данных СПМ (далее БД) готовых изделий.

2 База данных структурно-параметрических моделей готовых изделий

Ведение БД готовых изделий — это отдельная специальная функция АСВП. Информация для пополнения этой БД, как правило, берется из внешних источников. Опыт работы с такой информацией показывает, что основная проблема в использовании публичных данных состоит в необходимости приведения их к единой, непротиворечивой структуре и терминологии. Для этого в АСВП реализованы механизмы каталогизации и классификации информационных объектов, которые пользователь может настраивать в соответствии с решаемыми задачами.

На более высоком уровне одним из механизмов упорядочения информации становится создание базовых, **типовых** СПМ, с тем чтобы унифицировать состав, обозначения и определения конструктивных параметров изделий, размещаемых в БД. Только такие унифицированные данные могут в дальнейшем эффективно использоваться

и при анализе архивированной информации, и при создании математических моделей новых проектируемых изделий.

На рис. 1 представлен вид экрана ведения БД готовых изделий.

Понятие «готовые изделия» здесь толкуется в расширительном смысле. Наряду с моделями готовых, серийных изделий в БД может вноситься информация об изделиях, которые находятся в процессе изготовления. Атрибут модели «Этап ЖЦ» обозначает статус модели в БД, который может быть обозначен как «проектирование», «исследования», «испытания», «эксплуатация», «снят с эксплуатации», «типовая модель». Модель с обозначением ХХХ, которая выведена в первой строке таблицы на рис. 1, сформирована как «типовая модель», которая может быть использована в качестве шаблона при построении новых моделей данного типа изделий. Для каждого типа авиационных изделий в АСВП может быть создана своя типовая модель. Классификатор типов изделий по критерию назначения, который на рис. 1 представлен в левой панели экрана, создается в отдельном модуле АСВП пользователями системы и может отвечать конкретным задачам, решаемым на предприятии. В БД могут создаваться модели не только самолетов, но и других изделий, в частности авиационных двигателей и систем оборудования ЛА, которые могут

Тип	+	Обознач...	Производитель	Этап ЖЦ	Год	Область ^
A.2.1.1		XXX		типовая модель	2023	дозвук
A.2.1.1		Як-42Д	ОКБ Яковлева	эксплуатация	1987	дозвук
A.2.1.1.1	+	A300-600R	Airbus	эксплуатация	1974	дозвук
A.2.1.1.1	+	A310-200	Airbus	эксплуатация	1982	дозвук
A.2.1.1.1	+	A330-200	Airbus	эксплуатация	1997	дозвук
A.2.1.1.1	+	A340-300	Airbus	эксплуатация	1987	дозвук
A.2.1.1.1	+	A350-900	Airbus	эксплуатация	2006	дозвук
A.2.1.1.1	+	B747-100	Boeing	эксплуатация	1969	дозвук
A.2.1.1.1	+	B747-300	Boeing	эксплуатация	1983	дозвук
A.2.1.1.1	+	B767-300...	Boeing	эксплуатация	1981	дозвук
A.2.1.1.1	+	B777-200	Boeing	эксплуатация	1995	дозвук
A.2.1.1.1	+	B787-8	Boeing	эксплуатация	2009	дозвук
A.2.1.1.1	+	ИЛ-86	ОКБ Ильюши...	эксплуатация	1976	дозвук
A.2.1.1.1	+	ИЛ-96	ОКБ Ильюши...	эксплуатация	1988	дозвук

Рис. 1 База данных СПМ готовых изделий

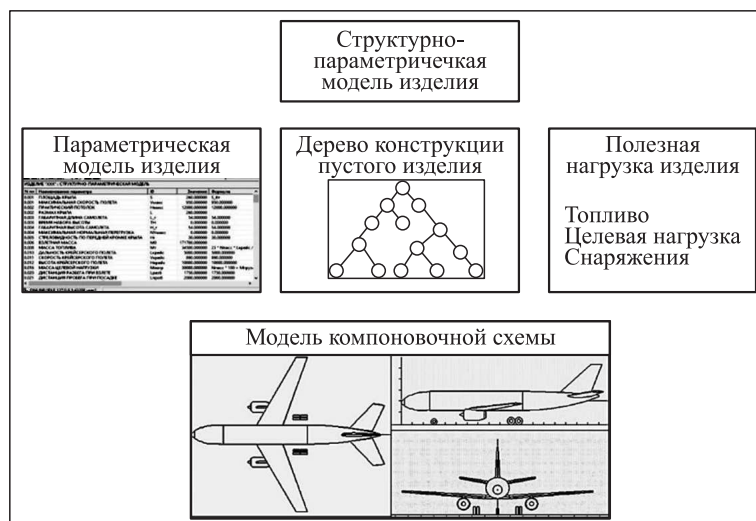


Рис. 2 Основные компоненты СПМ самолета

быть использованы при разработке моделей самолетов.

Включение новой модели в БД реализуется специальной командой модуля в интерактивном диалоге или путем копирования другой модели, чаще модели-шаблона, находящейся в БД. Содержательная информация об изделии, его структуре, конструктивных параметрах и характеристиках создается в процессе разработки модели. На рис. 2 показаны основные компоненты СПМ самолета.

3 Параметрические модели на этапе формирования облика самолета

В публичном пространстве — в различных авиационных справочниках и энциклопедиях — цифровая информация о самолетах сводится к значениям основных параметров и характеристик, которые дают общее представление о самолете, его размерах и эксплуатационных характеристиках. Такие параметры можно определить как основные тактико-технические характеристики (ТТХ), определяющие параметрический облик самолета. К этим параметрам относятся, во-первых, геометрические размеры самолета — размах крыла $L_{кр}$, площадь крыла S , габаритная длина фюзеляжа $L_{ф}$. Во-вторых, это основные весовые параметры: расчетная взлетная масса самолета $M_{взл}$, масса пустого изделия $M_{пуст}$, масса топлива $M_{топ}$, масса полезной нагрузки $M_{пол.нагр} = M_{топ} + M_{нагр}$. И третья группа параметров — это летно-технические характеристики самолета: максимальная скорость

полета $V_{макс}$, практический потолок $H_{макс}$, крейсерская дальность полета $L_{крейс}$, дистанция взлета $L_{взл}$ и посадки $L_{пос}$, стартовая тяга двигателей силовой установки P_0 и, возможно, некоторые другие характеристики. Все эти и другие известные из внешних источников параметры, определяющие облик самолета, могут быть включены в СПМ готового изделия при создании этой модели в БД.

Кроме этих основных параметров в практике проектирования используются различные производные величины, относительные или удельные конструктивные параметры, такие как безразмерная величина удлинения крыла $(L_{кр}^2)/S$, относительная стартовая масса топлива $M_{топ}/M_{взл}$, взлетная нагрузка на крыло $M_{взл}g/S$, взлетная тяговооруженность $P_0/(M_{взл}g)$ и другие параметры, вычисляемые по основным параметрам ТТХ. Удельные и относительные параметры более консервативны, чем абсолютные величины, изменяются в более узких диапазонах, которые в определенной степени характеризуют тип самолета. Это позволяет более эффективно проводить сравнительный анализ при создании новых проектов.

В АСВП пользователи при создании СПМ могут вводить любые параметры, описывающие изделие. При вводе нового параметра или при его модификации необходимо задать несколько атрибутов, которые показаны на рис. 3.

Множество параметров СПМ может состоять из большого числа записей, поэтому с помощью двухуровневой нумерации можно устанавливать удобный для обзора порядок следования параметров при их выводе в списке. Идентифицирующие реквизиты параметров вводятся пользователем по своему усмотрению (соблюдая общепринятые правила,

Рис. 3 Вводимые атрибуты параметров СПМ

терминологию и обозначения, что, впрочем, для обозначений, принятых в литературе, не всегда возможно). Под статусом параметра здесь понимается способ получения и использования его значений. Параметры могут быть задаваемыми пользователями, вычисляемыми по формулам, варьируемыми при параметрических расчетах или трактуемыми как критерии при оптимизации СПМ. Для вычисляемых параметров в поле «Формула/значение» вводится арифметическое выражение, используемое в качестве операндов, идентификаторы других параметров, встроенные функции и константы. Для задаваемых или варьируемых параметров задается числовое значение. В СПМ параметры ограничены своими минимальными и максимальными числовыми значениями. Если были выявлены ошибки в расчетных формулах или если значения параметра оказались вне диапазона допустимых значений, то в поле «Комментарий» выводится соответствующая диагностика.

4 Анализ требований к вновь создаваемым самолетам

В БД АСВП, как было сказано выше, может храниться не только архивная информация о существующих готовых изделиях авиационной техники,

но и могут создаваться модели новых проектируемых самолетов. В отличие от готовых изделий при создании СПМ для новых проектов самолетов некоторые параметры, определяющие их ТТХ, которые для готовых изделий брались из внешних источников, еще не известны, их еще предстоит вычислить в процессе работы над моделью. Для новых проектов разработка СПМ в АСВП начинается с ввода требований к тем ТТХ, которые сформулированы в ТТЗ к проектируемому изделию. Состав и уровень требований определяются на этапе внешнего проектирования [2].

На рис. 4 приведен типичный перечень характеристик, требования к которым обычно предъявляются в ТТЗ. Он отражает основные динамические характеристики ЛА и его характеристики как средства транспортировки различных объектов целевой нагрузки. Однако современные самолеты могут выполнять задачи, выходящие за рамки простого летающего транспортного средства. В большинстве случаев специальные задачи, соответствующие назначению самолета, обеспечиваются наличием на его борту систем, ориентированных на выполнение этих задач. В приведенных на рис. 4 требованиях это учтено в задании массы специального оборудования, перечень которого задается в ТТЗ.

Остальные параметры ТТХ самолета, такие как площадь несущей поверхности, габаритные параметры, а также взлетная масса, масса пустого изделия, масса топлива и, возможно, другие параметры, определяющие облик проектируемого изделия на момент создания, в БД его модели неизвестны, если, конечно, модель не копирует какой-нибудь прототип. Их значения, а точнее, более-менее достоверные диапазоны их возможных значений должны быть вычислены в результате анализа требований ТТЗ.

В процессе анализа требований ТТЗ в АСВП решаются две основные задачи. Первая задача — это сравнение требований ТТЗ нового проекта с уровнем ТТХ современных серийных или опытных изделий, СПМ которых хранятся в БД. Вместе с этой задачей решается задача выявления возможных прототипов для начала проектирования

Рис. 4 Требования к ТТХ самолета

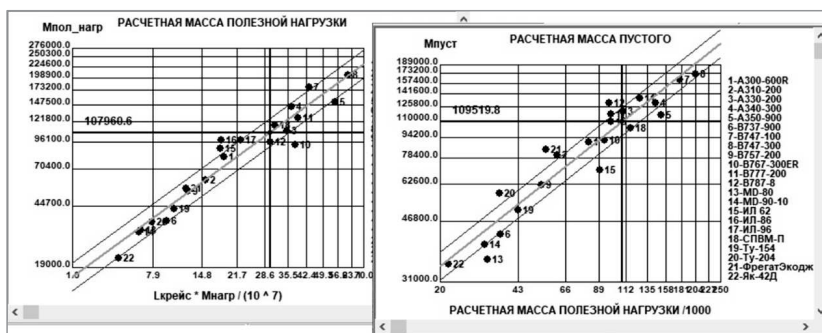


Рис. 5 Анализ массы пустого ЛА по выборке из БД готовых изделий

нового самолета. Вторая задача, которая решается на этапе анализа требований ТТЗ и которая может использовать результаты решения первой задачи, — определение диапазонов значений основных конструктивных параметров СПМ, в пределах которых следует вести дальнейшие исследования по формированию облика будущего проекта. Это задача построения области существования ЛА в пространстве параметров СПМ. В процессе решения первой задачи в АСВП формируется выборка из БД готовых изделий. Формирование выборок может вестись по разным критериям: по типу самолета, по производителю, по области применения, по близости характеристик изделий выборки к заявленным в ТТЗ требованиям. Если ввести какую-нибудь меру такой близости, то она может служить критерием для построения выборки. Выборки прототипов естественно использовать для определения диапазонов значений основных ТТХ, а также других конструктивных параметров, которые проектировщик использует при описании СПМ проекта.

В практике проектирования самолетов нашло широкое применение использование информации о готовых изделиях для построения парных или многофакторных зависимостей между конструктивными параметрами по выборкам прототипов проектируемого изделия. Чаще всего такие построенные зависимости используются для весового анализа при расчетах первого приближения массы самолета. Приведем пример такого анализа средствами АСВП¹. Основными параметрами ТТЗ, влияющими на весовые характеристики ЛА, служат масса целевой нагрузки $M_{нагр}$ и крейсерская дальность полета $L_{крейс}$, которая определяет необходимую массу заправляемого топлива $M_{топ}$ и которая в первом приближении должна быть вычислена в процессе анализа ТТЗ. На рис. 5 приведен при-

мер анализа зависимостей массы пустого самолета от массы полезной нагрузки по построенной выборке пассажирских самолетов. Полезная нагрузка самолета, определенная выше как сумма массы целевой нагрузки и заправляемого топлива, зависит от конкретного полетного задания. Для пассажирских самолетов имеет смысл рассматривать полетное задание в режиме крейсерского полета с полной загрузкой требуемого в ТТЗ числа пассажиров. Для определения стартовой полезной нагрузки для крейсерского полета введем новый параметр $LM = L_{крейс} M_{нагр}$, отражающий характеристику выполненного задания по транспортировке целевой нагрузки по заданному маршруту. Этот параметр легко вычислим по описаниям готовых изделий и по ограничениям ТТЗ. Построив выборку из БД реально существующих самолетов, можно проанализировать, существует ли корреляция между LM и $M_{пол_нагр}$, по крайней мере на этой выборке. И далее: существует ли связь между массой пустого самолета и массой полезной нагрузки.

На рис. 5 с точностью до статистической погрешности такие связи прослеживаются, и исходя из этого можно построить эмпирические модели для определения первого приближения массы пустого самолета, расчетной взлетной массы, расчетной стартовой массы заправляемого топлива и других весовых параметров самолета. Заметим, что масса размещаемой на борту целевой нагрузки и масса заправляемого топлива напрямую влияют и на параметры компоновочной схемы, в частности на ее габаритные размеры. Так, длина фюзеляжа пассажирских самолетов зависит от числа пассажиров (разумеется, и от компоновки пассажирских салонов), от объема фюзеляжных топливных баков.

На рис. 6 приведены в качестве примера результаты анализа по той же выборке, что и для весового

¹ В приведенном и в других примерах настоящей статьи используются данные из тестовой БД готовых изделий АСВП. Все эти данные брались из публичных источников интернета, при этом авторы допускают возможность каких-либо неточностей в цифровых значениях. Поэтому эти примеры не могут быть использованы в реальном проектировании и представлены исключительно в демонстрационных целях.

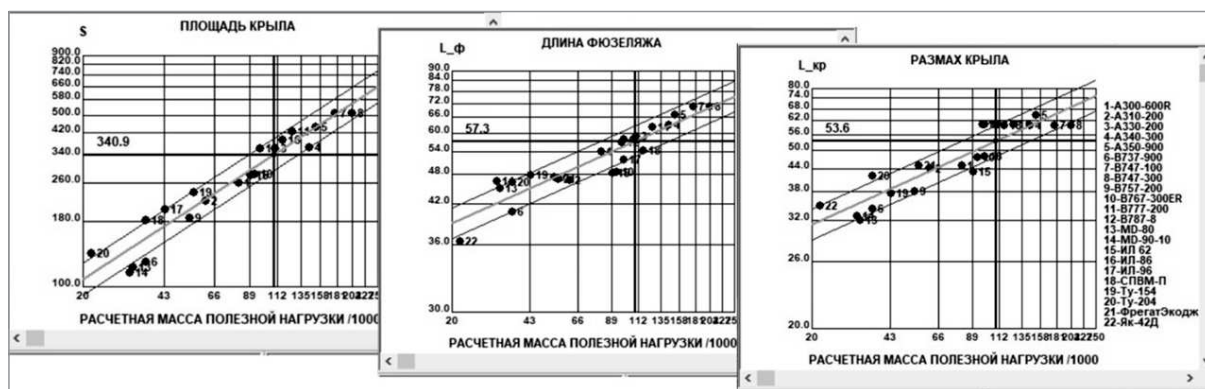


Рис. 6 Анализ зависимостей габаритов ЛА от массы полезной нагрузки

анализа, зависимостей основных конструктивных параметров, определяющих геометрические размеры пассажирского самолета, от того же параметра — массы полезной нагрузки.

На графике явно прослеживается связь между массой полезной нагрузки и анализируемыми параметрами. И что немаловажно, по этим графикам могут быть получены реальные диапазоны параметров, в рамках которых должны вестись дальнейшие исследования проекта.

Данные примеры приведены для демонстрации того, как на ранних этапах проектирования весьма эффективно могут использоваться эмпирические модели, основанные на анализе данных по различным выборкам прототипов проектируемых изделий. На начальных стадиях проектирования диапазоны изменения параметров проектируемого изделия определяются чаще всего исходя из опыта проектирования и выбора одного или нескольких реально существующих прототипов текущего проекта.

Эмпирические зависимости, полученные в результате анализа выбранной группы прототипов, подобные тем, которые приведены выше, могут быть использованы при синтезе компоновочной схемы.

5 Параметрическая модель компоновочной схемы

Как уже говорилось выше, основная задача этапа формирования облика самолета — построение его компоновочной схемы. Под компоновочной схемой в проектировании понимается отображение общего вида самолета в трех главных проекциях (см. рис. 2).

При дальнейшей детализации в компоновочную схему вносятся отображения двигателей силовой

установки, модели которых могут создаваться и храниться в БД АСВП, а также модели других самолетных систем и специального оборудования, экипажа, топлива и другой целевой нагрузки. В АСВП реализовано традиционное отображение компоновочной схемы в тех деталях, которые описаны в структурно-параметрической модели. Но «рисование» компоновочной схемы — это не только дань традиционным взглядам на процесс проектирования самолетов, это важный механизм визуального контроля корректности задания числовых параметров СПМ. В СПМ, реализованной в АСВП, компоновочная схема представлена деревом конструкции и набором параметров, определяющих геометрию, конструкцию и взаимное расположение агрегатов. Дерево конструкции — это иерархическая структура, задающая отношения включения между конструктивными элементами изделия, которые служат узлами этой иерархии. Для большинства современных типов самолетов верхние уровни дерева конструкции стандартны и состоят из основных агрегатов и систем — крыла, фюзеляжа, оперения, взлетно-посадочных устройств, бортового оборудования и силовой установки со своими системами, в том числе двигателями и топливными баками. Именно эти агрегаты, их параметры и их размещение определяют компоновочную схему самолета.

Принято выделять несколько аспектов конструкции компоновочной схемы, определяющих ее тип. Наиболее существенную роль в компоновке играет тип балансировочной схемы ЛА, который характеризует размещение относительно крыла агрегатов оперения, наличие заднего и переднего горизонтального оперения, число килей и другие детали компоновки несущих поверхностей. Форма основной несущей поверхности (крыла), наличие развитых передних наплывов крыла также влияют на положение аэродинамического фокуса

ИЗДЕЛИЕ "XXX": СТРУКТУРНО-ПАРАМЕТРИЧЕСКАЯ МОДЕЛЬ		
N nn	ПРОСТРАНСТВО ПАРАМЕТРОВ ИЗДЕЛИЯ	ID
2.000	ТИП КОМПОНОВОЧНОЙ СХЕМЫ САМОЛЕТА	РАЗДЕЛ_2
2.100	ОСНОВНЫЕ ПАРАМЕТРЫ КРЫЛА	РАЗДЕЛ_2.1
2.200	ОСНОВНЫЕ ПАРАМЕТРЫ ФЮЗЕЛЯЖА	РАЗДЕЛ_2.2
2.300	ОСНОВНЫЕ ПАРАМЕТРЫ ГО	РАЗДЕЛ_2.3
2.300	ОСНОВНЫЕ ПАРАМЕТРЫ ОПЕРЕНИЯ	РАЗДЕЛ_2.3
2.400	ОСНОВНЫЕ ПАРАМЕТРЫ СИЛОВОЙ УСТАНОВКИ	РАЗДЕЛ_2.4
2.500	ОСНОВНЫЕ ПАРАМЕТРЫ ШАССИ	РАЗДЕЛ_2.5
2.100	ОСНОВНЫЕ ПАРАМЕТРЫ КРЫЛА	РАЗДЕЛ_2.1
2.101	ПЛОЩАДЬ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	S_6r
2.102	УДЛИНЕНИЕ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	La_6r
2.103	СУЖЕНИЕ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	Eta_6r
2.104	СТРЕЛОВИДНОСТЬ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	Ni_6r
2.300	ОСНОВНЫЕ ПАРАМЕТРЫ ГО	РАЗДЕЛ_2.3
2.301	ОТНОСИТЕЛЬНАЯ ПЛОЩАДЬ ГО	S_ro
2.302	УДЛИНЕНИЕ ГО	La_ro
2.303	СУЖЕНИЕ ГО	Eta_ro
2.304	СТРЕЛОВИДНОСТЬ ПО ПК ГО	Ni_ro
2.400	ОСНОВНЫЕ ПАРАМЕТРЫ СИЛОВОЙ УСТАНОВКИ	РАЗДЕЛ_2.4
2.401	ДЛИНА ДВИГАТЕЛЯ	L_дв
2.402	МАКСИМАЛЬНЫЙ ДИАМЕТР ДВИГАТЕЛЯ	D_макс
2.403	ДИАМЕТР СОПЛА ДВИГАТЕЛЯ	D_сопла
2.404	X ДВИГАТЕЛЯ	X_дв
2.405	Y ДВИГАТЕЛЯ	Y_дв
2.000	ТИП КОМПОНОВОЧНОЙ СХЕМЫ САМОЛЕТА	РАЗДЕЛ_2
2.001	ТИП БАЛАНСИРОВОЧНОЙ СХЕМЫ (нормальная, с ПГО, у...	T_бал
2.002	ПОЛОЖЕНИЕ КРЫЛА ПО ВЫСОТЕ (высокоплан, среднепо...	T_pl
2.003	СХЕМА ОПЕРЕНИЯ (один киль, Т-образная схема, два кила)	T_op
2.004	ЧИСЛО ДВИГАТЕЛЕЙ СИЛОВОЙ УСТАНОВКИ	N_дв
2.005	ТИП ДВИГАТЕЛЕЙ (ТРД, ТРДФ, ТРДД, ТРДДФ, ПВРД, ТВД)	T_дв
2.006	РАЗМЕЩЕНИЕ ДВИГАТЕЛЕЙ (в мотогонд, в ф-же, у ф-жа)	T_су
2.007	ТИП ВОЗДУХОЗАБОРНИКА (лобовой, боковой нижний)	T_вз
2.008	ТИП ТЕЛЕЖКИ ГЛАВНОЙ ОПОРЫ ШАССИ (1-0,2-1,3-2,1x1-...	T_гош
2.200	ОСНОВНЫЕ ПАРАМЕТРЫ ФЮЗЕЛЯЖА	РАЗДЕЛ_2.2
2.201	ДЛИНА ФЮЗЕЛЯЖА	L_ф
2.202	УДЛИНЕНИЕ ГОЛОВНОЙ ЧАСТИ ФЮЗЕЛЯЖА	La_гч
2.202	ПОЛОЖЕНИЕ МИДЕЛЯ ФЮЗЕЛЯЖА	X_мид_ф
2.203	КОЭФФИЦИЕНТ ФОРМЫ МИДЕЛЯ ФЮЗЕЛЯЖА	k_мид_ф
2.330	ОСНОВНЫЕ ПАРАМЕТРЫ ВО	РАЗДЕЛ_2.3
2.340	ОТНОСИТЕЛЬНАЯ ПЛОЩАДЬ ВО	S_во
2.341	УДЛИНЕНИЕ ВО	La_во
2.342	СУЖЕНИЕ ВО	Eta_во
2.343	СТРЕЛОВИДНОСТЬ ПО ПК ВО	Ni_во
2.500	ОСНОВНЫЕ ПАРАМЕТРЫ ШАССИ	РАЗДЕЛ_2.5
2.501	КООРДИНАТА X ГЛАВНОЙ ОПОРЫ ШАССИ	X_гош
2.502	БАЗА ШАССИ	B_ш
2.503	КОЛЕСА ШАССИ	K_ш
2.504	ДИАМЕТР КОЛЕС ГОШ	D_гк

Рис. 7 Основные параметры компоновочной схемы самолета

и, следовательно, на балансировку ЛА. Традиционное членение фюзеляжа на головную, центральную и хвостовую части имеет свои особенности и иногда встречает затруднение в классификации. Например, для самолетов так называемых интегральных схем с плавным переходом конструкции от фюзеляжа к крылу в зависимости от значений параметров корневую часть несущей поверхности можно относить к фюзеляжу или к центроплану. На рис. 7 приведен список параметров компоновочной схемы самолета, реализованный средствами АСВП в рамках одного из возможных сценариев построения СПМ изделия.

Сценарий построения компоновочной схемы предполагает в первую очередь выбор балансировочной схемы самолета и расположение крыла относительно фюзеляжа по высоте. Под типом балансировочной схемы понимают расположение основной несущей поверхности относительно оперения в продольной оси самолета. Выбор балансировочной схемы определяет способ управления положением самолета в продольном канале. Отклонением органов управления горизонтального оперения создаются кабрирующие или пикирующие моменты, и самолет изменяет свою траекторию в вертикальной плоскости. Существуют «нормальная» балансировочная схема — когда горизонтальное оперение располагается в хвостовой части самолета, позади крыла, и схема «утка», предполагающая размещение горизонтального оперения в зоне носовой части фюзеляжа со смещением назад крылом. Иногда применяется смешанная схема — с задним и передним горизонтальным оперением, а иногда вообще горизонтального оперения нет, и самолет балансируется с помощью механизации крыла — такую схему называют «бесхвостка».

Выбор типа балансировочной схемы в определенной степени влияет на параметрическое задание основной несущей поверхности — крыла, и в первую очередь на так называемую «базовую трапецию». Базовая трапеция крыла — это не одна, а две симметричные относительно оси самолета трапеции с корневой хордой, лежащей на этой оси.

Для задания формы в плане базовой трапеции достаточно четырех основных параметров: размаха базовой трапеции L_{6T} , длины корневой хорды B_0 , длины концевой формы B_k и угла стреловидности по передней кромке Ni_{6T} . Наряду с этими базовыми параметрами задают площадь базовой трапеции S_{6T} , удлинение La_{6T} и сужение Eta_{6T} базовой трапеции. Параметры базовой трапеции связаны между собой соотношениями:

$$S_{6T} = \frac{(B_0 + B_k)L_{6T}}{2}; \quad La_{6T}m = \frac{L_{6T}^2}{S_{6T}}; \quad Eta_{6T} = \frac{B_0}{B_k}.$$

Положение базовой трапеции задается координатами носка корневой хорды X_{6T} и Y_{6T} .

Полная модель крыла существенно сложнее и включает дополнительно еще целый ряд параметров, которые могут задаваться или вычисляться по формулам, задаваемым при определении этих параметров. У современных самолетов часто передняя кромка крыла имеет сложную нелинейную форму. Это характерно для так называемых «интегральных» компоновочных схем с плавным сочленением крыла и фюзеляжа. В рамках АСВП такие компоновочные схемы аппроксимируются сложным составным крылом с кусочно-линейными передней и задней кромками. На рис. 8 показана форма в плане крыла, состоящего из отъемной части крыла (ОЧК) и корневой части.

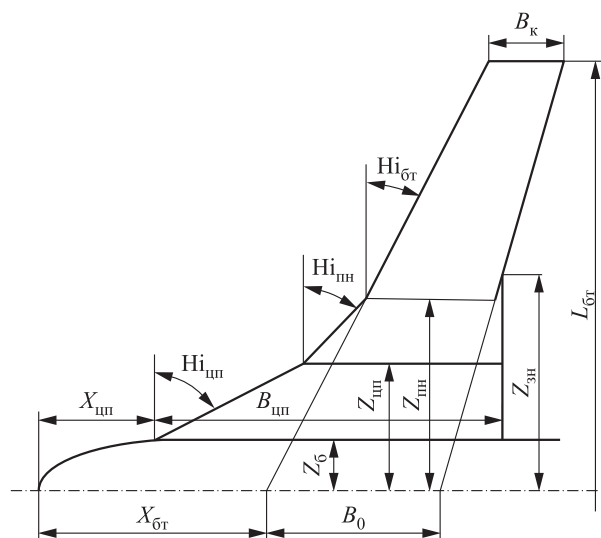


Рис. 8 Форма в плане крыла

Корневую часть крыла иногда называют центропланом. Однако, чтобы не путать ее с *подфюзеляжным* центропланом, служащим частью силовой конструкции, будем называть корневую часть крыла *околофюзеляжным* центропланом.

Параметры составного крыла кроме параметров базовой трапеции включают углы стреловидности переднего и заднего наплывов ОЧК ($H_{шт}$ и $H_{зшт}$), координаты точек излома у передней и задней кромки ОЧК ($Z_{шт}$ и $Z_{зшт}$), а также параметры *околофюзеляжного* центроплана ($Z_б$, $Z_{шт}$, $B_{шт}$, $H_{шт}$ и $H_{зшт}$). Заметим, что на рис. 8 параметры $H_{зшт}$ и $H_{зшт}$ не обозначены и имеют нулевые значения.

Кроме параметров, определяющих форму крыла в плане, для многих компоновочных задач и задач весового проектирования необходимо задать параметры профиля крыла и положение его в пространстве. Положение крыла по высоте относительно фюзеляжа определяется в зависимости от типа компоновки крыла — центроплан, высокоплан, низкоплан. Однако такое определение в не-

которых случаях требует уточнения, т.е. задания координаты $Y_{кр}$ и угла V-образности крыла — угла поворота плоскости крыла вокруг продольной оси самолета относительно плоскости земли в стояночном положении. Что касается профиля крыла, то он может выбираться из библиотек стандартных профилей. Основным параметр профиля, используемый в АСВП, от которого зависят многие компоновочные и весовые характеристики крыла, — это его относительная толщина. Этот параметр может задаваться отдельно для ОЧК и околофюзеляжного центроплана. Приведенные параметры крыла и соотношения между ними представляют собой в АСВП параметрическую модель крыла, фрагмент которой приведен на рис. 9.

Аналогичным образом строятся модели других агрегатов компоновочной схемы — фюзеляжа, оперения, шасси, силовой установки. Фюзеляж в реализованном сценарии построения компоновочной схемы задается своей длиной, миделем, а также удлинениями и габаритными размерами сечений головной и хвостовой частей. Вертикальное и горизонтальное оперение, включая переднее горизонтальное оперение, задаются аналогично крылу своими базовыми трапециями и координатами их размещения, привязанными к самолетной системе координат.

Модели синтезируемых в АСВП компоновочных схем ЛА позволяют перейти к построению **параметрических** трехмерных каркасных габаритных моделей проектируемых изделий. Трехмерные каркасные габаритные модели ЛА и его агрегатов на начальных этапах проектирования позволяют более точно вычислять характерные площади поверхностей и располагаемые объемы этих агрегатов — параметры, существенным образом влияющие на их массово-инерционные характеристики. Кроме того, по каркасным моделям можно аппроксимировать трехмерные поверхности, создавая параметрическое трехмерное представление проектируемого изделия, перестройка которого не требует боль-

ДЕРЕВО КОНСТРУКЦИИ ЛЕТАТЕЛЬНОГО АПП.		ПАРАМЕТРЫ АГРЕГАТА ДЕРЕВА КОНСТРУКЦИИ		
Наименование	N пп	Наименование параметра	ID	Формула
КРЫЛО	2.101	ПЛОЩАДЬ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	S_бт	50
БАЗОВАЯ ТРАПЕЦИЯ	2.102	РАЗМАХ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	Lкр	15
ОТЪЕМНАЯ ЧАСТЬ КРЫЛА (ОЧК)	2.103	УДЛИНЕНИЕ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	La_бт	$L_{кр} * L_{кр} / S_{бт}$
ЦЕНТРОПЛАН	2.104	СУЖЕНИЕ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	Eta_бт	3.5
ОКОЛОФЮЗЕЛЯЖНАЯ ЧАСТЬ ЦП	2.105	СТРЕЛОВИДНОСТЬ ПК БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	Hi_бт	40
ПОДФЮЗЕЛЯЖНАЯ ЧАСТЬ ЦП	2.106	ДЛИНА КОРНЕВОЙ ХОРДЫ БАЗОВОЙ ТРАПЕЦИИ	B_0	$S_{бт} * Eta_{бт} * 2.0 / ((Eta_{бт} + 1.0) * B_{шт})$
РАСЧЕТНАЯ МОДЕЛЬ СХО КРЫЛА	2.108	ДЛИНА КОНЦЕВОЙ ХОРДЫ БАЗОВОЙ ТРАПЕЦИИ	B_конц	$B_0 / Eta_{бт}$
РАСЧЕТНАЯ МОДЕЛЬ S_{y_al} КРЫЛА	2.109	ДЛИНА САХ БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	ВСАХ	$2.0 * B_0 * (1.0 + 1.0 / (Eta_{бт} * (Eta_{бт} + 1.0)))$
ФЮЗЕЛЯЖ	2.110	ПОЛОЖЕНИЕ САХ ПО X БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	XСАХ	$ZCAХ * Cos(XIR) / Sin(XIR) + X0$
ОПЕРЕНИЕ	2.111	ПОЛОЖЕНИЕ САХ НП ПО Z БАЗОВОЙ ТРАПЕЦИИ КРЫЛА	ZСАХ	$WK * ((ETA + 2.0) / (ETA + 1.0)) / 6.0$
ВПУ	2.112	ПЛОЩАДЬ ОМЫВАЕМОЙ ПОВЕРХНОСТИ ОЧК	S_ом_к...	$2.0 * (S_{бт} - S_{шт}) * Sqrt(1 + C_{y_шт})$

Рис. 9 Параметрическая модель крыла

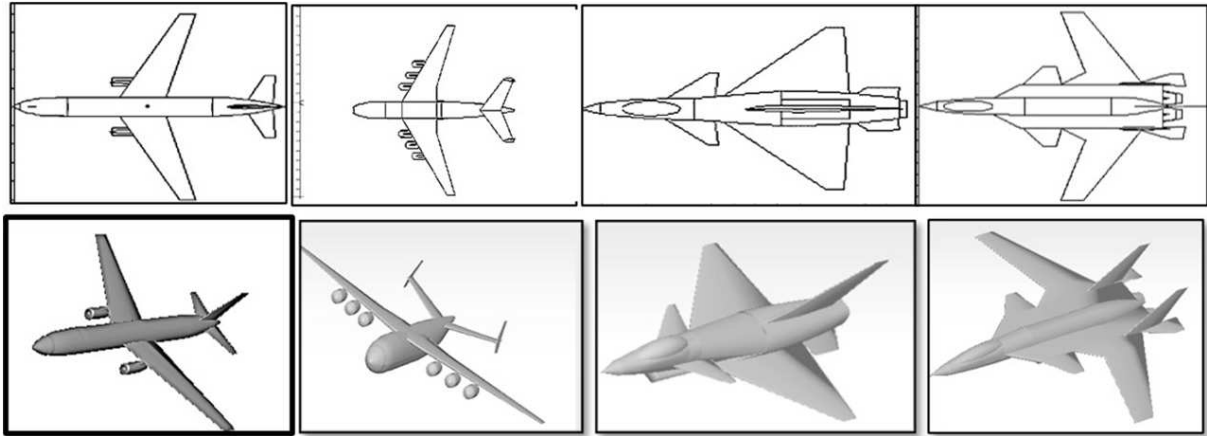


Рис. 10 Варианты компоновочных схем ЛА

ших вычислений. Такое представление служит не только эффектной иллюстрацией этапа формирования облика, но и позволяет решать такие задачи, как построение распределенных массово-инерционных характеристик конструкции ЛА на начальных этапах проектирования, когда полноценного электронного макета изделия еще не существует.

На рис. 10 показаны различные варианты параметрических компоновочных схем, построенных в АСВП, и переход от них к 3D-аппроксимациям облика самолета. Эти модели позволяют решать задачи формирования облика достаточно широкого класса самолетов разного назначения и предоставляют достаточно информации, чтобы на инженерном уровне проводить анализ основных характеристик ЛА и оценивать, может ли данная компоновочная схема быть взята за основу для дальнейшей проработки проекта. Задачи анализа компоновочных схем требуют построения другого класса моделей, в которых затрагиваются такие области знаний, как аэродинамика, динамика полета, теория двигателей и др. На начальных этапах проектирования необходима разработка инженерных упрощенных моделей для приближенных расчетов, которые позволяли бы решать на этих этапах задачи оптимизации компоновочных схем самолетов.

6 Заключение

Описанные в статье сценарии построения СПМ компоновочных схем самолетов не являются единственно возможными. Автоматизированная сис-

тема весового проектирования предоставляет гибкий инструмент для построения моделей широкого класса проектов самолетов. И не только самолетов. Она служит своего рода демонстрацией технологии «Генератора проектов» [5], применявшейся при создании многих информационно-вычислительных систем. Разработке этой технологии много лет посвятил ее создатель, коллега и товарищ авторов статьи, Николай Иванович Широков, недавно ушедший из жизни.

Литература

1. Вышинский Л. Л., Флеров Ю. А., Широков Н. И. Автоматизированная система весового проектирования самолетов // Информатика и её применения, 2018. Т. 12. Вып. 1. С. 18–30. doi: 10.14357/19922264180103. EDN: YTTTBQ.
2. Шейнин В. М., Козловский В. И. Весовое проектирование и эффективность пассажирских самолетов. — М.: Машиностроение, 1977. Т. 1. 343 с.
3. Егер С. М., Лисейцев И. К., Самойлович О. С. Основы автоматизированного проектирования самолетов. — М.: Машиностроение, 1986. 232 с.
4. Вышинский Л. Л., Самойлович О. С., Флеров Ю. А. Программный комплекс формирования облика летательных аппаратов // Задачи и методы автоматизированного проектирования в авиастроении. — М.: ВЦ АН СССР, 1991. С. 24–42.
5. Вышинский Л. Л., Гринев И. Л., Флеров Ю. А., Широков А. Н., Широков Н. И. Генератор проектов — инструментальный комплекс для разработки «клиент-серверных» систем // Информационные технологии и вычислительные системы, 2003. № 1-2. С. 6–25.

Поступила в редакцию 28.09.23

SYNTHESIS MODELS OF LAYOUT SCHEME IN THE TASK OF FORMING AN AIRCRAFT IMAGE

L. L. Vyshinsky and Yu. A. Flerov

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The article discusses the problems of synthesizing mathematical models of an aircraft at the initial design stage. The central task of the initial design stage is to develop a structural-parametric model of the product and construct its layout diagram. The structural-parametric model of the product describes the space of the main design parameters of the aircraft and the layout diagram represents its structure, the set of main units and systems of the aircraft, their relative arrangement, and structural connections between the units. The layout diagram of the aircraft serves as the basis for further detailing of the project. At the initial design stage, the permissible ranges of values of the design parameters of the units of the layout diagram are determined in which the search for optimal design solutions is carried out. The paper presents a description of structural-parametric models and layout diagrams of a wide class of aircraft which can be implemented using an automated weight design system.

Keywords: mathematical modeling; design automation; aircraft; weight design; project generator

DOI: 10.14357/19922264240109

EDN: DSPGKV

References

1. Vyshinskiy, L. L., Yu. A. Flerov, and N. I. Shirokov. 2018. Avtomatizirovannaya sistema vesovogo proektirovaniya samoletov [Automated system of weight design of aircraft]. *Informatika i ee Primeneniya — Inform Appl.* 12(1):18–30. doi: 10.14357/19922264180103. EDN: YTTRBQ.
2. Sheynin, V. M., and V. I. Kozlovskiy. 1977. *Vesovoe proektirovanie i effektivnost' passazhirskikh samoletov* [Weight design and efficiency of passenger aircraft]. Moscow: Mashinostroenie. Vol. 1. 343 p.
3. Eger, S. M., I. K. Liseytshev, and O. S. Samoylovich. 1986. *Osnovy avtomatizirovannogo proektirovaniya samoletov* [Fundamentals of aircraft automated design]. Moscow: Mashinostroenie. 232 p.
4. Vyshinsky, L. L., O. S. Samoylovich, and Yu. A. Flerov. 1991. Programmnyy kompleks formirovaniya oblika letatel'nykh apparatov [Program complex for forming the appearance of aircraft]. *Zadachi i metody avtomatizirovannogo proektirovaniya v aviastroenii* [Tasks and methods of computer-aided design in aircraft industry]. Moscow: Computing Center of the USSR Academy of Sciences. 24–42.
5. Vyshinskiy, L. L., I. L. Grinev, Yu. A. Flerov, A. N. Shirokov, and N. I. Shirokov. 2003. Generator proektov — instrumental'nyy kompleks dlya razrabotki “klient-servernykh” sistem [The project generator — tool complex for development of “client–server” systems]. *Informatsionnye tekhnologii i vychislitel'nye sistemy* [J. Information Technologies Computing Systems] 1-2:6–25.

Received September 28, 2023

Contributors

Vyshinsky Leonid L. (b. 1941) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vyshinsky@mail.ru

Flerov Yuri A. (b. 1942) — Corresponding Member of the Russian Academy of Sciences, Doctor of Science in physics and mathematics, professor, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; fler@ccas.ru

О ПОРОЖДЕНИИ СИНТЕТИЧЕСКИХ ПРИЗНАКОВ НА ОСНОВЕ ОПОРНЫХ ЦЕПЕЙ И ПРОИЗВОЛЬНЫХ МЕТРИК В РАМКАХ ТОПОЛОГИЧЕСКОГО ПОДХОДА К АНАЛИЗУ ДАННЫХ. ЧАСТЬ 1. ВКЛЮЧЕНИЕ В ФОРМАЛИЗМ ЭМПИРИЧЕСКИХ ФУНКЦИЙ РАССТОЯНИЯ*

И. Ю. Торшин¹

Аннотация: Анализ формализма топологической теории распознавания на основе фундаментальных понятий функционального анализа позволил предложить ранее не исследованные подходы к определению решеточных оценок. В частности, использование опорных цепей для анализа булевых решеток, формируемых над регулярными по Журавлёву множествами прецедентов, указало на новое направление исследований, заключающееся в замене оценок элементов решеток на определенного рода функции и/или векторы. Данное расширение формализма также позволяет проводить систематическое исследование известных в литературе полумпирических функционалов расстояния для решения прикладных задач. Обоснованы перспективные направления дальнейшего развития формализма, включающие введение функционалов, редуцирующих описания множеств булевой решетки к скалярным оценкам, и развитие математического аппарата для анализа решеток, в котором вместо оценок фигурируют операции над соответствующими функциями. Последнее направление интересно тем, что позволяет вводить расстояния на решетке без использования оценок.

Ключевые слова: топологический анализ данных; теория решеток; алгебраический подход Ю. И. Журавлёва; функциональный анализ

DOI: 10.14357/19922264240110

EDN: RIVOXR

1 Введение

Топологическая теория распознавания стала развитием алгебраического подхода к распознаванию, предложенного научной школой Ю. И. Журавлёва, и предназначена для решения плохо формализованных задач распознавания, классификации и прогнозирования [1]. Одна из основных целей данной теории — разработка методов систематического порождения и отбора синтетических признаков описаний объектов, которые бы характеризовались большей информативностью, чем исходные признаки [2]. В работе [3] было показано, что веса таких признаков можно эффективно настраивать посредством ранговой оптимизации; разработан формализм для порождения признаков описаний на основе параметризуемых решеточных оценок [4].

В рамках развиваемого формализма каждый объект x из множества исходных описаний N_0 объектов $\mathbf{X} = \{x_1, \dots, x_{N_0}\}$, $\mathbf{X} \subseteq S$, описываемый n признаками посредством функций $\Gamma_k : S \rightarrow I_k$

(где $I_k = \{\lambda_{k_1}, \lambda_{k_2}, \dots\}$ — множества значений признаков описаний), представлен множествами $\{\Gamma_k^{-1}(\Gamma_k(x))\}$, $k = \overline{1, n+l}$, где l — число целевых (таргетных) переменных. Значение t -й целевой (таргетной) переменной объекта x , $t = \overline{n+1, n+l}$, вычисляется как $\Gamma_t(x)$. Множество прецедентов над пространством допустимых признаков описаний объектов $J_{об}$ определяется как

$$Q = \varphi(\mathbf{X}) = \{D(x_\alpha) | x_\alpha \in \mathbf{X}\}$$

посредством $D : S \rightarrow J_{об}$ и $\varphi(\mathbf{X}) = \{D(x_\alpha) | x_\alpha \in \mathbf{X}\}$, $D(x_\alpha) = (\Gamma_1(x_\alpha) \times \dots \times \Gamma_k(x_\alpha) \times \dots \times \Gamma_{n+l}(x_\alpha))_\Delta$. При регулярности множеств \mathbf{X}/Q ($\forall x \in \mathbf{X}, x = D^{-1}(D(x))$) \mathbf{X} изоморфно Q и обоим множествам сопоставлены топология $T(\mathbf{X}) = \{\emptyset, \{\mathbf{X}\}, a \cup b, a \cap b : a, b \in U(\mathbf{X}) = \{\Gamma_k^{-1}(\lambda_{k_b})\}$ и булева решетка $L(T(\mathbf{X})) = \{a \vee b, a \wedge b : a, b \in T(\mathbf{X})\}$ [5].

Рассмотрим топологический подход к распознаванию с точки зрения фундаментальных понятий теории функций — рефлексивных и транзитивных

* Работа выполнена при поддержке гранта РНФ (проект № 23-21-00154) с использованием инфраструктуры Центра коллективного пользования «Высокопроизводительные вычисления и большие данные» (ЦКП «Информатика») ФИЦ ИУ РАН (г. Москва).

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, ty135@yahoo.com

бинарных отношений между множествами. В математике применяются два таких отношения: симметричное отношение эквивалентности множеств и антисимметричное отношение (частичного) порядка [6].

Прецедентному соотношению между значениями признака $\Gamma_k(x)$ и t -й целевой переменной, заданному множеством Q на решетке $L(T(\mathbf{X}))$, соответствует множество пар $\{(\{\Gamma_k^{-1}(\Gamma_k(x_i)), k = \overline{1, n}\}, \Gamma_t^{-1}(\Gamma_t(x_i))), i = \overline{1, N_0}\}$. На основании представленных в регулярном множестве Q прецедентов любой алгоритм распознавания («машинного обучения») строит некоторую модель соотношения между множествами $\{\Gamma_k^{-1}(\Gamma_k(x_i))\}$ и $\Gamma_t^{-1}(\Gamma_t(x_i))$.

Сфокусируемся на двух произвольных множествах $a = \Gamma_k^{-1}(\Gamma_k(x_i))$ и $b = \Gamma_t^{-1}(\Gamma_t(x_i))$. Очевидно, что говорить о выполнимости отношения эквивалентности или порядка между множествами a и b в общем случае не приходится: ведь эквивалентность соответствует идентичности значения k -го признака значению t -й целевой переменной, а частичный порядок — ядерной эквивалентности (т. е. эквивалентности a подмножеству b или наоборот). Такие случаи тривиальны и соответствуют «легко решаемым» задачам распознавания.

В то же время существующие в решетке $L(T(\mathbf{X}))$ отношения порядка порождают супремум $a \vee b$ и инфимум $a \wedge b$ множеств a и b . Поэтому в топологической теории распознавания вводятся более сложные функционалы над $a, b, a \vee b$ и $a \wedge b$, расширенно описывающие соотношения между произвольными множествами a и b в терминах расстояний. При выполнимости четырех аксиом метрики такие функционалы $\rho_L : L^2 \rightarrow R^+$ формируют метрическое пространство значений признаков $M_L(L(T(\mathbf{X})), \rho_L)$.

Простейшей метрикой служит функционал

$$\rho_0(a, b) = \frac{v[a \vee b] - v[a \wedge b]}{N_0},$$

где $v : L \rightarrow R^+$ — изотонная оценка на $L(T(\mathbf{X}))$. Возможны и более сложные варианты определения ρ_L при введении параметрических оценок [4] или использовании известных в литературе метрик, так что в общем случае имеется ряд метрик $\rho_m, m = \overline{1, m_0}$. Как правило, метрики ρ_m нормируются на интервал значений $[0 \dots 1]$.

Возвращаясь к рассмотрению фундаментальных отношений теории функций, можно сделать вывод о том, что метрика $\rho_L(a, b)$ служит функционалом, численно оценивающим выполнимость отношения эквивалентности между a и b на основе отношений порядка (заданных в форме $a \vee b, a \wedge b$).

Действительно, $\rho_L(a, b) = 0$ соответствует строгому выполнению отношения эквивалентности a и b , а $\rho_L(a, b) = 1$ — максимально возможному расстоянию между a и b (например, $\rho_0(a, b) = 1$ только для множеств \emptyset и $\{\mathbf{X}\}$, находящихся на концах максимальных цепей решетки $L(T(\mathbf{X}))$).

Таким образом, в рамках топологической теории распознавания соотношение между множествами $\{\Gamma_k^{-1}(\Gamma_k(x_i))\}$ и $\Gamma_t^{-1}(\Gamma_t(x_i))$ моделируется соответствующими массивами расстояний, порождаемыми той или иной метрикой ρ_m . Рассмотрим способы вычисления таких расстояний.

2 О различных способах вычисления метрик на решетке $L(T(\mathbf{X}))$

В литературе известны три принципиально различных способа определения метрических расстояний:

- (1) метрики на основе операций над множествами;
- (2) метрики над пространством векторов;
- (3) метрики над пространством функций.

Рассмотрим три этих подхода в применении к описанным выше конструктам топологической теории распознавания.

Метрики на основе операций над множествами.

Во введении были упомянуты метрики оценки расстояния между $a, b \in L(T(\mathbf{X}))$, вводимые как функционалы над $a \vee b$ (соответствует $a \cup b$) и $a \wedge b$ ($a \cap b$), оценками высот элементов в $L(T(\mathbf{X}))$ и другими теоретико-множественными операциями над множествами a и b . В работе [4] для порождения метрик используются взвешенные решеточные оценки

$$v_\alpha = \sum_{i=0, |\alpha|} \omega_i v_{\alpha_i}$$

на основе изотонных оценок v_{α_i} . Формирование наборов α может проводиться на основе «информативности» $\alpha_i \in \alpha$ методами метрического анализа данных [7] или на основе различных подцепей целевых числовых переменных.

В то же время известны многочисленные функционалы, носящие эмпирический характер и непосредственно оперирующие множествами: расстояния Танимото, Рэнда, Рассела–Рао, Симпсона, Брауна–Бланке, Роджера–Танимото, Фэйта, дисперсии, образов, Q_0 , Пирсона; различные варианты расстояний Тверского, Сокала–Сниса, Гоуэра–Лезандра, Юле и др. [8]. Метрические свойства

этих функционалов могут быть продемонстрированы посредством аналитических выводов или комбинаторного анализа на множествах прецедентов.

Векторные метрики. Альтернативно методу взвешенных решеточных оценок на основании набора α и оценок v_{α_i} для произвольного множества $a \in L(T(\mathbf{X}))$ может быть вычислен вектор

$$\vec{v}_\alpha[a] = (v_{\alpha_1}[a], v_{\alpha_2}[a], \dots, v_{\alpha_i}[a], \dots), \quad v_{\alpha_i}[a] \in R^+,$$

и введены метрики на пространстве векторов \vec{v}_α посредством известных подходов: l_1 -метрика, l_p -метрики Минковского, расстояния Пенроуза, Манхэттена, Лоренца, Кларка, Хеллинджера, Уайттеккера, симметрическое χ^2 , Махаланобиса (в том числе с настраиваемыми весами), расстояния пересечения, Ружечки, Робертса, Элленберга, Глисона, Мотыки, Брея–Куртиса, Канберры, Кульчинского и корреляционные расстояния (ковариационное, корреляционное, косинус, угловое, хордовое, подобности, Мориситы–Хорна, Спирмана, Кендалла) [8].

Метрики над пространством функций. Функциональный анализ и теория вероятностей предоставляют широкий инструментарий для определения расстояний между функциями с одинаковой областью определения: функционалы Колмогорова (в том числе максимальное отклонение D), фон Мизеса, Реньи, метрики (интегральная L_1 , инженерная, разделения, подобности среднего гармонического), расстояния Чебышёва, Степанова, варианты расстояний Золотарёва, Круглова, Бурби–Рао, Бхаттачарья, Чизара (включая вариации расстояний Кульбака–Лейблера, χ^2 , Хеллинджера) [8].

Очевидно, что некоторые из этих функций расстояний заведомо не относятся к метрикам. Например, в расстоянии Кульбака–Лейблера нарушена аксиома симметричности; оценка выполнимости аксиомы треугольника для каждой из этих функций требует отдельного исследования. Метризация рассматриваемых функций расстояния может осуществляться посредством введения дополнительных конструкций в определение функции. В частности, при отсутствии симметричности для функции $d(x, y)$ могут быть введены конструкции $\min(d(x, y), d(y, x))$, $\max(d(x, y), d(y, x))$ и др. Основной проблемой все же остается «привязка» этих подходов к разрабатываемому решеточному формализму. Для этого напомним ряд важных понятий.

Определение 1. Решеточный терм, или *изотонная оценка*, $v : L \rightarrow R^+$ над $L(T(\mathbf{X}))$ — функция, для которой выполнено *условие оценки* (**уО**: $\forall_L a, b : v[a] + v[b] = v[a \wedge b] + v[a \vee b]$) и *условие изотонности*

(**уИ**: $\forall_L a, b : a \supseteq b \Rightarrow v[a] \geq v[b]$). Для изотонной $v[\]$ гарантировано существование метрики $\rho_0(a, b)$.

Определение 2. *Однородными функциями* будем называть произвольные функции с одинаковыми областями определения и значений.

Определение 3. Пусть задано конечное множество чисел $A = \{a_1, a_2, \dots, a_i, \dots, a_n\}$, $a_i \in R$. Определим оператор $\hat{\phi}(x)$ для формирования эмпирической функции распределения (э. ф. р.) чисел в множестве A как $\hat{\phi}(x)A = \sup\{B \subseteq A \mid \forall a \in B : a \leq x\} / |A|$, $x \in R$, так что $\hat{\phi}(-\infty)A = 0$, $\hat{\phi}(+\infty)A = 1$. Для краткости $\hat{\phi}(x)A$ будем также записывать как $\hat{\phi}A$.

Определение 4. $\hat{\mu}$ — оператор вычисления математического ожидания значения $x \in A$ по э. ф. р. $\hat{\phi}A$ как

$$\hat{\mu}\hat{\phi}A = \frac{1}{m} \sum_{j=1}^m x_j \left(\hat{\phi}(x_j)A \hat{\phi}(x_{j-1})A \right),$$

где $m = |\hat{z}A|$; $x_j = \hat{v}^+(j)\hat{z}A$, а произвольное $x_0 < \inf(A)$, $x_0 \in R$; \hat{z} — оператор формирования множества значений множества A , $\hat{z}A = B \subseteq A \mid \forall a \in A : a \in B, \forall a, b \in B : a \neq b$; \hat{v}^+ — оператор упорядочения множества по возрастанию; $\hat{v}^+(j)A$ — j -й элемент множества \hat{v}^+A .

3 Анализ решетки $L(T(\mathbf{X}))$ с использованием функций $\hat{\phi}A$ на основе опорных цепей

Рассмотрим подходы к анализу решетки посредством э. ф. р. Определения 2–4 существенно расширяют формализм решеточных оценок, позволяя (1) проектировать $L(T(\mathbf{X}))$ в соответствующую решетку э. ф. р. посредством некоторой заранее выбранной (опорной) цепи; (2) измерять расстояния между этими э. ф. р.; (3) вводить новые разновидности оценок множеств (см. определение 1). В качестве опорной цепи может быть выбрана, в частности, цепь, соответствующая числовой таргетной переменной.

Теорема 1. *Выберем произвольную максимальную цепь A_t в качестве «опорной» для дальнейших построений. При условии регулярности множеств в \mathbf{X}/Q каждому элементу $L(T(\mathbf{X}))$ сопоставлена э. ф. р. из множества однородных э. ф. р.*

Доказательство. При выполнении условия регулярности для \mathbf{X}/Q решетка $L(T(\mathbf{X}))$ — булева (теорема 3 в [9]). Цепи в $L(T(\mathbf{X}))$ соответствуют тем или иным числовым признаковым описаниям (теорема 1 в [4]), так что произвольная (максимальная) цепь A_t в $L(T(\mathbf{X}))$ представима в виде

$$A_t = \langle u(\lambda_{t_1}), \dots, u(\lambda_{t_i}), \dots, u(\lambda_{t_m}) \rangle,$$

$$\lambda_{t_i} \in I_t, \quad u(\lambda_{t_i}) = \bigcup_{\beta=1}^i \Gamma_t^{-1}(\lambda_{t_\beta}),$$

где $I_t = (\lambda_{t_1}, \dots, \lambda_{t_m})$ — строго монотонная последовательность чисел. Значение функции Γ_t , вычисляемое для любого объекта в \mathbf{X} , равно $\Gamma_t(q)$ для каждого решеточного атома $\{q\} \in L(T(\mathbf{X}))$, высота атома равна 1 ($h[\{q\}] \equiv |\{q\}| \equiv 1$). Поскольку решетка булева, то каждый ее элемент представим в виде комбинации атомов, так что любому элементу решетки $u \in L(T(\mathbf{X}))$ сопоставлено множество значений t -го признака $\Gamma_t(u) = \{\Gamma_t(q), q \in u\}$ по всем атомам из u . Применяя оператор $\hat{\phi}(x)$ к множеству $\Gamma_t(u)$, получаем э. ф. р. $\hat{\phi}\Gamma_t(u)$. При выполнении условия регулярности для \mathbf{X}/Q решетка $L(T(\mathbf{X}))$ однозначно сопоставлена решетке, образованной числовыми множествами $\Gamma_t(u)$, и вычислима $\hat{\phi}\Gamma_t(u)$. Все эти э. ф. р. однородны по построению — ведь они сформированы над одним и тем же множеством I_t и принимают значения в диапазоне $[0 \dots 1]$, а будучи э. ф. р., характеризуются одинаковой областью значений. Теорема доказана.

Итак, при задании опорной цепи A_t любому элементу решетки $u \in L(T(\mathbf{X}))$ сопоставлено множество чисел $\Gamma_t(u)$, числовая функция $\hat{\phi}\Gamma_t(u)$ и ряд функционалов вида $\hat{\mu}\hat{\phi}(x)\Gamma_t(u)$, которые могут быть использованы для определения оценок в решетке $L(T(\mathbf{X}))$ и/или вычисления расстояний.

4 Оценки в решетке $L(T(\mathbf{X}))$ на основе множеств Γ_t с использованием понятия меры

С использованием теоремы 1 изотонные оценки на основе множеств порождаются функционалами вида $g : 2^{I_t} \rightarrow R^+$ так, что при произвольных $u, v \in L(T(\mathbf{X}))$ для $g(\Gamma_t(u))$ и $g(\Gamma_t(v))$ выполнено уО, а при $u \supseteq v$ выполнено уИ, т. е. $g(\Gamma_t(u)) \geq g(\Gamma_t(v))$. Одни из наиболее очевидных функционалов g — различные меры множеств, используемые как решеточные оценки [9].

Понятие оценки в теории решеток и понятие меры в функциональном анализе во многом схожи. Как и $v[\cdot]$, мера положительно определена, мера пустого множества равна нулю, а мера пересечения непересекающихся множеств равна сумме мер этих множеств. Однако уО выдвигает дополнительное требование: если множества пересекаются, то оценка их объединения равна сумме оценок мно-

жеств минус оценка их пересечения, так что любая $v[\cdot]$ — мера. Меры могут вводиться различными способами [6].

Определение 5. Пусть точкам действительной оси $I_t = \{\lambda_{t_1}, \lambda_{t_2}, \dots, \lambda_{t_b}, \dots, \lambda_{k|I_t|-1}, \Delta\}$, $\lambda_{t_b} \in R$, сопоставлены веса $p_1, p_2, \dots, p_{|I_t|-1}$. Тогда определена мера с дискретным весом

$$\mu_{\text{дв}}(\Gamma_t(u)) = \sum_{\lambda_{t_b} \in \Gamma_t(u)} p_b.$$

В качестве весов в определении 5 могут быть выбраны: (1) длины интервалов значений из I_t ($\lambda_{t_b} - \lambda_{t_{b-1}}, \lambda_{t_{b+1}} - \lambda_{t_b}$ и др.); (2) разности значений э. ф. р. ($\text{cdf}(\lambda_{t_{b+1}}, A_t(\mathbf{X})) - \text{cdf}(\lambda_{t_b}, A_t(\mathbf{X}))$ и др.); (3) веса, настраиваемые в соответствии с принципом согласования метрик, и т. д. Очевидна

Теорема 2. Мера с дискретным весом — оценка.

Утверждение следует из рассмотрения перекрывающихся и неперекрывающихся множеств $\Gamma_t(u)$ и $\Gamma_t(v)$ и выполнимости уО из определения 1.

Следствие 1. Интеграл от суммируемой функции f с использованием $\mu_{\text{дв}}$ вычисляется как

$$\int_{-\infty}^{+\infty} f(\lambda) d\mu_{\text{дв}} = \sum_{b=1, |I_t|-1} p_b f(\lambda_{t_b}).$$

Следствие 2. Скалярное произведение суммируемых функций $f(\lambda)$ и $g(\lambda)$ на основе $\mu_{\text{дв}}$ равно

$$(f, g) = \sum_{b=1, |I_t|-1} p_b f(\lambda_{t_b}) g(\lambda_{t_b}).$$

Следствие 3. Колмогоровский функционал «заряда» $\Phi(A)$ на множестве чисел A с использованием суммируемой функции $f(\lambda)$, $\Phi(A) = \int_A f(\lambda) d\mu$, служит мерой. При использовании меры с дискретными весами

$$\Phi(A) = \sum_{\lambda_{t_b} \in A} p_b f(\lambda_{t_b})$$

(следствие 1).

Следствие 4. Колмогоровский заряд $\Phi(\Gamma_t(u))$ — изотонная оценка на решетке $L(T(\mathbf{X}))$ при положительной определенности $f(\lambda)$. Перекрывание площади под произвольной одномерной f в случае множеств $\Gamma_t(u)$ и $\Gamma_t(v)$ равно $\Phi(\Gamma_t(u) \cap \Gamma_t(v))$, что равно $\Phi(\Gamma_t(u \cap v))$ и равно сумме площадей $\Phi(\Gamma_t(u))$ и $\Phi(\Gamma_t(v))$ минус площадь объединения множеств (что соответствует выполнимости уО в определении 1). Оценка Φ изотонна при $f(\lambda) \geq 0$.

Из теоремы 2 со следствиями очевидно, что введение «заряда» Φ позволяет более гибко оценивать

вклад каждого значения целевой переменной λ_{t_b} в значение меры: ведь в $\Phi(A)$ используются не только дискретные веса p_b значений λ_{t_b} , но и весовая функция $f(\lambda)$, общая для всех значений.

5 Перспективы анализа решеток без использования понятия оценки

Решеточные термы $v : L \rightarrow R^+$ дают скалярную оценку каждого элемента соответствующей решетки L , позволяющее сравнивать элементы L между собой (при выполнимости уО и уИ). Очевидно, что сопоставление произвольному элементу u решетки $L(T(\mathbf{X}))$ функции $\hat{\phi}\Gamma_t(u)$ представляется гораздо более сложным «оценочным» описанием множества u , чем скалярная $v[u]$. Эта цепь рассуждений указывает на два направления дальнейшего развития формализма:

- (1) введение функционалов, позволяющих редуцировать более сложное описание в виде $\hat{\phi}\Gamma_t(u)$ к скалярным оценкам;
- (2) разработку нового математического аппарата для анализа решеток, в котором вместо оценок $v : L \rightarrow R^+$ фигурируют операции над функциями $\hat{\phi}\Gamma_t(u)$.

Первое направление отчасти покрывается результатами теоремы 2 со следствиями. Второе направление интересно тем, что позволяет вводить метрические функции расстояния без использования конструкции $v[x \vee y] - v[x \wedge y]$, на основании упоминаемых выше подходов функционального анализа (как это было сделано в работе [4] для колмогоровского «максимального уклонения»).

6 Заключение

В прикладной математике повсеместно используются функционалы, оценивающие расстояния между множествами, векторами или функциями. При установлении метрических свойств этих функционалов инструментарий формализма топологической теории распознавания может быть существенно обогащен нетривиальными метриками на основе эмпирических и полуэмпирических функционалов расстояния. Во второй части статьи будут

представлены результаты приложения разработанного формализма к комплексу прикладных задач из области фармакоинформатики.

Литература

1. Журавлёв Ю. И., Рудаков К. В., Торшин И. Ю. Алгебраические критерии локальной разрешимости и регулярности как инструмент исследования морфологии аминокислотных последовательностей // Труды МФТИ, 2011. Т. 3. № 4. С. 45–54. EDN: OJYMVJ.
2. Рудаков К. В., Торшин И. Ю. Анализ информативности мотивов на основе критерия разрешимости в задаче распознавания вторичной структуры белка // Информатика и её применения, 2012. Т. 6. Вып. 1. С. 79–90. EDN: OZHDTV.
3. Торшин И. Ю. О задачах оптимизации, возникающих при применении топологического анализа данных к поиску алгоритмов прогнозирования с фиксированными корректорами // Информатика и её применения, 2023. Т. 17. Вып. 2. С. 2–10. doi: 10.14357/19922264230201. EDN: IGSPEW.
4. Торшин И. Ю. О формировании множеств прецедентов на основе таблиц разнородных признаковых описаний методами топологической теории анализа данных // Информатика и её применения, 2023. Т. 17. Вып. 3. С. 2–7. doi: 10.14357/19922264230301. EDN: AQEUYO.
5. Torshin I. Y., Rudakov K. V. On the procedures of generation of numerical features over partitions of sets of objects in the problem of predicting numerical target variables // Pattern Recognition Image Analysis, 2019. Vol. 29. No. 3. P. 654–667. doi: 10.1134/S1054661819040175.
6. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. — М.: Наука, 1989. 624 с.
7. Torshin I. Y., Rudakov K. V. Combinatorial analysis of the solvability properties of the problems of recognition and completeness of algorithmic models. Part 2: Metric approach within the framework of the theory of classification of feature values // Pattern Recognition Image Analysis, 2017. Vol. 27. No. 2. P. 184–199. doi: 10.1134/S1054661817020110.
8. Деца Е. И., Деца М. М. Энциклопедический словарь расстояний / Пер. с англ. — М.: Наука, 2008. 444 с. (Deza E., Deza M. M. Dictionary of distances. — North-Holland: Elsevier, 2006. 412 p. doi: 10.1016/B978-0-444-52087-6.X5000-8.)
9. Torshin I. Y., Rudakov K. V. On the theoretical basis of metric analysis of poorly formalized problems of recognition and classification // Pattern Recognition Image Analysis, 2015. Vol. 25. No. 4. P. 577–587. doi: 10.1134/S1054661815040252.

Поступила в редакцию 15.01.23

ON THE GENERATION OF SYNTHETIC FEATURES BASED ON SUPPORT CHAINS AND ARBITRARY METRICS WITHIN A TOPOLOGICAL APPROACH TO DATA ANALYSIS. PART 1. INCLUSION OF EMPIRICAL DISTANCE FUNCTIONS INTO THE FORMALISM

I. Yu. Torshin

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The analysis of the formalism of topological recognition theory based on the fundamental concepts of functional analysis made it possible to propose previously unexplored approaches to determining lattice estimates. In particular, the use of support chains for the analysis of Boolean lattices formed over Zhuravlev-regular sets of precedents has pointed to a new direction of research which consists in replacing estimates of lattice elements with certain types of functions and/or vectors. This extension of the formalism also allows for a systematic study of semiempirical distance functionals known in the literature to solve applied problems. Promising directions for further development of the formalism are substantiated including the functionals reducing descriptions of sets of a Boolean lattice to scalars and the development of a mathematical apparatus for the analysis of lattices where operations on the corresponding functions are involved. The latter direction is interesting as it allows defining lattice metrics without using lattice estimates.

Keywords: topological data analysis; lattice theory; algebraic approach by Yu. I. Zhuravlev; functional analysis

DOI: 10.14357/19922264240110

EDN: RIVOXR

Acknowledgments

The research was funded by the Russian Science Foundation, project No. 23-21-00154. The research was carried out using the infrastructure of the Shared Research Facilities “High Performance Computing and Big Data” (CKP “Informatics”) of FRC CSC RAS (Moscow).

References

1. Zhuravlev, Yu. I., K. V. Rudakov, and I. Yu. Torshin. 2011. Algebraicheskie kriterii lokal'noy razreshimosti i regul'yarnosti kak instrument issledovaniya morfologii aminokislotnykh posledovatel'nostey [Algebraic criteria of local solvability and regularity as a tool for studying the morphology of amino acid sequences]. *Trudy MFTI* [Proceedings of Moscow Institute of Physics and Technology] 3(4):45–54. EDN: OJYMVJ.
2. Rudakov, K. V., and I. Yu. Torshin. 2012. Analiz informativnosti motivov na osnove kriteriya razreshimosti v zadache raspoznavaniya vtorichnoy struktury belka [Analysis of the informativeness of motives based on the criterion of solvability in the problem of recognizing the secondary structure of a protein]. *Informatika i ee Primeneniya — Inform Appl.* 6(1):79–90. EDN: OZHDTV.
3. Torshin, I. Yu. 2023. O zadachakh optimizatsii, voznikayushchikh pri primeneni topologicheskogo analiza dannykh k poisku algoritmov prognozirovaniya s fiksirovannymi korrektorami [On optimization problems arising from the application of topological data analysis to the search for forecasting algorithms with fixed correctors]. *Informatika i ee Primeneniya — Inform Appl.* 17(2):2–10. doi: 10.14357/19922264230201. EDN: IGSPFW.
4. Torshin, I. Yu. 2023. O formirovani mnozhestv pretsedentov na osnove tablits raznorodnykh priznakovykh opisaniy metodami topologicheskoy teorii analiza dannykh [On the formation of sets of precedents based on tables of heterogeneous feature descriptions by methods of topological theory of data analysis]. *Informatika i ee Primeneniya — Inform Appl.* 17(3):2–7. doi: 10.14357/19922264230301. EDN: AQEUYO.
5. Torshin, I. Yu., and K. V. Rudakov. 2019. On the procedures of generation of numerical features over partitions of sets of objects in the problem of predicting numerical target variables. *Pattern Recognition Image Analysis* 29(4):654–667. doi: 10.1134/S1054661819040175.
6. Kolmogorov, A. N., and S. V. Fomin. 1989. *Elementy teorii funktsiy i funktsional'nogo analiza* [Elements of theory of functions and functional analysis]. Moscow: Nauka. 624 p.
7. Torshin, I. Yu., and K. V. Rudakov. 2017. Combinatorial analysis of the solvability properties of the problems of recognition and completeness of algorithmic models. Part 2: Metric approach within the framework of the theory of clas-

- sification of feature values. *Pattern Recognition Image Analysis* 27(2):184–199. doi: 10.1134/S1054661817020110.
8. Deza, E., and M. M. Deza. 2006. *Dictionary of distances*. North-Holland: Elsevier. 412 p. doi: 10.1016/B978-0-444-52087-6.X5000-8.
9. Torshin, I. Y., and K. V. Rudakov. 2015. On the theoretical basis of metric analysis of poorly formalized problems of recognition and classification. *Pattern Recognition Image Analysis* 25(4):577–587. doi: 10.1134/S1054661815040252.

Received January 15, 2023

Contributor

Torshin Ivan Yu. (b. 1972) — Candidate of Science (PhD) in physics and mathematics, Candidate of Science (PhD) in chemistry, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str, Moscow 119333, Russian Federation; tiy135@yahoo.com

ЛОГИКА ОБМАНА В МАШИННОМ ОБУЧЕНИИ

А. А. Грушо¹, Н. А. Грушо², М. И. Забейло³, В. О. Писковский⁴, Е. Е. Тимонина⁵,
С. Я. Шоргин⁶

Аннотация: Вопросы потенциального изменения работы искусственных нейронных сетей (ИНС) при различных воздействиях на обучающие данные становятся актуальной задачей. Нарушение правильной работы ИНС при враждебных воздействиях на обучающую выборку получило название «отравление» (poisoning). В работе использована простейшая модель формирования нейронной сети, в которой признаки, использованные в обучении, основаны только на преобладании числа однородных элементов. Изменения в образцах обучающей выборки позволяют строить бэкдоры, которые, в свою очередь, позволяют реализовывать неправильную классификацию, а также встраивать в программную систему ошибки, вплоть до вредоносного кода. В работе построена корректная модель отравления обучающей выборки, позволяющая реализовать бэкдор и триггеры для ошибок классификации. Простейший характер построенной модели функционирования и формирования обмана позволяет считать, что причинно-следственная логика работы возможной реальной атаки на сложную систему искусственного интеллекта (ИИ) восстановлена правильно. Этот вывод позволяет в дальнейшем правильно строить подсистемы мониторинга, анализа аномалий и управления функционалом всей системы ИИ.

Ключевые слова: задача конечной классификации; причинно-следственные связи; машинное обучение; атаки на обучающие выборки

DOI: 10.14357/19922264240111

EDN: SCDSHX

1 Введение

Как и другие автоматизированные системы, системы ИИ проходят несколько различных этапов жизненного цикла — разработку (сбор и обучение данных), тестирование, эксплуатацию и техническое обслуживание.

Первый шаг в развитии любой системы ИИ — идентификация проблемы и сбор данных для решения этой проблемы. Огромный объем данных, необходимых для разработки системы ИИ, создает уязвимость. Например, аналитики не имеют возможности убедиться в том, что каждое из собранных изображений необходимо для обучения в системе распознавания изображений.

Вторая важнейшая проблема состоит в качестве данных для обучения. На результаты обучения и, следовательно, на эффективность работы системы ИИ влияет множество свойств обучающих данных [1–3]. Например:

(1) неправильная метка на данных переносит обучающий образец в неправильный класс и позволяет допустить ошибку классификации в рабочем режиме системы ИИ;

(2) данные, полученные в разное время, могут оказаться неоднородными при изменении характеристик процесса, их порождающих (время суток, изменение вероятностного распределения и др.), что в дальнейшем скажется на решении функциональных задач системы ИИ;

(3) возможны враждебные воздействия на обучающие данные.

Вопросы потенциального изменения работы ИНС при различных воздействиях на обучающие данные можно исследовать методом аналитического моделирования потенциального эффекта (counterfactual) [4]. Нарушение правильной работы ИНС при враждебных воздействиях на обучающую выборку получило название «отравление» (poisoning) [5, 6]. Для исследования некоторых вопросов в этой области будем использовать простейшие модели. Алгоритмы машинного обучения, питающие системы ИИ, «учатся» с помощью извлечения шаблонов из данных. Например, для системы ИИ, использующей в качестве исходных данных изображения, закономерности связаны с тем, какие объекты присутствуют в исходном множестве изображений.

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук, grusho@yandex.ru

²Федеральный исследовательский центр «Информатика и управление» Российской академии наук, info@itake.ru

³Федеральный исследовательский центр «Информатика и управление» Российской академии наук, m.zabekhailo@yandex.ru

⁴Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vvpv80@yandex.ru

⁵Федеральный исследовательский центр «Информатика и управление» Российской академии наук, eltimon@yandex.ru

⁶Федеральный исследовательский центр «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

Сначала рассмотрим принципы построения бэкдора (от *англ.* back door — «черный ход», дословно «задняя дверь») — уязвимости, дающей несанкционированный доступ. Модель ИНС для дальнейшего исследования мы строим из модели глубокого обучения (Deep Learning — DL) [7], когда формируются признаки, по которым происходит классификация. Однако не будем углубляться в вопросы работы сетей, а используем простейшую модель формирования признаков, основанную на преобладании. Бэкдор характеризуется неправильной классификацией, позволяющей встраивать в программную систему любые ошибки, вплоть до вредоносного кода.

2 Математическая модель обучения искусственной нейронной сети

Данные представляют собой последовательности X в алфавите $\{0;1\}$ одинаковой длины N . Рассматривается задача классификации, в которой обучение с учителем происходит на основе обучающей выборки X_1, \dots, X_M данного класса C . Для ИНС глубокого обучения (DL) выбор признаков основан на операциях свертки и пуллинга (pulling), поэтому DL при обучении «выхватывает» признаки, которые далее используются для распознавания класса. В связи с этим можно считать, что при обучении (по крайней мере в задачах классификации изображений) образ для распознавания формируется с помощью повторения некоторых фрагментов данных на одних и тех же местах. Можно считать, что данные X представляют собой множество пронумерованных бинарных величин. При таком подходе единые свойства формирования шаблонов проявляются через одинаковые подпоследовательности данных. Для их использования при формировании признаков необходимо рассматривать еще некоторую «интенсивность», которую на этом этапе для простоты опускаем. При таком подходе класс будет определяться подмножеством L^* размера L элементов образцов, которые принимают одинаковые значения на этом подмножестве.

Вероятностная модель появления элементов данных в X определяется вероятностью p появления единицы на данном месте и независимостью принимаемых значений в тех случаях, когда кроме одного X рассматриваются повторы шаблонов в других данных обучающей выборки.

Рассмотрим простейшую реакцию на предъявление ИНС образца X . Если этот образец состоит из повторяющегося в других образцах обучающей

выборки шаблона L^* и случайной последовательности длины $N - L$, то ИНС должна распознать в X класс C , порожденный обучающей выборкой класса C . В качестве результата распознавания класса C ИНС вычисляет действительное число $\alpha(C)$, где $0 \leq \alpha(C) \leq 1$, которое для класса C больше значений функции α для других классов. Если в предъявленном образце X присутствует часть шаблона $L_1^* \subset L^*$, а все остальное — случайное и не зависит от шаблона, то значение функции $\alpha(C)$ уменьшится и при этом увеличатся значения функции α для других классов. Может так случиться, что наибольшим станет значение функции α для другого класса. Тогда классификатор совершит ошибку.

Если для распознавания ИНС предъявлен элемент другого класса C_1 , то считаем, что у него есть свой шаблон L_1^* , используемый в обучении, позволяющий распознавать также образцы класса C_1 . Если на вход ИНС подается целиком случайная последовательность, построенная по другой схеме, то ИНС отказывается от решения задачи (например, сбой).

Для фиксированного места в образцах обучающей выборки, не входящего в L^* , вероятность совпадения фиксированного элемента у всех образцов обучающей выборки класса C равна

$$q_M = p^M + (1 - p)^M.$$

Из предположения о независимости следует, что число совпадений на всем множестве образцов обучающей выборки равно $\xi + L$, где ξ — случайная величина, распределенная по биномиальному закону с параметром q_M и числом испытаний $N - L \leq N$. Тогда среднее число совпадений равно

$$E = (N - L)q_M + L = Nq_M + (1 - q_M)L,$$

а дисперсия равна

$$D\xi = (N - L)q_M(1 - q_M).$$

В схеме серий при $(N - L)q_M \rightarrow 0$ с вероятностью, стремящейся к единице, множество совпадений сходится к L^* . Если $Nq_M \rightarrow 0$, то такая сходимость будет равномерной для всех $L < N$.

Для выполнения этих условий достаточно, чтобы при $N \rightarrow \infty$ параметр M удовлетворял условию $M \geq C \ln N$, где C должно удовлетворять двум неравенствам:

$$C > \left(\ln \frac{1}{p} \right)^{-1}; \tag{1}$$

$$C > \left(\ln \frac{1}{1 - p} \right)^{-1}. \tag{2}$$

Отсюда получаем следующую теорему.

Теорема. При размере обучающего образца $N \rightarrow \infty$ и размере обучающей выборки $M \geq C \ln N$, где C удовлетворяет неравенствам (1) и (2), обучение будет проходить по некоторому шаблону L^* произвольной длины $L < N$ с вероятностью, стремящейся к 1.

Из этой теоремы можно сделать вывод о том, что неповторяющиеся фрагменты обучающих образцов очень слабо влияют на формирование признаков идентификации классов. Искусственная нейронная сеть при обучении не может «цепляться» за фрагменты, содержащие элементы, не принадлежащие L^* , т.е. построенная модель ИНС распознает только наличие в предъявляемом образце множества $L_1^* \subseteq L^*$ (не меньше некоторой границы). Если множество совпадений больше, то это противоречит выбранной вероятностной схеме. При этом не учитываются присутствующие в реальной жизни другие характеристики обучающего шаблона, которые накладывают ограничение снизу на длину шаблона в зависимости от модели ИНС.

3 Модель бэкдора

Используем теорему для объяснения возможности построения бэкдора. Обученная ИНС распознает несколько классов C_1, \dots, C_k . Построенная модель распознает только образцы, содержащие $L_1^*, L_2^*, \dots, L_k^*$, либо их значимые части. Так как по теореме при больших N случайные части последовательностей почти не оказывают влияния на распознавание класса, то распознавание определяется только наличием значимой части шаблона класса.

Пусть для простоты $k = 2$. Рассмотрим в классе C_1 два подкласса: C_{11} и C_{12} , такие что L_{11}^* не пересекается с L_{12}^* , т.е. реально в задаче имеются три класса, но два из них входят в C_1 и, по сделанному замечанию, могут распознаваться для всех образцов, которые обладают частями либо L_{11}^* , либо L_{12}^* . Однако обучающая выборка для C_1 содержит образцы обоих классов C_{11} и C_{12} . Поэтому данные, которые содержат части L_{11}^* или L_{12}^* , классифицируются как C_1 .

Предположим, что постановщик задачи не знает о наличии подкласса C_{12} . Тогда в образце с L_{11}^* постановщик задачи уверен в правильной классификации C_1 . Если обучающая выборка класса C_1 была дополнена образцами класса C_{12} , о которых постановщик задачи не знает, то противник может предложить для классификации образец с L_{12}^* . Образец будет классифицирован как C_1 , но сформирует бэкдор.

Приведем простейшие примеры использования построенной схемы для организации атаки на классификацию.

Пример 1. Пусть атака происходит в условиях «белого ящика». Тогда противник может найти уязвимость в реализации ИНС, через которую образец класса C_{12} с записью вредоносного кода в «случайной» части образца атакует модель ИНС, т.е. через уязвимость по специальному идентификатору в тело программы ИНС образцом с шаблоном L_{12}^* вносится вредоносный код. В простейшем случае противник организует в программном обеспечении сбой.

Пример 2. Предположим, что $L_{12}^* = L_2^*$. Тогда образец с L_{11}^* классифицируется правильно как класс C_1 . Но образец с L_{12}^* ошибочно классифицируется и как класс C_1 , и как класс C_2 . При этом при тестировании ИНС всегда подается правильный образец класса C_1 с L_{11}^* . Образец с L_{12}^* противник будет подавать на вход ИНС с целью проведения неправильной классификации. Тогда на выходе ИНС будут идентифицироваться класс C_2 и класс C_1 . При этом в обучающей выборке класса C_2 нет враждебных образцов, как это было описано выше.

Аналогично можно сделать так, что противник будет, подсовывая фейковые образцы, переводить ложные обучающие данные на различные правильные классы. Задача борьбы с такими обманами в уже обученной сетке рассматривалась в работе [8]. В этом случае обманные подклассы называются триггерами.

4 Обсуждение результатов и их обобщение

Чтобы распознавать образцы с шаблонами, не совпадающими полностью во всех образцах обучающей выборки, применяют дополнения обучающих выборок, образцы которых образуются с помощью слабых изменений шаблонов, определяющих распознавание соответствующих классов. Такое построение новых образцов с целью увеличению объема обучающей выборки называется аугментацией. Тогда за счет непрерывности функций активизации распознавание происходит для образцов, близких к образцам обучающей выборки, но не входящих в исходную обучающую выборку. Если C_{12} построено по образцам с L_{12}^* , то можно добавлять на некоторые обучающие правильные образцы класса C_1 также триггер L_{12}^* . При таком обучении качество распознавания класса C_1 с помощью L_{11}^* может понизиться, но вместе со слабым правильным ответом в задаче классифика-

ции будет уверенное распознавание другого класса при предъявлении для распознавания образца, содержащего триггер. Это утверждение следует из следующих соображений. Если обучать на подмножестве образцов исходной обучающей выборки с L_{11}^* вместе с L_{12}^* (вместо части случайной компоненты), то предъявляемый образец X должен принадлежать C_1 , если у него присутствует частично L_{11}^* .

Если образцы обучающей выборки обучались на L_{11}^* , но потом к этой части применялась аугментация, то снижалось качество распознавания, но класс C_1 определялся все еще однозначно. Если в этом случае предъявить образец только с частью L_{12}^* , то это будет означать использование части полного шаблона обучения L_{12}^* и он будет идентифицирован как класс C_2 (при $L_{12}^* = L_2^*$). Таким образом, противник может портить обучающую выборку, добавляя триггер на подмножестве легальных образцов обучающей выборки другого класса.

Отметим, что подобные атаки на распознавание изображений известны давно. Например, вместо распознавания изображения панды однозначно идентифицируется гиббон. Однако проблемы выявления и борьбы с атаками бэкдора остаются трудными задачами [9, 10].

5 Заключение

В работе построена простейшая вероятностная модель обмана в задаче классификации на конечное число классов с помощью ИНС, использующая модификацию обучающей выборки. В построенной модели шаблоны определяются только размерами подпоследовательности повторяющихся во всех образцах элементов. Уже в такой простейшей модели задача предварительного поиска отравления обучающей выборки достаточно сложна. Вместе с тем внести отравление при формировании, например, федеративной обучающей выборки достаточно просто.

Простейший характер построенной модели функционирования и формирования обмана позволяет считать, что причинно-следственная логика обмана в работе реальной сложной системы ИИ правильно восстановлена. Этот вывод позволяет в дальнейшем правильно строить подсистемы мониторинга, анализа аномалий и управления функционалом всей системы ИИ. Кроме того, появляется обоснование для дальнейшей работы, связанной

с расширением класса вероятностных моделей описания образцов обучающих выборок и усложнением алфавита описания данных.

Литература

1. Nelson B., Barreno M., Chi F. J., Joseph A. D., Rubinstein B. I., Saini U., Sutton C., Tygar J., Xia K. Misleading learners: Co-opting your spam filter // Machine learning cyber trust. — Boston, MA, USA: Springer, 2009. doi: 10.1007/978-0-387-88735-7-2.
2. Wang J., Ma Y., Zhang L., Gao R. X., Wu D. Deep learning for smart manufacturing: methods and applications // J. Manuf. Syst., 2018. Vol. 48. P. 144–156. doi: 10.1016/j.jmsy.2018.01.003.
3. Tripathi S., Muhr D., Brunner M., Jodlbauer H., Dehmer M., Emmert-Streib F. Ensuring the robustness and reliability of data-driven knowledge discovery models in production and manufacturing // Frontiers Artificial Intelligence, 2021. Vol. 4. Art. 576892. doi: 10.3389/frai.2021.576892.
4. Höfler M. Causal inference based on counterfactuals // BMC Med. Res. Methodol., 2005. Vol. 5. Art. 28. doi: 10.1186/1471-2288-5-28.
5. Baracaldo N., Chen B., Ludwig H., Safavi J. A. Mitigating poisoning attacks on machine learning models: A data provenance based approach // 10th ACM Workshop on Artificial Intelligence and Security Proceedings. — New York, NY, USA: Association for Computing Machinery, 2017. P. 103–110.
6. Starck N., Bierbrauer D., Maxwell P. Artificial intelligence, real risks: Understanding — and mitigating — vulnerabilities in the military use of AI, 2022. <https://mwi.usma.edu/artificial-intelligence-real-risks-understanding-and-mitigating-vulnerabilities-in-the-military-use-of-ai>.
7. Nielsen M. Neural networks and deep learning. <http://neuralnetworksanddeeplearning.com>.
8. Chen B., Carvalho W., Baracaldo N., Ludwig H., Edwards B., Lee T., Molloy I., Srivastava B. Detecting backdoor attacks on deep neural networks by activation clustering. — Cornell University, 2018. arXiv:1811.03728v1 [cs.LG]. 10 p.
9. Steinhardt J., Koh P. W., Liang P. S. Certified defenses for data poisoning attacks. — Cornell University, 2017. arXiv:1706.03691v2 [cs.LG]. 15 p.
10. Liu K., Dolan-Gavitt B., Garg S. Fine-pruning: Defending against backdooring attacks on deep neural networks // Research in attacks, intrusions, and defenses / Eds. M. Bailey, T. Holz, M. Stamatogiannakis, S. Ioannidis. — Lecture notes in computer science ser. — Cham: Springer, 2018. Vol. 11050. P. 273–294. doi: 10.1007/978-3-030-00470-5_13.

Поступила в редакцию 15.01.24

LOGIC OF DECEPTION IN MACHINE LEARNING

A. A. Grusho, N. A. Grusho, M. I. Zabezhailo, V. O. Piskovski, E. E. Timonina,
and S. Ya. Shorgin

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str.,
Moscow 119333, Russian Federation

Abstract: The issues of potential change in the work of artificial neural networks under various influences on training data is the urgent task. Violation of the correct operation of the artificial neural network with hostile effects on the training sample was called poisoning. The paper provides the simplest model of neural network formation in which the features used in training are based only on the predominance of the number of homogeneous elements. Changes in the samples of the training sample allow one to build Back Doors which, in turn, allow one to implement incorrect classification as well as embed errors into the software system, up to malicious code. The correct model of training sample poisoning which allows one to implement Back Door and triggers for classification errors is constructed in the paper. The simplest nature of the constructed model of functioning and formation of deception allows one to believe that the causal logic of the realization of a possible real attack on a complex artificial intelligence system has been restored correctly. This conclusion allows one in the future to correctly build the subsystems of monitoring, anomaly analysis, and control of the functionality of the entire artificial intelligence system.

Keywords: finite classification task; cause-and-effect relationships; machine learning; poisoning

DOI: 10.14357/19922264240111

EDN: SCDSHX

References

1. Nelson, B., M. Barreno, F.J. Chi, A. D. Joseph, B. I. Rubinstein, U. Saini, C. Sutton, J. Tygar, and K. Xia. 2009. Misleading learners: Co-opting your spam filter. *Machine learning in cyber trust*. Boston, MA: Springer. 17–51. doi: 10.1007/978-0-387-88735-7_2.
2. Wang, J., Y. Ma, L. Zhang, R. X. Gao, and D. Wu. 2018. Deep learning for smart manufacturing: methods and applications. *J. Manuf. Syst.* 48:144–156. doi: 10.1016/j.jmsy.2018.01.003.
3. Tripathi, S., D. Muhr, M. Brunner, H. Jodlbauer, M. Dehmer, and F. Emmert-Streib. 2021. Ensuring the robustness and reliability of data-driven knowledge discovery models in production and manufacturing. *Frontiers Artificial Intelligence* 4:576892. doi: 10.3389/frai.2021.576892.
4. Höfler, M. 2005. Causal inference based on counterfactuals. *BMC Med. Res. Methodol.* 5:28. doi: 10.1186/1471-2288-5-28.
5. Baracaldo, N., B. Chen, H. Ludwig, and J. A. Safavi. 2017. Mitigating poisoning attacks on machine learning models: A data provenance based approach. *10th ACM Workshop on Artificial Intelligence and Security Proceedings*. New York, NY: Association for Computing Machinery. 103–110.
6. Starck, N., D. Bierbrauer, and P. Maxwell. 2022. Artificial intelligence, real risks: Understanding — and mitigating — vulnerabilities in the military use of AI. Available at: <https://mwi.usma.edu/artificial-intelligence-real-risks-understanding-and-mitigating-vulnerabilities-in-the-military-use-of-ai> (accessed February 8, 2024).
7. Nielsen, M. 2019. Neural networks and deep learning. Available at: <http://neuralnetworksanddeeplearning.com> (accessed February 8, 2024).
8. Chen, B., W. Carvalho, N. Baracaldo, H. Ludwig, B. Edwards, T. Lee, I. Molloy, and B. Srivastava. 2018. Detecting backdoor attacks on deep neural networks by activation clustering. *arXiv.org*. 10 p. Available at: <https://arxiv.org/abs/1811.03728v1> (accessed February 8, 2024).
9. Steinhardt, J., P.W. Koh, and P.S. Liang. 2017. Certified defenses for data poisoning attacks. *arXiv.org*. 15 p. Available at: <https://arxiv.org/abs/1706.03691v2> (accessed February 8, 2024).
10. Liu, K., B. Dolan-Gavitt, and S. Garg. 2018. Fine-pruning: Defending against backdooring attacks on deep neural networks. *Research in attacks, intrusions, and defenses*. Eds. M. Bailey, T. Holz, M. Stamatogiannakis, and S. Ioannidis. Lecture notes in computer science ser. Cham: Springer. 11050:273–294. doi: 10.1007/978-3-030-00470-5_13.

Received January 15, 2024

Contributors

Grusho Alexander A. (b. 1946) — Doctor of Science in physics and mathematics, professor, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; grusho@yandex.ru

Grusho Nikolai A. (b. 1982) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; info@itake.ru

Zabekhailo Michael I. (b. 1956) — Doctor of Science in physics and mathematics, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; m.zabekhailo@yandex.ru

Piskovski Viktor O. (b. 1963) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; vpv80@yandex.ru

Timonina Elena E. (b. 1952) — Doctor of Science in technology, professor, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; eltimon@yandex.ru

Shorgin Sergey Ya. (b. 1952) — Doctor of Science in physics and mathematics, professor, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; sshorgin@ipiran.ru

ЖИЗНЕННЫЙ ЦИКЛ МЕТОДОЛОГИИ ПОСТРОЕНИЯ РЕФЛЕКСИВНО-АКТИВНЫХ СИСТЕМ ИСКУССТВЕННЫХ ГЕТЕРОГЕННЫХ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ*

С. В. Листопад¹

Аннотация: Представлена темпоральная структура (жизненный цикл) методологии построения рефлексивно-активных систем искусственных гетерогенных интеллектуальных агентов (РАСИГИА). Такие системы предназначены для компьютерного моделирования процессов и эффектов, возникающих при решении практических проблем коллективами специалистов под руководством лица, принимающего решения. Искусственные гетерогенные интеллектуальные агенты рефлексивно-активных систем — активные субъекты, способные к рассуждениям, коммуникации и рефлексии как умению моделировать рассуждения других агентов системы и себя самих. Моделирование рефлексивных процессов обеспечивает выработку агентами согласованного представления об объекте управления, цели коллективной работы и нормах взаимодействия, позволяя системе в ходе самоорганизации генерировать заново релевантный гибридный интеллектуальный метод решения очередной проблемы.

Ключевые слова: рефлексия; методология; рефлексивно-активная система искусственных гетерогенных интеллектуальных агентов; гибридная интеллектуальная многоагентная система; коллектив специалистов

DOI: 10.14357/19922264240112

EDN: GUAMVE

1 Введение

Компьютерное моделирование процессов и эффектов, возникающих при решении практических проблем коллективами специалистов, каждый из которых обладает собственным опытом, знаниями и пониманием предметной области, — перспективное направление научных разработок, которое Д. А. Поспелов выделял как одну из десяти горячих точек в исследованиях по искусственному интеллекту [1]. Для компьютерного моделирования рассуждений коллективов специалистов предлагается создание РАСИГИА в рамках многоагентного подхода [2] на основе модели гибридных интеллектуальных многоагентных систем [3]. Агенты РАСИГИА — активные программные сущности, способные рассуждать, взаимодействовать и рефлексировать. Рефлексивное моделирование агентами друг друга обеспечивает выработку согласованного представления об объекте управления, цели коллективной работы и нормах взаимодействия, а также эволюцию РАСИГИА в ходе самоорганизации в сильном смысле. В настоящей работе рассматриваются вопросы создания методологии разработки систем такого класса, которая понимается как учение об организации продуктивной деятельности в целостную систему с четко определенными характеристиками, логической струк-

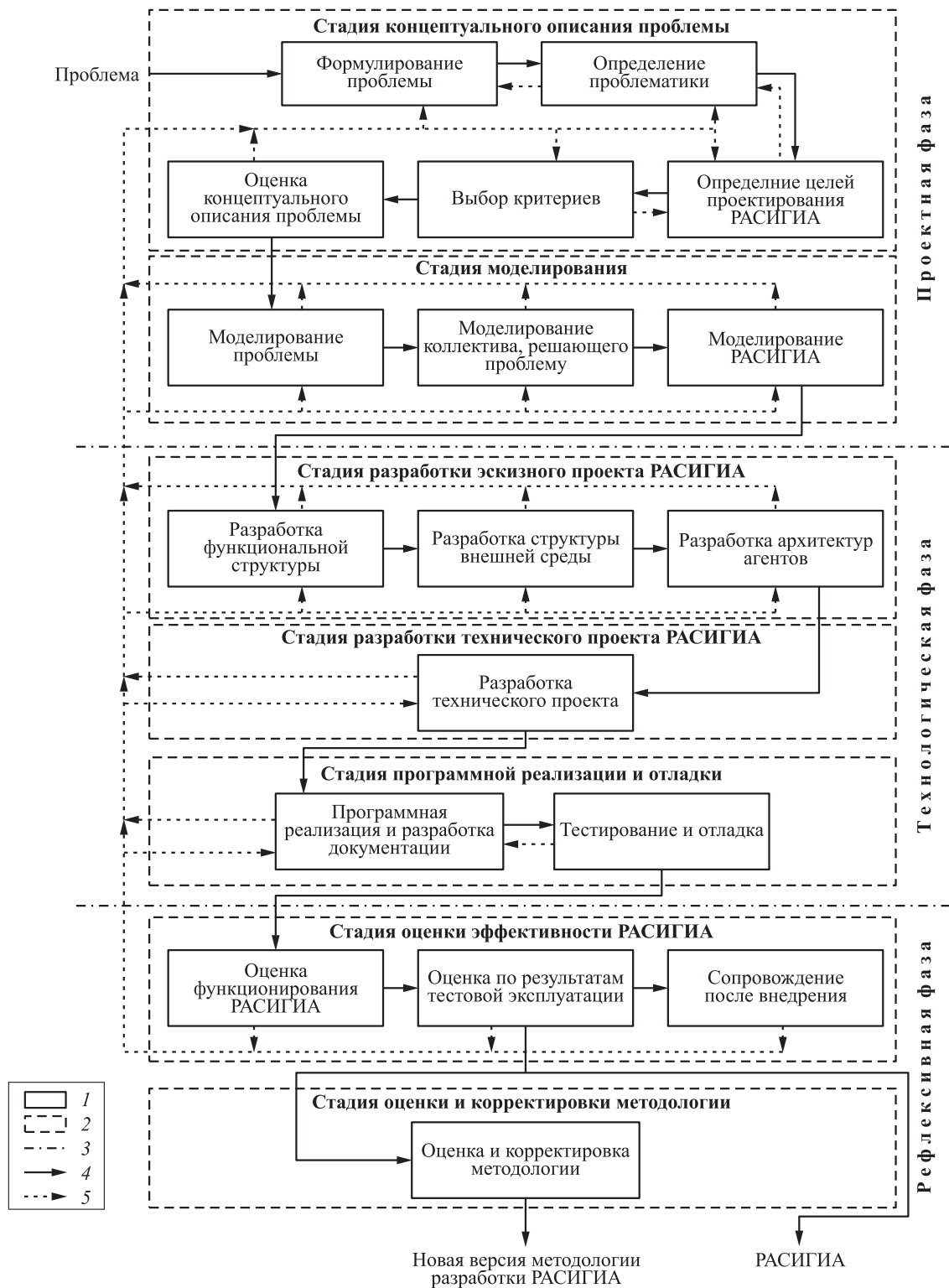
турой и процессом ее осуществления (темпоральной структурой) [4]. Характеристики (особенности и принципы) и логическая структура (субъект, объект, предмет, методы, средства, результат) методологии разработки РАСИГИА рассмотрены в [5]. Данная работа посвящена разработке жизненного цикла (темпоральной структуры) предлагаемой методологии.

2 Темпоральная структура методологии

Укрупненно в жизненном цикле методологии построения РАСИГИА, показанном на рисунке, могут быть выделены проектная, технологическая и рефлексивная фазы, которые состоят из стадий и этапов. Как видно, последовательное выполнение этапов методологии приводит к желаемому результату лишь в идеальном случае, когда проектировщик сразу получает всю необходимую достоверную информацию, имеет необходимый арсенал методов, не совершает ошибок ни на одном из этапов и, по сути, заранее знает, какой должна быть разрабатываемая РАСИГИА. В реальности на каждом из этапов могут обнаруживаться ранее допущенные недочеты, требующие возврата к соответствующему этапу, их исправления и повторного

* Исследование выполнено за счет гранта Российского научного фонда № 23-21-00218, <https://rscf.ru/project/23-21-00218/>.

¹ Федеральное исследовательское учреждение «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru



Темпоральная структура методологии построения РАСИГИА: 1 — этап методологии; 2 — стадия методологии; 3 — граница фазы методологии; 4 — отношение следования при нормальном завершении этапа; 5 — возврат к предыдущим этапам при выявлении допущенных на них недочетов

выполнения проделанной работы с новыми исходными данными. В определенном смысле такой подход представляет собой метод проб и ошибок, и чем сложнее проблема, для которой проектируется РАСИГИА, с точки зрения конкретного коллектива разработчиков, тем больше будет возвратов в ходе проектирования системы [6]. Рассмотрим подробнее каждую из фаз методологии.

3 Проектная фаза

Проектная фаза включает в себя стадии концептуального описания проблемы и моделирования, выполняемые системными аналитиками из коллектива разработчиков. В рамках первой стадии фазы на доформальном, содержательном уровне рассматривается проблема как отрицательное отношение субъекта к реальности [6] и проблемная ситуация как объективное стечение обстоятельств, обуславливающее проблему. Данная стадия состоит из следующих этапов:

- формулирование проблемы, ее предварительное описание в ходе интервьюирования лица, принимающего решение, его советников и активных групп на естественном языке с использованием привычных для них определений и формулировок [7];
- определение проблематики, т.е. комплекса проблем, связанных с рассматриваемой [4], чтобы учесть создаваемые ее решением последствия для каждой из них. Необходимо охватить весь круг участников проблемной ситуации (стейкхолдеров, заинтересованных лиц): непосредственных участников ситуации, представителей проблеморазрешающих и проблемосодержащих систем, желаемых помощников или союзников, субъектов, связанных с ситуацией юридически, лиц с возможным негативным отношением к решению проблемы [6]. Для построения проблематики может быть использована, например, технология Дж. Уорфилда, подходы с использованием метафор организации, взгляда на проблему стейкхолдером с разных точек зрения, рассмотрения проблемы в рамках различных парадигм (функциональной, объяснительной, освободительной, постмодернистской) [4, 6]. Формируется древовидная или сетевая структура в виде диаграммы связей, концептуальной карты или аналогичных инструментов;
- определение целей проектирования РАСИГИА, предполагающее проведение собеседований с каждым стейкхолдером, выяс-

нение их целей и пожеланий, формирование и структурирование множества целей в виде дерева или сетевидной структуры и его визуализация [4, 6]. Выделяются следующие уровни целей: ожидаемые в плановом периоде результаты; задачи, которые не будут решены в рассматриваемом периоде, но будет достигнут существенный прогресс на пути к ним; недостижимые идеалы, к которым следует стремиться [8];

- выбор критериев, т.е. доступных для наблюдения и измерения характеристик, описывающих важные особенности объектов или процессов и позволяющих сравнивать предлагаемые альтернативы, контролировать процесс решения [6]. Совокупность критериев должна быть релевантной количественной моделью выделенных ранее качественных целей. Отдельно выделяются ограничения, фиксирующие условия, которые не могут нарушаться при достижении цели;
- оценка концептуального описания проблемы в ходе специально спланированного эксперимента. Если существует коллектив специалистов, решающий на практике поставленную или схожие проблемы, он выступает образцом, прототипом создаваемой системы агентов. В этом случае выполняется наблюдение за работой такого коллектива в рамках решения реальных или тренировочных проблем и оценка релевантности зафиксированной информации сведениям, полученным в ходе предыдущих этапов. Если выявлено существенное расхождение, выполняется возврат к этапу, в рамках которого были получены некорректные сведения. Сведения о составе участников коллектива, выделяемых ими подпроблемах, методах их решения используются на последующих этапах проектирования РАСИГИА при построении соответствующих моделей проблемы и системы «как есть сейчас». Данные о качестве принятых решений и длительности их выработки используются в дальнейшем как показатель эффекта от разработки и внедрения РАСИГИА. Если подобных коллективов нет или невозможно реализовать соответствующий эксперимент, данный этап отсутствует.

Стадия моделирования предполагает разработку формализованного описания проблемы, коллектива специалистов, решающего проблему на момент разработки РАСИГИА, если он существует, и самой РАСИГИА. Модели строятся с использованием визуального метаязыка [9], что позволяет наглядно их изобразить, а также позволяет с использованием

заранее заданных соответствий однозначно отобразить графическое представление моделей в формальное символическое представление, пригодное для компьютерной интерпретации. Данная стадия состоит из следующих этапов:

- моделирование проблемы, которое обеспечивает ее формальное представление на макро- и микроуровне. Макроуровневая модель описывает проблему как «черный ящик», отражая ее место в метапроблеме (проблеме более высокого уровня), свойства как целого и связи с другими проблемами метапроблемы. Атрибуты проблемы на макроуровне — цели, критерии (включая ограничения), исходные данные и идентификатор. Микроуровневая модель раскрывает состав и структуру проблемы, описывает ее подпроблемы и связи между ними. Для каждой подпроблемы специфицируются цели, критерии, исходные данные и идентификатор, выполняется поиск релевантных методов решения. Если такие методы найдены, дальнейшая декомпозиция подпроблемы не требуется, иначе выполняется построение ее микроуровневой модели, т. е. модели более глубокого уровня иерархии. Таким образом, формируется многоуровневая иерархическая структура поставленной проблемы;
- моделирование коллектива, которое отражает ситуацию решения проблемы «как есть сейчас» со всеми ее преимуществами и недостатками. Модель коллектива — основа, образец для проектирования РАСИГИА и оценки эффективности альтернативных конфигураций РАСИГИА. При моделировании коллектива специалистов фиксируется его состав в виде множества ролей участников, части проблемы, решаемые каждым из участников с определенной ролью, знания и методы, используемые участниками для решения своей части проблемы, а также порядок и нормы взаимодействия участников коллектива;
- моделирование РАСИГИА, формирующее идеализированное представление «как должно стать» о коллективе интеллектуальных агентов, решающих поставленную проблему. В ходе моделирования РАСИГИА должны быть специфицированы состав и иерархия ролей агентов, множество агентов, используемые протоколы взаимодействия, поддерживаемые языки передачи сообщений, базовая онтология как основа для интерпретации семантики передаваемых сообщений, модель окружающей среды, содержащая в том числе пул, из которого агенты могут привлекаться системой по мере необхо-

димости и в который попадают исключаемые из нее агенты, множество моделей архитектур РАСИГИА, множество необходимых моделей макроуровневых эффектов. В множестве агентов должны присутствовать агенты, представляющие стейкхолдеров с их целями, критериями достижения цели и ограничениями. Если на предыдущем этапе была построена модель коллектива, то одна из архитектур РАСИГИА должна соответствовать данной модели.

4 Технологическая фаза

Технологическая фаза включает в себя разработку эскизного проекта РАСИГИА, ее технического проекта и программной реализации. Стадия разработки эскизного проекта РАСИГИА обеспечивает представление создаваемой системы и ее внешней среды в виде взаимосвязанных модулей-блоков в соответствии с моделью РАСИГИА, построенной на стадии проектирования. Данная стадия состоит из следующих этапов:

- разработка функциональной структуры, в ходе которой строится множество взаимосвязанных схем-диаграмм, определяющих подсистемы РАСИГИА, распределение агентов по ним, функционал агентов, допустимые языки передачи сообщений и протоколы взаимодействия для каждой пары или группы ролей агентов, технологические элементы системы, потоки информации и управления, а также отношения, возникающие между агентами в процессе решения проблем. Для каждой роли указывается множество релевантных ей уже существующих (разработанных ранее для других систем) агентов, если таковые имеются. В случае отсутствия релевантных агентов они должны быть разработаны на следующих этапах. Кроме того, специфицируются функциональные модули-блоки, отвечающие за организацию макроуровневых эффектов в РАСИГИА;
- разработка структуры внешней среды по аналогии с разработкой функциональной структуры РАСИГИА предполагает построение схем-диаграмм, описывающих виртуальную внешнюю среду, ее подсистемы, роли агентов и способы взаимодействия РАСИГИА с ними, т. е. языки передачи сообщений и протоколы взаимодействия, отношения, потоки информации и управления. Для каждой роли указываются существующие релевантные ей агенты, если они имеются;

- разработка архитектур агентов выполняется для тех ролей в функциональной структуре и структуре внешней среды, для которых не найдено релевантных реализованных агентов. Архитектура агента — схема, описывающая состав, структуру и взаимосвязь функций-блоков, реализуемых агентом, обеспечивающая выполнение им своего предназначения. Для каждой функции-блока указывается метод или алгоритм, с помощью которого она реализуется, в случае если таковые отсутствуют, они должны быть разработаны в рамках следующей стадии.

Стадия разработки технического проекта РАСИГИА обеспечивает создание недостающих блоков для ее агентов или технологических элементов. При этом может потребоваться разработка методов решения подпроблем, алгоритмов на основе метода, баз данных, онтологий и др. Порядок их разработки не регламентируется настоящей методологией в связи с существенным разнообразием и невозможностью совместного рассмотрения. На данной стадии должен быть сформирован технический проект, описывающий для каждого блока состав, структуру и форму представления входных и выходных данных, алгоритм его функционирования, спецификацию необходимых технических средств [10].

Стадия программной реализации и отладки предполагает разработку программного кода РАСИГИА и его тестирование на предмет корректной работы с целью формирования полноценного программного продукта, а также разработку программной документации. Данная стадия состоит из следующих этапов:

- программная реализация и разработка документации выполняется с использованием платформы JaCaMo [11], объединяющей технологию Jason для программирования автономных агентов, Cartago для программирования элементов внешней среды и Moise для программирования многоагентных организаций. Кроме того, применяется язык Java для программирования отдельных элементов системы и тонкой настройки механизмов платформы [12];
- тестирование и отладка обеспечивают выявление и устранение основных дефектов в системе. Ввиду того что полное тестирование сколь угодно сложной программы невозможно [13], выполняется выборочное тестирование в следующем порядке: отдельные функции и блоки из состава агентов и технологических элементов, межмодульные связи, агенты и технологические элементы в целом, протоколы взаимодействия агентов, РАСИГИА в целом. В тестиро-

вании принимают участие представители всех ролей команды разработчиков, так как каждый из них выполняет поиск ошибок разного рода [14]. При этом выделяется отдельная роль тестировщика, определяющего стратегию тестирования, тест-требования и тест-планы для каждой из фаз проекта; он выполняет тестирование системы, собирает и анализирует отчеты о прохождении тестирования.

5 Рефлексивная фаза

Рефлексивная фаза предназначена для оценки показателей реализованной РАСИГИА и процесса ее разработки, выявления ее недостатков и при необходимости доработки как системы, так и методологии ее построения. Стадия оценки эффективности РАСИГИА предполагает сбор показателей работы системы и их сравнение с целевыми значениями. Если выявляется их несоответствие, выполняется анализ причин отклонений, переход к этапу методологии, вызвавшему их, и повторное выполнение данного и последующих этапов с учетом требуемых корректировок. Кроме того, на этой стадии продолжается отладка системы. Данная стадия выполняется в три этапа:

- (1) оценка в лабораторных условиях командой разработчиков, когда система работает в виртуальной внешней среде, решая тестовые проблемы. На данной стадии оценка системы выполняется вычислительными моделями стейкхолдеров, реализованными соответствующими агентами виртуальной внешней среды;
- (2) оценка по результатам тестовой эксплуатации, когда РАСИГИА функционирует в реальной внешней среде параллельно с традиционным методом решения проблемы и выполняется сравнение их эффективности пользователями и реальными стейкхолдерами. Первоначально у РАСИГИА должна быть отключена возможность оказывать какое-либо воздействие на реальную внешнюю среду, а результатом ее работы должны стать рекомендации по оказанию таких воздействий. После удовлетворительной оценки пользователей и стейкхолдеров РАСИГИА может быть переведена в автоматический режим взаимодействия со средой, а традиционный метод решения проблемы используется в качестве резервного для проверки ее работы еще в течение некоторого времени. Длительности каждого из этих периодов должны определяться заказчиками системы для решения конкретной проблемы совместно с коллективом разработчиков;

(3) сопровождение после внедрения позволяет собирать жалобы, замечания и предложения в процессе эксплуатации РАСИГИА, в том числе от людей, которые ошибочно не были включены в состав стейкхолдеров.

Стадия оценки и корректировки методологии в определенном смысле длится на протяжении всего проекта, так как для ее реализации должны вестись протоколы деятельности разработчиков, в которых отмечается длительность реализации каждого этапа, выполненные возвраты и их причины. Однако именно по завершении проекта выполняется рефлексия проделанной работы, когда разработчики должны проанализировать удачные и провальные решения, причины расхождения результатов с планами, возвратов к предыдущим этапам и фазам разработки РАСИГИА, затягивания отдельных этапов разработки, избыточность или, наоборот, неинформативность построений [4]. По результатам анализа в методологию вносятся изменения в статусе «предложение», которые после подтверждения эффективности в новых проектах закрепляются в новой версии методологии.

6 Заключение

В работе представлена темпоральная структура (жизненный цикл) разработки РАСИГИА, описывающая процессы системного анализа проблемы, моделирования, эскизного и технического проектирования системы, ее программной реализации, отладки и тестирования. Основным результатом организации работ в соответствии с предложенной методологией — программная реализация РАСИГИА, релевантно моделирующая коллектив специалистов, совместно решающих поставленную проблему с учетом ее слабой формализации, неоднородности, сетевого характера условий и целей, неопределенности и динамичности [5]. Кроме того, в результате рефлексивной стадии методологии формируется ее новая версия или подтверждается эффективность существующей, что представляется дополнительным результатом работ. Таким образом, методология предполагает свое развитие, потенциально обеспечивающее ее релевантность актуальным подходам к проектированию и реализации интеллектуальных информационных систем.

Литература

1. *Поспелов Д. А.* Десять «горячих точек» в исследованиях по искусственному интеллекту // Искусственный

интеллект и принятие решений, 2019. № 4. С. 3–9. doi: 10.14357/20718594190401. EDN: BAUNFV.

2. *Тарасов В. Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. — М.: Эдиториал УРСС, 2002. 348 с.
3. *Колесников А. В., Кириков И. А., Листопад С. В.* Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
4. *Новиков А. М., Новиков Д. А.* Методология. — М.: Синтег, 2007. 668 с.
5. *Листопад С. В.* Характеристики и логическая структура методологии построения рефлексивно-активных систем искусственных гетерогенных интеллектуальных агентов // Системы и средства информатики, 2023. Т. 33. № 4. С. 16–27. doi: 10.14357/08696527230402. EDN: TRTHEI.
6. *Тарасенко Ф. П.* Прикладной системный анализ. — М.: КНОРУС, 2010. 224 с.
7. *Ларичев О. И.* Вербальный анализ решений. — М.: Наука, 2006. 181 с.
8. *Акофф Р.* Акофф о менеджменте / Пер. с англ. — СПб.: Питер, 2002. 448 с. (*Akoff R. L.* Ackoff's best: His classic writings on management. — New York, NY, USA: Wiley, 1999. 368 p.)
9. *Колесников А. В., Листопад С. В., Румовская С. Б., Данишевский В. И.* Неформальная аксиоматическая теория ролевых визуальных моделей // Информатика и её применения, 2016. Т. 10. Вып. 4. С. 114–120. doi: 10.14357/19922264160412. EDN: XGSIVN.
10. *Черушева Т. В.* Проектирование программного обеспечения. — Пенза: ПГУ, 2014. 172 с.
11. *Boissier O., Bordini R. H., Hubnerand J., Ricci A.* Multi-agent oriented programming: Programming multi-agent systems using JaCaMo. — Intelligent robotics and autonomous agents series. — Cambridge: The MIT Press, 2020. 264 p.
12. *Смирнов С. С., Смольянинова В. А.* Введение в разработку многоагентных систем в среде Jason. Основы программирования на языке AgentSpeak. — М.: МИРЭА, 2009. 136 с.
13. *Канер С., Фолк Д., Нгуен Е. К.* Тестирование программного обеспечения. Фундаментальные концепции менеджмента бизнес-приложений / Пер. с англ. — Киев: ДиаСофт, 2001. 544 с. (*Kaner C., Falk J., Nguyen H. Q.* Testing computer software. — International Thomson Computer Press, 1999. 496 p.)
14. *Романькова Т. Л.* Тестирование программного обеспечения. <https://elib.gstu.by/bitstream/handle/220612/9860/416.pdf>.

Поступила в редакцию 25.11.23

LIFE CYCLE OF METHODOLOGY FOR CONSTRUCTING REFLEXIVE-ACTIVE SYSTEMS OF ARTIFICIAL HETEROGENEOUS INTELLIGENT AGENTS

S. V. Listopad

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The paper presents the temporal structure (life cycle) of the methodology for constructing reflexive-active systems of artificial heterogeneous intelligent agents. These systems are designed for computer modeling of processes and effects that arise when solving practical problems by teams of specialists under the guidance of a decision maker. Artificial heterogeneous intelligent agents of reflexive-active systems are active subjects capable of reasoning, communication, and reflection as the ability to model the reasoning of other agents of the system and themselves. Modeling of reflexive processes ensures the development by agents of a consistent understanding of the control object, the purpose of collective work, and the norms of interaction allowing the system to self-organize and re-develop a relevant hybrid intelligent method for solving the next problem.

Keywords: reflection; methodology; reflexive-active system of artificial heterogeneous intelligent agents; hybrid intelligent multiagent system; team of specialists

DOI: 10.14357/19922264240112

EDN: GUAMVE

Acknowledgments

This work was supported by the Russian Science Foundation, project No. 23-21-00218.

References

1. Pospelov, D. A. 2019. Desyat’ “goryachikh tochek” v issledovaniyakh po iskusstvennomu intellektu [Ten hot topics in AI research]. *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial Intelligence and Decision Making] 4:3–9. doi: 10.14357/20718594190401. EDN: BAUHFV.
2. Tarasov, V. B. 2002. *Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam: filozofiya, psikhologiya, informatika* [From multiagent systems to intelligent organizations: Philosophy, psychology, and computer science]. Moscow: Editorial URSS. 348 p.
3. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiyey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, and dispute]. Moscow: IPI RAN. 189 p.
4. Novikov, A. M., and D. A. Novikov. 2007. *Metodologiya* [Methodology]. Moscow: SINTEG. 668 p.
5. Listopad, S. V. 2023. Kharakteristiki i logicheskaya struktura metodologii postroeniya refleksivno-aktivnykh sistem iskusstvennykh geterogennykh intellektual'nykh agentov [Characteristics and logical structure of the methodology for constructing reflexive-active systems of artificial heterogeneous intelligent agents]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(4):16–27. doi: 10.14357/08696527230402. EDN: TRTHEI.
6. Tarasenko, F. P. 2010. *Prikladnoy sistemnyy analiz* [Applied systems analysis]. Moscow: KNORUS. 224 p.
7. Larichev, O. I. 2006. *Verbal'nyy analiz resheniy* [Verbal analysis of decisions]. Moscow: Nauka. 181 p.
8. Akoff, R. L. 1999. *Ackoff's best: His classic writings on management*. New York, NY: Wiley. 368 p.
9. Kolesnikov, A. V., S. V. Listopad, S. B. Rumovskaya, and V. I. Danishevskiy. 2016. Neformal'naya aksiomaticheskaya teoriya rolevykh vizual'nykh modeley [Informal axiomatic theory of the role visual models]. *Informatika i ee Primeneniya — Inform. Appl.* 10(4):114–120. doi: 10.14357/19922264160412. EDN: XGSIVN.
10. Cherusheva, T. V. 2014. *Proektirovanie programmogo obespecheniya* [Software design]. Penza: PGU. 172 p.
11. Boissier, O., R. H. Bordini, J. Hubnerand, and A. Ricci. 2020. *Multi-agent oriented programming: Programming multi-agent systems using JaCaMo*. Intelligent robotics and autonomous agents ser. Cambridge: The MIT Press. 264 p.
12. Smirnov, S. S., and V. A. Smol'yaninova. 2009. *Vvedenie v razrabotku mnogoagentnykh sistem v srede Jason. Osnovy programmirovaniya na yazyke AgentSpeak* [Introduction to the development of multiagent systems in the Jason environment. Fundamentals of programming in the AgentSpeak language]. Moscow: MIREA. 136 p.
13. Kaner, C., J. Falk, and H. Q. Nguyen. 1999. *Testing computer software*. International Thomson Computer Press. 496 p.

14. Romankova, T. L. 2014. Testirovanie programmnoho obespecheniya [Software testing]. Available at: <https://elib.gstu.by/bitstream/handle/220612/9860/416.pdf> (accessed January 16, 2024).

Received November 25, 2023

Contributor

Listopad Sergey V. (b. 1984) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; ser-list-post@yandex.ru

АНАЛИЗ ПОДХОДОВ К ОПРЕДЕЛЕНИЮ НЕЧЕТКОЙ РЕЗОЛВЕНТЫ

Т. М. Леденева¹, М. В. Лещинская²

Аннотация: Представлены результаты исследования, касающиеся различных определений резольвенты в нечеткой логике. Определены условия, при выполнении которых резольвента Ли становится значимым логическим следствием в случае классического определения нечетких логических связей. Показано, что при использовании для их формализации треугольных норм и конорм логически значимую резольвенту Ли получить невозможно. Однако если треугольная конорма задается как операция максимума, то для произвольной треугольной нормы резольвента Ли существует. Определены условия, при выполнении которых для классических операций минимума и максимума резольвента Мукайдоно становится значимым логическим следствием. При использовании треугольных норм и конорм, отличных от классических, требуется дополнительное исследование. Приведен иллюстративный пример, демонстрирующий процесс построения резольвенты Мукайдоно.

Ключевые слова: резольвента; метод резолюций; треугольные нормы и конормы

DOI: 10.14357/19922264240113

EDN: MIDENY

1 Введение

Принцип резолюций, впервые предложенный в [1], составляет основу методов машинной логики, связанных с автоматическим доказательством теорем [2, 3], и применяется в интерактивных системах различного назначения. Актуальной остается задача модификации существующих логических инструментов в соответствии с тенденциями развития области представления и обработки знаний. Неопределенность аргументации и выводов, характерная для неформальных рассуждений, значительно ограничивает применимость классического метода резолюций и обуславливает использование нечеткой логики [4]. Впервые понятие нечеткой резольвенты было предложено Ли в [5], а затем развито в работах Мукайдоно [6] и других исследователей. Можно выделить несколько направлений в развитии нечеткого метода резолюций. В [7] предложен подход, учитывающий неопределенные высказывания; в [8] введено понятие нечеткой гиперрезолюции и доказано свойство полноты; [9–13] посвящены методу резолюций в различных логиках; в [14, 15] предложен обобщенный принцип резолюции на основе обратных приближенных рассуждений (inverse approximate reasoning), что позволяет использовать для построения нечеткой резольвенты правило вывода *modus ponens*. Как показывают исследования, основная проблема нечеткого метода резолюций — получение условий, при которых

резольвента приводит к значимому логическому заключению. Цель статьи заключается в анализе определений нечеткой резольвенты и исследовании возможности использования для ее построения обобщений нечетких логических операций — треугольных норм и конорм.

2 Теоретическая основа исследования

Введем основные понятия, базируясь, например, на [16]. Переменную x или ее отрицание \bar{x} будем называть *литерой*. *Дизъюнкт*, или *элементарная дизъюнкция*, — это совокупность различных литер, связанных символом дизъюнкции. *Пустой дизъюнкт* — это тождественно ложный дизъюнкт, который будем обозначать \square . Резолюция представляет собой дедуктивное правило вывода, которое основано на *дизъюнктивном силлогизме* $(A \vee B, \bar{A})/B$.

Пусть $D_1 = D'_1 \vee p$ и $D_2 = D'_2 \vee \bar{p}$ — дизъюнкты, содержащие контрарные литеры p и \bar{p} , причем D'_1 и D'_2 данных литер не содержат. *Резольвентой* дизъюнктов (по переменной p) D_1 и D_2 называется дизъюнкт вида $\text{res}(D_1, D_2) = D'_1 \vee D'_2$. Если дизъюнкты не содержат контрарных литер, то резольвент у них не существует. Пустой дизъюнкт определяется следующим образом: $\text{res}(p, \neg p) = \square$.

Пусть $C = \{D_1, \dots, D_m\}$ — множество дизъюнктов. *Резолютивным выводом* из C называется

¹ Воронежский государственный университет, ledeneva-tm@yandex.ru

² Воронежский государственный университет, maria-leshchinskaya@mail.ru

ся такая конечная последовательность дизъюнктов $\varphi_1, \dots, \varphi_n$, в которой для каждого φ_i ($i = \overline{1, n}$) выполняется одно из условий: (1) $\varphi_i \in C$; (2) существуют $j, k < i$ такие, что $\varphi_i = \text{res}(\varphi_j, \varphi_k)$. Дизъюнкт D считается *резольтивно выводимым* из множества C , если существует резольтивный вывод из C , последней формулой которого является D . Согласно теореме о полноте резольвий, множество C противоречиво в том и только в том случае, когда существует резольтивный вывод из C , заканчивающийся пустым дизъюнктом \square .

Метод резольвий используется для решения следующих задач, которые сводятся к доказательству противоречивости специальным образом построенного множества C : проверка правильности рассуждений; доказательство выводимости формулы из множества гипотез; доказательство того, что формулу можно считать теоремой; доказательство противоречивости заданного множества формул.

Чтобы ввести нечеткую резольвенту, рассмотрим необходимые понятия из нечеткой логики [4], которую можно определить как алгебру $\langle [0; 1], \wedge, \vee, \neg \rangle$, при этом для формализации логических связей \vee и \wedge используются различные операции. Наиболее распространенный подход предполагает, что \vee реализуется как \max , а \wedge — как \min при использовании стандартного отрицания. Заметим, что для операций \max и \min , называемых классическими, не выполняются законы коммутативности, поэтому в нечеткой логике высказывание, в котором a и \bar{a} встречаются одновременно, оказывается содержательным в отличие от классической логики. В общем случае для формализации \wedge и \vee используются треугольные нормы T и двойственные им (в смысле законов де Моргана) треугольные конормы S . Различные семейства треугольных норм и конорм представлены в [17]. Новые параметрические семейства треугольных норм и конорм получены в [18, 19]. Следует отметить, что для классической пары не выполняется свойство коммутативности, а для остальных помимо коммутативности не выполняются свойства идемпотентности и дистрибутивности, что обуславливает необходимость проявления аккуратности при выполнении всевозможных преобразований логических формул.

Каждому элементарному высказыванию поставим в соответствие оценку его истинности $t(x) \in [0; 1]$, при этом чем ближе значение $t(x)$ к 0, тем в большей степени высказывание считается *ложным*; чем ближе $t(x)$ к 1, тем в большей степени высказывание считается *истинным*; если $t(x) = 0,5$, то высказывание считается *неопределенным* (соответствует максимальной неопределенности). В нечеткой логике высказывание x принимается,

если степень его истинности больше степени истинности высказывания *не* $x(\bar{x})$, т. е. $t(x) > 1 - t(x)$, а следовательно, $t(x) > 0,5$.

Степенью доверия к $t(x)$ называется величина $c_x = 2(t(x) - 0,5) \in [-1; 1]$, а ее модуль $|c_x|$ характеризует уровень определенности высказывания x : если $t(x) = 0,5$, то $c_x = 0$ [6].

Пусть F — некоторая формула. При заданной интерпретации I значение истинности каждой переменной определяется однозначно, поэтому, выбрав подходящую формализацию логических связей, можно получить значение истинности формулы F . Формула F называется *выполнимой* в интерпретации I , если $t(F) > 0,5$, и *невыполнимой*, если $t(F) < 0,5$. Если $t(F) = 0,5$, то интерпретация I одновременно удовлетворяет и опровергает F .

Если $F = \{F_1, \dots, F_m\}$ — множество формул, то $t_I(F) = t_I(F_1 \wedge \dots \wedge F_m)$.

3 Нечеткая резольвента и ее свойства

Будем считать, что

$$\begin{aligned} t(a \vee b) &= \max\{t(a), t(b)\}; \\ t(a \wedge b) &= \min\{t(a), t(b)\}; \\ t(\bar{a}) &= 1 - t(a). \end{aligned}$$

Пусть $D_1 = p \vee D'_1$, $D_2 = \bar{p} \vee D'_2$, $\text{res}(D_1, D_2) = D'_1 \vee D'_2$. В [5] нечеткая резольвента определяется в виде

$$t(\text{res}(D_1, D_2)) = t(D'_1 \vee D'_2) = \max\{t(D'_1), t(D'_2)\}$$

и называется *резольвентой Ли* (L-резольвентой).

В классической логике действует следующий постулат [20]: истинность множества формул остается неизменной, если к ним добавить их резольвенту, т. е. $t(D_1 \wedge D_2) = t(D_1 \wedge D_2 \wedge \text{res}(D_1, D_2))$, и поэтому

$$d(D_1 \wedge D_2) \leq t(\text{res}(D_1, D_2)). \quad (1)$$

Резольвента называется *логически значимой* или *значимым логическим следствием*, если она удовлетворяет (1). Найдем условия, при которых выполняется соотношение (1), а следовательно, резольвента становится логически значимой.

Положим $t(D_1 \wedge D_2) = \min\{t(D_1), t(D_2)\} = a$. Без ограничения общности будем считать, что

$$t(D_1) = t(p \vee D'_1) = \max\{t(p), t(D'_1)\} = a; \quad (2)$$

$$\begin{aligned} t(D_2) &= t(\bar{p} \vee D'_2) = \max\{t(\bar{p}), t(D'_2)\} = \\ &= \max\{1 - t(p), t(D'_2)\} > a. \quad (3) \end{aligned}$$

В соответствии с введенными предположениями возможны ситуации:

– для (2) имеем:

$$(а) t(p) = a, t(D'_1) < a;$$

$$(б) t(p) < a, t(D'_1) = a;$$

– для (3) имеем:

$$(в) \begin{cases} (в1) a < t(D'_2) < 1 - t(p); \\ (в2) t(D'_2) < a < 1 - t(p); \end{cases}$$

$$(г) \begin{cases} (г1) a < 1 - t(p) < t(D'_2); \\ (г2) 1 - t(p) < a < t(D'_2). \end{cases}$$

Анализ комбинаций приведенных ситуаций позволил сформулировать следующие утверждения.

Утверждение 1. Пусть $t(p) < 0,5$, тогда неравенство (1) выполняется, а следовательно, резольвента становится значимым логическим следствием, если имеет место одна из следующих ситуаций:

(1) для одного из дизъюнктов D'_1 или D'_2 степень истинности больше $t(p)$, а степень истинности другого принадлежит промежутку $(\min \{t(D'_1), t(D'_2)\}, 1 - t(p))$;

(2) для одного из дизъюнктов D'_1 или D'_2 степень истинности меньше $t(p)$, а степень истинности другого принадлежит промежутку $(t(p), 1 - t(p))$;

(3) $t(D'_1) < t(D'_2)$ и $t(D'_2) > 1 - t(p)$;

(4) $t(D'_1) < t(D'_2)$ и $t(D'_1) \in (1 - t(p), t(D'_2))$.

Утверждение 2. Если $t(p) > 0,5$, $t(D'_1) < t(D'_1)$ и $t(D'_2) > t(p)$, то неравенство (1) выполняется.

Утверждение 3. Если для переменной p и каждого из дизъюнктов D'_1 и D'_2 степень истинности меньше $0,5$, то неравенство (1) не выполняется.

Утверждение 4. L -резольвента становится выполнимой формулой, если

(1) $t(p) < 0,5$ и $\max \{t(D_1), t(D_2)\} \in [0,5; 1]$;

(2) $t(p) > 0,5$ и $\max \{t(D_1), t(D_2)\} \in [t(p); 1]$.

Таким образом, L -резольвента становится значимым логическим следствием, если степень ее истинности или степень истинности по крайней мере одного из дизъюнктов больше $0,5$. На основе приведенных рассуждений был построен алгоритм для нахождения логически значимой резольвенты.

Теперь попытаемся обобщить результаты на другие типы логических связок. Рассмотрим случай, когда конъюнкция формализуется треугольной нормой T , дизъюнкция — треугольной конормой S , так что пара (T, S) образует пару двойственных операций относительно стандартного отрицания $n(x) = 1 - x$. Исследуем условие (1) для произвольной пары (T, S) , учитывая, что для нее не

выполняются законы дистрибутивности, коммутативности и идемпотентности. В этом случае

$$t(D_1) = t(D'_1 \vee p) = S \{t(D'_1), t(p)\};$$

$$t(D_2) = t(D'_2 \vee \bar{p}) = S \{t(D'_2), t(\bar{p})\};$$

$$t(\bar{D}_1) = t(\overline{D'_1 \vee p}) = T \left(t(\bar{D}'_1), t(\bar{p}) \right);$$

$$t(\bar{D}_2) = t(\overline{D'_2 \vee \bar{p}}) = T \left(t(\bar{D}'_2), t(p) \right);$$

$$t(\text{res}(D_1, D_2)) = t(D'_1 \vee D'_2) = S \{t(D'_1), t(D'_2)\}.$$

Тогда, учитывая, что для любых T и S имеют место неравенства $T(x, y) \leq \min\{x, y\}$ и $\max\{x, y\} \leq S(x, y)$, получим

$$\begin{aligned} t(D_1 \wedge D_2) &= T(t(D_1), t(D_2)) = \\ &= 1 - S(1 - t(D_1), 1 - t(D_2)) = \\ &= 1 - S(t(\bar{D}_1), t(\bar{D}_2)) = \\ &= 1 - S \left(T \left(t(\bar{D}'_1), t(\bar{p}) \right), T \left(t(\bar{D}'_2), t(p) \right) \right) = \\ &= 1 - S(1 - S(t(D'_1), t(p)), 1 - S(t(D'_2), t(\bar{p}))) = \\ &= T(S(t(D'_1), t(p)), S(t(D'_2), t(\bar{p}))) \leq \\ &\leq \min(S(t(D'_1), t(p)), S(t(D'_2), t(\bar{p}))) \leq \\ &\leq \max(S(t(D'_1), t(p)), S(t(D'_2), t(\bar{p}))) \leq \\ &\leq S(S(t(D'_1), t(p)), S(t(D'_2), t(\bar{p}))) = \\ &= S(S(t(D'_1), t(D'_2)), S(t(p), t(\bar{p}))). \end{aligned}$$

Заметим, что в качестве результата данного выражения можно получить резольвенту $S(t(D'_1), t(D'_2)) = t(\text{res}(D_1, D_2))$, если $S(t(p), t(\bar{p})) = 0$, так как для треугольных конорм $S(x, 0) = x$ для всех $x \in [0, 1]$. Согласно определению, соотношение $S(t(p), t(\bar{p})) = 0$ выполняется, если одновременно $t(p) = 0$ и $t(\bar{p}) = 0$, что невозможно. Если $S(t(p), t(\bar{p})) = S(t(D'_1), t(D'_2))$, то для идемпотентных конорм неравенство (1) будет выполнено. Но единственной идемпотентной парой остается пара (\min, \max) [17]. Таким образом, доказано

Утверждение 5. Для произвольной пары (T, S) двойственных треугольных операций, отличных от классических \min и \max , L -резольвента не существует.

Рассмотрим комбинацию произвольной треугольной нормы T и $S = \max$:

$$t(\text{res}(D_1, D_2)) = t(D'_1 \vee D'_2) = \max(t(D'_1), t(D'_2));$$

$$t(D_1) = t(D'_1 \vee p) = \max\{t(D'_1), t(p)\};$$

$$t(D_2) = t(D'_2 \vee \bar{p}) = \max\{t(D'_2), t(\bar{p})\}.$$

Найдем

$$\begin{aligned} t(D_1 \wedge D_2) &= T(t(D_1), t(D_2)) \leq \\ &\leq \min\{t(D_1), t(D_2)\} = \end{aligned}$$

$$\begin{aligned}
 &= \min \{ \max \{ t(D'_1), t(p) \}, \max \{ t(D'_2), t(\bar{p}) \} \} = \\
 &= \max \{ \min \{ t(D'_1), t(D'_2) \}, \min \{ t(D'_1), t(\bar{p}) \}, \\
 &\quad \min \{ t(D'_2), t(p) \}, \min \{ t(p), t(\bar{p}) \} \} \leq \\
 &\leq \max \{ \max \{ t(D'_1), t(D'_2) \}, \min \{ t(D'_1), t(\bar{p}) \}, \\
 &\quad \min \{ t(D'_2), t(p) \}, \min \{ t(p), t(\bar{p}) \} \}.
 \end{aligned}$$

Заметим, что результатом последнего выражения будет

$$\max \{ t(D'_1), t(D'_2) \} = t(\text{res}(D_1, D_2)),$$

а следовательно, неравенство (1) будет выполнено, если одновременно будут иметь место следующие неравенства:

$$\begin{cases} \max \{ t(D'_1), t(D'_2) \} \geq \min \{ t(D'_1), t(\bar{p}) \}; \\ \max \{ t(D'_1), t(D'_2) \} \geq \min \{ t(D'_2), t(p) \}; \\ \max \{ t(D'_1), t(D'_2) \} \geq \min \{ t(p), t(\bar{p}) \}. \end{cases}$$

С учетом того что существуют всего 24 случая различных упорядочений значений $t(D'_1)$, $t(D'_2)$, $t(p)$ и $t(\bar{p})$, установлено, что одновременно данные неравенства не выполняются, если

$$\max \{ t(D'_1), t(D'_2) \} < \min \{ t(p), t(\bar{p}) \}.$$

Так как по закону де Моргана

$$\min \{ t(p), t(\bar{p}) \} = 1 - \max \{ t(p), t(\bar{p}) \},$$

то предыдущее неравенство равносильно неравенству

$$\max \{ t(D'_1), t(D'_2) \} + \max \{ t(p), t(\bar{p}) \} < 1.$$

Таким образом, доказано следующее

Утверждение 6. Пусть для формализации конъюнкции используется норма T , а дизъюнкции — \max . Если $\max \{ t(D'_1), t(D'_2) \} + \max \{ t(p), t(\bar{p}) \} < 1$, то неравенство (1) не выполняется. В остальных случаях неравенство (1) выполняется, при этом резольвента становится значимым логическим следствием, если $\max \{ t(D'_1), t(D'_2) \} > 0,5$.

Заметим, что L-резольвента может быть или не быть значимым логическим следствием в зависимости от степени истинности ее слагаемых компонент и переменной p . В [6] Мукайдоно получил обобщение L-резольвенты, сопряженное со степенью доверия.

Пусть $D_1 = p \vee D'_1$ и $D_2 = \bar{p} \vee D'_2$ — дизъюнкты. Нечеткая резольвента Мукайдоно — M-резольвента дизъюнктов D_1 и D_2 , обозначаемая $\text{res}(D_1, D_2)_{c_p}$, где $c_p = 2(\max \{ t(p), t(\bar{p}) \} - 0,5)$ — степень доверия к переменной p , вычисляется по формуле:

$$\text{res}(D_1, D_2)_{c_p} = \text{res}(D_1, D_2) \vee (p \wedge \bar{p}),$$

$$\text{res}(D_1, D_2) = D'_1 \vee D'_2.$$

Множество дизъюнктов C считается невыполнимым тогда и только тогда, когда существует вывод пустого дизъюнкта \square с ненулевой степенью доверия [6].

Используя пару (T, S) двойственных операций, степень истинности M-резольвенты запишем в виде:

$$\begin{aligned}
 t(\text{res}(D_1, D_2)_{c_p}) &= \\
 &= S \{ t(\text{res}(D_1, D_2)), T \{ t(p), t(\bar{p}) \} \}.
 \end{aligned}$$

Учитывая рассуждения при доказательстве утверждения 5, используя закон де Моргана и свойство монотонности S , в предположении, что $1 - T(t(\bar{p}), t(p)) \leq T(t(\bar{p}), t(p))$, получим

$$\begin{aligned}
 t(D_1 \wedge D_2) &\leq S(S(t(D'_1), t(D'_2)), S(t(p), t(\bar{p}))) = \\
 &= S(S(t(D'_1), t(D'_2)), 1 - T(t(\bar{p}), t(p))) \leq \\
 &\leq S(S(t(D'_1), t(D'_2)), T(t(p), t(\bar{p}))) = \\
 &= S(t(\text{res}(D_1, D_2)), T(t(p), t(\bar{p}))) = \\
 &= t(\text{res}(D_1, D_2)_{c_p}),
 \end{aligned}$$

а следовательно, неравенство (1) выполняется, при этом степень истинности переменной p и треугольная норма T должны быть такими, чтобы $T(t(\bar{p}), t(p)) \geq 0,5$. В данном случае необходимо исследование семейств треугольных норм, позволяющих обеспечить выполнение данного неравенства, и в общем случае какой-то вывод сделать невозможно. При $T = \min$ неравенство $\min \{ t(\bar{p}), t(p) \} \geq 0,5$ выполняется только как равенство при $t(\bar{p}) = t(p) = 0,5$, что дает возможность работать с неопределенными высказываниями. Кроме того, для поиска подходящей треугольной нормы можно рассматривать семейства, для которых «действует» отрицание, не являющееся инволюцией.

Пусть $T = \min$ и $S = \max$, тогда резольвента имеет вид:

$$\begin{aligned}
 t(\text{res}(D_1, D_2)_{c_p}) &= \\
 &= \max \{ \max \{ t(D'_1), t(D'_2) \}, \min \{ t(p), t(\bar{p}) \} \}.
 \end{aligned}$$

Рассмотрим

$$\begin{aligned}
 t(D_1, D_2) &= \min \{ t(D_1), t(D_2) \} = \\
 &= \min \{ \max \{ t(D'_1), t(p) \}, \max \{ t(D'_2), t(\bar{p}) \} \} = \\
 &= \max \{ \min \{ t(D'_1), t(D'_2) \}, \min \{ t(D'_1), t(\bar{p}) \}, \\
 &\quad \min \{ t(D'_2), t(p) \}, \min \{ t(p), t(\bar{p}) \} \} \leq \\
 &\leq \max \{ \max \{ t(D'_1), t(D'_2) \}, \min \{ t(D'_1), t(\bar{p}) \}, \\
 &\quad \min \{ t(D'_2), t(p) \}, \min \{ t(p), t(\bar{p}) \} \} =
 \end{aligned}$$

$$= \max \left\{ \underbrace{\max \{ \max \{ t(D'_1), t(D'_2) \}, \min \{ t(p), t(\bar{p}) \} \}}_{t(\text{res}(D_1, D_2)_{c_p})}, \right. \\ \left. \max \{ \min \{ t(D'_1), t(\bar{p}) \}, \min \{ t(D'_2), t(p) \} \} \right\}.$$

Заметим, что

$$\max \{ \min \{ t(D'_1), t(\bar{p}) \}, \min \{ t(D'_2), t(p) \} \} = \\ = \min \left\{ \underbrace{\max \{ t(D'_1), t(D'_2) \}}_{t(\text{res}(D_1, D_2))}, \underbrace{\max \{ t(D'_1), t(p) \}}_{t(D_1)}, \right. \\ \left. \underbrace{\max \{ t(D'_2), t(\bar{p}) \}, \max \{ t(p), t(\bar{p}) \}}_{t(D_2)} \right\} = \\ = \min \left\{ \underbrace{\min \{ t(D_1), t(D_2) \}}_{t(D_1 \wedge D_2)}, \right. \\ \left. \min \{ t(\text{res}(D_1, D_2)), \max \{ t(p), t(\bar{p}) \} \} \right\}.$$

Таким образом, получено следующее неравенство:

$$t(D_1, D_2) \leq \max \left\{ t(\text{res}(D_1, D_2)_{c_p}), \right. \\ \min \{ t(D_1 \wedge D_2), t(\text{res}(D_1, D_2)) \}, \\ \left. \max \{ t(p), t(\bar{p}) \} \right\}.$$

Здесь $\text{res}(D_1, D_2)$ — L-резольвента. Тогда если L-резольвента логически значима, т.е. выполняется неравенство $t(D_1 \wedge D_2) \leq t(\text{res}(D_1, D_2))$, то

$$t(D_1 \wedge D_2) \leq \max \left\{ t(\text{res}(D_1, D_2)_{c_p}), \right. \\ \left. \min \{ t(D_1 \wedge D_2), \max \{ t(p), t(\bar{p}) \} \} \right\}.$$

Если $t(D_1 \wedge D_2) \leq \max \{ t(p), t(\bar{p}) \}$, то

$$t(D_1 \wedge D_2) \leq \max \{ t(\text{res}(D_1, D_2)_{c_p}), t(D_1 \wedge D_2) \}$$

и неравенство (1) выполняется.

Если $t(D_1 \wedge D_2) > \max \{ t(p), t(\bar{p}) \}$, то

$$t(D_1 \wedge D_2) \leq \\ \leq \max \left\{ t(\text{res}(D_1, D_2)_{c_p}), \max \{ t(p), t(\bar{p}) \} \right\}$$

и неравенство (1) выполняется, если

$$\max \{ t(p), t(\bar{p}) \} \leq t(\text{res}(D_1, D_2)_{c_p}).$$

Таким образом, доказано следующее

Утверждение 7. Если L-резольвента — значимое логическое следствие и

$$\max \{ t(p), t(\bar{p}) \} \leq t(\text{res}(D_1, D_2)_{c_p}),$$

то M-резольвента $t(\text{res}(D_1, D_2)_{c_p})$ удовлетворяет неравенству (1), а следовательно, также становится значимым логическим следствием.

```

Input: C — исходное множество дизъюнктов
Input: k = 0, Rk = ∅ (текущее множество резольвент)
Output: Результат работы алгоритма: {C is inconsistent, C is unknown}
// упорядочить литеры, образующие контрарные пары, по убыванию степени доверия cp
calculateConfidenceLevel(cp); sortConfidenceLevels(cp);
// вычисление резольвент
while (∃ Di ∈ C and ∃ Dj ∈ C) do
// если существует значимая резольвента, то добавить ее в Rk+1
  if (∃ res(Di, Dj) is significant) do
    searchAndChoose(Di, Dj, cp); Rk+1 = Rk ∪ {res(Di, Dj)}
  // проверка: существует ли в Rk+1 пустой дизъюнкт □
  if (□ ∈ Rk+1) then C is inconsistent; end program;
  // если в Rk+1 нет □, то изменить множество C
  else C = C \ Di ∪ C \ Dj ∪ res(Di, Dj); k = k + 1;
  end if
end while
// если в C нет дизъюнктов для получения резольвенты, то противоречивость C неизвестна
C is unknown;
end program

```

Рис. 1 Алгоритм построения M-резольвенты: C — множество дизъюнктов, невыполнимость (противоречивость) которого исследуется

4 Нечеткий метод резолюций

Нечеткий метод резолюций реализуется согласно классической схеме [16] на основе определений L-резольвенты или M-резольвенты, при этом в первом случае вычисляются только значимые L-резольвенты.

При нахождении M-резольвенты используются степени доверия литер: чем больше c_p , тем построение резольвенты из дизъюнктов, содержащих литеру p , происходит раньше.

Рассмотрим более подробно резолютивный вывод на основе M-резольвенты (рис. 1). Для компактного изложения алгоритма введем следующие функции:

- (1) функция $\text{calculateConfidenceLevel}(c_p)$ вычисляет степень доверия каждой литеры (результат в c_p);
- (2) функция $\text{sortConfidenceLevels}(c_p)$ сортирует последовательность степеней доверия литер в порядке убывания;
- (3) функция $\text{searchAndChoose}(D_i, D_j, c_p)$ осуществляет поиск дизъюнктов, содержащих литеру с максимальным значением степени доверия c_p .

Рассмотрим пример. Докажем или опровергнем противоречивость следующего множества дизъюнктов: $\{W(x) \vee O(x), Z(y) \vee \neg W(y), Y(A), \neg Z(A), \neg O(z) \vee \neg Y(z)\}$, где $t(W) = 0,5$; $t(O) = 0,4$; $t(Z) = 0,3$; $t(Y) = 0,8$. Результаты работы алгоритма, основанного на вычислении M-резольвенты, представлены на рис. 2.

5 Заключение

Метод резолюций прежде всего известен как инструмент для автоматического доказательства теорем. В моделях обработки знаний, основанных на правиле дедукции, проблема формулируется в виде совокупности утверждений-гипотез и целевого утверждения — теоремы, справедливость которой следует установить или опровергнуть на основе данных гипотез, аксиом и правил вывода. Данная схема может быть положена в основу различных автоматизированных процедур принятия решений в медицине, экономике, финансовой сфере, судебной практике. Метод резолюций используется для моделирования итерационных вычислений в рамках дедуктивного синтеза программ. Теория рассуждений, основанная на методе резолюций, стала активно развивающимся направлением искусственного интеллекта. Классический метод резолюций базируется на двужначной логике, но зачастую для многих рассуждений характерна неопределенность знаний, обусловленная различными факторами. Поэтому многие исследователи пытаются обобщить метод резолюций на случай неклассических логик и сделать его таким же эффективным, как в классической логике. В данной статье предложены различные обобщения известных определений нечетких резольвент с использованием треугольных норм и конорм, которые формализуют нечеткие логические связки *и* и *или*. Проведенное исследование показывает, что для того чтобы нечеткая резольвента была значимым логическим следствием, необходимо выполнение определенных условий, которые уменьшают неопределенность и делают правило нечеткого вывода, опирающееся на

$C = \{W(x) \vee O(x), Z(y) \vee \neg W(y), Y(A), \neg Z(A), \neg O(z) \vee Y(z)\}$ $c_p(Y) = 2(0,8 - 0,5) = 0,6, c_p(Z) = 2(0,7 - 0,5) = 0,4$ $c_p(O) = 2(0,6 - 0,5) = 0,2, c_p(W) = 2(0,5 - 0,5) = 0$
$k = 0, R_0 = \emptyset; R_0$ не содержит пустого дизъюнкта; в C есть унифицируемые литералы
$D_i = Y(A), D_j = \neg O(z) \vee \neg Y(z)$
$\neg O(A) \vee \neg Y(A) = \neg O(A), R_1 = \{\neg O(A)\}$
$k = 1; R_1$ не содержит пустого дизъюнкта; в C есть унифицируемые литералы
$C = \{W(x) \vee O(x), Z(y) \vee \neg W(y), \neg Z(A), \neg O(A)\}$
$D_i = Z(y) \vee \neg W(y), D_j = \neg Z(A)$
$\neg W(A) \vee Z(A) = \neg W(A), R_2 = \{\neg O(A), \neg W(A)\}$
$k = 2, R_2$ не содержит пустого дизъюнкта, в C есть унифицируемые литералы
$C = \{W(x) \vee O(x), \neg O(A), \neg W(A)\}$
$D_i = W(x) \vee O(x), D_j = \neg O(A), R_3 = \{\neg O(A), \neg W(A), W(A)\}$
$k = 3, R_3$ не содержит пустого дизъюнкта; в C есть унифицируемые литералы
$C = \{\neg W(A), W(A)\}$
$D_i = \neg W(A), D_j = W(A), R_4 = \{\neg O(A), \neg W(A), W(A), \square\}$
R_4 содержит пустой дизъюнкт — исходное множество C противоречиво.

Рис. 2 Работа алгоритма в интерпретации Мукайдоно

резольционный принцип, осмысленным в различных интерпретациях.

Литература

1. *Robinson J. A.* A machine-oriented logic based on the resolution principle // *J. ACM*, 1965. Vol. 12. Iss. 1. P. 23–41. doi: 10.1145/321250.321253.
2. *Newborn M.* Automated theorem proving: Theory and practice. — Berlin, Heidelberg: Springer-Verlag, 2001. 231 p.
3. *Ламберов Л. Д.* Практика компьютерных доказательств и человеческое понимание: эпистемологическая проблематика // *Вестник Пермского университета. Философия. Психология. Социология*, 2021. № 1. С. 5–19. doi: 10.17072/2078-7898/2021-1-5-19. EDN: GTOYOE.
4. *Новак В., Перфильева И., Мочкорж И.* Математические принципы нечеткой логики. — М.: Физматлит, 2006. 252 с.
5. *Lee R. C. T.* Fuzzy logic and the resolution principle // *J. ACM*, 1972. Vol. 19. Iss. 1. P. 109–119. doi: 10.1145/321679.321688.
6. *Mukaidono M.* Fuzzy inference of resolution style // *Fuzzy set and possibility theory* / Ed. R. R. Yager. — New York, NY, USA: Pergamon Press, 1988. P. 224–231.
7. *Dubois D., Prade H.* Necessity and resolution principle // *IEEE T. Syst. Man Cybern.*, 1987. Vol. 17. Iss. 3. P. 474–478. doi: 10.1109/TSMC.1987.4309063.
8. *Guller D.* Hyperresolution for Gödel logic with truth constants // *Fuzzy Set. Syst.*, 2019. Vol. 363. P. 1–65. doi: 10.1016/j.fss.2018.09.008.
9. *Tammet T.* A resolution theorem prover for intuitionistic logic // *Automated deduction — Cade-13* / Eds. M. A. McRobbie, J. K. Slaney. — Lecture notes in computer science ser. — Berlin, Heidelberg: Springer, 1996. Vol. 1104. P. 2–16. doi: 10.1007/3-540-61511-3_65.
10. *Viedma M. A. C., Morales R. M., Sanchez I. N.* Fuzzy temporal constraint logic: A valid resolution principle // *Fuzzy Set. Syst.*, 2001. Vol. 117. Iss. 2. P. 231–250. doi: 10.1016/S0165-0114(99)00099-8.
11. *Habiballa H.* Resolution principle and fuzzy logic // *Fuzzy logic — algorithms, techniques, and implementations* / Ed. E. Dadios. — London: IntechOpen, 2012. P. 55–74.
12. *Nguyen T. M. T., Tran D. A. K.* Resolution method in linguistic propositional logic // *Int. J. Advanced Computer Science Applications*, 2016. Vol. 7. Iss. 1. P. 672–678. doi: 10.14569/IJACSA.2016.070191.
13. *Samokhvalov Y.* Proof of theorems in fuzzy logic based on structural resolution // *Cybern. Syst. Anal.*, 2019. Vol. 55. P. 207–219. doi: 10.1007/s10559-019-00125-8.
14. *Raha S., Ray K. S.* Approximate reasoning based on generalized disjunctive syllogism // *Fuzzy Set. Syst.*, 1994. Vol. 61. Iss. 2. P. 143–151. doi: 10.1016/0165-0114(94)90230-5.
15. *Mondal B., Raha S.* Approximate reasoning in fuzzy resolution // *Int. J. Intelligence Science*, 2013. Vol. 3. Iss. 2. P. 86–98. doi: 10.4236/ijis.2013.32010.
16. *Леденева Т. М., Лецинская М. В.* Метод резолюций и стратегии поиска опровержений // *Вестник ВГУ. Сер. Системный анализ и информационные технологии*, 2021. № 1. С. 98–111. doi: 10.17308/sait.2021.1/3374. EDN: RBAVMW.
17. *Klement E. P., Mesiar R., Pap E.* Triangular norms. Position paper II: General constructions and parameterized families // *Fuzzy Set. Syst.*, 2004. Vol. 145. P. 411–438. doi: 10.1016/S0165-0114(03)00327-0.
18. *Ledeneva T.* Additive generators of fuzzy operation in the form of linear fractional function // *Fuzzy Set. Syst.*, 2020. Vol. 386. P. 1–24. doi: 10.1016/j.fss.2019.03.005.
19. *Ledeneva T.* New family of triangular norms for decreasing generators in the form of a logarithm of a linear fractional function // *Fuzzy Set. Syst.*, 2022. Vol. 427. P. 37–54. doi: 10.1016/j.fss.2020.11.020.
20. *Вагин В. Н., Головина Е. Ю., Загорянская А. А., Фомина М. В.* Достоверный и правдоподобный вывод в интеллектуальных системах / Под ред. В. Н. Вагина, Д. А. Поспелова. — М.: Физматлит, 2008. 712 с. EDN: MUWRTJ.

Поступила в редакцию 24.04.23

ANALYSIS OF APPROACHES TO DEFINING FUZZY RESOLVENT

T. M. Ledeneva and M. V. Leshchinskaya

Voronezh State University, 1 Universitetskaya Sq., Voronezh 394010, Russian Federation

Abstract: The article presents the results of a study concerning various definitions of the resolvent in fuzzy logic. Conditions are defined under which the Lee resolvent is a significant logical consequence in the case of the classical definition of fuzzy logical connectives. It is shown that when using triangular norms and conorms for their formalization, it is impossible to obtain a logically significant Lee resolvent. However, if the triangular conorm is defined as max, then the Lee resolvent exists for any triangular norm. Conditions are defined under which the Mukaidano resolvent is a significant logical consequence for classical min and max operations. When using

triangular norms and conorms other than classical ones, further research is required. An illustrative example is provided demonstrating the process of constructing the Mukaidano resolvent.

Keywords: resolvent; resolution method; triangular norms and conorms

DOI: 10.14357/19922264240113

EDN: MIDEHY

References

1. Robinson, J. A. 1965. A machine-oriented logic based on the resolution principle. *J. ACM* 12(1):23–41. doi: 10.1145/321250.321253.
2. Newborn, M. 2000. *Automated theorem proving: Theory and practice*. New York, NY: Springer. 231 p.
3. Lamberov, L. D. 2021. Praktika komp'yuternykh dokazatel'stv i chelovecheskoe ponimanie: epistemologicheskaya problematika [Computer proofs practice and human understanding: Epistemological issues]. *Vestnik Permskogo universiteta. Filosofiya. Psikhologiya. Sotsiologiya* [Perm University Herald. Philosophy. Psychology. Sociology] 1:5–19. doi: 10.17072/2078-7898/2021-1-5-19. EDN: GTOYOE.
4. Novak, V., I. Perfil'eva, and I. Mochkorzh. 2006. *Matematicheskie printsipy nechetkoy logiki* [Mathematical principles of fuzzy logic]. Moscow: Fizmatlit. 252 p.
5. Lee, R. C. T. 1972. Fuzzy logic and the resolution principle. *J. ACM* 19(1):109–119. doi:10.1145/321679.321688.
6. Mukaidono, M. 1988. Fuzzy inference of resolution style. *Fuzzy set and possibility theory*. Ed. R. R. Yager. New York, NY: Pergamon Press. 224–231.
7. Dubois, D., and H. Prade. 1987. Necessity and resolution principle. *IEEE T. Syst. Man Cyb.* 17(3):474–478. doi: 10.1109/TSMC.1987.4309063.
8. Guller, D. 2019. Hyperresolution for Gödel logic with truth constants. *Fuzzy Set. Syst.* 363:1–65. doi: 10.1016/j.fss.2018.09.008.
9. Tammet, T. 1996. A resolution theorem prover for intuitionistic logic. *Automated deduction — Cade-13*. Eds. M. A. McRobbie and J. K. Slaney. Lecture notes in computer science ser. Berlin, Heidelberg: Springer. 1104:2–16. doi: 10.1007/3-540-61511-3.65.
10. Viedma, M. A. C., R. M. Morales, and I. N. Sanchez. 2001. Fuzzy temporal constraint logic: A valid resolution principle. *Fuzzy Set. Syst.* 117(2):231–250. doi: 10.1016/S0165-0114(99)00099-8.
11. Habiballa, H. 2012. Resolution principle and fuzzy logic. *Fuzzy logic algorithms, techniques, and implementations*. Ed. E. Dadios. London: Intech. 55–74.
12. Nguyen, T. M. T., and D. A. K. Tran. 2016. Resolution method in linguistic propositional logic. *Int. J. Advanced Computer Science Applications* 7(1):672–678. doi: 10.14569/IJACSA.2016.070191.
13. Samokhvalov, Y. 2019. Proof of theorems in fuzzy logic based on structural resolution. *Cybern. Syst. Anal.* 55:207–219. doi: 10.1007/s10559-019-00125-8.
14. Raha, S., and K. S. Ray. 1994. Approximate reasoning based on generalized disjunctive syllogism. *Fuzzy Set. Syst.* 61(2):143–151. doi: 10.1016/0165-0114(94)90230-5.
15. Mondal, B., and S. Raha. 2013. Approximate reasoning in fuzzy resolution. *Int. J. Intelligence Science* 3(2):86–98. doi: 10.4236/ijis.2013.32010.
16. Ledeneva, T. M., and M. V. Leshchinskaya. 2021. Metod rezolyutsiy i strategii poiska oproverzheniy [The resolution inference and strategies for finding refutation]. *Vestnik VGU. Ser. Sistemnyy analiz i informatsionnye tekhnologii* [Proceedings of Voronezh State University. Ser. Systems Analysis and Information Technologies] 1:98–111. doi: 10.17308/sait.2021.1/3374. EDN: RBAVMW.
17. Klement, E. P., R. Mesiar, and E. Pap. 2004. Triangular norms. Position paper II: General constructions and parameterized families. *Fuzzy Set. Syst.* 145(3):411–438. doi: 10.1016/S0165-0114(03)00327-0.
18. Ledeneva, T. 2020. Additive generators of fuzzy operations in the form of linear fractional functions. *Fuzzy Set. Syst.* 386:1–24. doi: 10.1016/j.fss.2019.03.005.
19. Ledeneva, T. 2022. New family of triangular norms for decreasing generators in the form of a logarithm of a linear fractional function. *Fuzzy Set. Syst.* 427:37–54. doi: 10.1016/j.fss.2020.11.020.
20. Vagin, V. N., E. Yu. Golovina, A. A. Zagoryanskaya, and M. V. Fomina. 2008. *Dostovernyy i pravdopodobnyy vyvod v intellektual'nykh sistemakh* [Reliable and plausible conclusion in intelligent systems]. Eds. V. N. Vagin and D. A. Pospelov. Moscow: Fizmatlit. 712 p. EDN: MUWRTJ.

Received April 24, 2023

Contributors

Ledeneva Tatyana M. (b. 1959) — Doctor of Science in technology, professor, Department of Computational Mathematics and Applied Information Technologies, Faculty of Applied Mathematics, Informatics and Mechanics, Voronezh State University, 1 Universitetskaya Sq., Voronezh 394010, Russian Federation; ledeneva-tm@yandex.ru

Leshchinskaya Maria V. (b. 1995) — PhD student, Faculty of Applied Mathematics, Informatics and Mechanics, Voronezh State University, 1 Universitetskaya Sq., Voronezh 394010, Russian Federation; maria-leshchinskaya@mail.ru

Борисов Андрей Владимирович (р. 1965) — доктор физико-математических наук, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Босов Алексей Вячеславович (р. 1969) — доктор технических наук, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Вышинский Леонид Леонидович (р. 1941) — кандидат физико-математических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Николай Александрович (р. 1982) — кандидат физико-математических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Забжайло Михаил Иванович (р. 1956) — доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ковалёв Сергей Протасович (р. 1972) — доктор физико-математических наук, ведущий научный сотрудник Института проблем управления им. В. А. Трапезникова Российской академии наук

Кривенко Михаил Петрович (р. 1946) — доктор технических наук, профессор, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кудрявцев Алексей Андреевич (р. 1978) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломо-

носова; старший научный сотрудник Московского центра фундаментальной и прикладной математики

Леденева Татьяна Михайловна (р. 1959) — доктор технических наук, профессор кафедры вычислительной математики и прикладных информационных технологий факультета прикладной математики, информатики и механики Воронежского государственного университета

Лери Марина Муксумовна (р. 1969) — кандидат технических наук, научный сотрудник Института прикладных математических исследований Федерального исследовательского центра «Карельский научный центр Российской академии наук»

Лещинская Мария Владимировна (р. 1995) — аспирант факультета прикладной математики, информатики и механики Воронежского государственного университета

Листопад Сергей Викторович (р. 1984) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Малашенко Юрий Евгеньевич (р. 1946) — доктор физико-математических наук, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Назарова Ирина Александровна (р. 1966) — кандидат физико-математических наук, научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Павлов Юрий Леонидович (р. 1949) — доктор физико-математических наук, профессор, главный научный сотрудник Института прикладных математических исследований Федерального исследовательского центра «Карельский научный центр Российской академии наук»

Писковский Виктор Олегович (р. 1963) — кандидат физико-математических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Синицын Игорь Николаевич (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук; профессор кафедры «Мехатроника и теоретическая механика» Московского авиационного института (национального исследовательского университета)

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Торшин Иван Юрьевич (р. 1972) — кандидат физико-математических наук, кандидат химических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Флёров Юрий Арсениевич (р. 1942) — член-корреспондент РАН, доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шестаков Олег Владимирович (р. 1976) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова; старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук; ведущий научный сотрудник Московского центра фундаментальной и прикладной математики

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Правила подготовки рукописей 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. По специальному решению редколлегии отдельные статьи могут печататься на английском 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 статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации предоставленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.
5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или ЛАТЭХ (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.

7. 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 10 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете: <http://www.ipiran.ru/journal/template.doc>

8. Статья должна содержать следующую информацию на *русском и английском языках*:

- название статьи;
- Ф.И.О. авторов, на английском можно только имя и фамилию;
- место работы, с указанием почтового адреса организации и электронного адреса каждого автора;
- сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах: http://www.ipiran.ru/journal/issues/2013_07_01/authors.asp и http://www.ipiran.ru/journal/issues/2013_07_01_eng/authors.asp;
- аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
- ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами;
- источники финансирования работы (ссылки на гранты, проекты, поддерживающие организации и т. п.).

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний. Если источник имеет DOI и/или EDN, то их необходимо указывать.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала;
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Ниже приведены примеры ссылок на различные виды публикаций в списке “References”.

Описание статьи из журнала:

Zagurenko, A. G., V. A. Korotovskikh, A. A. Kolesnikov, A. V. Timonov, and D. V. Kardymon. 2008. Tekhniko-ekonomicheskaya optimizatsiya dizayna gidrorazryva plasta [Technical and economic optimization of the design of hydraulic fracturing]. *Neftyanoe khozyaystvo [Oil Industry]* 11:54–57.

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Описание статьи из электронного журнала:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Описание статьи из продолжающегося издания (сборника трудов):

Astakhov, M. V., and T. V. Tagantsev. 2006. Eksperimental'noe issledovanie prochnosti soedineniy “stal’–kompozit” [Experimental study of the strength of joints “steel–composite”]. *Trudy MGTU “Matematicheskoe modelirovanie slozhnykh tekhnicheskikh sistem” [Bauman MSTU “Mathematical Modeling of Complex Technical Systems” Proceedings]*. 593:125–130.

Описание материалов конференций:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursoberegayushchie tekhnologii nedropol'zovaniya i povysheniya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

Описание книги (монографии, сборники):

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publ. 352 p.

Latyshev, V. N. 2009. *Tribologiya rezaniya. Kn. 1: Friksionnye protsessy pri rezanii metallov* [Tribology of cutting. Vol. 1: Frictional processes in metal cutting]. Ivanovo: Ivanovskii State Univ. 108 p.

Описание переводной книги (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Тимошенко С. П., Янг Д. Х., Уивер У. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (*Timoshenko S. P., Young D. H., Weaver W. Vibration problems in engineering. — 4th ed. — New York, NY, USA: Wiley, 1974. 521 p.*)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. New York: Wiley. 521 p.

Описание неопубликованного документа:

Laturov, A. R., M. M. Khasanov, and V. A. Baikov. 2004 (unpubl.). *Geologiya i dobycha (NGT GiD)* [Geology and production (NGT GiD)]. Certificate on official registration of the computer program No. 2004611198.

Описание интернет-ресурса:

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

Описание диссертации или автореферата диссертации:

Semenov, V. I. 2003. *Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor* [Mathematical modeling of the plasma in the compact torus]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. *Tekhnologiya razrabotki semanticheskogo slovary informatsionnogo monitoringa* [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. *Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv* [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publ. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. *Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya* [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.

11. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- указывать в тексте письма название статьи, авторов и журнал, в который направляется статья;
- использовать attach (присоединение);
- в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.

12. Журнал «Информатика и её применения» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

Адрес редакции журнала «Информатика и её применения»:
Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН
Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05
e-mail: ieep@frccsc.ru (Стригина Светлана Николаевна)
<http://www.ipiran.ru/journal/issues/>

Requirements for manuscripts submitted to Journal “Informatics and Applications”

Journal “Informatics and Applications” (Inform. Appl.) publishes theoretical, review, and discussion articles on the research and development in the field of informatics and its applications.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

The topics covered include the following areas:

- theoretical fundamentals of informatics;
- mathematical methods for studying complex systems and processes;
- information systems and networks;
- information technologies; and
- architecture and software of computational complexes and networks.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

Agreement on the transfer of rights to publish:

“We, the undersigned authors of the manuscript “. . .”, pass to the Founder and the Editorial Board of the Journal “Informatics and Applications” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.

Author(s) signature(s): (name(s), address(es), date).

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If the authors receive their article for correction after reviewing, it does not mean that the article is approved for publication. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L^AT_EX (.tex) formats, and additionally in the .pdf format. All documents may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.
7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site: <http://www.ipiran.ru/journal/template.doc>.

8. The articles should enclose data both in **Russian and English**:

- title;
- author’s name and surname;
- affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
- data on authors according to the format (see site):

http://www.ipiran.ru/journal/issues/2013_07_01/authors.asp and

http://www.ipiran.ru/journal/issues/2013_07_01_eng/authors.asp;

- abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae;
 - indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.
Important! Keywords must not be sentences; and
 - Acknowledgments.
9. References. Russian references have to be presented both in English translation and Latin transliteration (refer <http://www.translit.net/ru/bgn/>).
- Please take into account the following examples of Russian references appearance:
- Article in journal:**
Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.
- Journal article in electronic format:**
Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).
- Article from the continuing publication (collection of works, proceedings):**
Astakhov, M. V., and T. V. Tagantsev. 2006. Eksperimental'noe issledovanie prochnosti soedineniy "stal'-kompozit" [Experimental study of the strength of joints "steel-composite"]. *Trudy MGTU "Matematicheskoe modelirovanie slozhnykh tekhnicheskikh sistem"* [Bauman MSTU "Mathematical Modeling of Complex Technical Systems" Proceedings]. 593:125–130.
- Conference proceedings:**
Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursoberegayushchie tekhnologii nedropol'zovaniya i povysheniya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.
- Books and other monographs:**
Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publ. 352 p.
- Dissertation and Thesis:**
Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.
- State standards and patents:**
GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publ. 10 p.
Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.
- References in Latin transcription are presented in the original language.
References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.
10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
- the journal title and author's name in the "Subject" field;
 - the article title, authors' names, and the journal title, whereto the paper is being submitted, in the text of the email;
 - an article and additional materials have to be attached using the "attach" function; and
 - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "Informatics and Applications" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

Editorial Board address:

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia
Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05
e-mail: iiep@frccsc.ru (to Svetlana Strigina)
<http://www.ipiran.ru/english/journal.asp>