



Российская Академия Наук

**А
Т
АВТОМАТИКА
И ДАЛЕКОУПРАВЛЕНИЕ
AUTOMATION AND REMOTE CONTROL**

Журнал основан в 1936 году
Выходит 12 раз в год

7

ИЮЛЬ

Москва

2025

Учредители журнала:

Отделение энергетики, машиностроения, механики и процессов управления РАН,
Институт проблем управления им. В.А. Трапезникова РАН (ИПУ РАН),
Институт проблем передачи информации им. А.А. Харкевича РАН (ИППИ РАН)

Главный редактор:

Галяев А.А.

Заместители главного редактора:

Рубинович Е.Я., Хлебников М.В.

Ответственный секретарь:

Самохин А.С.

Редакционный совет:

Васильев С.Н., Желтов С.Ю., Каляев И.А., Кузнецов О.П., Кулешов А.П.,
Куржанский А.Б., Мартынюк А.А. (Украина), Пешехонов В.Г., Попков Ю.С.,
Федосов Е.А., Черноусько Ф.Л.

Редакционная коллегия:

Алескеров Ф.Т., Арутюнов А.В., Бахтадзе Н.Н., Бобцов А.А., Виноградов Д.В.,
Вишневский В.М., Воронцов К.В., Граничин О.Н., Жилиякова Н.Ю.,
Каравай М.Ф., Кибзун А.И., Краснова С.А., Крищенко А.П., Кузнецов Н.В.,
Кушнер А.Г., Лазарев А.А., Ляхов А.И., Маликов А.И., Матасов А.И.,
Меерков С.М. (США), Мещеряков Р.В., Миллер Б.М., Михальский А.И.,
Моржин О.В., Мунасыпов Р.А., Назин А.В., Немировский А.С. (США),
Новиков Д.А., Олейников А.Я., Пакшин П.В., Пальчунов Д.Е.,
Поляков А.Е. (Франция), Протасов В.Ю., Рапопорт Л.Б., Родионов И.В.,
Сельвесюк Н.И., Соболевский А.Н., Степанов О.А., Филимонюк Л.Ю.,
Фрадков А.Л., Хоров Е.М., Цыбаков А.Б. (Франция), Чеботарев П.Ю.,
Чхартишвили А.Г., Щербаков П.С.

Адрес редакции: 117997, Москва, Профсоюзная ул., 65

Тел./факс: 8 (495) 198-17-20, доб. 1443

Электронная почта: redacsia@ipu.ru, ait@pran.ru

Сайт: <https://aitras.ru>

Зав. редакцией *Е.А. Мартехина*

Москва

ФГБУ «Издательство «Наука»

Управление в технических системах

© 2025 г. Ю.Г. БУЛЫЧЕВ, д-р техн. наук (profbulychev@yandex.ru)
(АО «Концерн Радиоэлектронные технологии», Москва)

ГИБРИДНЫЙ МЕТОД КОМПЛЕКСИРОВАНИЯ РАЗНЕСЕННЫХ ДАТЧИКОВ ДЛЯ АДАПТИВНОЙ ИДЕНТИФИКАЦИИ СТАЦИОНАРНЫХ И ПОДВИЖНЫХ ОБЪЕКТОВ В УСЛОВИЯХ АПРИОРНОЙ НЕОПРЕДЕЛЕННОСТИ

В рамках совместного применения классического статистического подхода, основанного на вероятностных моделях, а также нетрадиционного подхода, ориентированного на принципы непрерывности, размножения и кластеризации, для совокупности разнесенных в пространстве датчиков (произвольного типа) развит новый метод адаптивной идентификации стационарных и подвижных объектов. Задача решается применительно к условиям существенной априорной неопределенности, связанной, например, с наличием в измерениях плохо формализуемых аномальных ошибок измерений, сбоями в работе и выходе из строя отдельных датчиков или значительными изменениями условий наблюдения объекта идентификации. Приводятся модели, критерий и алгоритм идентификации, устойчивый к такой неопределенности. В качестве примера сравнительного анализа использована триангуляционная система датчиков, для которой показан эффект, достигаемый за счет адаптации. Даны практические рекомендации.

Ключевые слова: существенная априорная неопределенность, датчик, информационно-измерительная система, объект идентификации, комплексирование, гибридный метод, адаптация, аномальные ошибки измерений, принципы непрерывности, размножения и кластеризации, целевая функция, семейство частных оценок, конкурирующий кластер, алгоритм идентификации, апостериорные веса, результирующая оценка.

DOI: 10.31857/S0005231025070017, EDN: JQXEВH

1. Введение

К настоящему времени накоплен большой теоретический и практический задел для построения различных информационно-измерительных систем (ИИС), состоящих из разнесенных в пространстве датчиков (однородных, неоднородных, стационарных и мобильных), используемых для совместного решения задачи идентификации состояния различных неподвижных и подвижных объектов. Подобные системы находят широкое применение в различных областях: медицина, техническая диагностика, радиоастрономия, радиолокация (активная и пассивная), радиотехническая разведка, полигонные испытания, охрана и т.д.

Среди ИИС существует целый класс систем, для которых характерны следующие условия функционирования: многоэтапность обработки данных, режим реального времени, существенная неопределенность и большие риски (ущерб) от некачественных оценок и принимаемых на их основе решений. Неопределенность может быть связана с отсутствием достоверных знаний о законах распределения ошибок измерений, пропусками в наблюдениях, появлением сбоев и отказов аппаратуры, наличием аномальных ошибок измерений (АОИ) и т.д. (см., например, [1–5]). Говоря о возможных рисках, предполагается, что важно учитывать не только возможный средний ущерб, но и потери от некачественного решения в каждом частном случае, т.е. оперируя с фиксированной выборкой измерений, в том числе и малого объема. С подобными выборками приходится сталкиваться на этапах вторичной и третичной обработки данных, например, в активных и пассивных радиолокационных системах (при формировании косвенных измерений пеленгов, временных задержек, наклонных дальностей, радиальных скоростей и др.). От качества таких измерений зависит эффективность решения задачи оценивания местоположения и параметров движения наблюдаемых объектов в рамках известных методов координатометрии (триангуляционного, разностно-дальномерного, триангуляционно-разностно-дальномерного и др.).

Основу алгоритмов, связанных с обработкой данных в условиях неопределенности, составляют классические вероятностные методы максимума апостериорной плотности вероятности или максимального правдоподобия, как правило, в расширенном варианте [1, 6–8]. Данный вариант предполагает включение всех неизвестных параметров неопределенности в расширенный вектор состояния, однако процедура расширения приводит на практике к резкому росту вычислительных затрат, известному эффекту «размазывания точности» и проблемам со сходимостью и устойчивостью формируемых оценок. Указанные трудности классических методов еще более усугубляются, если учесть, что для преодоления неопределенности, как правило, вводится семейство различных гипотез (многоальтернативный подход [6]), каждая из которых должна быть либо отвергнута, либо получить свое подтверждение. Анализ показывает, что для условий неопределенности и жестких требованиях к ИИС классические вероятностные методы зачастую малопригодны, однако их можно эффективно использовать на этапе математического моделирования для исследования потенциальных возможностей проектируемых ИИС.

Существуют и другие методы оптимального и квазиоптимального оценивания, которые можно использовать для условий неопределенности. К ним относятся: расширенный метод наименьших квадратов (РМНК) [7], методы минимума различных геометрических и кинематических невязок [9], метод инвариантов [10], всевозможные методы линейной и нелинейной фильтрации (например, калмановской) [1, 6, 7, 11], различные адаптивные методы с подстройкой параметров [12, 13], робастные [6, 14, 15], рандомизированные [16, 17], кластерные [17, 18] и нейросетевые [19] методы, а также

большое число эвристических методов, применяемых в специализированных ИИС (см., например, [20–25]). При выполнении ряда ограничений на условия наблюдения указанные методы позволяют преодолевать факторы неопределенности, например, бороться с АОИ. Однако в условиях существенной неопределенности данные методы сталкиваются с различными непредвиденными ограничениями и жесткими требованиями к ИИС, которые значительно сужают сферу их применения.

Это зачастую связано с трудностью учета геометрического фактора, зависящего от размещения в пространстве семейства датчиков и наблюдаемого объекта. Например, для угломерных систем возникают триангуляционно-некорректные области наблюдения, характерные даже для случая хороших измерений. Кроме того, задачи пространственного оценивания, как правило, сталкиваются с выраженной нелинейностью, с применением неограниченных функций (например, некоторых тригонометрических, необходимых для описания связи оцениваемых и измеряемых параметров) и др.

Указанные выше методы, как правило, гарантируют сходимость в среднем (либо по ансамблю реализаций, либо по одной достаточно длинной реализации), но не гарантируют качество формируемой оценки по одной выборке малого объема. Кроме того, в процедурах оценивания в основном используется весовая обработка, в которой расчет весовых коэффициентов осуществляется на основании большого объема априорной информации. В условиях указанной неопределенности такая информация зачастую может отсутствовать или оказаться недостоверной.

Общая тенденция к построению ИИС, способных противостоять различным факторам неопределенности, состоит в комплексировании различных методов оценивания с учетом границ их применимости (в том числе и тех методов, которые указаны выше). Каждый из методов выбранного семейства в некотором смысле должен быть «ортогонален» по точности другим методам этого семейства и реализовывать свои потенциальные возможности для определенных условий наблюдения. Гибридный метод, реализующий принцип комплексирования, должен сочетать в себе достоинства как традиционного статистического подхода (обеспечивающего оптимальность в среднем), так и нетрадиционного подхода (обеспечивающего оптимальность в частном), в максимальной степени учитывающего специфику применяемой ИИС и указанные условия неопределенности.

Для нетрадиционного подхода можно использовать кластерно-вариационный метод [18], который зарекомендовал себя в качестве эффективного инструмента преодоления априорной неопределенности, как в среднем, так и частном. Он базируется на принципе размножения единичных оценок (т.е. рассмотрении множества вариантов решения одной и той же задачи для различных наборов измерений, формируемых на базе исходного входного наблюдения), принципе кластеризации (разбиении данных оценок на несколько конкурирующих кластеров) и алгоритме построения результирующей оценки

с учетом этой кластеризации. К недостаткам указанного метода можно отнести следующее: рассматривается лишь триангуляционная система угломерных (т.е. однородных) датчиков; объект идентификации полагается стационарным, т.е. речь идет только об оценке местоположения; построение результирующей оценки не учитывает разноточность единичных оценок в пределах одного кластера (это не соответствует важному принципу адаптивности алгоритма идентификации для более широкого спектра условий наблюдения).

Ниже развивается обобщенный кластерно-вариационный метод идентификации (ОКВМИ), который полностью устраняет все отмеченные недостатки кластерно-вариационного метода [18]. Метод должен опираться на принципы, понятные специалистам в области построения и применения указанного класса ИИС, которые позволят обеспечить формирование гарантированных оценок (в рамках принятых ограничений) не только местоположения, но и параметров движения объектов идентификации в условиях существенной неопределенности. Формируемые весовые коэффициенты (для обработки измерений) должны учитывать, как априорную, так и апостериорную информацию.

2. Основные определения, модели и ограничения. Постановка задачи

По аналогии с [18] описание предлагаемого метода в основном будет рассматриваться для случая одного объекта, что позволит опустить громоздкие обозначения и выкладки. Обобщение метода на случай нескольких объектов рассматривается в разделе 5.

В декартовой системе координат XYZ рассматривается система, состоящая из множества датчиков D_m , $m = \overline{1, M}$, при этом положение каждого датчика задается вектором $\mathbf{p}_m = [\rho_{xm}, \rho_{ym}, \rho_{zm}]^T$. Объект характеризуется вектором оцениваемых постоянных параметров $\boldsymbol{\lambda} = [\lambda_p, p = \overline{1, P}]^T$, а все измеряемые параметры объединяются в вектор $\mathbf{s} = [s_j, j = \overline{1, J}]^T$. Для случая подвижного объекта координаты вектора $\boldsymbol{\lambda}$ соответствуют неизвестным спектральным коэффициентам при соответствующих базисных функциях выбранного функционального пространства. Для стационарного объекта в состав $\boldsymbol{\lambda}$ входят декартовы координаты объекта. В зависимости от типа используемых датчиков в качестве координат вектора \mathbf{s} могут, например, выступать значения следующих величин: азимута и угла места, наклонной дальности, радиальной скорости, уровня (амплитуды или мощности), разности фаз и временных задержек принимаемого сигнала (а также различных производных от указанных величин), относящиеся к различным точкам пространства и моментам времени.

Входное наблюдение характеризуется вектором

$$(1) \quad \mathbf{h} = \mathbf{s} + \boldsymbol{\Delta} + \boldsymbol{\xi},$$

где $\Delta = [\Delta_j, j = \overline{1, J}]^T$ – неизвестная ошибка измерений, $\xi = [\xi_j, j = \overline{1, J}]^T$ – шум измерений с неизвестным законом распределения, но нулевым математическим ожиданием и заданной корреляционной матрицей \mathbf{K}_ξ .

Для описания ошибки Δ воспользуемся следующим подходом. Пусть J^Δ соответствует реальному количеству тех ненулевых координат вектора Δ , которые являются АОИ и приводят к ненадежным результатам идентификации. Считаем, что $J^\Delta \leq J_{\max}^\Delta$, где $J_{\max}^\Delta \in \{0, 1, \dots\}$. Кроме того, выполняется условие:

$$(2) \quad J_{\min}^{\mathbf{h}} + J_{\max}^\Delta \leq J,$$

где $J_{\min}^{\mathbf{h}}$ – минимальное количество координат вектора \mathbf{h} , которого достаточно для качественного решения задачи идентификации объекта при условии, что АОИ отсутствуют.

В дальнейшем на базе \mathbf{h} будут формироваться различные наборы измерений $\mathbf{h}_i = [h_{ij}, j = \overline{1, J_i}]^T$ (где $h_{ij} \in \{h_j, j = \overline{1, J}\}$, $J_i \leq J$, $\mathbf{h}_i \neq \mathbf{h}_r \forall i \neq r$, $i, r = \overline{1, I}$), для которых вводится ограничение

$$(3) \quad J_{\min}^{\mathbf{h}} \leq J_i \leq J.$$

При таком подходе к описанию ошибки Δ часть ее ненулевых координат может принимать произвольные значения (в том числе и аномальные), они могут располагаться в векторе Δ самым произвольным образом и для их описания не существует какой-либо универсальной и удовлетворительной модели. Единственным способом формализации ошибки Δ является введение некоторых количественных ограничений (типа (2), (3)), которые соответствуют практике эксплуатации конкретной ИИС.

На вектор параметров λ накладываются традиционные ограничения

$$(4) \quad \lambda \in \Lambda \subset R^P,$$

где $\Lambda = \{\Lambda_{\min}^p \leq \lambda_p \leq \Lambda_{\max}^p\}$, $\Lambda_{\min}^p \in R^1$ и $\Lambda_{\max}^p \in R^1$.

Считается известным оператор Ψ , такой что

$$(5) \quad \Psi : \mathbf{s} \rightarrow \lambda,$$

т.е. $\lambda = \Psi(\mathbf{s})$, при этом существует обратный оператор

$$(6) \quad \Psi^{-1} : \lambda \rightarrow \mathbf{s},$$

для которого $\mathbf{s} = \Psi^{-1}(\lambda)$.

Разрабатываемый ОКВМИ базируется на трех основных принципах: непрерывности, размножения и кластеризации. Первый принцип связан с утверждением о непрерывной зависимости качества решения задачи идентификации от параметров входного наблюдения достаточной размерности.

В частности, последовательное уменьшение параметра J (путем исключения тех или иных координат вектора \mathbf{h}) приводит к плавной эволюции точности результирующей оценки. Второй принцип связан с утверждением о возможности формирования множества частных оценок (как хороших, так и плохих), при этом множество хороших оценок имеет (с учетом условий (2)–(4)) достаточную мощность. Третий принцип связан с утверждением о том, что при выполнении условий (2)–(4) все возможные частные оценки можно кластеризовать, при этом среди конкурирующих кластеров есть оптимальный, который служит основой для формирования надежной результирующей оценки.

Согласно принципу размножения основному вектору \mathbf{h} ставится в соответствие семейство частных векторов $\{\mathbf{h}_i, i = \overline{1, I}\}$:

$$(7) \quad \mathbf{h}_i = \mathbf{s}_i + \mathbf{\Delta}_i + \boldsymbol{\xi}_i,$$

где $\mathbf{s}_i = [s_{ij}, j = \overline{1, J_i}]^T$, $\mathbf{\Delta}_i = [\Delta_{ij}, j = \overline{1, J_i}]^T$, $\boldsymbol{\xi}_i = [\xi_{ij}, j = \overline{1, J_i}]^T$.

Выбор значений I и J_1, \dots, J_I должен отвечать условиям (2) и (3), а кроме того, должно гарантироваться наличие среди $\{\mathbf{h}_i, i = \overline{1, I}\}$ векторов, не содержащих АОИ.

Каждому вектору \mathbf{h}_i можно поставить в соответствие некоторый оператор Ψ^* классического оптимального оценивания (например, оператор метода максимального правдоподобия)

$$(8) \quad \Psi^* : (\mathbf{h}_i, \mathbf{K}_{\xi_i}) \rightarrow \boldsymbol{\lambda}_i^*,$$

где $\boldsymbol{\lambda}_i^* = \Psi^*(\mathbf{h}_i, \mathbf{K}_{\xi_i})$ – частная оценка, характеризуемая корреляционной матрицей $\mathbf{K}_{\xi_i}^*$ ошибок оценивания, \mathbf{K}_{ξ_i} – корреляционная матрица шума $\boldsymbol{\xi}_i$.

Для вычисления элементов матрицы $\mathbf{K}_{\xi_i}^*$ используются известные формулы (см., например, [1, 3, 4]), вид которых полностью определяется типом ИИС (триангуляционная, разностно-дальномерная, угломерно-мощностная и т.д.), ее геометрией и условиями наблюдения, а также значением матрицы \mathbf{K}_{ξ_i} .

Применительно к каждой оценке $\boldsymbol{\lambda}_i^*$ с помощью известного оператора Θ (также зависящего от типа, геометрии и параметров ИИС) формируется оценка вектора вторичных измерений $\left(\mathbf{s}_i^* = [s_{ij}^*, j = \overline{1, J_i}]^T\right)$:

$$(9) \quad \Theta : \boldsymbol{\lambda}_i^* \rightarrow \mathbf{s}_i^*,$$

где $\mathbf{s}_i^* = \Theta(\boldsymbol{\lambda}_i^*)$.

Следует помнить, что в общем случае вектор \mathbf{s}_i^* имеет размерность J_i , не совпадающую с размерностью J вектора \mathbf{s} . При вычислении \mathbf{s}_i^* учитываются декартовы координаты только тех датчиков ИИС, которые используются в формировании частного вектора \mathbf{h}_i .

Применение вторичных измерений позволяет отсеять плохие частные оценки $\boldsymbol{\lambda}_i^*$ путем проверки условия

$$(10) \quad |s_{ij}^* - h_{ij}| \leq \varepsilon_{ij} \quad \forall j = \overline{1, J_i},$$

где ε_{ij} – максимальная величина невязки, вычисляемая с учетом типа ИИС и матрицы \mathbf{K}_i^* .

Если условие (10) не выполняется или оценка λ_i^* не попадает в ограничение (4), то данная частная оценка отсеивается. Оставшиеся после отсева частные оценки обозначим через $\lambda_{[n]}^*$, $n = \overline{1, N}$, где $\lambda_{[n]}^* \in \{\lambda_1^*, \dots, \lambda_J^*\}$. Очевидно, что множество таких оценок не пусто, если выполнены условия (2) и (3).

Определение 1. Скалярная ошибка Δ_{ij} , являющаяся одной из составляющих вектора Δ_i , называется аномальной, если она приводит к нарушению условия (10). В противном случае ошибка Δ_{ij} называется нормальной.

Определение 2. Любая группа скалярных ошибок (из двух и более), являющихся составляющими вектора Δ_i , называется аномальной (даже если каждая из этих составляющих сама по себе является нормальной), если она приводит к нарушению условия (10). В противном случае группа называется нормальной.

Определение 3. Под коэффициентом «засорения» исходной выборки \mathbf{h} понимается величина $k_{\mathbf{h}} = 100(k_{\Delta}/J)\%$, где k_{Δ} – суммарное количество аномальных скалярных ошибок и скалярных ошибок, входящих в аномальные группы.

Сформированное множество $\mathbf{X}^* = \{\lambda_{[1]}^*, \dots, \lambda_{[N]}^*\}$ разбивается на конкурирующие кластеры \mathbf{K}_q , $q = \overline{1, Q}$ (для этого может применяться любой из известных алгоритмов кластеризации, например, [18, 26, 27]). Оценки $\lambda_{[1]}^*, \dots, \lambda_{[N]}^*$ последовательно (по шагам) объединяются в группы: сначала самые близкие, а затем все более отдаленные друг от друга. На первом шаге каждая оценка $\lambda_{[n]}^*$ рассматривается как отдельный кластер.

Алгоритм кластеризации устанавливает отображение: $\Upsilon: \mathbf{X}^* \rightarrow \{\mathbf{K}_q, q = \overline{1, Q}\}$, при этом должно выполняться условие

$$(11) \quad \mathbf{X}^* = \bigcup_{q=1}^Q \mathbf{K}_q,$$

где $\mathbf{K}_k \cap \mathbf{K}_r = \emptyset$, $k, r \in \{1, \dots, Q\}$, $k \neq r$ и $\mathbf{K}_q \neq \emptyset \forall q = \overline{1, Q}$.

Для каждого кластера будем использовать представление $\mathbf{K}_q = \left\{ \lambda_{[q,n]}^*, n = \overline{1, N_q} \right\}$, где $\lambda_{[q,n]}^*$ – частная оценка, которая попала в кластер \mathbf{K}_q и ей присвоен номер n . Каждой $\lambda_{[q,n]}^*$ соответствует оценка вектора вторичных полномерных измерений $\mathbf{s}_{[q,n]}^* = \left[s_{j[q,n]}^*, j = \overline{1, J} \right]^T$ (по аналогии с (9)):

$$(12) \quad \mathbf{s}_{[q,n]}^* = \Theta \left(\lambda_{[q,n]}^* \right).$$

Под полномерностью понимается то, что размерности векторов \mathbf{s} и $\mathbf{s}_{[q,n]}^*$ совпадают.

Выбор семейства кластеров $\{K_1, \dots, K_Q\}$ подробно обсуждается в [18], где предложен новый адаптивный алгоритм кластеризации, в том числе и при неизвестном Q . Он позволяет разбить все семейство $\{\lambda_{[1]}^*, \dots, \lambda_{[N]}^*\}$ на кластеры автономно (без участия оператора), достаточно прост для компьютерной реализации и не требует больших вычислительных затрат. Результаты применения данного алгоритма показали его высокую эффективность и хорошие асимптотические свойства.

Требуется на базе (1)–(12) разработать в оптимальной постановке адаптивный ОКВМИ, устойчивый к ошибке Δ и учитывающий при построении результирующей оценки $\lambda^* = [\lambda_p^*, p = \overline{1, P}]^T$ разноточность частных оценок, входящих в конкурирующие кластеры. Метод должен включать в себя обоснование целевой функции, критерия, алгоритма выбора оптимального кластера и построения оптимальной результирующей оценки.

3. Обоснование и построение целевой функции. Критерий и алгоритм идентификации

В случае, когда $\Delta = 0$ и априорная информация о шуме ξ ограничивается знанием корреляционной матрицы K_ξ , предпочтение отдается традиционному статистическому методу, в котором используется весовая матрица $W_\xi = K_\xi^{-1}$ (это обеспечивает оптимизацию в среднем). В рамках разрабатываемого ОКВМИ, применительно к условиям существенной неопределенности, необходимо сформировать новую целевую функцию, которая учитывала бы все параметры, входящие в ограничения (2)–(4) и характеризующие конкурирующие кластеры. В данной функции должны фигурировать весовые коэффициенты, которые в отличие от любого оптимального статистического метода вычисляются апостериорно по одной фиксированной выборке измерений (это обеспечивает оптимизацию в частном).

Для решения задачи идентификации используем следующую безразмерную нормированную целевую функцию:

$$(13) \quad F(\lambda, q) = J^{-1} \sum_{j=1}^J \left\{ \left[\frac{s_j(\lambda) - h_j}{\delta_j} \right]^2 w(h_j, q) \right\}, \quad 0 \leq F(\lambda, q) \leq 1,$$

где $s_j(\lambda)$ – известная зависимость измеряемого параметра s_j от λ , $w(h_j, q)$ – апостериорный вес измерения h_j , $0 \leq w(h_j, q) \leq 1$, $\delta_j > 0$ – коэффициент, ограничивающий величину $|s_j(\lambda) - h_j|$ для всех просматриваемых (модельных) значений $\lambda \in \Lambda$ и обеспечивающий соответствующую нормировку.

Для нахождения весов заметим, что применительно к кластеру K_q и его любой частной оценке $\lambda_{[q,n]}^*$ необходимо проверять условие (по аналогии с (10))

$$(14) \quad \left| s_{[q,n]}^* - h_j \right| \leq \varepsilon_{j[q,n]}.$$

Левая часть в (14) характеризует невязку между первичными h_j и вторичными $s_{j[q,n]}^*$ измерениями для всех возможных значений j , q и n . Для учета ограничения на данную невязку воспользуемся индикаторной функцией

$$(15) \quad \varphi(p) = \begin{cases} 1 - p, & 0 \leq p \leq 1, \\ \varphi(p) = 0, & p > 1. \end{cases}$$

Принимая $p = (s_{j[q,n]}^* - h_j)^2 / \varepsilon_{j[q,n]}^2$, весовые коэффициенты, учитывающие все принятые ограничения задачи идентификации, а также результаты размножения частных оценок и их кластеризации, находятся так

$$(16) \quad w(h_j, q) = N_q^{-1} \sum_{n=1}^{N_q} \varphi \left(\frac{(s_{j[q,n]}^* - h_j)^2}{\varepsilon_{j[q,n]}^2} \right), \quad 0 \leq w(h_j, q) \leq 1,$$

где N_q^{-1} – играет роль нормирующего множителя.

Формула (16) позволяет учесть количество N_q частных оценок в каждом кластере K_q , пороги $\varepsilon_{j[q,n]}$ допустимых погрешностей вторичных измерений и наличие грубых ошибок измерений (одиночных и групповых) путем введения индикаторной функции $\varphi(p)$. Коэффициент $w(h_j, q)$ показывает удельный вклад измерения h_j в образование всех отметок кластера K_q . Его можно назвать коэффициентом соответствия измерения h_j элементам кластера K_q . Чем больше значение коэффициента, тем больше доверия к первичному измерению h_j , с которым связан этот коэффициент.

С учетом (14)–(16) искомая целевая функция принимает вид

$$(17) \quad F(\boldsymbol{\lambda}, q) = (JN_q)^{-1} \sum_{j=1}^J \sum_{n=1}^{N_q} \left[\frac{s_j(\boldsymbol{\lambda}) - h_j}{\delta_j} \right]^2 \varphi \left(\frac{(s_{j[q,n]}^* - h_j)^2}{\varepsilon_{j[q,n]}^2} \right).$$

Критерий оптимальной идентификации сводится к следующему

$$(18) \quad \boldsymbol{\lambda}^* = \arg \min_{\boldsymbol{\lambda}} F(\boldsymbol{\lambda}, q^*),$$

$$(19) \quad q^* = \arg \max_q w(\mathbf{h}, q) = \arg \max_q \left\{ J^{-1} \sum_{j=1}^J w(h_j, q) \right\}, \quad 0 \leq w(\mathbf{h}, q) \leq 1,$$

где $w(\mathbf{h}, q)$ – интегральный безразмерный нормированный весовой коэффициент.

Минимизация решающей функции $F(\boldsymbol{\lambda}, q)$, по векторному аргументу $\boldsymbol{\lambda}$ приводит к уравнению

$$(20) \quad \left[\frac{\partial F(\boldsymbol{\lambda}, q)}{\partial \boldsymbol{\lambda}} \right]^T = \mathbf{0}.$$

Раскрывая в (20) частные производные, получим систему скалярных уравнений

$$(21) \quad \sum_{j=1}^J \left\{ \left(\frac{s_j(\boldsymbol{\lambda}) - h_j}{\delta_j} \right) w(h_j, q) \frac{\partial s_j(\boldsymbol{\lambda})}{\partial \lambda_p} \right\} = 0, \quad p = \overline{1, P},$$

где частные производные раскрываются с учетом известной зависимости $\mathbf{s} = \boldsymbol{\Psi}^{-1}(\boldsymbol{\lambda})$.

Системе (21) соответствуют частные оценки $\boldsymbol{\lambda}^*(q)$, $q = \overline{1, Q}$. В качестве результирующей оценки $\boldsymbol{\lambda}^* \in \{\boldsymbol{\lambda}^*(1), \dots, \boldsymbol{\lambda}^*(Q)\}$ выбирается та частная оценка $\boldsymbol{\lambda}^*(q^*)$ (где $q^* \in \{1, \dots, Q\}$), которая согласно (19) удовлетворяет критерию (для фиксированного \mathbf{h})

$$(22) \quad w(\mathbf{h}, q^*) > w(\mathbf{h}, q) \quad \forall q = \overline{1, Q}, \quad q \neq q^*.$$

Выбором значений параметров, фигурирующих в условиях (2)–(4), (10), а также вводя ограничения на диапазоны изменения значений I , Q и \mathbf{K}_ξ , всегда можно обеспечить единственность и требуемую точность решения задачи идентификации. При правильном планировании измерительного эксперимента (см., например, [7]) практически всегда обеспечивается надежное обнаружение результирующего кластера \mathbf{K}_{q^*} , содержащего наибольшее число частных оценок, образованных достоверными наблюдениями \mathbf{h}_i . В свою очередь, параметры $\varepsilon_{j[q,n]}$ существенно влияют на точность идентификации.

Введение ограничения (4) позволяет отсеять так называемые некорректные области идентификации, в которых может нарушаться условие единственности решения. Например, для триангуляционной ИИС, состоящей из двух пеленгаторов, речь идет об областях, для которых треугольник, образуемый базой и двумя линиями визирования, имеет слишком маленький угол при его вершине (в ней находится объект). В этих случаях линии визирования почти параллельны и даже незначительные ошибки пеленгования, не говоря уже об АОИ, могут приводить к существенным ошибкам координатометрии. Кроме того, среди конкурирующих кластеров могут появиться такие, которые находятся на одном направлении, разнесены на большие расстояния и имеют приблизительно одинаковое число элементов. Для борьбы с такой некорректностью либо вводят ограничения (4), либо используют дополнительные измерения от некоторой внешней системы, которая в совокупности с рассматриваемой ИИС устраняет некорректные области. Это также относится к задачам оптимального планирования эксперимента.

Если ограничиться одним кластером, снять принятые ограничения, а весовым коэффициентам $w(h_1, q), \dots, w(h_J, q)$ поставить в соответствие обратные величины дисперсий ошибок первичных измерений (распределенных по нормальному закону), то приведенное решение (18)–(21) будет соответствовать известной максимально правдоподобной оценке для случая отсутствия АОИ.

В классической статистической постановке задача идентификации решается на базе всех измерений (независимо от того, хорошие они или плохие),

подлежащих совместной оптимальной обработке с учетом заранее назначаемых априорных весов. Развиваемый гибридный метод также оперирует со всеми измерениями, но учитывает как априорную информацию (в виде весовых матриц $\mathbf{W}_{\xi_i} = \mathbf{K}_{\xi_i}^{-1}$, используемых при построении частных оценок), так и апостериорную информацию (в виде весов $w(h_1, q), \dots, w(h_J, q)$ и $w(\mathbf{h}, q)$), которая требуется для нахождения оптимального кластера и построения результирующей оценки.

Сам алгоритм решения задачи идентификации в кластерной оптимальной постановке сводится к следующему:

1) Проводится кластеризация частных оценок путем построения кластеров K_q , $q = \overline{1, Q}$, (рекомендации по кластеризации даны в [18]).

2) Для каждого кластера K_q подсчитываются количество входящих в него отметок (N_q), частные и интегральный весовые коэффициенты $w(h_1, q), \dots, w(h_J, q)$ и $w(\mathbf{h}, q)$ соответственно).

3) С использованием критерия (19) решается задача нахождения номера q^* оптимального кластера K_{q^*} .

4) Для кластера K_{q^*} решается система уравнений (21), выбирая в качестве начального условия частную оценку, принимаемую за центр этого кластера. В итоге, с учетом (21), получаем результирующую оценку $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}^*(q^*) = [\lambda_p(q^*), p = \overline{1, P}]^T$ вектора параметров $\boldsymbol{\lambda}^*$ идентифицируемого объекта.

Замечание 1. Если воспользоваться известным приближенным подходом к построению максимально правдоподобной оценки, основанным на линейной аппроксимации невязок $s_j(\boldsymbol{\lambda}) - h_j$, то вместо (21) можно получить соответствующую систему линейных алгебраических уравнений.

Реализация предлагаемого алгоритма (с учетом подготовки исходных данных) связана со следующими вычислительными затратами $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$, где Γ_1 – затраты на построение семейства частных оценок, Γ_2 – затраты на кластеризацию этих оценок и выбор результирующего кластера, Γ_3 – затраты на построение результирующей оценки. Реализация указанных операций в специализированных вычислительных средах не представляет особых сложностей и позволяет обеспечить режим реального времени.

Соотношения (13)–(22) составляют математическую основу ОКВМИ.

4. Применение метода к триангуляционной информационно-измерительной системе

Рассмотрим триангуляционную ИИС, в которой в качестве датчиков ($\mathcal{D}_m, m = \overline{1, M}$) используются двухканальные пеленгаторы, измеряющие азимут α_m и угол места β_m стационарного излучающего объекта. В этом случае: $P = 3$, $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \lambda_3]^T = [x, y, z]^T$, $J = 2M$, $\mathbf{s} = [\alpha_m, \beta_m, m = \overline{1, M}]^T$ и $\mathbf{h} = [h_j, j = \overline{1, J}]^T = [h_m^\alpha, h_m^\beta, m = \overline{1, M}]^T$, $\delta_j = \pi$.

В качестве $\mathbf{s} = \Psi^{-1}(\boldsymbol{\lambda})$ применяются известные формулы

$$(23) \quad \begin{cases} \alpha_m(\boldsymbol{\lambda}) = \arccos \left\{ (x - x_m) [(x - x_m)^2 + (y - y_m)^2]^{-1/2} \right\}, \\ \beta_m(\boldsymbol{\lambda}) = \arcsin \left\{ (z - z_m) [(x - x_m)^2 + (y - y_m)^2 + (z - z_m)^2]^{-1/2} \right\}. \end{cases}$$

На базе (23) формируется система уравнений (21).

Условие (14) преобразуется к виду

$$(24) \quad \begin{cases} \left| \alpha_{m[q,n]}^* - h_m^\alpha \right| \leq \varepsilon_{m[q,n]}^\alpha, \\ \left| \beta_{m[q,n]}^* - h_m^\beta \right| \leq \varepsilon_{m[q,n]}^\beta, \end{cases}$$

а целевая функция

$$(25) \quad F(\boldsymbol{\lambda}, q) = \sum_{m=1}^M \left[\left(\frac{\alpha_m(\boldsymbol{\lambda}) - h_m^\alpha}{\pi} \right)^2 w(h_m^\alpha, q) + \left(\frac{\beta_m(\boldsymbol{\lambda}) - h_m^\beta}{\pi} \right)^2 w(h_m^\beta, q) \right],$$

$$(26) \quad \begin{cases} w(h_m^\alpha, q) = N_q^{-1} \sum_{n=1}^{N_q} \varphi \left(\frac{(\alpha_{m[q,n]}^* - h_m^\alpha)^2}{(\varepsilon_{m[q,n]}^\alpha)^2} \right), \\ w(h_m^\beta, q) = N_q^{-1} \sum_{n=1}^{N_q} \varphi \left(\frac{(\beta_{m[q,n]}^* - h_m^\beta)^2}{(\varepsilon_{m[q,n]}^\beta)^2} \right). \end{cases}$$

Интегральный весовой коэффициент вычисляется так

$$(27) \quad w(\mathbf{h}, q) = M^{-1} \sum_{m=1}^M [w(h_m^\alpha, q) + w(h_m^\beta, q)].$$

В качестве примера рассмотрим плоскостную задачу, принимая $\rho_{zm} = 0$ и $\lambda_3 = z = 0$, т.е. $\beta_m = 0 \forall m = \overline{1, M}$. Матрицу первичных азимутальных ошибок примем диагональной, т.е. $\mathbf{K}_\xi = \text{diag}[\sigma_{\alpha 1}^2, \dots, \sigma_{\alpha M}^2]$. В этом случае на базе известного оператора Ψ^* можно сформировать семейство частных оценок $\boldsymbol{\lambda}_i^* = [x_i^*, y_i^*]^T$, где

$$\begin{cases} x_i^* = (B_i E_i - C_i D_i)(A_i C_i - B_i^2)^{-1}, \\ y_i^* = (B_i D_i - A_i E_i)(A_i C_i - B_i^2)^{-1}. \end{cases}$$

Здесь $A_i = \sum_{m=1}^{M_i} \cos^2 \alpha_{mi} / \gamma_{mi}$, $B_i = 2^{-1} \sum_{m=1}^{M_i} \sin 2\alpha_{mi} / \gamma_{mi}$, $C_i = \sum_{m=1}^{M_i} \sin^2 \alpha_{mi} / \gamma_{mi}$, $D_i = \sum_{m=1}^{M_i} \varphi_{mi} / \gamma_{mi}$, $E_i = \sum_{m=1}^{M_i} \mu_{mi} / \gamma_{mi}$, $\gamma_{mi} = \tilde{R}_{mi}^2 \sigma_{\alpha mi}^2$, $\varphi_{mi} = \rho_{xmi} \cos^2 \alpha_{mi} - 2^{-1} \rho_{ymi} \sin 2\alpha_{mi}$, $\mu_{mi} = 2^{-1} \rho_{xmi} \sin 2\alpha_{mi} - \rho_{ymi} \sin^2 \alpha_{mi}$, \tilde{R}_{mi} – ориентировочная дальность от объекта до датчика \mathcal{L}_{mi} , $m \in \{1, \dots, M_i\}$.

Применительно к триангуляционной ИИС строятся эллипсы ошибок оценивания, которые с учетом частных оценок и геометрии триангуляционной ИИС используются для расчета констант $\varepsilon_{m[q,n]}^\alpha$ применительно к каждому кластеру и входящих в него частных оценок.

Соотношения (25)–(27) составляют основу одноэтапного ОКВМИ для триангуляционной ИИС. В отличие от известного кластерно-вариационного метода триангуляционного оценивания [18], разработанный метод учитывает при формировании апостериорных весов тот факт, что правые части неравенств (24) зависят как от номера q конкурирующего кластера, так и от номера n входящей в него частной оценки.

Можно также использовать двухэтапный ОКВМИ для триангуляционной ИИС в двух вариантах. Первый вариант основан на том, что есть возможность реализации двухэтапного ОКВМИ сначала только на азимутальных каналах (первый этап), что позволяет сформировать семейство плоскостных частных оценок и кластеров, а затем определить оценки x^* и y^* двух декартовых координат объекта (x и y). Точке (x^*, y^*) в трехмерном пространстве соответствует своя линия положения объекта. Объединяя эту линию и всевозможные конусы положения, соответствующие углам места, строим недостающую оценку z^* пространственной координаты z (второй этап).

Второй вариант основан на том, что применительно только к азимутальным каналам с использованием коэффициентов соответствия $w(h_m^\alpha, q)$ находятся недостоверные азимутальные каналы, измерения которых исключаются из дальнейшей обработки. Далее на оставшихся приоритетных азимутальных и всех угломестных каналах реализуется ОКВМИ в полном объеме. В этом случае возрастает количество исследуемых частных оценок (по сравнению с первым вариантом), что обеспечивает повышение качества триангуляционного оценивания в условиях неопределенности.

Очевидно, что двухэтапный ОКВМИ для двух рассмотренных вариантов не исчерпывает потенциальных возможностей оптимального ОКВМИ в условиях неопределенности, но является более экономичным с вычислительной точки зрения. Это связано, в первую очередь, с существенным сокращением числа частных триангуляционных оценок и конкурирующих кластеров.

5. Обобщение метода на случай многих объектов

Распространение развиваемого ОКВМИ на этот случай зависит от назначения, принципов построения и организации вычислительного процесса в конкретной ИИС. Можно выделить два основных варианта реализации ОКВМИ на случай многих объектов.

Вариант 1. Когда разбиение измерений на классы по признаку принадлежности к данному объекту выделено в отдельный этап (задача отождествления измерений). Такая усеченная постановка задачи весьма распространена на практике при соответствующей декомпозиции вычислительного про-

цесса. В этом случае ОКВМИ применяется к каждому выделенному классу. Для этого формируются векторы измерений $\mathbf{h}_{(l)} = [h_{(l)j}, j = \overline{1, J_{(l)}}]^T$, где $l \in \overline{1, L}$ – номер объекта (класса). Для l -го класса (по аналогии с разделом 3) формируются элементы $\mathbf{K}_{(l)q}$, $M_{(l)q}$, $w_{(l)}(h_j, q)$, $w_{(l)}(\mathbf{h}, q)$, $q_{(l)}^*$ и на их основе результирующие оценки $\lambda_{(l)}^* = [\lambda_{(l)j}^*, j = \overline{1, J}]^T$, $l = \overline{1, L}$. Вариант 1, наиболее эффективный с вычислительной точки зрения, позволяет организовать L параллельных каналов обработки измерений, однако он не исчерпывает потенциальных возможностей совместной обработки доступных измерений.

Примечание 1. Если в ходе решения задачи отождествления допущены неверные решения, то ложные измерения, попавшие в чужой класс, можно рассматривать как измерения, содержащие грубые ошибки. Предлагаемый ОКВМИ позволяет эффективно бороться с такими ошибками независимо от их природы.

Примечание 2. Возможно привлечение для решения задачи отождествления вспомогательных измерений (например, в радиолокации и радиотехнической разведке к ним можно отнести несущую частоту, период следования импульсов, их длительность, вид внутриимпульсной модуляции и т.д.), которыми «нагружены» основные измерения. Совместное использование основных и вспомогательных измерений зачастую необходимо для более глубокой идентификации объектов.

Вариант 2. В этом случае ОКВМИ применяется сразу ко всем измерениям с образованием кластеров \mathbf{K}_q , $q = \overline{1, Q}$. Первоначально требуется определить номера $q_{(l)}^* \in \{1, 2, \dots, Q\}$ приоритетных (оптимальных) кластеров $\mathbf{K}_{q_{(l)}^*}$, что достигается введением критерия $w_{(l)}(\mathbf{h}, q_{(l)}^*) \geq \gamma$, где γ – заданный порог распознавания объекта, $\gamma > 0$. Далее для всех этих кластеров строятся результирующие оценки $\lambda_{(l)}^* = [\lambda_{(l)j}^*, j = \overline{1, J}]^T$, $l = \overline{1, L}$.

Вариант 2 является достаточно затратным с вычислительной точки зрения, так как приводит к большому числу кластеров Q . Достоинство варианта состоит в том, что задача идентификации для случая многих объектов реализует потенциальные возможности ОКВМИ в условиях неопределенности.

6. Сравнительный анализ

Рассмотрим триангуляционную ИИС, датчики D_m которой расположены на окружности: $\mathbf{p}_m = [\rho_{xm}, \rho_{ym}, \rho_{zm}]^T$, $\rho_{xm} = 10^4 \cos[2\pi(m-1)/M]$, $\rho_{ym} = 10^4 \sin[2\pi(m-1)/M]$, $\rho_{zm} = 0$, $M = 5$, $m = \overline{1, 5}$. Для стационарного объекта исследовалось множество его возможных положений, при этом каждому положению соответствовал свой вектор $\lambda_k = [x_k, y_k, z_k]^T$, $x_k = 5 \times 10^4 \cos(2\pi k/K)$, $y_k = 5 \times 10^4 \sin(2\pi k/K)$, $z_k = 3 \times 10^3$, $k = \overline{1, K}$, $K = 180$. Здесь и далее координаты объекта и датчиков задаются в метрах, азимут, угол места и ошибки пеленгования в радианах. Для каждого k принима-

лось, что шумы измерений различных равноточных датчиков независимы, распределены по нормальному закону с нулевым математическим ожиданием, а корреляционная матрица датчика \mathbf{D}_m имеет диагональный вид $\mathbf{K}_m = \text{diag}[\sigma_\alpha^2, \sigma_\beta^2]$, где $\sigma_\alpha = \pi/360$ и $\sigma_\beta/360$. Ошибки формировались с помощью датчика случайных чисел. Номера недостоверных каналов (не более, чем половина всех азимутальных и не более, чем половина всех угломестных каналов), содержащих АОИ, выбирались случайным образом. В каналах формировалась результирующая ошибка измерений (это сумма АОИ и допустимой случайной ошибки), принадлежащая интервалам возможных значений $(3\sigma_\alpha, \pi/6)$ для азимутального канала и $(3\sigma_\beta, \pi/6)$ – для угломестного.

Для сравнительного анализа рассматривались четыре алгоритма, соответствующие известному кластерно-вариационному методу (КВМ, см. [28]) и ОКВМИ (развиваемый метод): A_1 – алгоритм КВМ с заранее заданным количеством кластеров Q , A_2 и A_3 – соответственно алгоритмы одноэтапного и двухэтапного КВМ с оптимальным выбором количества кластеров Q при заданном начальном условии Q_0 , A_4 – алгоритм, соответствующий ОКВМИ. Требовалось сравнить эти алгоритмы по точности и оперативности.

Процедура оценивания проводилась для каждого k с последующим усреднением по 100 экспериментам. В алгоритме A_1 принято $Q = 7$, а в алгоритмах A_2 и A_3 в качестве начального условия принято $Q_0 = 2$. Для сравниваемых алгоритмов использовались две числовые характеристики: $S(A_i)$ – интегральная характеристика точности (выражается в метрах), $T(A_i)$ – характеристика вычислительной оперативности (выражается в секундах).

Для интегральной характеристики точности имеем

$$S(A_i) = \sum_{k=1}^{180} S_k(A_i) = \frac{2\pi}{180} \sum_{k=1}^{180} \tau_k(A_i),$$

где $\tau_k(A_i) = \left\| \bar{\lambda}_k^*(A_i) - \lambda_k \right\|_2$ – частная невязка, соответствующая k -му положению объекта,

$$\bar{\lambda}_k^*(A_i) = \sum_{p=1}^{100} \lambda_{kp}^*(A_i) / 100$$

– усредненная оценка вектора λ_k (для 100 экспериментов), $\lambda_{kp}^*(A_i)$ – единичная оценка (для k -го положения объекта и p -го эксперимента).

Для сравнительного анализа алгоритмов A_i далее используются относительная интегральная характеристика

$$\delta S(A_i) = 100S(A_i)S^{-1}(A_1)[\%]$$

и относительная вычислительная оперативность

$$\delta T(A_i) = 100T(A_i)T^{-1}(A_4)[\%],$$

принимая во внимание, что алгоритм A_1 менее точный, а A_4 требует больше временных затрат по сравнению с другими алгоритмами. Результаты моделирования представлены в таблице сравнения алгоритмов.

Таблица

Алгоритмы (A_i)	A_1	A_2	A_3	A_4
$\delta S(A_i)$, %	100	50	67	39
$\delta T(A_i)$, %	81	92	14	100

Видим, что в аномальных условиях функционирования триангуляционной ИИС модифицированный алгоритм A_4 , учитывающий апостериорное изменение весов в зависимости от положения объекта, номера конкурирующего кластера и положения каждой частной оценки в этом кластере, существенно лучше (в плане точности) по сравнению с альтернативными алгоритмами. Моделирование также показало (по аналогии с [18]), что при данных аномальных условиях наблюдения метод максимального правдоподобия не применим ни в простом, ни в расширенном варианте. Так, ошибка координатометрии на базе рассматриваемой ИИС на ряде направлений достигает 100%.

7. Заключение

Предложенный ОКВМИ позволяет распространить идею кластерно-вариационного оценивания с простейшей задачи определения местоположения стационарного объекта в триангуляционной ИИС на более сложную задачу идентификации вектора параметров подвижного объекта для ИИС произвольного типа, оперирующих не только с однородными, но и разнородными измерениями, относящимися к произвольным точкам пространства и времени. Метод позволяет строить устойчивую к влиянию грубых ошибок измерений (например, аномальных одиночных и групповых) результирующую оценку вектора параметров в условиях существенной априорной неопределенности. ОКВМИ как самостоятельно, так и в совокупности с традиционными статистическими подходами может быть эффективно использован при совершенствовании существующих и разработке перспективных ИИС. В случае ограниченности вычислительных ресурсов и/или большом количестве датчиков и высокой размерности вектора наблюдений рассмотренный двухэтапный вариант реализации ОКВМИ может позволить существенно повысить оперативность идентификации.

Современный прогресс в области построения ЭВМ параллельного действия (особенно на новых принципах действия [29, 30]) позволяет надеяться, что предложенный принцип размножения частных оценок и их кластеризации в оптимальном варианте не станет препятствием для практической реализации ОКВМИ применительно к перспективным ИИС реального времени. Однако далеко не для всех существующих ИИС возможен такой подход, который на выборках большого объема может потребовать слишком большого числа каналов параллельной обработки данных и существенных вычислительных затрат.

СПИСОК ЛИТЕРАТУРЫ

1. Черняк В.С. Многопозиционная радиолокация. М.: Радио и связь, 1993.
2. Куприянов А.И., Петренко П.Б., Сычев М.П. Теоретические основы радиоэлектронной разведки. М.: Изд-во МГТУ им. Н.Э. Баумана, 2010.
3. Бердышев В.П., Гарин Е.Н., Фомин А.Н. Радиолокационные системы. Красноярск: Изд-во Сибир. федер. ун-та, 2021.
4. Верба В.С., Татарский Б.Г. Основы теории радиолокационных систем и комплексов. М.: Техносфера, 2024.
5. Zekavat S., Buehrer R. Handbook of Position Location: Theory Practice and Advances. Second ed. Hoboken. New Jersey: Wiley-IEEE Press, 2019.
6. Репин В.Г., Тартаковский Г.П. Статистический синтез при априорной неопределенности и адаптация информационных систем. М.: Сов. радио, 1977.
7. Жданюк Б.Ф. Основы статистической обработки траекторных измерений. М.: Сов. радио, 1978.
8. Фомин А.Ф., Новоселов О.Н., Плющев А.В. Отбраковка аномальных результатов измерений. М.: Энергоатомиздат, 1985.
9. Бульчев Ю.Г., Бульчев В.Ю., Ивакина С.С. и др. Обоснование методов оптимального оценивания параметров движения цели в триангуляционной измерительной системе // ТиСУ. 2015. № 4. С. 94–110.
10. Бульчев Ю.Г., Бульчев В.Ю., Ивакина С.С. и др. Классификация инвариантов пассивной локации и их применение // ТиСУ. 2015. № 6. С. 133–143.
11. Wang X., Wang A., Wang D., et al. A modified Sage-Husa adaptive Kalman filter for state estimation of electric vehicle servo control system // Energy Reports. 2022. V. 8. No. 5. P. 20–27.
12. Иванов А.В., Шишкин В.Ю., Бойков Д.В. и др. Адаптивные алгоритмы обработки информации в навигационных комплексах подвижных наземных объектов // РЭ. 2021. Т. 66. № 8. С. 760–771.
13. Калянов А.А., Лукин О.В., Цыганова Ю.В. Об алгоритме адаптивной фильтрации параметров движения объекта // Автоматизация процессов управления. 2023. № 1 (71). С. 75–87.
14. Peng L., Wenhui W., Junda Q., et al. Robust Generalized Labeled MultiBernoulli Filter and Smoother for Multiple Target Tracking using Variational Bayesian // KSII Transactions on Internet and Information Systems. 2022. V. 16. No. 3. P. 908–928.
15. Gao G., Gao B., Gao S., et al. A Hypothesis Test-Constrained Robust Kalman Filter for INS/GNSS Integration with Abnormal Measurement // IEEE Transactions on Vehicular Technology. 2023. V. 72. No. 2. P. 1662–1673.
16. Граничин О.Н., Поляк Б.Т. Рандомизированные алгоритмы оценивания и оптимизации при почти произвольных помехах. М.: Наука, 2003.
17. Граничин О.Н., Шлымов Д.С., Аврос Р. и др. Рандомизированный алгоритм нахождения количества кластеров // АиТ. 2011. № 4. С. 86–98.
18. Бульчев Ю.Г., Чепель Е.Н. Оптимизация кластерно-вариационного метода построения многопозиционной пеленгационной системы для условий априорной неопределенности // АиТ. 2023. № 4. С. 96–114.
19. Татузов А.Л. Нейронные сети в задачах радиолокации. М.: Радиотехника, 2009.

20. Мансур М.Э., Степанов О.А. Алгоритмы комплексной обработки в задаче коррекции показаний навигационных систем при наличии нелинейных измерений // Изв. Тульского ГУ. Технические науки. 2016. № 6. С. 89–102.
21. Гаджиев Ч.М. Метод отбраковки аномальных измерений для многомерных динамических систем // Автометрия. 2003. Т. 39. № 4. С. 39–46.
22. Шэнь К., Шахтарин Б.И., Неусыпкин Б.И. и др. Алгоритмические методы коррекции навигационной информации с использованием спутниковой радионавигационной системы в условиях аномальных измерений // РЭ. 2019. Т. 64. № 1. С. 31–37.
23. Gu T., Luo Z., Guo T., et al. New Reconstruction Method for Measurement Data with Multiple Outliers // IEEE Transact. Instrument. Measurement. 2022. V. 71. P. 1–9.
24. Ji C., Song C., Li S., et al. An Online Combined Compensation Method of Geomagnetic Measurement Error // IEEE Sensor. J. 2022. V. 22. No. 14. P. 14026–14037.
25. Иванов А.В., Шишкин В.Ю., Бойков Д.В. и др. Адаптивные алгоритмы обработки информации в навигационных комплексах подвижных наземных объектов // РЭ. 2021. Т. 66. № 8. С. 760–771.
26. Мандель И.Д. Кластерный анализ. М.: Финансы и статистика, 1988.
27. Паклин Н.Б., Орешков В.И. Кластерные силуэты / Системный анализ в проектировании и управлении: Сб. тр. XX Междунар. научно-практ. конф.; Санкт-Петербург, 29 июня–1 июля, СПб., 2016. С. 314–321.
28. Булычев Ю.Г., Чепель Е.Н. Мультиструктурный метод триангуляционного оценивания параметров движения излучающей цели в условиях априорной неопределенности // ТиСУ. 2019. № 6. С. 26–42.
29. Ежова Н.А., Соколинский Л.Б. Обзор моделей параллельных вычислений // Вестник ЮУрГУ. Серия «Вычислительная математика и информатика». 2019. Т. 8. № 3. С. 58–91.
30. Иванов А.И., Штилевая С.Г. О квантовых параллельных вычислениях // Вестник Балт. ун-та им. Канта. Серия «Физико-математические и технические науки». 2021. № 2. С. 95–99.

Статья представлена к публикации членом редколлегии Н.Н. Бахтадзе.

Поступила в редакцию 14.07.2024

После доработки 25.02.2025

Принята к публикации 27.02.2025

Управление в технических системах

© 2025 г. В.В. ХУТОРЦЕВ, д-р техн. наук (hvv.56@mail.ru)
(Ростовский-на-Дону научно-исследовательский институт радиосвязи)

ПРОГРАММНОЕ УПРАВЛЕНИЕ ПОИСКОМ ОБЪЕКТОВ В УСЛОВИЯХ ОГРАНИЧЕНИЙ НА КОЛИЧЕСТВО ЛИНИЙ ОБРАБОТКИ В НЕЗАВИСИМЫХ КАНАЛАХ ПОИСКОВОЙ СИСТЕМЫ

Рассмотрена задача поиска объектов наблюдения, математическая модель появления которых удовлетворяет закономерностям пуассоновского потока. Ее решение получено для случая, когда количество линий обработки в каналах многоканальной поисковой системы ограничено. В условиях высокой интенсивности появления объектов из потока это может привести к появлению очередей на их обработку. В качестве модели, используемой для определения закона поиска, рассмотрен набор систем массового обслуживания с независимыми входящими потоками, для которых системы дифференциальных уравнений Колмогорова описывают динамику изменения вероятностных характеристик их состояний в условиях ограничения на количество линий обработки в поисковых каналах. Рассмотрены две разновидности формулировки оптимизационной задачи – для вероятностного и временного критериев качества. Синтезирована итерационная процедура для формирования законов распределения интенсивностей поисковых усилий в каналах поисковой системы. Приведены примеры.

Ключевые слова: поиск объектов наблюдения, многоканальная поисковая система, распределение интенсивностей поиска, система дифференциальных уравнений Колмогорова.

DOI: 10.31857/S0005231025070028, **EDN:** JRJAYX

1. Введение

Задачи рационального распределения поисковых усилий в поисковых системах (ПС) [1–13] образуют один из важных классов задач общей теории управления наблюдениями [14–17]. Их эффективное решение обеспечивает уменьшение вероятности наличия необнаруженных объектов наблюдения (ОН) в зоне обзора ПС, сокращение времени их поиска, повышение достоверности обнаружения ОН и т.д. В частности, к таким задачам могут быть отнесены поиск и обнаружение космического мусора, поток частиц которого может рассматриваться как пуассоновский поток [18], мониторинг автомобильного трафика на многополосных автострадах, контроль воздушного движения в районах крупных аэропортов, мониторинг использования воздушного пространства мегаполисов и т.д.

Закономерности решения указанного класса задач могут быть использованы при управлении другими системами и сетями массового обслуживания, включая управление водными ресурсами, энергосетями, процессами передачи данных в телекоммуникационных системах и т.д. [19–26].

Наиболее общим предположением относительно структуры математической модели, описывающей ОН в процессе поиска, является предположение о произвольном (в том числе неограниченном) их количестве и появлении в области обзора в соответствии с закономерностями пространственно-временного случайного потока [27–30], как правило пуассоновского.

Исследованию задачи поиска в последнем случае посвящены работы [6–10]. Их общей чертой является отсутствие ограничений на количество линий обработки в каналах ПС. При определенных условиях это позволяло существенно упростить математические модели, лежащие в основе формирования законов распределения поисковых усилий. Однако для реальных ПС, в каналах которых количество линий обработки конечно, рассмотренные подходы к формированию законов распределения интенсивностей поисковых усилий использованы быть не могут. Это обусловлено особенностями математических моделей (уравнений Колмогорова), описывающих вариацию во времени вероятностных характеристик состояний каналов ПС в терминах задачи обслуживания в условиях ограничений на количество линий обработки в каждом из каналов [31–34].

В связи с этим актуальным становится вопрос определения управления распределением поисковых усилий для многоканальных ПС параллельного типа с независимыми каналами, у которых количество линий обработки в каждом канале ограничено.

2. Анализ структуры математической модели поиска

Пусть объекты наблюдения появляются в области обзора X поисковой системы в соответствии с закономерностями пространственно-временного пуассоновского потока φ [19–22]. Здесь $X \subset R^n$ ($1 \leq n \leq 3$), (x_1, \dots, x_n) – декартова система координат в X .

Предположим, что область обзора X разбита на I подобластей X_i , $i = \overline{1, I}$, каждую из которых обслуживает соответствующий i -й канал ПС.

Будем полагать, что $X = \bigcup_i X_i$, $X_i \cap X_j = \emptyset$, $i = \overline{1, I}$, $j = \overline{1, I}$, $i \neq j$. Тогда любые два потока, определяемые из φ как

$$(2.1) \quad \varphi_i(t) = \varphi(X_i, t), \quad \varphi_j(t) = \varphi(X_j, t), \quad i = \overline{1, I}, \quad j = \overline{1, I}, \quad i \neq j$$

являются пуассоновскими и независимыми [27, 30].

В соответствии с [8, 9, 27, 34, 35] меры интенсивностей, или интенсивности временных пуассоновских потоков $\varphi_i(t)$, $i = \overline{1, I}$, соответственно $\xi_i(t)$, являются известными непрерывными и ограниченными детерминированными

функциями и характеризуют временные пуассоновские потоки $\varphi_i(t)$, $i = \overline{1, I}$, как потоки с переменными параметрами [27–29].

Для подобласти X_i вероятность появления одного очередного ОН в течение интервала времени $[t, t + \Delta t]$ определяется соотношением [32, 33]

$$(2.2) \quad \xi_i(t)\Delta t + o(\Delta t),$$

тогда как вероятность появления более одного ОН равна $o(\Delta t)$. Здесь $o(\Delta t)$ – остаток порядка малости выше, чем Δt .

Введем ограничение: для i -го канала ПС, $i = \overline{1, I}$, количество линий обработки ограничено и равно a_i .

Определим интервал времени поиска как $\Omega = [0, \bar{t}]$ и предположим, что в течение этого интервала i -й канал ПС обеспечивает интенсивность поиска $\lambda_i(t) \geq 0$.

Тогда в случае, когда в подобласти X_i присутствует k ОН, вероятность того, что в течение времени $[t, t + \Delta t]$ хотя бы один из них будет найден, составляет [8–10, 32, 33]

$$(2.3) \quad \begin{aligned} k\lambda_i(t)\Delta t + o(\Delta t), & \quad k \leq a_i, \\ a_i\lambda_i(t)\Delta t + o(\Delta t), & \quad k > a_i. \end{aligned}$$

Иными словами при $k > a_i$ вероятность обнаружения ОН на $[t, t + \Delta t]$ с увеличением значений k перестает расти.

Интенсивности поиска $\lambda_i(t)$ в подобластях X_i , $i = \overline{1, I}$, полагаются неизвестными неотрицательными непрерывными ограниченными на $[0, \bar{t}]$ функциями времени, подлежащими определению в результате решения оптимизационной задачи.

Задача поисковой системы применительно к подобласти X_i может быть интерпретирована как задача ее обслуживания при потоке ОН, образующем нагрузку пуассоновского типа с мерой интенсивности $\xi_i(t)$. Для описания задач этого класса, как правило, используется математический аппарат процессов размножения и гибели [31–33]. Для описания закономерностей изменения характеристик таких процессов служат бесконечномерные системы линейных обыкновенных дифференциальных уравнений Колмогорова относительно вероятностей пребывания процесса в каждом из возможных состояний, образующие совокупность задач Коши [32, 33, 36]:

$$(2.4) \quad \begin{aligned} \dot{P}_{i0} &= -\xi_i(t)P_{i0} + \lambda_i(t)P_{i1}, \\ \dot{P}_{ik} &= -(\xi_i(t) + k\lambda_i(t))P_{ik} + \xi_i(t)P_{ik-1} + (k+1)\lambda_i(t)P_{ik+1}, \\ & \quad (k = 1, \dots, a_i - 1), \\ \dot{P}_{ik} &= -(\xi_i(t) + a_i\lambda_i(t))P_{ik} + \xi_i(t)P_{ik-1} + a_i\lambda_i(t)P_{ik+1}, \quad k \geq a_i, \\ P_{i0}(0) &= 1, \quad P_{ik}(0) = 0, \quad i = \overline{1, I}, \quad t \in [0, \bar{t}], \end{aligned}$$

где P_{i0} – вероятность отсутствия необнаруженных ОН в X_i или вероятность события, связанного с незадействованием ни одной из линий обработки i -го канала ПС при $t \in \Omega$; P_{ik} – вероятность нахождения в X_i k необнаруженных ОН или вероятность события, связанного с задействованием k линий обработки i -го канала ПС, если $1 \leq k < a_i$, или a_i линий обработки, если $k \geq a_i$; функции $\{\lambda_i(t), \xi_i(t), i = \overline{1, I}\}$ на интервале Ω полагаются неотрицательными, непрерывными и ограниченными.

Задача Коши (2.4) отражает физический смысл задачи поиска, проводимого независимыми каналами ПС с конечным числом линий обработки и осуществляемого последовательно во времени по мере появления ОН в подобластях обзора $X_i, i = \overline{1, I}$.

Использование математических моделей (2.4) для определения закона управления распределением поисковых усилий в ПС является затруднительным в силу их бесконечной размерности. Кроме того, структуры указанных моделей не позволяют провести их свертку, например в терминах математических ожиданий количества необнаруженных ОН в каждом из каналов ПС, как, например, это было сделано в [10].

Под элементами обслуживания i -го канала будем понимать его линии обработки. Тогда очереди на обработку поисковой информации в i -м канале не возникают, если

- не занят ни один из его элементов обслуживания или не включена в обработку ни одна из линий обработки, что соответствует $k = 0$;
- заняты k элементов обслуживания ($k = 1, \dots, a_i$) или включены в обработку k линий обработки.

Указанные состояния i -го канала можно представить множеством

$$(2.5) \quad S_i = \left\{ s_{i0}, s_{i1}, \dots, s_{ik} \Big|_{k=a_i} \right\}, \quad i = \overline{1, I},$$

где каждое из подмножеств s_{ik} ($k = 0, 2, \dots, a_i$) включает $C_{a_i}^k = \frac{a_i!}{k!(a_i-k)!}$ элементов.

3. Постановка задачи

Обозначим через $p_i, i = \overline{1, I}$, вероятность события, связанного с однократным выходом i -го канала ПС из соответствующего множества (2.5) при условии, что в начальный момент времени количество занятых элементов обслуживания соответствовало условию $0 \leq k \leq a_i$. Отметим, что под однократным выходом i -го канала ПС ($i = \overline{1, I}$) из соответствующего множества состояний (2.5) подразумевается выход из S_i ровно один раз.

Такие события являются совместными и независимыми, а вероятность их суммы определяется соотношением [37]

$$(3.1) \quad p_{\Sigma}(t) = p_1(t) + p_2(t)(1 - p_1(t)) + \dots + p_I(t) \prod_{i=1}^{I-1} (1 - p_i(t)).$$

Определим первый критерий качества функционирования поисковой системы соотношением

$$(3.2) \quad \Upsilon_1 = p_{\Sigma}(\bar{t}) + \eta \int_0^{\bar{t}} \lambda(t)^T \lambda(t) dt \rightarrow \min_{\lambda \in \Lambda},$$

где $\lambda(t) = [\lambda_1(t) \dots \lambda_I(t)]^T$ – вектор интенсивностей поиска; $\eta \in R^1$, $\eta > 0$ – весовой коэффициент; Λ – множество допустимых законов управления или законов распределения интенсивностей поиска.

Под множеством допустимых законов управления понимаются множество неотрицательных непрерывных и ограниченных на $[0, \bar{t}]$ функций времени.

Критерий (3.2) определим как вероятностный. Он предполагает минимизацию суммы двух компонент. Первая компонента характеризует значение в конечный момент времени наблюдения \bar{t} вероятности выхода состояний хотя бы одного из поисковых каналов ПС из соответствующего множества (2.5), вторая – эквивалентна аналогу энергетических затрат ПС на проведение поиска.

Рассмотрим второй критерий качества. Для этого введем обозначения для математических ожиданий длительностей пребывания каналов ПС в множествах состояний, исключающих возникновение очередей на обработку поисковой информации: $m_i = M[\tau_i]$, $i = \overline{1, I}$, где τ_i – случайная величина, определяющая длительность пребывания i -го канала ПС в соответствующем множестве состояний (2.5). Структуру второго критерия качества зададим соотношением

$$(3.3) \quad \Upsilon_2 = I\bar{t} - \sum_{i=1}^I m_i + \eta \int_0^{\bar{t}} \lambda(t)^T \lambda(t) dt \rightarrow \min_{\lambda \in \Lambda}.$$

Критерий (3.3) определим как временной. Он предполагает минимизацию в конечный момент наблюдения суммы двух компонент. Первая компонента характеризует сумму математических ожиданий длительностей пребывания каналов ПС в соответствующих множествах состояний (2.5), взятую со знаком минус, вторая, как и в (3.2), эквивалентна аналогу энергетических затрат ПС на проведение поиска.

Поставим задачу провести синтез законов управлений распределениями интенсивностей поисковых усилий $\lambda_i(t)$, $i = \overline{1, I}$, $t \in \Omega$ в (2.4) между каналами многоканальной поисковой системы с конечным количеством линий обработки в каждом канале, обслуживающими непересекающиеся подобласти X_i , $i = \overline{1, I}$, области обзора ПС X , для вероятностного (3.2) и временного (3.3) критериев качества.

4. Синтез управления распределениями интенсивностей поисковых усилий в многоканальной ПС с конечным количеством линий обработки в каждом канале для вероятностного критерия качества

Рассмотрим процедуру синтеза управления поиском объектов наблюдения из пространственно-временного пуассоновского потока φ в многоканальной поисковой системе с независимыми каналами и конечным количеством линий обработки в каждом канале ПС применительно к вероятностному критерию качества (3.2). Для этого покажем справедливость следующего утверждения.

Утверждение 1. Вероятность события, связанного с однократным выходом i -го канала ПС ($i = \overline{1, I}$) из соответствующего множества S_i состояний (2.5) при условии, что при $t = 0$ состояния этого канала принадлежали S_i , определяется соотношением

$$(4.1) \quad p_i(t) = \bar{P}_{i a_i+1}(t),$$

где значение вероятности $\bar{P}_{i a_i+1}(t)$ определяется из решения системы дифференциальных уравнений

$$(4.2) \quad \begin{aligned} \dot{\bar{P}}_{i0}(t) &= -\xi_i(t)\bar{P}_{i0}(t) + \lambda_i(t)\bar{P}_{i1}(t), \\ \dot{\bar{P}}_{ik}(t) &= -(\xi_i(t) + k\lambda_i(t))\bar{P}_{ik}(t) + \xi_i(t)\bar{P}_{ik-1}(t) + (k+1)\lambda_i(t)\bar{P}_{ik+1}(t), \\ &\quad (k = 1, \dots, a_i - 1), \\ \dot{\bar{P}}_{ia_i}(t) &= -(\xi_i(t) + a_i\lambda_i(t))\bar{P}_{ia_i}(t) + \xi_i(t)\bar{P}_{ia_i-1}(t), \\ \dot{\bar{P}}_{ia_i+1}(t) &= \xi_i(t)\bar{P}_{ia_i}(t), \quad i = \overline{1, I}, \end{aligned}$$

полученного при начальных условиях

$$(4.3) \quad \sum_{k=0}^{a_i} \bar{P}_{ik}(0) = 1, \quad \bar{P}_{i a_i+1}(0) = 0.$$

Доказательство утверждения приведено в Приложении.

Необходимо отметить, что уравнения (4.2) образуют совокупность фиктивных динамических систем [15] с векторным управлением $\lambda = [\lambda_1 \lambda_2 \dots \lambda_I]^T$. Начальные условия (4.3) определяют принадлежность при $t = 0$ состояний i -го канала множеству S_i и соответствуют отсутствию очереди на его входах.

Таким образом, для (3.1) получим

$$(4.4) \quad \begin{aligned} p_\Sigma(\bar{t}) &= \bar{P}_{1 a_1+1}(\bar{t}) + \bar{P}_{2 a_2+1}(\bar{t})(1 - \bar{P}_{1 a_1+1}(\bar{t})) + \dots + \\ &\quad + \bar{P}_{I a_I+1}(\bar{t}) \prod_{i=1}^{I-1} (1 - \bar{P}_{i a_i+1}(\bar{t})). \end{aligned}$$

С учетом (3.2), (4.2) определим гамильтониан

$$(4.5) \quad H = \sum_{i=1}^I \psi_i^T A_i P_i + \eta \lambda^T \lambda,$$

где $\psi_i = \psi_i(t) \in R^{a_i+2}$ – вектор сопряженных переменных; $P_i(t) = [\bar{P}_{i0}(t) \bar{P}_{i1}(t) \dots \bar{P}_{i a_i+1}(t)]^T$; $A_i \in R^{(a_i+2) \times (a_i+2)}$; $A_i = \xi_i B_{i1} + \lambda_i B_{i2}$; структура матриц B_{i1} , B_{i2} , следующая из (4.2), определена в Приложении.

Специфической особенностью матриц B_{i1} , B_{i2} является равенство нулю сумм элементов каждого столбца.

Из (4.5) получим системы уравнений относительно $\psi_i(t)$, $i = \overline{1, I}$:

$$(4.6) \quad \dot{\psi}_i = -\frac{\partial}{\partial P_i} H = -(\xi_i B_{i1}^T + \lambda_i B_{i2}^T) \psi_i, \quad t \in \Omega.$$

С учетом (4.4) краевые условия для (4.6) можно представить в виде

$$(4.7) \quad \psi_i(\bar{t}) = \frac{\partial}{\partial P_i(\bar{t})} P_{\Sigma}(\bar{t}) = \begin{bmatrix} 0 & 0 & \dots & \prod_{j=1, j \neq i}^I (1 - \bar{P}_{j a_j}(\bar{t})) \end{bmatrix}^T, \quad i = \overline{1, I}.$$

Из условия минимума (4.5) по λ получим

$$(4.8) \quad \frac{\partial}{\partial \lambda_i} H = \psi_i^T B_{i2} P_i + 2\eta \lambda_i = 0, \quad i = \overline{1, I}.$$

С учетом (4.8) определим структуру оптимального управления

$$(4.9) \quad \lambda_i = \frac{-\psi_i^T B_{i2} P_i}{2\eta}, \quad i = \overline{1, I}.$$

Для решения двухточечной краевой задачи (ДТКЗ) (4.2), (4.6), (4.7) воспользуемся методом последовательных приближений Крылова–Черноустько [14, 15, 39]. Конкретизируя (4.3), будем полагать, что $\bar{P}_{i0}(0) = 1$.

Пусть на q -м шаге итерационной процедуры сформирован закон управления поиском (вектор интенсивностей поисковых усилий λ^q).

Метод последовательных приближений предполагает выполнение следующих операций.

1. Проводится решение уравнений (4.2) на интервале времени Ω в прямом времени и формируются значения векторов переменных P_i^q , $i = \overline{1, I}$, соответствующих распределению интенсивностей поиска λ_i^q , $i = \overline{1, I}$.

2. По полученным решениям в соответствии с (4.7) формируются конечные условия для сопряженных переменных $\psi_i^q(\bar{t})$, $i = \overline{1, I}$.

3. Проводится решение системы уравнений (4.6) на интервале времени Ω в обратном времени с конечными условиями (4.7) и определяются значения вектора сопряженных переменных $\psi_i^q(t)$, $t \in \Omega$, $i = \overline{1, I}$, соответствующие управлениям λ_i^q , $i = \overline{1, I}$, и вектору переменных P_i^q , $i = \overline{1, I}$.

4. По полученным значениям векторов $P_i^q(t)$, $\psi_i^q(t)$, $i = \overline{1, I}$ в соответствии с (4.9) рассчитывается промежуточное значение вектора управлений, соответствующее $(q+1)$ -му шагу итерационной процедуры

$$(4.10) \quad \tilde{\lambda}_i^{q+1} = \frac{-(\psi_i^q)^T B_{i2} P_i^q}{2\eta}, \quad i = \overline{1, I}, \quad t \in \Omega.$$

5. Исходя из принципа частичного обновления управления [5, 8, 28] с учетом значений распределения интенсивностей поисковых усилий λ_i^q , $i = \overline{1, I}$, полученных на предыдущем шаге, проводится расчет законов управления интенсивностями поиска для $(q + 1)$ -й итерации

$$(4.11) \quad \lambda_i^{q+1}(t) = \varepsilon^q \tilde{\lambda}_i^{q+1}(t) + (1 - \varepsilon^q) \lambda_i^q(t), \quad i = \overline{1, I}, \quad t \in \Omega,$$

где $\varepsilon^q \in (0, 1)$.

Параметр ε^q обозначает степень обновления законов управления поиском $\lambda_i^q(t)$, $i = \overline{1, I}$, $t \in \Omega$. Он определяется из условия минимума целевой функции критерия (3.2) на соответствующем шаге итерационной процедуры.

Далее в качестве исходного для $(q + 2)$ -го шага используется управление $\lambda_i^{q+1}(t)$, и итерационная процедура 1–5 повторяется.

Необходимо отметить, что для начального шага итерационной процедуры ($q = 0$) начальное распределение интенсивностей поиска λ_i^0 , $i = \overline{1, I}$, выбирается из множества Λ допустимых законов управления.

Оптимальное управление (оптимальный закон распределения интенсивностей поисковых усилий) определяется соотношением

$$(4.12) \quad \lambda_{i\text{оп}}(t) = \lim_{q \rightarrow \infty} \lambda_i^q(t), \quad i = \overline{1, I}, \quad t \in \Omega.$$

Отметим, что на каждом шаге итерационной процедуры частичная вариация управления (4.11), осуществляемая на основании решения оптимизационной задачи, направлена на уменьшение значений целевого функционала. Значения $\varepsilon^q \in (0, 1)$ в (4.11) выбираются из условия его максимального уменьшения. С другой стороны, в соответствии с (3.2) соответствующий этому критерию целевой функционал ограничен снизу (целевой функционал, по крайней мере, не может быть отрицательным). Таким образом, предел (4.12) существует.

На практике, как правило, ограничиваются конечным числом итераций $q \leq Q$, где Q – номер шага, после которого вариации целевого функционала критерия (3.2) становятся незначительными. При этом полагают, что $\lambda_{i\text{оп}}(t) \simeq \lambda_i^Q(t)$, $i = \overline{1, I}$.

5. Синтез управления распределениями интенсивностей поисковых усилий в многоканальной ПС с конечным количеством линий обработки в каждом канале для временного критерия качества

Пусть значение верхней границы \bar{t} интервала наблюдения Ω выбрано так, что для плотностей распределения $w_i(t)$, $i = \overline{1, I}$, длительностей τ_i пребывания каналов ПС в соответствующих множествах (2.5) выполняются условия

$$(5.1) \quad \int_0^{\bar{t}} w_i(t) dt = 1, \quad \forall i, i = \overline{1, I},$$

т.е. плотности распределения финитны на Ω .

С учетом (5.1) преобразуем критерий (3.3).

Утверждение 2. При выполнении условия (5.1) математическое ожидание длительности τ_i однократного пребывания i -го канала ПС во множестве S_i (2.5) с учетом начальных условий (4.3) определяется соотношением

$$(5.2) \quad m_i = \int_0^{\bar{t}} (1 - \bar{P}_{ia_i+1}(t)) dt.$$

Доказательство утверждения приведено в Приложении.

С учетом (5.2) преобразуем критерий качества (3.3). В результате получим

$$(5.3) \quad \Upsilon_2 = \int_0^{\bar{t}} \left[\eta \lambda(t)^T \lambda(t) + \sum_{i=1}^I \bar{P}_{ia_i+1}(t) \right] dt \rightarrow \min_{\lambda \in \Lambda}.$$

Математическим моделям (4.2), (5.3) соответствует гамильтониан

$$(5.4) \quad H = \sum_{i=1}^I (\psi_i^T A_i P_i + \bar{P}_{ia_i+1}(t)) + \eta \lambda^T \lambda.$$

Уравнения для сопряженных переменных определяются соотношениями

$$(5.5) \quad \dot{\psi}_i = -\frac{\partial}{\partial P_i} H = -(\xi_i B_{i1}^T + \lambda_i B_{i2}^T) \psi_i - U_i, \quad i = \overline{1, I}, \quad t \in \Omega,$$

где значения матриц B_{i1} , B_{i2} , как и в первом случае, определяются в соответствии с (П.2); $U_i \in R^{a_i+2}$, $U_i = [0 \ 0 \ \dots \ 0 \ 1]^T$.

В соответствии с (5.3) краевые условия для (5.5) имеют вид $\psi_i(\bar{t}) = 0$, $i = \overline{1, I}$, а согласно (5.4) структура формируемых управлений совпадает с (4.9).

Для решения ДТКЗ (4.2), (5.5), как и для вероятностного критерия, необходимо использовать метод последовательных приближений Крылова–Черноусько. Порядок его применения аналогичен порядку, рассмотренному в предыдущем разделе.

Структура математических моделей в рассматриваемом случае позволяет проводить поканальную декомпозицию вычислительной процедуры оптимизации.

Необходимо отметить, что в случае, когда длительность интервала наблюдения не согласована со структурой хотя бы одной из плотностей распределения $w_i(t)$, $i = \overline{1, I}$ (т.е. равенство (5.1) выполняется лишь приближенно), получаемый результат переходит в категорию субоптимальных.

**6. Примеры синтеза управления распределениями
интенсивностей поисковых усилий в многоканальной ПС
с конечным числом линий обработки в каждом канале**

Пример 1. Рассмотрим двухканальную ПС ($I = 2$), у которой в первом канале три линии обработки, а во втором – четыре. Фиктивные динамические системы, соответствующие такой ПС, описываются математической моделью, состоящей из двух систем уравнений (4.2), для которых соответственно $a_1 = 3$ и $a_2 = 4$. Критерий качества (3.2) в этом случае приобретает вид

$$(6.1) \quad \Upsilon_1 = \bar{P}_{14}(\bar{t}) + \bar{P}_{25}(\bar{t}) - \bar{P}_{14}(\bar{t})\bar{P}_{25}(\bar{t}) + \eta \int_0^{\bar{t}} (\lambda_1^2(t) + \lambda_2^2(t)) dt \rightarrow \min_{\lambda_1, \lambda_2 \in \Lambda}.$$

Решение прямых систем уравнений (4.2) при $i = 1, 2$ проводилось с начальными условиями, соответствующими (4.3),

$$\begin{aligned} \bar{P}_{10}(0) &= \bar{P}_{20}(0) = 1, \\ \bar{P}_{11}(0) &= \bar{P}_{12}(0) = \bar{P}_{13}(0) = \bar{P}_{14}(0) = 0, \\ \bar{P}_{21}(0) &= \bar{P}_{22}(0) = \bar{P}_{23}(0) = \bar{P}_{24}(0) = \bar{P}_{25}(0) = 0 \end{aligned}$$

на интервале времени поиска $\Omega = [0, 5]$. Здесь и далее параметры представлены в безразмерных переменных.

Сопряженные системы (4.6) полностью определяются структурами матриц (П.2), для которых

$$(6.2) \quad B_{11}, B_{12} \in R^{5 \times 5}; \quad B_{21}, B_{22} \in R^{6 \times 6}.$$

Их решение проводилось в обратном времени с конечными условиями, соответствующими (4.7)

$$(6.3) \quad \begin{aligned} \psi_1(\bar{t}) &= [0 \ 0 \ 0 \ 0 \ [1 - \bar{P}_{25}(\bar{t})]]^T, \\ \psi_2(\bar{t}) &= [0 \ 0 \ 0 \ 0 \ 0 \ [1 - \bar{P}_{14}(\bar{t})]]^T. \end{aligned}$$

Интенсивности появления ОН в подобластях X_1, X_2 зоны поиска ПС полагались линейными функциями времени: $\xi_1(t) = 2 + 0,4t$, $\xi_2(t) = 3 - 0,1t$ (прямые соответственно 1, 2 на рис. 1).

Начальные распределения интенсивностей поисковых усилий λ_1^0, λ_2^0 полагались равномерными на интервале наблюдения (прямые 1, 2 на рис. 2). Решения уравнений на интервале времени поиска проводилось с шагом $\Delta = 10^{-3}$.

Зависимость вероятности $p_\Sigma(\bar{t})$ выхода состояний хотя бы одного из двух поисковых каналов ПС из множеств состояний соответственно

$$(6.4) \quad S_1 = \{s_{10}, s_{11}, \dots, s_{13}\}, \quad S_2 = \{s_{20}, s_{21}, \dots, s_{24}\}$$

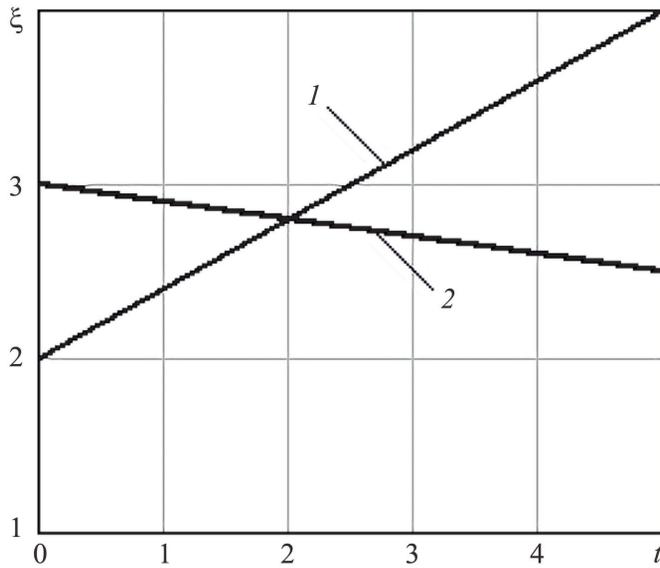


Рис. 1

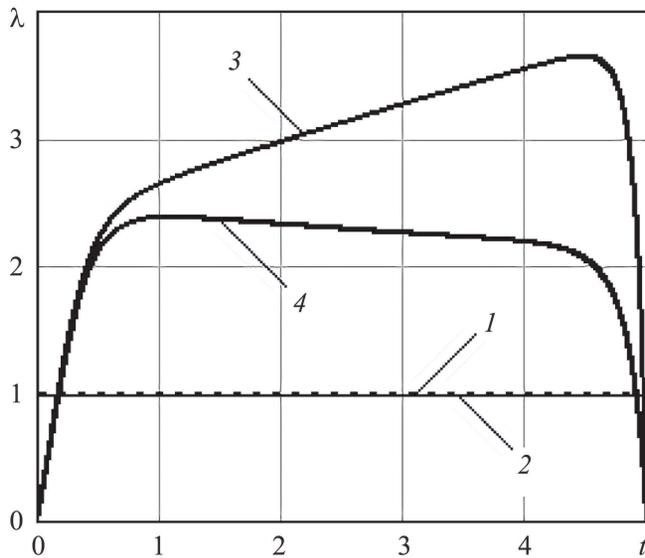


Рис. 2

от номера итерации q в процедуре последовательных приближений Крылова–Черноусько представлена на рис. 3.

Вариации вероятностной составляющей целевого функционала критерия (6.1) становятся незначительными уже при $q \geq Q = 7$. Структура оптимальных законов управления распределением интенсивностей поисковых усилий представлена на рис. 2. Для первого канала – кривая 3, для второго – кривая 4.

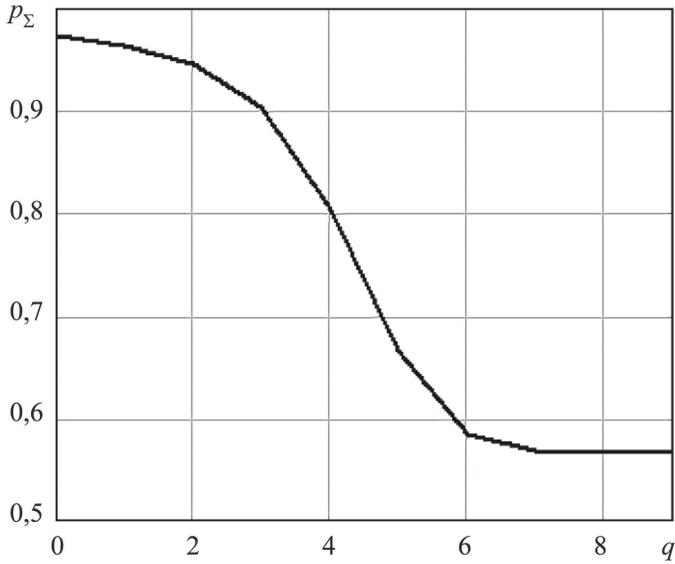


Рис. 3

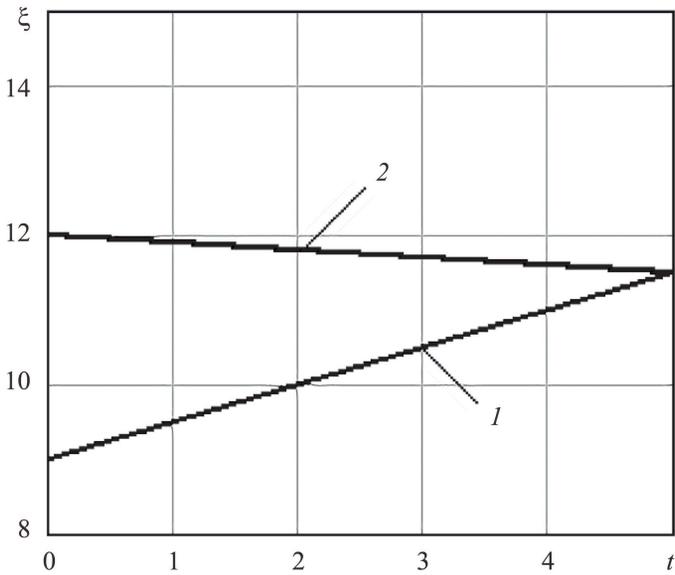


Рис. 4

Необходимо отметить, что характер полученных зависимостей согласуется с характером вариации интенсивностей появления ОН из потока (рис. 1) в соответствующих каналах ПС.

Относительный выигрыш от оптимизации применительно к вероятностному показателю составил $\delta = \frac{p_{\Sigma 0}(\bar{t}) - p_{\Sigma \text{оп}}(\bar{t})}{p_{\Sigma 0}(\bar{t})} \simeq 0,42$, где $p_{\Sigma 0}(\bar{t})$ – вероятность выхода состояний хотя бы одного из двух поисковых каналов ПС из мно-

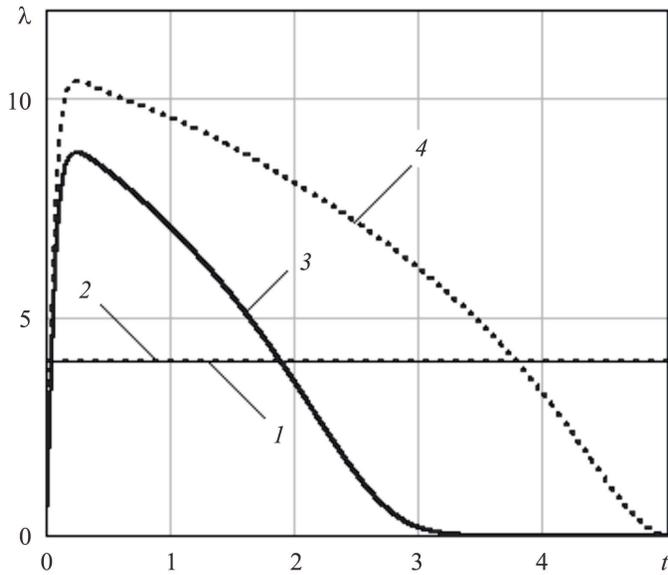


Рис. 5

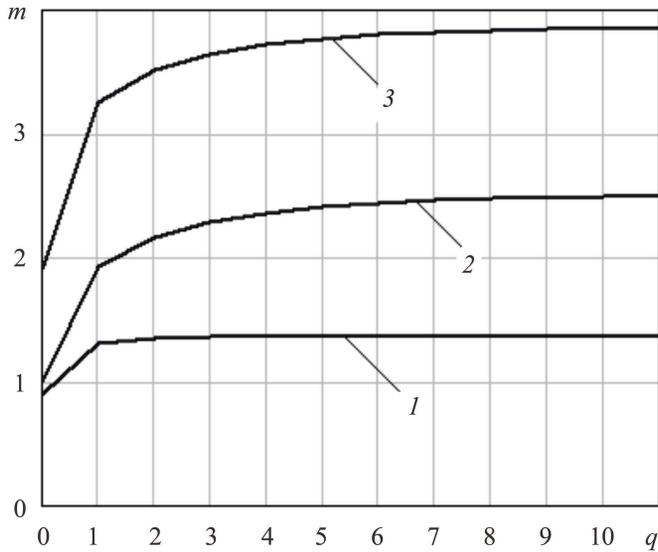


Рис. 6

жеств (6.4) при $\lambda_1^0, \lambda_2^0, p_{\Sigma\text{ОП}}(\bar{t})$ – аналогичный показатель при $\lambda_i^Q(t) \simeq \lambda_{i\text{ОП}}(t), i = 1, 2$.

Пример 2. Рассмотрим задачу синтеза управления распределениями интенсивностей поисковых усилий для рассмотренной в первом примере двухканальной поисковой системы на основе временного критерия качества (3.3)

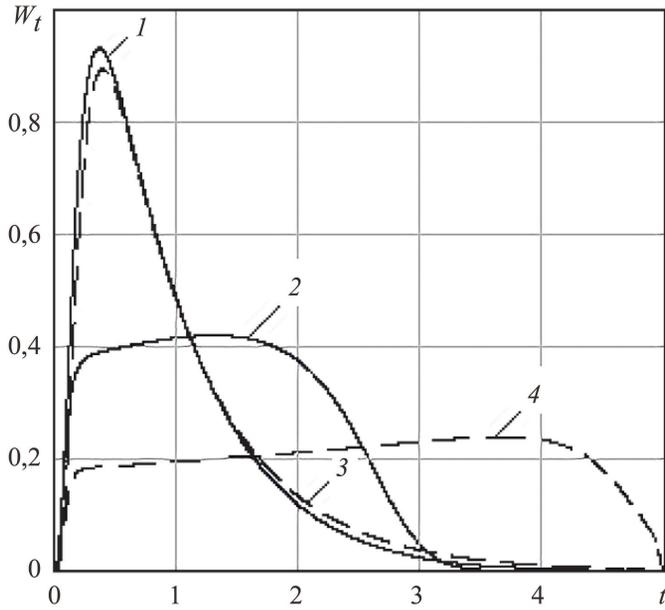


Рис. 7

в форме (5.3). Его конкретизация для рассматриваемого случая имеет вид

$$(6.5) \quad \Upsilon_2 = \int_0^{\bar{t}} [\eta(\lambda_1^2(t) + \lambda_2^2(t)) + \bar{P}_{14}(t) + \bar{P}_{25}(t)] dt \rightarrow \min_{\lambda_1, \lambda_2 \in \Lambda}.$$

Прямые системы уравнений остаются неизменными и определяются формой (4.2). Системы уравнения для сопряженных переменных соответствуют (5.5) с матрицами (6.2), векторами $U_1 \in R^5$, $U_2 \in R^6$ и нулевыми конечными условиями $\psi_i(\bar{t}) = 0$, $i = 1, 2$.

Интенсивности появления ОН в X_1 , X_2 полагались линейными функциями времени: $\xi_1(t) = 9 + 0,5t$, $\xi_2(t) = 12 - 0,1t$ (прямые соответственно 1, 2 на рис. 4).

Начальные распределения интенсивностей поисковых усилий λ_1^0 , λ_2^0 полагались равномерными на интервале наблюдения (прямые 1, 2 на рис. 5).

Графики зависимостей математических ожиданий m_1 (кривая 1) и m_2 (кривая 2) времен нахождения каналов ПС во множествах (6.4) и их суммы (кривая 3) от номера итерации q в процедуре последовательных приближений Крылова–Черноушко представлены на рис. 6. Изменения временных характеристик стали незначительными при $q \geq Q = 9$.

Меньший рост m_1 по сравнению с m_2 связан с меньшим количеством состояний во множестве S_1 по сравнению со множеством S_2 .

Зависимости оптимальных законов управления распределением интенсивностей поисковых усилий представлены на рис. 5 соответственно для первого

(кривая 3) и второго (кривая 4) каналов ПС. Меньшая интенсивность поиска в первом канале обусловлена меньшей интенсивностью появления в нем ОН (рис. 4).

Структура плотностей распределения вероятностей времени пребывания каналов ПС во множествах (6.4) представлены на рис. 7: для $t = 0$ кривая 1 и кривая 3 соответственно для первого и второго каналов и аналогично для $t = 5$ кривая 2 и кривая 4.

В результате оптимизации в условиях примера суммарное математическое ожидание нахождения каналов ПС во множествах (6.4), исключающих появление очередей на обработку, возросло более чем в 2 раза.

С учетом структуры рассмотренных математических моделей несложно показать, что формируемые в результате решения оптимизационных задач интенсивности поисковых усилий как для критерия (3.2), так и для критерия (3.3) при выполнении условий утверждения 1 всегда относятся к классу неотрицательных непрерывных и ограниченных функций.

7. Заключение

Рассмотренный подход к формированию законов управления поиском ОН из пространственно-временного пуассоновского потока в многоканальной ПС ориентирован на ряд практически важных случаев, когда количество линий обработки в каналах поисковой системы конечно. Для них применение методов, связанных со сверткой бесконечной системы уравнений Колмогорова, например, в терминах математического ожидания количества необнаруженных ОН, рассмотренных, в частности, в [10], не представляется возможным.

В основе предлагаемого подхода лежит переход от бесконечномерных систем дифференциальных уравнений, описывающих вероятностные характеристики состояний каналов ПС, к конечномерным вспомогательным системам. Их размерность определяется количеством линий обработки каждого канала.

Параметрами вспомогательных систем уравнений являются вероятностные характеристика пребывания линий обработки каналов ПС во множествах состояний, не приводящих к появлению очередей на обработку. Они дают возможность подойти к решению задачи синтеза управления распределением поисковых усилий как с точки зрения минимизации вероятности выхода из рассмотренных множеств состояний хотя бы одного канала ПС, так и с точки зрения максимизации математических ожиданий времени пребывания каналов в этих множествах.

Решение оптимизационной задачи первого типа для вероятностного критерия качества (3.2) является более сложным в силу структуры терминальной составляющей этого критерия. Ее размерность характеризуется суммарной размерностью всех вспомогательных систем уравнений (4.2) и определяется соотношением $\sum_{i=1}^I a_i + 2I$.

Во втором случае возможна поканальная декомпозиция общей процедуры формирования законов управления интенсивностями поиска.

В силу достаточно высокого уровня сложности возникающих оптимизационных задач их решение целесообразно проводить с использованием метода последовательных приближений Крылова–Черноузько с применением принципа частичного обновления управления. Приведенные примеры показали возможность эффективного их решения с помощью рассмотренного подхода и проиллюстрировали выигрыш, получаемый от оптимизации.

Необходимо отметить, что рассмотренный подход может быть обобщен на случай, когда интенсивности поиска ОН в различных каналах ПС не являются независимыми и удовлетворяют ограничению $\sum_{i=1}^I \lambda_i(t) = \lambda_\Sigma = \text{const}$.

ПРИЛОЖЕНИЕ

Доказательство утверждения 1

Поставим в соответствие системе уравнений (2.4) граф состояний [38], представленный на рис. 8.

Граф отражает последовательность вариации состояний i -го канала с ограниченным числом линий обработки. Выделим на нем состояния, соответствующие множеству

$$(П.1) \quad S_i = \{s_{i0}, s_{i1}, \dots, s_{ik}|_{k=a_i}\} \rightarrow \{\bar{s}_{i0}, \bar{s}_{i1}, b, \dots, \bar{s}_{ik}|_{k=a_i}\} = \bar{S}_i.$$

Для (П.1) построим вспомогательный граф состояний (рис. 9) как фрагмент исходного графа (рис. 8) с добавлением одного только поглощающего состояния \bar{s}_{ia_i+1} .

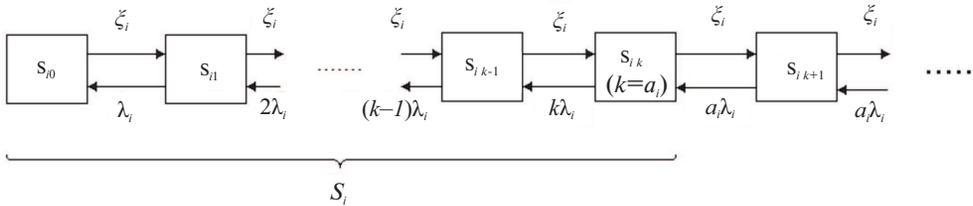


Рис. 8

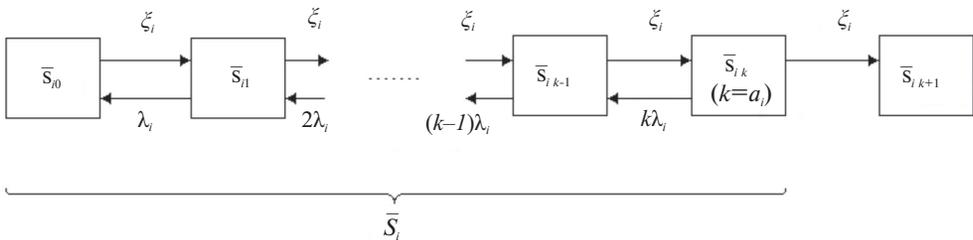


Рис. 9

Можно показать [38], что вариация во времени вероятностных характеристик $\bar{P}_{i0}(t), \bar{P}_{i1}(t), \dots, \bar{P}_{ia_i}(t), \bar{P}_{ia_i+1}(t)$ состояний $\{\bar{s}_{i0}, \bar{s}_{i1}, \dots, \bar{s}_{ik}|_{k=a_i}, \bar{s}_{ia_i+1}\}$, соответствующих графу, изображенному на рис. 9, описывается системой дифференциальных уравнений (4.2). Для начальных условий (4.3) вероятность события, связанного с однократным выходом из множества \bar{S}_i (соответственно и из множества S_i), определяется вероятностью перехода из состояния $\bar{s}_{ik}|_{k=a_i}$ в \bar{s}_{ia_i+1} , т.е. соответствует составляющей решения системы (4.2), определяемой вероятностью $\bar{P}_{ia_i+1}(t)$.

Утверждение доказано.

Структура матриц $B_{i1}, B_{i2} \in R^{(a_i+2) \times (a_i+2)}$

$$(II.2) \quad B_{i1} = \begin{bmatrix} -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & -1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & -1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

$$B_{i2} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 2 & 0 & \dots & 0 & 0 \\ 0 & 0 & -2 & 3 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & a_i & 0 \\ 0 & 0 & 0 & 0 & \dots & -a_i & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

Доказательство утверждения 2

Запишем последнее уравнение вспомогательной системы уравнений (4.2), соответствующей i -му каналу

$$(II.3) \quad \dot{\bar{P}}_{ia_i+1}(t) = \xi_i(t)\bar{P}_{ia_i}(t), \quad \bar{P}_{ia_i+1}(0) = 0.$$

Поскольку параметр $\bar{P}_{ia_i+1}(t)$ имеет смысл вероятности события, связанного с выходом i -го канала ПС из множества S_i (2.5), то его производная (II.3) определяет плотность распределения $w_i(t)$ случайной длительности τ_i пребывания этого канала в S_i . Иными словами, [38]

$$(II.4) \quad w_i(t) = \xi_i(t)\bar{P}_{ia_i}(t).$$

Домножим теперь левую и правую части (II.3) на t и проинтегрируем на интервале наблюдения Ω . В результате с учетом (II.4) получим

$$(II.5) \quad \int_0^{\bar{t}} t \dot{\bar{P}}_{ia_i+1}(t) dt = \int_0^{\bar{t}} t \xi_i(t) \bar{P}_{ia_i}(t) dt = \int_0^{\bar{t}} t w_i(t) dt.$$

Из (П.5) следует

$$(П.6) \quad \int_0^{\bar{t}} t \dot{\bar{P}}_{ia_i+1}(t) dt = m_i,$$

где $m_i = \int_0^{\bar{t}} t w_i(t) dt$ – математическое ожидание случайной длительности τ_i .

Интегрируя по частям левую часть (П.6), получим

$$(П.7) \quad t \bar{P}_{ia_i+1}(t) \Big|_0^{\bar{t}} - \int_0^{\bar{t}} \bar{P}_{ia_i+1}(t) dt = m_i.$$

Или с учетом условия (5.1)

$$(П.8) \quad \int_0^{\bar{t}} (1 - \bar{P}_{ia_i+1}(t)) dt = m_i.$$

Утверждение доказано.

СПИСОК ЛИТЕРАТУРЫ

1. Хеллман О. Введение в теорию оптимального поиска. М.: Наука, 1985.
2. Альведе Р., Вегнер И. Задачи поиска. М.: Мир, 1985.
3. Абчук В.А., Суздаль В.Г. Поиск объектов. М.: Сов. радио, 1977.
4. Аркин В.И. Задача оптимального распределения поисковых усилий // Теория вероятностей и ее применения. 1964. Т. 9. № 1. С. 179–180.
5. Болдырихин Н.В., Хуторцев В.В. Управление наблюдениями за потоками случайных процессов // АИТ. 2006. № 12. С. 43–55.
Boldyrikhin N.V., Khutortsev V.V. Control of Observations Over Random Processes Fluxes // Autom. Remote Control. 2006. V. 67. No. 12. P. 1900–1912.
6. Хуторцев В.В., Баранов И.В. Оптимизация управления наблюдениями в задаче дискретного поиска для пуассоновской модели потока объектов наблюдения // Радиотехника. 2010. № 3. С. 20–24.
7. Баранов И.В., Хуторцев В.В. Текущая оптимизация поиска объектов для модели распределенного пуассоновского потока их появления // Изв. РАН. ТИСУ. 2011. № 6. С. 3–13.
8. Хуторцев В.В. Оптимизация последовательно-параллельного поиска объектов для модели распределенного пуассоновского потока их появления // Изв. РАН. ТИСУ. 2019. № 1. С. 31–41.
9. Хуторцев В.В. Управление поиском объектов из пространственно-временного пуассоновского потока при неоднородной области обзора информационно-измерительной системы // АИТ. 2021. № 9. С. 133–149.
Khutortsev V.V. Controlled Search for Targets Arriving According to a Spatio-Temporal Poisson Point Process by an Information Measurement System with Inhomogeneous Scope // Autom. Remote Control. 2021. V. 82. No. 9. P. 1568–1580.

10. *Хуторцев В.В.* Управление поиском объектов наблюдения из пространственно-временного пуассоновского потока в многоканальной поисковой системе // *АиТ.* 2023. № 1. С. 84–97.
Khutortsev V.V. Control of the Search for Observation Objects from a Spatio-Temporal Poisson Flow in a Multi-Channel Search System // *Autom. Remote Control.* 2023. V. 84. No. 2. P. 143–152.
11. *Попович В.В., Кожемьяка В.Н., Гребцов Г.Н.* Теория обнаружения и поиска подвижных объектов. СПб.: Наука, 2016.
12. *Snayd A., Stevenson M.* Modelling stopping criteria for search results using Poisson processes // arXiv:1909.06239v1 [cs.IR] 13 sep 2019.
13. *Филимонов А.Б., Филимонов Н.Б., Тихонов В.Ю.* Планирование операций в задачах пространственного поиска объектов // *Изв. ЮФУ. Технические науки.* 2017. № 1–2. С. 185–197.
14. *Чернуоусько Ф.Л., Колмановский В.Б.* Оптимальное управление при случайных возмущениях. М.: Наука, 1980.
15. *Мальшиев В.В., Красильщиков М.Н., Карлов В.И.* Оптимизация наблюдения и управления летательных аппаратов. М.: Машиностроение, 1989.
16. *Григорьев Ф.Н., Кузнецов Н.А., Серебровский А.П.* Управление наблюдениями в автоматических системах. М.: Наука, 1986.
17. *Миллер В.М.* Обобщенная оптимизация в задачах управления наблюдениями // *АиТ.* 1991. № 10. С. 83–92.
Miller V.M. Generalized optimization in observation control problems // *Autom. Remote Control.* 1991. V. 52. No. 10. P. 1397–1404.
18. *Асланов В.С., Пироженко А.В., Волшенюк О.Л., Кислов А.В., Ящук А.В.* Определение времени выживания космической тросовой системы // *Изв. Самар. Науч. центра РАН.* 2010. Т. 12. № 4. С. 138–143.
19. *Мясников Д.В., Семенихин К.В.* Управление параметрами одноканальной системы массового обслуживания при наличии ограничений // *Изв. РАН. ТиСУ.* 2016. № 1. С. 66–85.
20. *Куракин С.З., Онуфрей А.Ю., Разумов А.В.* Исследование вариантов построения информационно управляющих систем на основе сетевых моделей систем массового обслуживания // *Информация и автоматизация.* 2024. Т. 6. Вып. 23. С. 1609–1642.
21. *Лихтциндер Б.Я., Моисеев В.И.* Групповые пуассоновские и гиперпуассоновские модели пакетного трафика // *i-Methods.* 2022. Т. 14. № 3. С. 1–14.
22. *Лихтциндер Б.Я., Привалов А.Ю.* Обобщение формул для моментов очереди при неординарном пуассоновском потоке для очередей пакетов в системах телекоммуникаций // *Проблемы передачи информации.* 2023. Т. 59. Вып. 4. С. 32–37.
23. *Бутов А.А., Галимов Л.А.* Оптимальное управление интенсивностью входящего потока многоканальной СМО с роутером при эпизодически наблюдаемой длине очереди на приборах // *Современные проблемы науки и образования.* 2015. № 2. С. 758–759.
24. *Сатрутдинова А.М.* Оптимизация работы электросетевого предприятия на основе систем массового обслуживания // *Вестник науки.* 2024. Т. 1. № 2(71). С. 97–99.

25. *Васильев А.П., Бандурин И.И.* Математические модели оптимальной структуры оперативного обслуживания электрических сетей // Вестник ИГЭУ. 2010. Вып. 2. С. 1–7.
26. *Миллер Б.М., Миллер Г.Б., Семенихин К.В.* Методы синтеза оптимального управления марковским процессом с конечным множеством состояний при наличии ограничений // АиТ. 2011. № 2. С. 111–130.
Miller B.M., Miller G.B., Semenikhin K.V. Methods to Design Optimal Control of Markov Process with Finite State Set in the Presence of Constraints // Autom. Remote Control. 2011. V. 72. No. 2. P. 571–582.
27. *Кингман Дж.* Пуассоновские процессы. М.: МЦНМО, 2007, 136 с.
28. *Daley D.J., Vere-Jones D.* An Introduction to the Theory of Point Processes. New York: Springer, 2013. 702 p.
29. *Last G.* Stochastic Analysis for Poisson Processes / *Peccati G. and Reitzner M.* (eds)
30. *Хуторцев В.В.* Плотность интенсивности пространственно-временного пуассоновского потока с нулевой вероятностью наступления событий на стохастических подмножествах его пространственной области определения // Математика и математическое моделирование. 2020. № 3. С. 15–28.
31. *Хинчин А.Я.* Математические методы теории массового обслуживания // Тр. МИАН СССР. 1955. Т. 49. С. 3–122.
32. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1. М.: Мир, 1984.
33. *Колмогоров А.Н.* Проблема ожидания / Теория вероятностей и математическая статистика: Сб. статей. М.: Наука, 1986. С. 106–111.
34. *Orsingher T., Cesari R., Mosco V.* An introduction to Poisson Processes and their generalizations. IVASS. Istituto Per la Vigilanza Sulle Assicurazioni, 2022.
35. *Rang G., Shen Z., Zuo H.* Hitting probabilities of weighten Poisson processes with different intensities and their subordinations // Acta Math. Sci. 2021. V. 41. No. B(1). P. 67–84.
36. *Feller W.* On Boundaries and Lateral Conditions for the Kolmogorov Differential Equations // Ann. Math. 1957. V. 65. No. 3. P. 527–570.
37. *Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф.* Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.
38. *Вентцель Е.С., Овчаров Л.А.* Теория случайных процессов и ее инженерные приложения. М.: Наука, 1991.
39. *Крылов И.А., Черноуцько Ф.Л.* О методе последовательных приближений для решения задач оптимального управления // ЖВМ и МФ. 1962. Т. 2. № 6. С. 142–153.

Статья представлена к публикации членом редколлегии А.И. Кибзуном.

Поступила в редакцию 28.11.2023

После доработки 08.02.2025

Принята к публикации 24.02.2025

Управление в технических системах

© 2025 г. Д.В. ШТАКИН (shtakin@iacp.dvo.ru),
А.А. ПЛОТНИКОВ (plotnikov_aa@iacp.dvo.ru),
О.Ю. СНЕГИРЕВ, канд. техн. наук (snegirevoleg@iacp.dvo.ru),
А.Ю. ТОРГАШОВ, д-р техн. наук (torgashov@iacp.dvo.ru)
(Институт автоматики и процессов управления ДВО РАН, Владивосток)

ПОСТРОЕНИЕ ВИРТУАЛЬНОГО АНАЛИЗАТОРА С НЕСКОЛЬКИМИ ВЫХОДАМИ НА ОСНОВЕ ПРЕДИКТОРА ОШИБОК ДЛЯ УПРАВЛЕНИЯ КАЧЕСТВОМ ЦЕЛЕВОГО ПРОДУКТА СЛОЖНОЙ РЕКТИФИКАЦИОННОЙ КОЛОННЫ¹

Рассматривается задача повышения точности виртуального анализатора (ВА) с несколькими выходными переменными. Показано, что введение предиктора векторного временного ряда ошибок обеспечивает учет динамической взаимозависимости компонент процесса и позволяет повысить точность ВА. Построение предиктора ошибок с несколькими выходами выполняется с использованием векторных авторегрессионных моделей и набора моделей авторегрессии с распределенным лагом, оптимальные структуры и параметры которых находятся численными методами. Проведено сравнение предложенного подхода к построению многомерного ВА с традиционными методами, основанными на последовательном построении одномерных по выходу ВА в системе управления показателями качества целевого продукта (легкой дизельной фракции) промышленной сложной ректификационной колонны. Эффективность предложенного подхода также продемонстрирована для класса адаптивных ВА.

Ключевые слова: виртуальный анализатор с несколькими выходами, предиктор ошибок с несколькими выходами, векторная авторегрессионная модель, авторегрессионная модель с распределенным лагом, адаптация, сложная ректификационная колонна.

DOI: 10.31857/S0005231025070039, EDN: JRKVG5

1. Введение

Виртуальные анализаторы (ВА) являются ключевым элементом современных систем управления технологическими процессами, отвечающим за обратную связь по качеству выходной продукции [1]. В нефтепереработке и нефтехимии одним из основных и энергоемких непрерывных технологических процессов является ректификация. В промышленных ректификационных колоннах происходит разделение исходного сырья на различные фракции нефтепродуктов. Как правило, разделение осуществляется в сложных колоннах,

¹ Работа выполнена в рамках государственного задания ИАПУ ДВО РАН по теме № FWW-2021-0003 (метод построения многомерного ВА с предиктором ошибок) и теме № FWW-2025-0002 (реализация предложенного метода и апробация на экспериментальных данных).

которые характеризуются наличием нескольких промежуточных циркуляционных орошений и отбором различных продуктов. Выделяют целевой или наиболее ценный продукт, стабилизация качества которого обеспечивает достижение значительного экономического эффекта для производства. Важными показателями качества целевого продукта, например легкой дизельной фракции технологической установки гидрокрекинга, являются ключевые точки фракционного состава (температуры начала кипения, 10, 50, 90, 95%), которые определяются с частотой 1–2 раза в сутки в заводской лаборатории, что недостаточно для решения задач оптимального управления в режиме реального времени. Поэтому внедрение ВА прежде всего обеспечивает оценку показателей качества получаемых продуктов на каждом такте управления. В состав ВА, как правило, входят статистические модели, которые связывают труднодоступные измеряемые показатели качества (выходные переменные) с легкодоступными измерениями (входными переменными) значений технологических параметров (температуры, давление, расходы потоков и др.). Эффективность управления технологическим процессом во многом зависит от точности моделей в составе ВА. В связи с этим задача повышения точности ВА является актуальной и напрямую связана с повышением эффективности управления производственными процессами [2, 3]. Можно отметить, что задачи определения зависимостей между показателями качества продуктов ректификационных колонн и набором входных переменных довольно широко исследуются для случая с одной выходной переменной [4–8]. Распространение также получили методы построения ВА на основе нейронных сетей. Работа [9] посвящена разработке динамического ВА на основе сверточной нейронной сети, учитывающей изменение характеристик анализируемого процесса во времени. Известно применение методов глубокого обучения [10] для построения ВА. Работа [11] посвящена методу обучения нейронной сети, совмещающего методы обратного распространения ошибки и частичных наименьших квадратов. Ключевым преимуществом нейросетевых методов является возможность аппроксимации нелинейных зависимостей с высокой точностью на обучающей выборке, но в то же время на тестовой выборке могут наблюдаться существенные ошибки.

Следует отметить актуальность задачи построения многомерных ВА (МВА). В сравнении с традиционными одномерными моделями многомерный регрессионный анализ лучше объясняет зависимость между оцениваемыми переменными и входными данными за счет учета коррелированности выходных переменных при рассмотрении сложных взаимосвязей между входными и выходными данными [12]. Задача построения МВА может решаться разработкой новых методов и моделей, как в работе [13], где описан метод построения многомерных ВА на основе многомерной цепи деревьев. Сущность этого метода заключается в построении ансамбля деревьев, оценивающих несколько выходных переменных. Также для построения МВА могут усовершенствоваться уже существующие методы; так, в [14] представлена модификация метода наименьших квадратов для машины опорных векторов (англ. LS-SVM,

Least Squares Support Vector Machine) для построения моделей с несколькими выходами и учета нелинейной взаимосвязи между различными выходными данными, а также предложен метод обучения такой модели.

Широко распространенным классом методов повышения точности ВА нестационарных объектов являются методы разработки адаптивных ВА. Известна работа [15], посвященная разработке ансамблевого адаптивного ВА на основе нескольких моделей гауссовских процессов. В [16] рассмотрены вопросы комбинирования нескольких механизмов адаптации для достижения наилучшей точности ВА. Не менее распространены методы построения адаптивных ВА на основе подхода «точно-в-срок» (JIT, англ. Just-In-Time) [17]. Наиболее важным достоинством методов построения адаптивных ВА является возможность учета развития процесса во времени не только на уровне обновления параметров модели, но и на уровне обновления структуры модели, что показано в работах [18, 19], посвященных ВА на основе подхода JIT, совмещенного с отбором входных переменных ВА в онлайн режиме. Следует отметить, что эффективность подхода JIT тем ниже, чем меньше данных в обучающей выборке, что является существенным ограничением в реальных условиях и не позволяет использовать данный подход для построения МВА в условиях значительных пропусков данных в обучающей выборке или при наличии малой обучающей выборки.

Существует класс методов повышения точности ВА посредством обновления свободного члена (ОСЧ) регрессионной модели. В [20] описана стратегия обновления свободного члена регрессионной модели, основанная на байесовском подходе, учитывающая среднее значение и стандартное отклонение ошибки ВА. Известна стратегия взвешенного ОСЧ регрессионной модели, рассматриваемая в [21]. Данная стратегия основана на учете предыдущих значений свободного члена регрессионной модели.

Общий недостаток рассмотренных выше методов и подходов – отсутствие учета взаимовлияния между оцениваемыми показателями качества, что в некоторых случаях может иметь существенное значение. Существует ряд работ, рассматривающих задачу ОСЧ как задачу прогноза ошибки ВА, что позволяет учитывать динамическое влияние ошибки на точность ВА. Например, в [22] стратегия ОСЧ регрессионной модели МВА рассматривается как задача прогноза ошибки ВА при помощи модели авторегрессии – скользящего среднего. Работа [23] посвящена построению прогнозирующего фильтра для адаптивного ВА. Следует отметить, что в [19–22] не учитывается взаимное влияние коррелированных ошибок ВА. В [24] решается задача прогноза вектора кросс-коррелированных ошибок ВА при помощи аналитических методов, но выбор порядка модели предиктора по формам авто- и кросс-спектров не позволяет реализовать структурную адаптацию предиктора в онлайн режиме на каждом такте дискретного времени, а также повышается сложность разработки ВА для случая n -мерного выхода.

В данной работе предлагается метод построения МВА с предиктором, отличающийся учетом взаимозависимости ошибок при помощи векторных авторегрессионных (ВАР) моделей [25] или моделей авторегрессии с распределенным лагом (АРПЛ) [26], построенных с использованием численных методов, что позволяет реализовать как структурную, так и параметрическую адаптацию предиктора ошибки ВА в каждый дискретный момент времени в замкнутом контуре системы управления показателями качества целевого продукта непрерывного технологического процесса.

2. Объект исследования и постановка задачи

Объектом исследования является сложная ректификационная колонна технологической установки гидрокрекинга для разделения подаваемого сырья на следующие продукты: бензиновая фракция (БФ), керосиновая фракция (КФ), легкая дизельная фракция (ЛДФ), тяжелая дизельная фракция (ТДФ), остаток гидрокрекинга (ОГ). ЛДФ в дальнейшем используется в качестве компонента дизельного топлива арктического (ДТА). На рис. 1 приве-

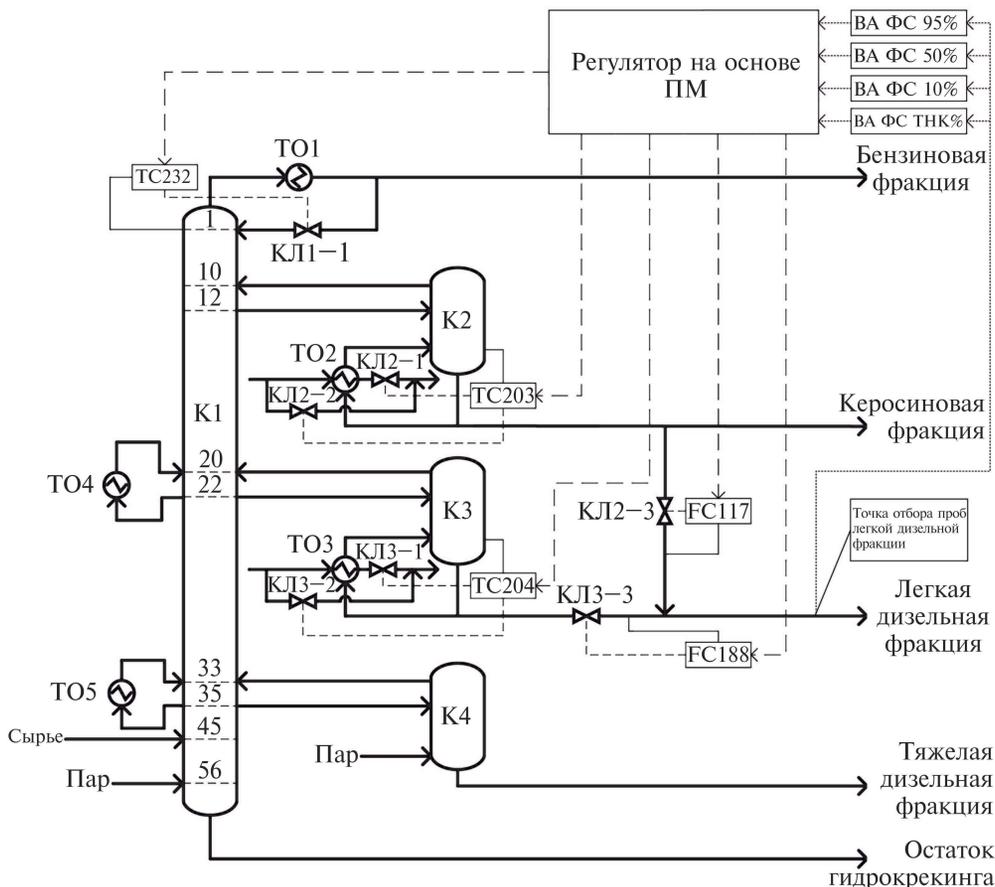


Рис. 1. Схема технологического объекта и системы управления.

дена схема технологического процесса ректификации и системы управления с регулятором на основе прогнозирующей модели (ПМ) по фракционному составу ЛДФ. Приняты следующие условные обозначения: К1 – сложная ректификационная колонна; К2 – отпарная колонна керосиновой фракции; К3 – отпарная колонна дизельной фракции; К4 – отпарная колонна тяжелой дизельной фракции; ТО1 – конденсатор паров бензиновой фракции; ТО2 – ребойлер колонны К2, ТО3 – ребойлер колонны К3, ТО4 – охладитель верхнего циркуляционного орошения; ТО5 – охладитель нижнего циркуляционного орошения; КЛ1-1 – клапан на байпасе ТО1; КЛ2-1 – клапан на выходе ТО2; КЛ2-2 – клапан на байпасе ТО2; КЛ2-3 – клапан на линии потока КФ на смешение с ЛДФ; КЛ3-1 – клапан на выходе ТО3; КЛ3-2 – клапан на байпасе ТО3; КЛ3-3 – клапан на линии потока ЛДФ на смешение с КФ. В табл. 1 представлен фрагмент матрицы коэффициентов усиления технологического объекта. В табл. 2 представлены используемые в промышленных условиях входные переменные ВА фракционного состава (ФС) ЛДФ: ФС температура начала кипения (ТНК), 10%, 50%, 95%.

Рассматривается задача оценки параметров МВА и построение модели предиктора коррелированных ошибок ВА. В общем виде модель МВА с предиктором коррелированных ошибок (МВА ПКО) имеет следующий вид:

$$(1) \quad \hat{\mathbf{Y}}_t^* = K \mathbf{X}_t + \tilde{\mathbf{b}} + \Phi \mathbf{E}_t^{past},$$

где $\hat{\mathbf{Y}}_t^* = (\hat{y}_{1,t}^* \dots \hat{y}_{j,t}^* \dots \hat{y}_{N,t}^*)^T$ – вектор оцениваемых значений показате-

телей качества; $K = \begin{pmatrix} k_{1,1} & \dots & k_{1,j} & \dots & k_{1,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ k_{h,1} & \dots & k_{h,j} & \dots & k_{h,m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ k_{N,1} & \dots & k_{N,j} & \dots & k_{N,m} \end{pmatrix}$ – матрица коэффициен-

тов модели МВА; $\mathbf{X}_t = (x_{1,t} \dots x_{j,t} \dots x_{m,t})^T$ – вектор входных переменных МВА; $\tilde{\mathbf{b}} = (b_1 \dots b_h \dots b_N)^T$ – вектор свободных членов модели МВА; m – число входных переменных МВА; N – число выход-

ных переменных МВА; $\Phi = \begin{pmatrix} \phi_{1,1|1} & \dots & \phi_{1,1|\alpha} & \dots & \phi_{1,N|1} & \dots & \phi_{1,N|\alpha} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ \phi_{N,1|1} & \dots & \phi_{N,1|\alpha} & \dots & \phi_{N,N|1} & \dots & \phi_{N,N|\alpha} \end{pmatrix}$ –

матрица коэффициентов предиктора коррелированных ошибок МВА; $\phi_{i,j|\alpha}$ – α -й коэффициент предиктора коррелированных ошибок МВА для учета влияния между j -й и i -й компонентами многомерного ряда ошибок; $\mathbf{E}_t^{past} = (e_{1,t-1} \dots e_{1,t-\alpha} \dots e_{N,t-1} \dots e_{N,t-\alpha})^T$ – входной вектор предиктора ошибок; α – порядок предиктора ошибок; $e_{i,t} = y_{i,t} - (b_i + [K]_i \mathbf{X}_t)$ – ошибка оценки МВА i -го показателя качества в момент времени t ; $[\cdot]_i$ – i -я строка матрицы.

МВА с предиктором коррелированных ошибок с предварительным «выбеливанием» (т.е. с пропуском через фильтр, выравнивающий спектральную

Таблица 1. Фрагмент матрицы коэффициентов усиления технологического объекта

ВА	ТС232 (темп. верха К1)	ТС203 (темп. в К2)	ТС204 (темп. в К3)	FC117 (расход КФ на смешение с ЛДФ)	FC188 (расход ЛДФ на смешение с КФ)
ФС ТНК	2,2	2,4	2,7	0,5	
ФС 10%	1,1	1,2	2,3	0,35	
ФС 50%				0,2	1,2
ФС 90%					1,7

Таблица 2. Описание измеряемых переменных технологического объекта

№	Обозначение	Описание	Ед.изм.
1	P131	Давление верха К1	МПа
2	P124	Давление низа К1	МПа
3	FC105	Расход острого орошения К1	м ³ /ч
4	FC106	Расход ВЦО К1	м ³ /ч
5	FC107	Расход НЦО К1	м ³ /ч
6	FC116	Расход керосиновой фракции с установки	м ³ /ч
7	FC117	Расход керосиновой фракции на смешение с ЛДФ	м ³ /ч
8	FC188	Расход легкой дизельной фракции на смешение КФ	м ³ /ч
9	ТС232	Температура верха К1	°С
10	T233	Температура возврата верхнего продукта К2 в К1	°С
11	T234	Температура возврата верхнего продукта К3 в К1	°С
12	T235	Температура возврата верхнего продукта К4 в К1	°С
13	T238	Температура бокового потока из К1 в К2	°С
14	T239	Температура бокового потока из К1 в К3	°С
15	T240	Температура бокового потока из К1 в К4	°С
16	T242	Температура в зоне над вводом сырья в К1	°С
17	T247	Температура на выходе из ТО2 в К2	°С
18	T248	Температура на выходе из К2 в ТО2	°С
19	T253	Температура на выходе из ТО3 в К3	°С
20	T254	Температура на выходе из К3 в ТО3	°С
21	ТС275	Температура возврата НЦО в К1	°С
22	ТС279	Температура возврата ВЦО в К1	°С
23	FY098	Задание по расходу сырья в К1	м ³ /ч

плотность) ряда ошибок при помощи АР модели (МВА ПКОВ) описывается выражением:

$$(2) \quad \hat{Y}_t^{**} = K X_t + \tilde{b} + W E_t^{past} + \tilde{\Phi} V_t^{past},$$

где $\mathbf{E}_t^{past} = (e_{1,t-\tilde{1}} \dots e_{1,t-q} \dots e_{N,t-\tilde{1}} \dots e_{N,t-q})^T$ – вектор ошибки модели МВА; $W = \begin{pmatrix} w_{1,1|1} & \dots & w_{1,1|q} & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & w_{2,2|1} & \dots & w_{2,2|q} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & w_{N,N|1} & \dots & w_{N,N|q} \end{pmatrix}$ – матрица коэффициентов выбеливающего предиктора порядка q ; $w_{i,j|q}$ – q -й коэффициент выбеливающего предиктора для учета влияния между j -й и i -й компонентами многомерного ряда ошибок; $\tilde{\Phi} = \begin{pmatrix} \tilde{\phi}_{1,1|\tilde{\alpha}} & \dots & \tilde{\phi}_{1,1|\tilde{\alpha}} & \dots & \tilde{\phi}_{1,N|\tilde{\alpha}} & \dots & \tilde{\phi}_{1,N|\tilde{\alpha}} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\phi}_{N,1|\tilde{\alpha}} & \dots & \tilde{\phi}_{N,1|\tilde{\alpha}} & \dots & \tilde{\phi}_{N,N|\tilde{\alpha}} & \dots & \tilde{\phi}_{N,N|\tilde{\alpha}} \end{pmatrix}$ – матрица коэффициентов предиктора коррелированных ошибок выбеливания порядка $\tilde{\alpha}$; $\mathbf{V}_t^{past} = (v_{1,t-\tilde{1}} \dots v_{1,t-\tilde{\alpha}} \dots v_{N,t-\tilde{1}} \dots v_{N,t-\tilde{\alpha}})^T$ – входной вектор предиктора ошибок; $v_{i,t} = e_{i,t} - [W]_i \mathbf{E}_t^{past}$ – ошибка выбеливания ошибок МВА i -го показателя качества.

Матрицы искомых параметров θ_1 для (1) и θ_2 для (2) имеют следующий вид:

$$(3) \quad \theta_1 = (K \quad \tilde{\mathbf{b}} \quad \Phi),$$

$$(4) \quad \theta_2 = (K \quad \tilde{\mathbf{b}} \quad W \quad \tilde{\Phi}).$$

Определение оценок неизвестных параметров θ_z ($z = 1, 2$) осуществляется путем решения оптимизационной задачи при заданном функционале качества J :

$$(5) \quad \hat{\theta}_z = \underset{\theta_z \in \mathbb{R}}{\operatorname{argmin}} J(\theta_z).$$

В качестве критерия J рассматривается средняя квадратичная ошибка МВА:

$$(6) \quad J = \frac{1}{n} \sum_{i=1}^n \left\| \mathbf{Y}_i - \hat{\mathbf{Y}}_i(\theta_z) \right\|^2,$$

где \mathbf{Y}_i и $\hat{\mathbf{Y}}_i$ – векторы измеренных и оцененных значений показателей качества целевого продукта соответственно; n – число наблюдений в выборке; $\|\cdot\|$ – евклидова норма.

3. Построение многомерного ВА

Схема предлагаемого МВА с предиктором в составе системы управления показана на рис. 2, где: MPC – регулятор на основе ПМ; $G(U)$ – технологический объект; U – вектор управляющих воздействий; $\hat{\mathbf{Y}}_t$ – оценка ВА (без

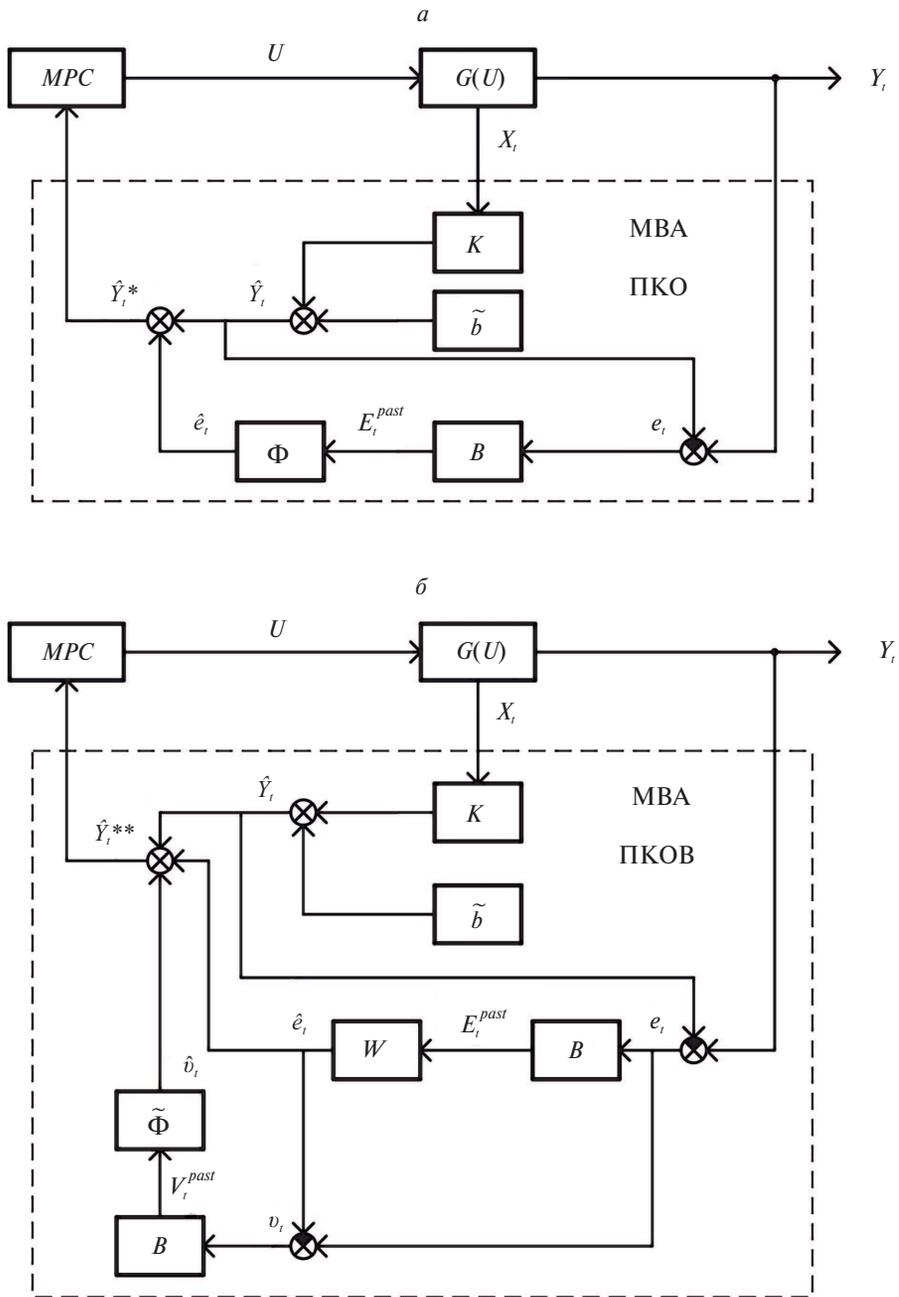


Рис. 2. Структурная схема системы управления с МВА с предиктором коррелированных ошибок: *a* – без предварительного выбеливания; *б* – с предварительным выбеливанием.

учета предиктора ошибки); e_t – вектор ошибок ВА; v_t – вектор ошибок выбеливания; \hat{e}_t – прогноз ошибки; \hat{v}_t – прогноз ошибки выбеливания; B – опе-

ратор сдвига назад на один квант времени. Основное преимущество предлагаемого МВА заключается в учете динамического влияния коррелированных ошибок оценок показателей качества целевого продукта сложной ректификационной колонны. Следует отметить, что предлагаемый предиктор является надстройкой над любой моделью в составе МВА и может быть использован в комбинации с другими моделями, например нейросетевыми, формирующими базовую составляющую оценки выхода \hat{Y}_t .

В выражениях (1) и (2) первые два слагаемых $(K\mathbf{X}_t + \tilde{\mathbf{b}})$ представляют модель множественной линейной регрессии (МЛР), которая широко распространена в промышленных условиях [27].

Поиск решения оптимизационной задачи (5) состоит из нескольких этапов. На первом этапе происходит поиск первых двух блоков матрицы θ_z (\hat{K} и $\hat{\mathbf{b}}$), отвечающих за модель в составе МВА:

$$(7) \quad \{\hat{K}, \hat{\mathbf{b}}\} = \operatorname{argmin}_{K, \tilde{\mathbf{b}} \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \left\| \mathbf{Y}_i - (K\mathbf{X}_i + \tilde{\mathbf{b}}) \right\|^2.$$

На втором этапе осуществляется построение предиктора ошибок МВА, т.е. определяется третий блок матриц θ_z . На этом этапе для МВА ПКО происходит поиск $\hat{\Phi}$ с учетом кросс-корреляций ошибок МВА:

$$(8) \quad \hat{\Phi} = \operatorname{argmin}_{\Phi \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \left\| \mathbf{e}_i - \Phi \mathbf{E}_i^{past} \right\|^2.$$

Для МВА ПКОВ находится \hat{W} , т.е. определяются параметры «выбеливающего» предиктора без учета кросс-корреляций ошибок МВА:

$$(9) \quad \hat{W} = \operatorname{argmin}_{W \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \left\| \mathbf{e}_i - W \mathbf{E}_i^{past} \right\|^2.$$

В связи с тем, что на предыдущем этапе для МВА ПКОВ не учитывается взаимозависимость рядов ошибок, появляется необходимость третьего этапа, в ходе которого происходит поиск оптимальных параметров предиктора ошибок «выбеливания» $\hat{\Phi}$ с учетом их кросс-корреляции:

$$(10) \quad \hat{\Phi} = \operatorname{argmin}_{\tilde{\Phi} \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n \left\| \mathbf{v}_i - \tilde{\Phi} \mathbf{V}_i^{past} \right\|^2.$$

В качестве модели «выбеливающего» предиктора W могут быть использованы известные авторегрессионные (АР) модели. Как правило, АР модели с одним выходом позволяют значительно повысить точность ВА, однако данный подход не лишен недостатков, а именно, не учитывается коррелированность ошибок ВА взаимосвязанных показателей качества.

Описанный в (1) МВА ПКО далее рассматривается для случая $N = 2$, а его параметры θ_1 найдены при помощи (7) и (8). В рамках данной работы ПКО для МВА реализован на основе следующей VAR модели:

$$(11) \quad (\hat{e}_{1,t} \quad \hat{e}_{2,t})^T = \Omega \mathbf{E}_t^{(q)},$$

где $\Omega = \begin{pmatrix} \omega_{1,1|1} & \dots & \omega_{1,1|j} & \dots & \omega_{1,1|q} & \omega_{1,2|1} & \dots & \omega_{1,2|j} & \dots & \omega_{1,2|q} \\ \omega_{2,1|1} & \dots & \omega_{2,1|j} & \dots & \omega_{2,1|q} & \omega_{2,2|1} & \dots & \omega_{2,2|j} & \dots & \omega_{2,2|q} \end{pmatrix}$ – матрица коэффициентов VAR модели;

$\mathbf{E}_t^{(q)} = (e_{1,t-1} \dots e_{1,t-j} \dots e_{1,t-q} \quad e_{2,t-1} \dots e_{2,t-j} \dots e_{2,t-q})^T$ – входной вектор VAR модели: q – порядок VAR модели.

Также в данной работе рассмотрен еще один способ реализации ПКО при помощи авторегрессионной модели с распределенным лагом (АРРЛ модели) [26] следующего вида:

$$(12) \quad \hat{e}_{1|t} = \mathbf{a}_1 E_{1|t}^{endo(1)} + \mathbf{l}_1 E_{2|t}^{exo(1)}, \quad \hat{e}_{2|t} = \mathbf{a}_2 E_{2|t}^{endo(2)} + \mathbf{l}_2 E_{1|t}^{exo(2)},$$

где $\mathbf{a}_1 = (a_{1,1} \dots a_{1,j} \dots a_{1,A_1})$ и $\mathbf{a}_2 = (a_{2,1} \dots a_{2,j} \dots a_{2,A_2})$ – коэффициенты регрессий эндогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 ; $\mathbf{l}_1 = (l_{1,1} \dots l_{1,j} \dots l_{1,L_1})$ и $\mathbf{l}_2 = (l_{2,1} \dots l_{2,j} \dots l_{2,L_2})$ – коэффициенты регрессий экзогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 ; $E_{1|t}^{endo(1)} = (e_{1,t-1} \dots e_{1,t-j} \dots e_{1,t-A_1})^T$ и $E_{2|t}^{endo(2)} = (e_{2,t-1} \dots e_{2,t-j} \dots e_{2,t-A_2})^T$ – входные векторы эндогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 ; $E_{2|t}^{exo(1)} = (e_{2,t-d_1} \dots e_{2,t-d_1-j} \dots e_{2,t-d_1-L_1+1})^T$ и $E_{1|t}^{exo(2)} = (e_{1,t-d_2} \dots e_{1,t-d_2-j} \dots e_{1,t-d_2-L_2+1})^T$ – входные векторы экзогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 ; A_1 и A_2 – порядки регрессий эндогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 ; d_1 и d_2 – запаздывания экзогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 ; L_1 и L_2 – порядки регрессий экзогенной составляющей моделей рядов ошибок \hat{e}_1 и \hat{e}_2 .

АРРЛ модель является более гибким инструментом учета коррелированности ошибок ВА в сравнении с VAR моделью, так как позволяет задать различные порядки регрессии для эндогенных и экзогенных составляющих, а также значение запаздывания для экзогенной составляющей. Предлагаемый подход к построению МВА включает возможность предварительного выбеливания (2) ряда ошибок ВА (коррекция ошибки ВА при помощи AR модели) перед применением VAR или АРРЛ моделей (МВА ПКОВ). Параметры МВА ПКОВ найдены при помощи (7), (9), (10). В этом случае VAR и АРРЛ модели используют для прогноза не ошибок ВА, а ошибок выбеливания:

$$(13) \quad v_{1,t} = e_{1,t} - \hat{e}_{1,t}, \quad v_{2,t} = e_{2,t} - \hat{e}_{2,t}.$$

Поэтому VAR (11) и АРРЛ (12) модели принимают вид (14) и (15) соответственно.

$$(14) \quad (\hat{v}_{1,t} \quad \hat{v}_{2,t})^T = \Omega \mathbf{V}_t^{(q)},$$

где $V_t^{(q)} = (v_{1,t-1} \dots v_{1,t-j} \dots v_{1,t-q} \quad v_{2,t-1} \dots v_{2,t-j} \dots v_{2,t-q})^T$ – входной вектор VAR модели.

$$(15) \quad \hat{v}_{1|t} = \mathbf{a}_1 V_{1|t}^{endo(1)} + \mathbf{l}_1 V_{2|t}^{exo(1)}, \quad \hat{e}_{2|t} = \mathbf{a}_2 V_{2|t}^{endo(2)} + \mathbf{l}_2 V_{1|t}^{exo(2)},$$

где $V_{1|t}^{endo(1)} = (v_{1,t-1} \dots v_{1,t-j} \dots v_{1,t-A_1})^T$ и $V_{2|t}^{endo(2)} = (v_{2,t-1} \dots v_{2,t-j} \dots v_{2,t-A_2})^T$ – входные векторы эндогенной составляющей моделей рядов ошибок \hat{v}_1 и \hat{v}_2 ;

$$V_{2|t}^{exo(1)} = (v_{2,t-d_1} \dots v_{2,t-d_1-j} \dots v_{2,t-d_1-L_1+1})^T \text{ и}$$

$V_{1|t}^{exo(2)} = (v_{1,t-d_2} \dots v_{1,t-d_2-j} \dots v_{1,t-d_2-L_2+1})^T$ – входные векторы экзогенной составляющей моделей рядов ошибок \hat{v}_1 и \hat{v}_2 .

Как правило, порядок AP модели временного ряда определяется при помощи визуального анализа его коррелограммы или автоспектра. Несмотря на широкую распространенность данного подхода, в настоящей работе для определения порядка AP модели используется численный метод, основанный на анализе оценки частной автокорреляции $\hat{\psi}_k$, определяемой по уравнению Юла–Уолкера:

$$(16) \quad \hat{\Psi}^{(k)} = P_k^{-1} \rho^{(k)},$$

где k – задержка; $\hat{\Psi}^{(k)} = (\hat{\psi}_1 \dots \hat{\psi}_j \dots \hat{\psi}_k)^T$ – вектор коэффициентов AP модели порядка k ; $\rho^{(k)} = (r_1 \dots r_j \dots r_k)^T$ – вектор оценок автокорреляции

до задержки k ; $P_k = \begin{pmatrix} 1 & r_1 & r_2 & \dots & r_{k-1} \\ r_1 & 1 & r_1 & \dots & r_{k-2} \\ r_2 & r_1 & 1 & \dots & r_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{k-1} & r_{k-2} & r_{k-3} & \dots & 1 \end{pmatrix}$ – теплицева матрица на

основе вектора $\rho^{(k)}$. Уравнение (16) связывает значения коэффициентов авторегрессионной модели и значения оценки автокорреляции [28]. Применение численных методов позволяет автоматизировать процесс определения порядка AP модели и реализовать структурную адаптацию предиктора. Способ определения порядка AP модели на основе анализа оценки $\hat{\psi}_k$ заключается в следующем. Значение оценки частной автокорреляции на задержке k приравнивается нулю, если $|\hat{\psi}_k| < 2\hat{\sigma}_{\hat{\psi}_k}$, где $\hat{\sigma}_{\hat{\psi}_k}$ – стандартная ошибка оценок частной автокорреляции, определяемая как $\hat{\sigma}_{\hat{\psi}_k} = \frac{1}{n}$. Затем в качестве порядка p используется максимальное значение задержки k , для которой $\hat{\psi}_k \neq 0$.

Частная автокорреляция используется также и для определения порядка q VAR модели. При помощи вышеописанного подхода определяются порядки авторегрессионных моделей для каждого ряда ошибок, включаемого в векторную модель, после чего в качестве порядка векторной модели q выбирается наименьший порядок из ранее полученных порядков авторегрессионных мо-

делей. Данный подход позволяет избежать включения в модель избыточных переменных.

Для определения значений A и L для АРРЛ модели можно использовать способ из [29], основанный на выявлении непостоянства знаков оценок параметров модели при усложнении структуры модели. Выполняется последовательное увеличение порядка модели до тех пор, пока соблюдается следующее условие:

$$(17) \quad \sum_{i=0}^{q-1} \operatorname{sgn} b_i^{(q-1)} \operatorname{sgn} b_i^{(q)} = q - 1,$$

где q – порядок модели: $\hat{x}_{t+1} = b_0^{(q)} x_t + b_1^{(q)} x_{t-1} + \dots + b_q^{(q)} x_{t-q} = \sum_{i=0}^q b_i^{(q)} x_{t-i}$. Аналогичным способом определяются порядки A и L АРРЛ модели. Представленные выше способы определения порядков АР, ВАР и АРРЛ моделей используются для реализации структурной адаптации предикторов ошибок в онлайн режиме, т.е. на каждом такте дискретного времени.

Определение коэффициентов АР модели осуществляется с использованием уравнения Юла–Уолкера:

$$(18) \quad \mathbf{W}_{YW} = P_\tau^{-1} \rho^{(\tau)},$$

где $\mathbf{W}_{YW} = (w_1 \dots w_j \dots w_p)^T$ – вектор коэффициентов АР модели порядка p ; $\rho^{(\tau)} = (r_1 \dots r_j \dots r_\tau)^T$ – вектор оценок автокорреляции до задержки $\tau = p$; P_τ – теплицева матрица на основе вектора $\rho^{(\tau)}$ определяется, как и в случае (16). Задачи (8), (10) для нахождения параметров ВАР и АРРЛ моделей решаются методом наименьших квадратов.

4. Апробация предложенного метода на экспериментальных данных и обсуждение результатов

Тестирование предлагаемого подхода к построению МВА с предиктором выполнено следующим образом. Для реализации предикторов ошибки, учитывающих коррелированность ошибок, ВА фракционного состава рассматриваются попарно: первая пара – ФС ТНК, ФС 10%; вторая пара – ФС 50%, ФС 95%. На рис. 3, 4 показано наличие динамической взаимосвязи между выбранными парами временных рядов ошибок ВА с помощью сглаженных выборочных оценок [30] квадрата спектра когерентности:

$$(19) \quad \overline{K}_{12}^2(\lambda) = \frac{\overline{A}_{12}^2(\lambda)}{\overline{C}_{11}(\lambda)\overline{C}_{22}(\lambda)}, \quad 0 \leq \lambda \leq F,$$

где $\overline{A}_{12}(\lambda) = \sqrt{\overline{L}_{12}^2(\lambda) + \overline{Q}_{12}^2(\lambda)}$, $0 < \lambda < F$ – сглаженная выборочная оценка взаимного амплитудного спектра;

$\overline{L}_{12}(\lambda) = 2 \left(\tilde{l}_{12}(0) + 2 \sum_{\kappa=1}^{L-1} \tilde{l}_{12}(\kappa) \beta(\kappa) \cos \frac{\pi \kappa \lambda}{F} \right)$, $0 \leq \lambda \leq F$ – сглаженная выборочная оценка коспектра;

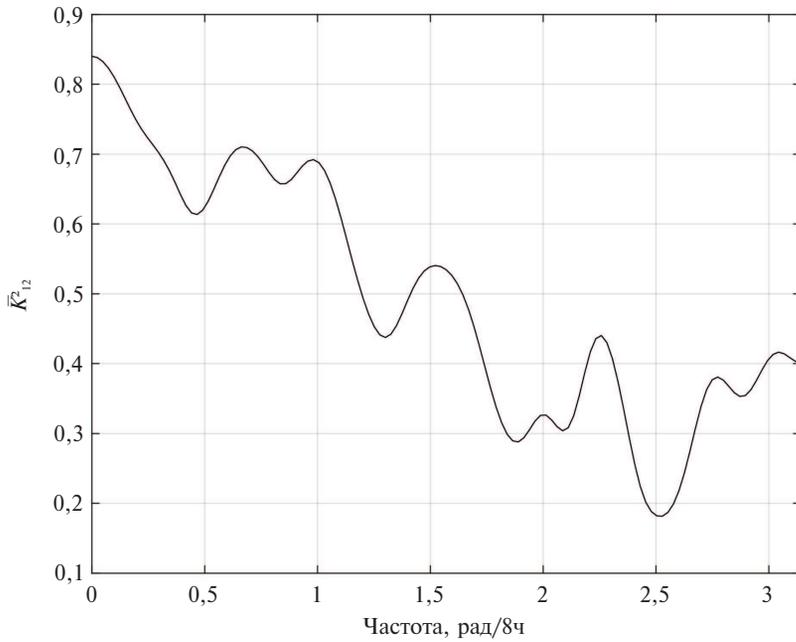


Рис. 3. Сглаженная выборочная оценка квадрата спектра когерентности ошибок ВА ФС ТНК и ФС 10% на обучающей выборке.

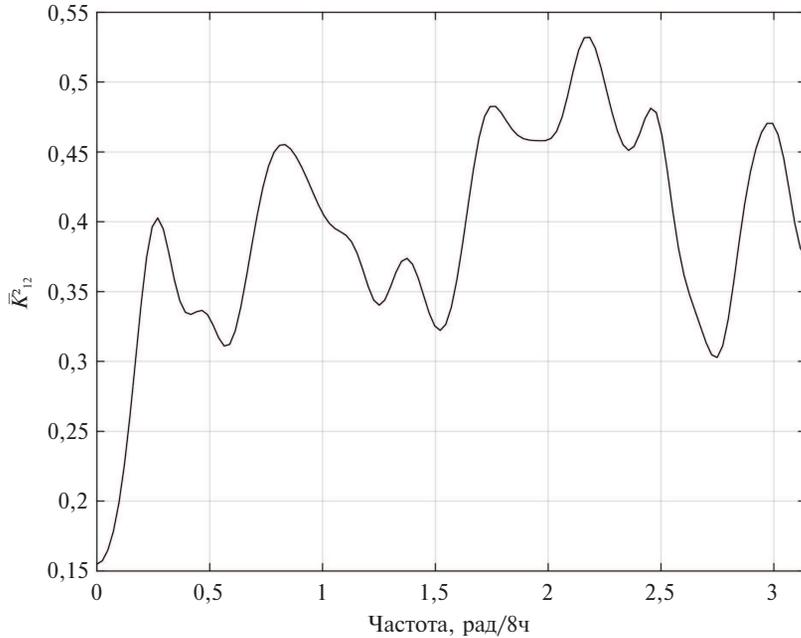


Рис. 4. Сглаженная выборочная оценка квадрата спектра когерентности ошибок ВА ФС 50% и ФС 95% на обучающей выборке.

$\bar{Q}_{12}(\lambda) = 4 \sum_{\kappa=1}^{L-1} \tilde{q}_{12}(\kappa) \beta(\kappa) \sin \frac{\pi \kappa \lambda}{F}$, $1 \leq \lambda \leq F-1$ – сглаженная выборочная оценка квадратурного спектра;

$\tilde{l}_{12}(\kappa) = 0,5 (c_{12}(\kappa) + c_{12}(-\kappa))$, $0 \leq \kappa \leq L-1$ – четная часть выборочной взаимной ковариационной функции;

$\tilde{q}_{12}(\kappa) = 0,5 (c_{12}(\kappa) - c_{12}(-\kappa))$, $0 \leq \kappa \leq L-1$ – нечетная часть выборочной взаимной ковариационной функции;

$c_{12}(\kappa) = \frac{1}{n} \sum_{t=1}^{n-\kappa} (e_{1,t} - \bar{e}_1)(e_{2,t+\kappa} - \bar{e}_2)$, $c_{12}(-\kappa) = \frac{1}{n} \sum_{t=1}^{n-\kappa} (e_{1,t+\kappa} - \bar{e}_1)(e_{2,t} - \bar{e}_2)$, $0 \leq \kappa \leq L-1$ – выборочная оценка взаимной ковариационной функции;

$\bar{C}_{11}(\lambda) = 2 \left(c_{11}(0) + 2 \sum_{\kappa=1}^{L-1} c_{11}(\kappa) \beta(\kappa) \cos \frac{\pi \kappa \lambda}{F} \right)$, $0 \leq \lambda \leq F$ – сглаженная выборочная спектральная оценка;

$$\beta(\kappa) = \begin{cases} 1 - \frac{|\kappa|}{M}, & |\kappa| \leq M \\ 0, & |\kappa| > M \end{cases} \text{ – сглаживающее окно Бартлетта;}$$

$c_{11}(\kappa) = \frac{1}{n} \sum_{t=1}^{n-\kappa} (e_{1,t} - \bar{e}_1)(e_{1,t+\kappa} - \bar{e}_1)$, $0 \leq \kappa \leq L-1$ – выборочная оценка автоковариационной функции;

$$\bar{e}_1 = \frac{1}{n} \sum_{t=1}^n e_{1,t} \text{ – среднее значение ошибки ВА;}$$

L – число запаздываний ковариационных функций, используемых в расчете;

F – максимальная частота для расчета спектральных оценок;

оценки $\bar{C}_{22}(\lambda)$ и $c_{22}(\kappa)$ рассчитываются аналогично $\bar{C}_{11}(\lambda)$ и $c_{11}(\kappa)$.

Значения амплитуды квадрата спектра когерентности, представленного на рис. 3, более 0,6 в низкочастотной области и свидетельствуют о наличии тесной взаимосвязи ошибок ВА ФС ТНК и ФС 10% и малой автокорреляции рядов ошибок данных показателей. График, приведенный на рис. 4, отражает наличие взаимосвязи ошибок ВА ФС 50% и ФС 95%, но в низкочастотной области имеется значительная автокорреляция ошибок ВА по одному из показателей, что выражается в малом значении амплитуды сглаженной выборочной оценки квадрата спектра когерентности в окрестности значения 0,15.

Сравнение предлагаемого МВА проводится с предиктором ошибок на основе АР модели (не учитывающей коррелированность ошибок ВА) и широко распространенной в промышленности процедурой ОСЧ модели МЛР в составе ВА. Предлагаемый подход реализован для адаптивных (АВА) и неадаптивных ВА. Адаптация реализована в рамках метода скользящего окна (СО). Также реализованы варианты предикторов ошибок со структурной адаптацией (СА) и параметрической адаптацией (ПА). Под параметрической адаптацией понимается повторный расчет, на сместившемся СО, параметров предикторов без изменения их порядка, определяемого методами, описанными в предыдущем разделе на первой итерации СО. Под структурной адаптацией понимается расчет параметров предикторов с определением их порядка при каждом смещении скользящего окна.

Имеющаяся общая выборка содержит 1382 наблюдений. Она разделена на обучающую выборку (ОБ), содержащую 967 наблюдений, и тестовую выборку (ТВ), содержащую 415 наблюдений. Для адаптивных ВА и предикторов размер ДО совпадает с размером ОБ неадаптивных ВА.

В табл. 3, 4 представлены результаты тестирования для неадаптивных и адаптивных ВА соответственно. В качестве критериев точности для j -го показателя качества ($j = 1, \dots, N$) использовались коэффициент детерминации R_j^2 и средняя абсолютная ошибка CAO $_j$:

$$(20) \quad R_j^2 = 1 - \frac{\sum_{i=1}^n (y_{j,i} - \hat{y}_{j,i})^2}{\sum_{i=1}^n (y_{j,i} - \bar{y}_j)^2},$$

$$(21) \quad \text{CAO}_j = \frac{1}{n} \sum_{i=1}^n |y_{j,i} - \hat{y}_{j,i}|,$$

где $y_{j,i}$ – значение лабораторного анализа j -го показателя качества; $\hat{y}_{j,i}$ – оцененное значение j -го показателя качества; $\bar{y}_j = \frac{1}{n} \sum_{i=1}^n y_{j,i}$.

Таблица 3. Результаты тестирования неадаптивных ВА

Тип ВА	Пара с коррелированными ошибками				Пара с коррелированными ошибками			
	ФС ТНК		ФС 10%		ФС 50%		ФС 95%	
	R^2	CAO	R^2	CAO	R^2	CAO	R^2	CAO
Используемые подходы в промышленности								
ВА МЛР	0,250	13,590	0,506	10,177	0,824	4,735	0,274	7,302
ВА ОСЧ	0,667	9,023	0,776	6,808	0,905	3,293	0,601	5,105
Предлагаемый метод в сравнении с АР моделью (без адаптации)								
ВА АР	0,872	5,108	0,851	4,974	0,928	2,606	0,375	6,418
ВА АРЛ	0,868	5,180	0,909	3,731	0,927	2,679	0,707	4,047
ВА ВАР	0,868	5,184	0,909	3,732	0,928	2,678	0,707	4,048
ВА АР АРЛ	0,873	5,063	0,839	5,166	0,928	2,606	0,397	6,298
ВА АР ВАР	0,872	5,108	0,851	4,975	0,928	2,607	0,375	6,416
Предлагаемый метод в сравнении с АР моделью (параметрическая адаптация)								
ВА АР ПА	0,887	4,558	0,829	5,373	0,933	2,443	0,372	6,378
ВА АРЛ ПА	0,883	4,654	0,914	3,510	0,933	2,473	0,715	3,910
ВА ВАР ПА	0,883	4,647	0,914	3,507	0,933	2,472	0,715	3,909
ВА АРПА АРЛ ПА	0,888	4,522	0,819	5,552	0,933	2,450	0,395	6,269
ВА АРПА ВАР ПА	0,887	4,551	0,829	5,375	0,933	2,441	0,373	6,375
Предлагаемый метод в сравнении с АР моделью (структурная адаптация)								
ВА АР СА	0,885	4,652	0,828	5,366	0,933	2,479	0,372	6,378
ВА АРЛ СА	0,883	4,654	0,914	3,510	0,933	2,473	0,715	3,910
ВА ВАР СА	0,883	4,607	0,914	3,485	0,933	2,472	0,715	3,909
ВА АРСА АРЛ СА	0,887	4,565	0,817	5,592	0,932	2,469	0,395	6,274
ВА АРСА ВАР СА	0,885	4,645	0,828	5,369	0,933	2,478	0,373	6,375

Таблица 4. Результаты тестирования адаптивных ВА

Тип ВА	Пара с коррелированными ошибками				Пара с коррелированными ошибками			
	ФС ТНК		ФС 10%		ФС 50%		ФС 95%	
	R^2	CAO	R^2	CAO	R^2	CAO	R^2	CAO
Используемые подходы в промышленности								
АВА МЛР	0,549	10,642	0,659	8,522	0,892	3,401	0,494	5,579
АВА ОСЧ	0,771	7,504	0,828	5,899	0,930	2,566	0,677	4,251
Предлагаемый метод в сравнении с АР моделью (без адаптации)								
АВА АР	0,877	4,904	0,885	4,131	0,938	2,279	0,545	5,328
АВА АРЛ	0,877	4,935	0,912	3,694	0,939	2,268	0,729	3,657
АВА ВАР	0,877	4,938	0,912	3,694	0,939	2,267	0,729	3,659
АВА АР АРЛ	0,879	4,861	0,877	4,268	0,938	2,272	0,560	5,225
АВА АР ВАР	0,877	4,905	0,885	4,131	0,938	2,280	0,546	5,325
Предлагаемый метод в сравнении с АР моделью (параметрическая адаптация)								
АВА АР ПА	0,879	4,857	0,885	4,137	0,939	2,247	0,545	5,334
АВА АРЛ ПА	0,879	4,886	0,913	3,671	0,939	2,248	0,728	3,661
АВА ВАР ПА	0,879	4,883	0,913	3,669	0,939	2,248	0,728	3,660
АВА АР ПА АРЛ ПА	0,880	4,810	0,877	4,286	0,938	2,245	0,561	5,217
АВА АР ПА ВАР ПА	0,879	4,854	0,885	4,137	0,939	2,246	0,545	5,332
Предлагаемый метод в сравнении с АР моделью (структурная адаптация)								
АВА АР СА	0,876	4,932	0,885	4,175	0,938	2,247	0,545	5,330
АВА АРЛ СА	0,879	4,886	0,913	3,671	0,939	2,248	0,728	3,661
АВА ВАР СА	0,882	4,802	0,915	3,647	0,939	2,248	0,728	3,660
АВА АР СА АРЛ СА	0,878	4,849	0,876	4,347	0,939	2,249	0,561	5,219
АВА АР СА ВАР СА	0,876	4,929	0,885	4,176	0,939	2,246	0,545	5,329

Следует отметить, что параметры моделей находились по критерию (6). Критерии (20) и (21) используются для сравнительного анализа из-за их широкой распространенности в системах мониторинга и автоматизации производственных процессов.

Исходя из табл. 3 и 4 для оценки ФС ТНК лучшие результаты по критериям точности достигаются при применении МВА ПКОВ вида (2). Для оценки ФС 10% лучшие результаты показывают МВА ПКО вида (1). Для ФС 95% лучшие результаты также показывают МВА ПКО вида (1). Отмечается высокая степень близости результатов МВА ПКО вида (1) на основе ВАР и АРЛ моделей и их значительное преимущество в сравнении с ВА с предиктором ошибок на основе АР модели.

Согласно рис. 4 и табл. 3, 4 незначительный эффект от использования ПКО для ВА ФС 50% (в сравнении с ПКО для других ВА) связан с недостаточно сильной зависимостью ошибок ВА ФС 50% от ошибок ВА ФС 95%, а также наличием значительной сериальной корреляции ошибок ВА ФС 50%.

Таблица 5. Изменение критериев точности для неадаптивных ВА

Тип ВА	Пара с коррелированными ошибками				Пара с коррелированными ошибками			
	ФС ТНК		ФС 10%		ФС 50%		ФС 95%	
	R^2	САО	R^2	САО	R^2	САО	R^2	САО
Базовый метод								
ВА МЛР	71,83%	66,73%	44,62%	65,76%	11,67%	48,41%	61,69%	46,47%
ВА ОСЧ	24,93%	49,88%	15,10%	48,81%	3,08%	25,81%	15,92%	23,44%
ВА АР	1,82%	11,47%	6,95%	29,94%	0,53%	6,26%	47,55%	39,10%
ВА АР ПА	0,11%	0,79%	9,32%	35,15%	0,00%	0,00%	47,90%	38,72%
ВА АР СА	0,36%	2,79%	9,39%	35,07%	0,08%	1,48%	47,90%	38,72%
Лучший результат								
	ВА АР ПА АРЛ ПА		ВА ВАР СА		ВА АР ПА		ВА ВАР ПА	
	R^2	САО	R^2	САО	R^2	САО	R^2	САО
	0,888	4,522	0,914	3,485	0,933	2,443	0,715	3,909

Таблица 6. Изменение критериев точности для адаптивных ВА

Тип ВА	Пара с коррелированными ошибками				Пара с коррелированными ошибками			
	ФС ТНК		ФС 10%		ФС 50%		ФС 95%	
	R^2	САО	R^2	САО	R^2	САО	R^2	САО
Базовый метод								
АВА МЛР	37,71%	54,88%	27,97%	57,20%	4,96%	33,99%	32,14%	34,39%
АВА ОСЧ	12,53%	36,00%	9,43%	38,17%	0,92%	12,53%	6,99%	13,90%
АВА АР	0,53%	2,08%	3,30%	11,70%	0,03%	1,51%	25,04%	31,30%
АВА АР ПА	0,34%	1,12%	3,25%	11,84%	-0,04%	0,10%	25,12%	31,38%
АВА АР СА	0,61%	2,64%	3,21%	12,65%	-0,01%	0,10%	25,09%	31,33%
Лучший результат								
	АВА ВАР СА		АВА ВАР СА		АВА АР ПА АРЛ ПА		АВА ВАР ПА	
	R^2	САО	R^2	САО	R^2	САО	R^2	САО
	0,882	4,802	0,915	3,647	0,938	2,245	0,728	3,660

В табл. 5 и 6 представлено изменение критериев точности в процентах для наилучшего результата относительно базовых методов (ВА, ВА ОСЧ, ВА АР, ВА АР ПА, ВА АР СА), не учитывающих кросс-корреляцию ошибок для неадаптивных и адаптивных ВА соответственно. Согласно табл. 5 и 6 наименьшее значение САО чаще достигается при помощи ПКО на основе ВАР модели с параметрической или структурной адаптацией (ВАР ПА, ВАР СА в таблицах выделены жирным).

Результаты, представленные в табл. 3–6, позволяют заключить что предлагаемый метод (ВА/АВА ВАР, АРРЛ, АР ВАР, АР АРРЛ) эффективен в сравнении с рассмотренными существующими методами. Для неадаптивных и адаптивных ВА сокращение САО в сравнении с ОСЧ составило в среднем 37 и 25,1% соответственно, т.е. для ФС ТНК – 49,9 и 36%, для ФС 10% – 48,8 и 38,2%, для ФС 50% – 25,8 и 12,5%, для ФС 90% – 23,4 и 13,9%. В сравнении с предиктором ошибки на основе АР сокращение САО для ФС ТНК – 11,5 и 2,1%, для ФС 10% – 29,9 и 11,7%, для ФС 50% – 6,3 и 1,5%, для ФС 90% – 39,1 и 31,3%, что в среднем 21,7% и 11,7% соответственно для неадаптивных и адаптивных ВА.

5. Заключение

Решена задача построения МВА с предиктором коррелированных ошибок. Показана эффективность предложенного подхода к построению МВА с предиктором, используемого в системе управления показателями качества фракционного состава легкой дизельной фракции сложной ректификационной колонны технологической установки гидрокрекинга. Проведенное сравнение предложенного МВА с ОСЧ и предиктором ошибки на основе АР модели без учета кросс-корреляции ошибок показало сокращение САО в среднем на 29,3 и 21% и увеличение коэффициента детерминации на 14 и 7% для неадаптивных и адаптивных ВА соответственно.

СПИСОК ЛИТЕРАТУРЫ

1. *Lawrence N.P., Damarla S.K., Kim J.W., et.al.* Machine learning for industrial sensing and control: A survey and practical perspective // *Control Engineering Practice*. 2024. V. 145. 105841. <https://doi.org/10.1016/j.conengprac.2024.105841>
2. *Бахтадзе Н.Н.* Виртуальные анализаторы (идентификационный подход) // *АиТ*. 2004. № 11. С. 3–24.
Bakhtadze N.N. Virtual Analyzers: Identification Approach // *Autom. Remote Control*. 2004. V. 65. No. 11. P. 1691–1709.
<https://doi.org/10.1023/B:AURC.0000047885.52816.c7>
3. *Логунов П.Л., Шаманин М.В., Кнеллер Д.В., et.al.* Усовершенствованное управление ТП: от контура регулирования до общезаводской оптимизации // *Автоматизация в промышленности*. 2015. № 4. С. 4–14.
Logunov P.L., Shamanin M.V., Kneller D.V., et.al. Advanced Process Control: From a PID loop up to refinery-wide optimization // *Autom. Remote Control*. 2020. V. 81. P. 1929–1943. <https://doi.org/10.1134/S0005117920100100>
4. *Шварцер В.И.* Определение характеристик ректификационной колонны по данным нормальной эксплуатации методом потенциальных функций // *АиТ*. 1968. № 6. С. 142–144.
5. *Mejdell T., Skogestad S.* Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression // *Industrial & Engineering Chemistry Research*. 1991. V. 30. Iss. 12. P. 2543–2555.
<https://doi.org/10.1021/ie00060a007>

6. *Torgashov A., Skogestad S.* The use of first principles model for evaluation of adaptive soft sensor for multicomponent distillation unit // *Chemical Engineering Research and Design*. 2019. V. 151. P. 70–78. <https://doi.org/10.1016/j.cherd.2019.08.017>
7. *Ferreira J., Pedemonte M., Torres A.I.* Development of a machine learning-based soft sensor for an oil refinery's distillation column // *Computers & Chemical Engineering*. 2022. V. 161. 107756. <https://doi.org/10.1016/j.compchemeng.2022.107756>
8. *Черешко А.А.* Виртуальные анализаторы качества на основе цифровых моделей // *Автоматизация в промышленности*. 2022. № 7. С. 33–38. <https://doi.org/10.25728/avtprom.2022.07.05>.
Chereshko A.A. Soft sensors based on digital models // *Autom. Remote Control*. 2023. V. 84. No. 7. P. 788–796. <https://doi.org/10.1134/s0005117923070044>
9. *Wang K., Shang C., Liu L.* Dynamic soft sensor development based on convolutional neural networks // *Industrial & Engineering Chemistry Research*. 2019. V. 58. No. 26. P. 11521–11531. <https://doi.org/10.1021/acs.iecr.9b02513>
10. *Xibilia M.G., Latino M., Marinkovic Z.* Soft sensors based on deep neural networks for applications in security and safety // *IEEE Transactions on Instrumentation and Measurement*. 2020. V. 69. No. 10. P. 7869–7876. <https://doi.org/10.1109/TIM.2020.2984465>
11. *Xuefeng Y.* Hybrid artificial neural network based on BP-PLSR and its application in development of soft sensors // *Chemometrics and Intelligent Laboratory Systems*. 2010. V. 103. No. 2. P. 152–159. <https://doi.org/10.1016/j.chemolab.2010.07.002>
12. *Li L., Li N., Wang X.* Multi-output soft sensor modeling approach for penicillin fermentation process based on features of big data // *Expert Systems with Applications*. 2023. V. 213. P. 119208. <https://doi.org/10.1016/j.eswa.2022.119208>
13. *Mastelini S.M., Da Costa V.G.T., Santana E.J.* Multi-output tree chaining: An interpretative modelling and lightweight multi-target approach // *J. Signal Proc. Syst.* 2019. V. 91. No. 2. P. 191–215. <https://doi.org/10.1007/s11265-018-1376-5>
14. *Xu S., An X., Qiao X.* Multi-output least-squares support vector regression machines // *Pattern Recognition Letters*. 2013. V. 34. No. 9. P. 1078–1084. <https://doi.org/10.1016/j.patrec.2013.01.015>
15. *Grbić R., Slišković D., Kadlec P.* Adaptive soft sensor for online prediction and process monitoring based on a mixture of Gaussian process models // *Computers & Chemical Engineering*. 2013. V. 58. P. 84–97. <https://doi.org/10.1016/j.compchemeng.2013.06.014>
16. *Bakirov R., Gabrys B., Fay D.* Multiple adaptive mechanisms for data-driven soft sensors // *Computers & Chemical Engineering*. 2017. V. 96. P. 42–54. <https://doi.org/10.1016/j.compchemeng.2016.08.017>
17. *Saporo A.* State of the art in the development of adaptive soft sensors based on just-in-time models // *Procedia Chemistry*. 2014. V. 9. P. 226–234. <https://doi.org/10.1016/j.proche.2014.05.027>
18. *Pan B., Jin H., Wang L., et al.* Just-in-time learning based soft sensor with variable selection and weighting optimized by evolutionary optimization for quality prediction of nonlinear processes // *Chemical Engineering Research and Design*. 2019. V. 144. P. 285–299. <https://doi.org/10.1016/j.cherd.2019.02.004>
19. *Fujiwara K., Kano M., Hasebe S., Takinami A.* Soft-sensor development using correlation-based just-in-time modeling // *AIChE J.* 2009. V. 55. No. 7. P. 1754–1765.

20. *Sangoi E., Sanseverinatti C.I., Clementi L.A., Vega J.R.* A Bayesian bias updating procedure for automatic adaptation of soft sensors // *Computers & Chemical Engineering*. 2021. V. 147. P. 107250.
<https://doi.org/10.1016/j.compchemeng.2021.107250>
21. *Xiong W., Zhang W., Xu B., Huang B.* JITL based MWGPR soft sensor for multi-mode process with dual-updating strategy // *Computers & Chemical Engineering*. 2016. V. 90. P. 260–267. <https://doi.org/10.1016/j.compchemeng.2016.04.033>
22. *Klimchenko V., Torgashov A., Shardt Y.A.W., Yang F.* Multi-output soft sensor with a multivariate filter that predicts errors applied to an industrial reactive distillation process // *Mathematics*. 2021. V. 9. No. 16. P. 1947.
<https://doi.org/10.3390/math9161947>
23. *Климченко В.В., Снегирев О.Ю., Шевлягина С.А., Торгашов А.Ю.* Разработка адаптивного виртуального анализатора с использованием прогнозирующего фильтра для нестационарного технологического процесса // *АиТ*. 2022. No. 12. С. 141–155.
Klimchenko V.V., Snegirev O.Yu., Shevlyagina S.A., Torgashov A.Yu. Developing an adaptive soft sensor using a predictive filter for a nonstationary process // *Autom. Remote Control*. 2022. V. 83. P. 1984–1994.
<https://doi.org/10.1134/S00051179220120104>
24. *Torgashov A., Snegirev O., Klimchenko V., Yang F.* Development of a multivariate quality estimator for an industrial fractionator in the presence of cross-correlation of output variables // *IFAC-PapersOnLine*. 2023. V. 56. No. 2. P. 7160–7165.
<https://doi.org/10.1016/j.ifacol.2023.10.596>
25. *Hamilton J.D.* Time series analysis. Princeton: Princeton University Press. 1994.
26. *Pesaran M.H.* An autoregressive distributed lag modelling approach to cointegration analysis // Department of Applied Economics, University of Cambridge, Cambridge. 1995. V. 9514. P. 371–413.
27. *Yeo W.S., Saptoro A., Kumar P., Kano M.* Just-in-time based soft sensors for process industries: A status report and recommendations // *J. Process Control*. 2023. V. 128. 103025. <https://doi.org/10.1016/j.jprocont.2023.103025>
28. *Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M.* Time series analysis: forecasting and control fifth ed. Hoboken, NJ: Wiley. 2015.
29. *Alt F.L.* Distributed Lags // *Econometrica*. 1942. V. 10. No. 2. P. 113–128.
<https://doi.org/10.2307/1905784>
30. *Jenkins G.M., Watts D.G.* Spectral analysis and its applications. San Francisco: Holden-Day, 1968.

Статья представлена к публикации членом редколлегии Н.Н. Бахтадзе.

Поступила в редакцию 31.01.2025

После доработки 30.04.2025

Принята к публикации 05.05.2025

Управление в социально-экономических системах

© 2025 г. М.И. ГЕРАСЬКИН, д-р экон. наук (innovation@ssau.ru)
(Самарский национальный исследовательский университет
им. академика С.П. Королева)

АЛГОРИТМ УПРАВЛЕНИЯ ДЕЙСТВИЯМИ ИГРОКА В ИГРЕ ТРИПОЛИИ ПРИ ЛИНЕЙНЫХ ФУНКЦИЯХ СПРОСА И ИЗДЕРЖЕК

Рассматривается игра трех фирм олигополии, т.е. ситуация триполии, в случае линейных функций спроса и издержек игроков. Исследуется рефлексивное поведение игроков, которое формализовано с помощью предположительных вариаций, т.е. предположений игроков о влиянии их действий на действие контрагента. Разработан метод вычисления суммы предположительных вариаций некоторого игрока, которая является оптимальной по функциям полезности остальных игроков (окружения). Предложен алгоритм управления действиями игрока со стороны окружения, реализация которого формирует заданный ментальный тип игрока и предопределяет его целенаправленное поведение. Проведено численное моделирование иерархической игры, в которой осуществляется информационное управление действиями игрока со стороны окружения (Центра) на примере параметров телекоммуникационного рынка России. Численный эксперимент подтвердил работоспособность алгоритма управления и увеличение полезности окружения в результате управления.

Ключевые слова: олигополия, предположительная вариация, иерархическая игра, оптимальное управление, телекоммуникационный рынок.

DOI: 10.31857/S0005231025070043, **EDN:** JRNBPV

1. Введение

Теоретико-игровая модель олигополии как сферы экономики, в которой однотипный товар продают несколько фирм, формализует взаимодействие этих фирм как игроков. Игроки, как правило, считаются ценополучателями, т.е. сталкиваются с равновесной рыночной ценой на агрегированной кривой спроса, и совершают игровые действия (выбирают стратегии) в виде объемов предложения товара. Ключевая черта игры состоит в том, что равновесная цена зависит от действия всех игроков, поэтому при выборе стратегии игроки должны предугадывать действия контрагентов (окружения). Предсказанное игроком изменение действия окружения в ответ на единичное изменение действия игрока получило название предположительной вариации. Априорная неинформированность игрока о предположительных вариациях окружения

предопределяет фундаментальную сложность игры олигополии. Как известно из классических моделей А. Курно [1] и Г. Штакельберга [2], оптимальные стратегии игроков зависят от двух факторов: обратной функции спроса на товар и функций издержек игроков. Поэтому чем проще форма этих функций, тем полнее и информативнее будет анализ исходов игры олигополии.

Игра олигополии, анализируемая при линейной функции спроса на товар и линейных функций издержек игроков [3–17], представляет собой наиболее типичный случай, рассматриваемый исследователями, поскольку такая модель игры является наиболее удобным инструментом изучения игровых стратегий. В частности, в такой постановке исследовались динамические игры олигополии на основе конечно-разностных уравнений реакций игроков [3–6], а также динамические игры с введением дифференциальных уравнений процесса изменения функций полезности игроков [7–9]. Изучалась разновидность динамической игры на основе фрактальных дифференциальных уравнений изменения действий игроков [10]. Рассматривались статические игры олигополии [11] и сопоставлялись равновесия Курно и Штакельберга.

Прикладные аспекты игры олигополии с линейными функциями спроса и издержек рассматривались в контексте выбора налоговой политики фирмы [12], при оценке последствий слияния фирм [13], в рамках проблемы входа новой фирмы в игру билатеральной олигополии [14], в случае продажи товарных наборов [15, 16], при внедрении новой технологии [17].

Вышеупомянутые исследования рассматривали неиерархическую теоретико-игровую модель олигополии, в которой игроки изначально равноправны, даже в случае лидерства по Штакельбергу: лидерство не обусловлено изначально установленной иерархией игроков, а возникает вследствие асимметрии информированности игроков в процессе игровой динамики. Уровни информированности формализуются с помощью рангов рефлексии r , а под рефлексией понимается [18] процесс выдвигания игроком предположений о стратегиях окружения. В результате этого процесса каждый игрок формирует множество окружающих его игроков-фантомов, существующих в его представлении, которые могут быть ведомыми на 1-м ранге рефлексии, лидерами по Штакельбергу первого уровня на 2-м ранге, лидерами по Штакельбергу $(r - 1)$ -го уровня на r -м ранге. Иными словами, на каждом ранге рефлексии игрок думает, что его окружают игроки рангом ниже. Соотнося эти теоретические рассуждения с практикой бизнеса, следует подчеркнуть, что принимающее решения лицо (ЛПР) на предприятии может не иметь понятия о лидерстве по Штакельбергу как научной категории. Но ЛПР при выборе действия на рынке по отношению к контрагентам должен учитывать их возможные реакции, т.е. должен рефлексировать. Поэтому если игрок верно предсказывает представления окружения, т.е. его фантомы совпадают с реальными игроками, то он в равновесии получает преобладание в соответствии с его уровнем лидерства [19]. Но в неиерархической модели игры лидер фактически не управляет окружающими игроками.

Однако в реальности нетрудно представить игровую ситуацию, в которой некоторая группа игроков может выбирать общую стратегию, направленную на повышение своих выигрышей за счет интересов одного игрока. Поэтому актуальным вопросом является изучение иерархической системы игроков олигополии, в которой группа игроков (Центр) управляет стратегией некоторого игрока посредством выбора стратегии группы, побуждающей игрока выбирать стратегию, оптимальную с позиций функций полезности группы. Процесс управления в этом случае также основан на вышеупомянутом свойстве рефлексивного поведения игроков. В данной статье эта проблема исследуется для модели игры трех игроков с линейными функциями спроса и издержек, что позволяет анализировать равновесия в явном виде.

2. Неиерархическая игра с линейными функциями спроса и издержек (линейная модель)

Неиерархическая модель рынка олигополии описывает поведение n игроков (фирм), поставляющих на рынок одинаковый товар, покупательский спрос на который характеризуется обратной функцией спроса $P(Q)$, убывающей по совокупному объему продаж Q ($P'_Q < 0$). Игроки равноправны и выбирают действия (стратегии) в виде объемов предложения товара Q_i , соотносясь со своими функциями издержек $C_i(Q_i)$, которые считаются возрастающими ($C'_{Q_i} > 0$), с целью максимизации функции полезности $\pi_i(Q, Q_i) = P(Q)Q_i - C_i(Q_i)$.

Рассмотрим игровую задачу олигополии в случае линейных функций спроса и издержек игроков

$$\begin{aligned} P(Q) &= a - bQ, \quad A > 0, \quad b > 0, \quad a \gg b, \\ C_i(Q_i) &= B_{0i} + B_i Q_i, \quad B_{0i} \geq 0, \quad B_i > 0, \end{aligned}$$

где Q_i – действие i -го игрока; a, b, B_{0i}, B_i – постоянные коэффициенты, выражающие параметры функции спроса (максимальную цену a и темп снижения цены b) и параметры функции издержек (постоянные издержки B_{0i} и предельные издержки B_i). Суммарное действие игроков равно $Q = \sum_{i \in N} Q_i$.

В этом случае цели игроков выражаются в максимизации их функций полезности π_i следующим образом:

$$(1) \quad \max_{Q_i \geq 0} \pi_i(Q, Q_i) = \max_{Q_i \geq 0} [(a - bQ)Q_i - B_{0i} - B_i Q_i], \quad i \in N = \{1, \dots, n\},$$

где N – множество игроков; n – количество игроков.

Вычисление равновесия Нэша Q_i^* в неиерархической игре $\Gamma = \langle N, \{Q_i, i \in N\}, \{\pi_i, i \in N\} \rangle$ базируется на необходимых условиях оптимальности $\frac{\partial \pi_i(Q_i^*, \rho_{ij})}{\partial Q_i} = 0, i, j \in N$, в которых фигурирует предположительная вариация i -го игрока. Для системы целевых функций (1) необходимые условия опти-

мальности имеют вид:

$$(2) \quad a - bQ - b(1 + S_i^r)Q_i - B_i = 0, \quad i \in N, \quad S_i^r = \sum_{j \in N \setminus i} \rho_{ij}^r,$$

где S_i^r – сумма предположительных вариаций i -го игрока на r -м ранге рефлексии. Как следует из уравнений (2), величина *суммы предположительных вариаций* каждого игрока, которую в дальнейшем будем называть *СПВ*, играет ключевую роль для расчета равновесия в игре, поскольку остальные параметры в (2) можно считать общим знанием, а величина СПВ игрока может быть неизвестна его окружению.

Если ввести обозначение параметра типа игрока $\alpha_i = \frac{a - B_i}{b}$, то систему (2) можно представить в более удобном виде

$$(2,a) \quad \alpha_i - (2 + S_i^r)Q_i - Q_{j\Sigma} = 0, \quad i \in N,$$

где $Q_{j\Sigma}$ – суммарное действие окружения i -го игрока; окружение обозначено как обобщенный игрок j .

Сравнительный анализ (1) и (2,a) позволяет вывести зависимость максимальной прибыли игрока i от его равновесного действия, т.е. функцию $\pi_i^* = \pi_i(Q_i^*)$, для случая рефлексивного поведения игроков.

Утверждение 1. Зависимость максимума функции полезности игрока от его равновесного действия при рефлексивном поведении игроков имеет вид

$$(3) \quad \pi_i^* = b(1 + S_i^r)(Q_i^*)^2 - B_{0i}.$$

Доказательство. Из (1) следует, что

$$\pi_i = (a - b(Q_i + Q_{j\Sigma})) - B_i Q_i - B_{0i} = b(\alpha_i - (Q_i + Q_{j\Sigma}))Q_i - B_{0i};$$

с другой стороны, из (2,a) следует

$$\alpha_i - Q_i^* - Q_{j\Sigma}^* = (1 + S_i^r)Q_i^*$$

подстановка этого выражения в предыдущее приводит к (3).

Анализ уравнений (2) показывает, что вектор равновесных действий $\mathbf{Q}^* = \{Q_i^*, i \in N\}$ как решение этой системы зависит от вектора СПВ $\mathbf{S} = \{S_i, i \in N\}$. Поэтому если у некоторого игрока i величина СПВ будет такой, что в равновесии максимизируются функции полезности остальных игроков (окружения), то можно сказать, что окружение целенаправленно управляет этим игроком, т.е. возникает иерархически игра.

3. Иерархическая игра в линейной модели триполии

Рассмотрим игру олигополии в случае трех игроков (т.е. случай триполии). В рассматриваемой модели один игрок i противостоит всем остальным игрокам, т.е. окружению, которое имеет общую цель добиться от этого игрока действия, выгодного окружению. Поэтому далее будем называть этого игрока «управляемый игрок» или «объект управления – игрок» (*ОУ-игрок*). Будем обобщенно обозначать всех игроков окружения индексом « j ».

Опишем систему информированности при рефлексивном поведении игроков в случае иерархической игры в виде следующих предположений относительно поведения и информированности игроков.

1) Все игроки имеют полную информированность о функциях полезности и действиях друг друга, выбирают действия одновременно, а управляемый игрок и окружение также независимы друг от друга.

2) Игроки окружения j информированы о предположительных вариациях друг друга и не информированы о значениях СПВ ОУ-игрока; они выбирают одинаковые стратегии (действия, обозначенные через q) и имеют одинаковую величину СПВ, которую обозначим через s . Кроме того, будем считать игроков окружения идентичными по параметру типа α_j , $j \in N \setminus i$, рассматривая среднее значение этого параметра как α . Соответственно значения целевых функций окружения будут одинаковы, обозначим их символом π . Игроки окружения могут скоординированно изменять значение СПВ в зависимости от СПВ управляемого игрока, причем эта зависимость обратная, т.е. с ростом СПВ управляемого игрока (снижением его уровня лидерства) СПВ окружения уменьшается (растет уровень лидерства окружения). Формализуем это предположение следующим образом:

$$(4) \quad \begin{aligned} q &= Q_j, \quad \alpha = \frac{1}{n-1} \sum_{j \in N \setminus i} \alpha_j, \quad \pi = \pi_j, \\ s &= S_j, \quad s' = \frac{\partial S_j}{\partial S_i} < 0, \quad s' = \text{const} \forall j \in N \setminus i. \end{aligned}$$

3) Управляемый игрок i в каждый момент игры выбирает действие независимо от действий окружения, учитывая предшествующие действия окружения. ОУ-игрок не информирован о значениях СПВ окружения, но может оценить их косвенно, по соотношению действий окружения и своих действий путем восстановления оптимальных реакций. ОУ-игрок определяет для себя значение СПВ как лидер по Штакельбергу, т.е. дифференцируя функцию оптимальной реакции окружения.

4) Оба контрагента (ОУ-игрок и окружение) не информированы об истинном уровне лидерства друг друга, т.е. не знают СПВ своего оппонента в игре.

Последнее предположение является ключевым для организации процесса информационного управления на основе манипуляции действиями ОУ-игрока. Выполняя целенаправленное действие, окружение может создать у

ОУ-игрока представление о том, что СПВ окружение изменилась, хотя на самом деле такое действие может быть единичным, а кривая реакции и СПВ окружения остаются неизменными.

Базируясь на этих предположениях, сформулируем иерархическую модель игры: если управляемый игрок выбирает оптимальное для себя действие с учетом имеющихся данных о значениях СПВ окружения, то окружение максимизирует свои функции полезности, побуждая ОУ-игрока выбрать наилучшее для окружения СПВ, т.е. для окружения ставится следующая задача:

$$(5) \quad \max_{S_i} \pi^*(q(s, S_i), Q_i(s, S_i)), \quad s = s^0 = \text{const},$$

при том что ОУ-игрок выбирает стратегию по условию

$$(6) \quad \max_{Q_i} \pi_i(Q(s, S_i), q(s, S_i)),$$

где s^0 – значение СПВ окружения в начальный момент игры.

Рассмотрим процесс управления поведением игрока i со стороны окружения, для чего определим СПВ этого игрока \bar{S}_i , которая оптимальна по функции полезности окружения (5), т.е. $\bar{S}_i = \arg \max_{S_i} \pi(q(s, S_i), Q_i(s, S_i))$.

Утверждение 2. Значение СПВ управляемого игрока, максимизирующее функцию полезности окружения (5), вычисляется из решения уравнения

$$(7) \quad 2(1+s)q_{S_i}^{*'} + q^*s' = 0$$

при условии

$$(7,а) \quad s' < -\frac{(1+s)q^*}{2q_{S_i}^{*'}}.$$

Доказательство. Из условия оптимальности первого порядка, примененного к функции (3) для игроков окружения, т.е. к целевой функции игрока окружения π^* , получим:

$$\pi_{S_i}^{*'} = 2b(1+s)q_{s_i}^{*'}q^* + b(q^*)^2s' = 0,$$

откуда следует (7). Условие оптимальности второго порядка

$$(7,б) \quad \pi_{S_i S_i}^{*''} = b\{2s'q_{s_i}^{*'}q^* + 2(1+s)(q_{S_i S_i}^{*''}q^* + q^{*2}) + 2q_{S_i}^{*'}q^*s' + q^{*2}s''\} < 0$$

преобразуем, предположив слабое влияние изменения СПВ на сдвиг равновесия $q_{S_i S_i}^{*''}$ и учитывая $s'' = 0$ согласно (4); тогда

$$(7,в) \quad 2s'q_{s_i}^{*'} + (1+s)q^* < 0.$$

Поскольку из анализа динамики уровней лидерства по Штакельбергу следует $q_{s_i}^{*'} > 0$ и в линейной модели олигополии $1+s > 0$, то максимум достигается согласно (7,б) при $s' < 0$, как было введено в предположении (4), в частности, при условии (7,а).

Отметим, что уравнение (7) соответствует полученной ранее [20] формуле для задачи олигополии общего вида $2(1 + S_i^r)Q_j^*Q_{jS_i}^*P'_Q + \left((1 + S_i^r)P''_{QS_i} + P'_Q \frac{\partial S_i^r}{\partial S_i}\right)Q_j^{*2} = 0$, поскольку в рассматриваемой модели линейных функций спроса и издержек $P'_Q = -b$, $P''_{QS_i} = 0$.

4. Методы формирования оптимального управления

Уравнение (7) записано для каждого игрока окружения. Искомое значение \bar{S}_i в (7) входит в выражения для $q_{s_i}^*$, q^* , которые необходимы для вычисления этого неизвестного. Поэтому далее рассмотрим конкретный случай триполии, в котором без ограничения общности будем считать второго игрока объектом управления ($i = 2$), а первого и третьего игроков – окружением ($j = 1, 3$).

Определим формулы расчета равновесия в рассматриваемой игре, а также решение задачи управления в иерархической системе (5).

Утверждение 3. В игре триполии, если игроки $j = 1, 3$ являются окружением, а игрок $i = 2$ является объектом управления, то

i) функции оптимальных реакций управляемого игрока и окружения следующие:

$$(8,а) \quad q = \frac{\alpha - Q_2}{3 + s}, \quad Q_2 = \frac{\alpha_2 - 2q}{2 + S_2};$$

ii) выражения равновесных действий имеют вид:

$$(8,б) \quad q^* = \frac{\alpha_2 - \alpha(2 + S_2)}{2 - (2 + S_2)(3 + s)}, \quad Q_2^* = \frac{2\alpha - \alpha_2(3 + s)}{2 - (2 + S_2)(3 + s)};$$

iii) значение СПВ управляемого игрока, максимизирующее функцию полезности окружения, вычисляется из решения уравнения

$$(8,в) \quad 2(1 + s)\frac{zq^* - \alpha}{y} + q^*s' = 0,$$

при условии

$$(8,г) \quad \frac{\pi_{S_i S_i}^{*''}}{b} = \frac{4q^*s'}{y} \left(\frac{zq^*}{2} + (1 + s)q^* - \alpha \right) + 2(1 + s)q^{*2} < 0,$$

где $y = 2 - (2 + S_2)(3 + s)$, $z = s'(2 + S_2) + 3 + s$.

Доказательство. В рассматриваемом случае система (2,а) имеет вид

$$\begin{cases} \alpha_1 - (2 + S_1)Q_1 - Q_3 - Q_2 = 0, \\ \alpha_2 - (2 + S_2)Q_2 - Q_1 - Q_3 = 0, \\ \alpha_3 - (2 + S_3)Q_3 - Q_1 - Q_2 = 0, \end{cases} \Rightarrow \begin{cases} 2\alpha - 2(2 + s)q - 2Q_2 - 2q = 0, \\ \alpha_2 - (2 + S_2)Q_2 - 2q = 0. \end{cases}$$

Отсюда следуют функции оптимальных реакций второго игрока и окружения (8,а), разрешение которых дает выражения равновесных действий (8,б). Введем обозначения для упрощения записи: $y = 2 - (2 + S_2)(3 + s)$, $z = s'(2 + S_2) + 3 + s$. В этом случае производная $q_{S_2}^{* \prime}$ вычисляется следующим образом:

$$q_{S_2}^{* \prime} = \frac{\alpha_2 z - \alpha[y + z(2 + S_2)]}{y^2} = \frac{z[\alpha_2 - \alpha(2 + S_2)]}{y^2} - \frac{\alpha y}{y^2} = \frac{zq^* - \alpha}{y}.$$

Поэтому уравнение (7) имеет вид (8,в).

Условие экстремума целевой функции окружения (7,а) второго порядка, учитывая $s'' = 0$, определим при следующем выражении $q_{S_2 S_2}^{* \prime \prime}$ (т.е. в более точном виде, чем (7,а)):

$$q_{S_2 S_2}^{* \prime \prime} = \frac{(z'q^* + zq_{S_2}^{* \prime})y - y'(zq^* - \alpha)}{y^2} = \frac{z'q^* + 2zq_{S_2}^{* \prime}}{y}.$$

Подставим это выражение в (7,б), и преобразуем с учетом $z' = 2s'$:

$$\begin{aligned} \frac{\pi_{S_i S_i}^{* \prime \prime}}{b} &= q^* \left(4s' \frac{zq^* - \alpha}{y} + 2(1 + s) \frac{2s'q^* + 2zq_{S_2}^{* \prime}}{y} \right) + 2(1 + s)q^{*2} = \\ &= \frac{4q^*}{y} \left(z[s'q^* + (1 + s)q_{S_2}^{* \prime}] + s'[(1 + s)q^* - \alpha] \right) + 2(1 + s)q^{*2}, \end{aligned}$$

в последнем выражении согласно (7) сделаем замену $(1 + s)q_{S_2}^{* \prime} = -\frac{s'q^*}{2}$, поэтому получим достаточное условие максимума в виде (8,г).

Результаты, сформулированные в утверждении 3, позволяют найти \bar{S}_2 .

Разработаем алгоритм достижения целевого для окружения значения \bar{S}_2 . Пусть в игре в некоторый начальный момент 0 установилось равновесие (q^0, Q_2^0) при некоторых СПВ управляемого игрока S_2^0 и окружения s^0 , значения которых неизвестны контрагентам. Определим фантомную реакцию окружения, при которой ОУ-игрок выберет СПВ, равную \bar{S}_2 . Обозначим символом s_f значение СПВ, которое должно быть у окружения, чтобы ОУ-игрок определил свою СПВ как \bar{S}_2 , и будем называть эту величиной *фантомной СПВ окружения*.

Утверждение 4. В игре триполии необходимым и достаточным условием для того, чтобы при фантомной СПВ окружения ($j = 1, 3$) управляемый игрок $i = 2$ установил свою реакцию по \bar{S}_2 , является равенство

$$(9) \quad s_f = -\frac{1}{\bar{S}_2} - 3.$$

Доказательство. Если ОУ-игрок совершает действие Q_2 , а в ответ получает от окружения действие q , то он рассчитывает свою СПВ, дифференцируя функцию реакции окружения. Фантомную реакцию окружения $q_f(Q_2)$

найдем, подставив в (8,а) неизвестную величину s_f : $q_f = \frac{\alpha - Q_2}{3 + s_f}$. Тогда $S_2 = q'_{Q_2} = \left(\frac{\alpha - Q_2}{3 + s_f} \right)'_{Q_2} = -\frac{1}{3 + s_f}$. Поэтому если цель окружения побудить ОУ-игрока выбрать \bar{S}_2 , то из $\bar{S}_2 = -\frac{1}{3 + s_f}$ получим необходимую для этого СПВ окружения (9).

На основе результата утверждения 4 организуем следующий процесс управления, в котором рассмотрим моменты времени $0, t, t + 1, t + 2$ и т.д., показанные как верхние индексы действий игроков, $t > 0$.

1. Окружение в момент t совершает действие, вычисленное по его фантомной функции реакции:

$$q^t = \frac{\alpha - Q_2^0}{3 + s_f} = -\bar{S}_2(\alpha - Q_2^0).$$

2. ОУ-игрок в момент $t + 1$ вычисляет СПВ по этому действию, которая согласно (9) равна \bar{S}_2 , и переходит на новую функцию реакции $Q_2^{t+1} = \frac{\alpha_2 - 2q}{2 + \bar{S}_2}$, по которой выбирает ответное действие:

$$Q_2^{t+1} = \frac{\alpha_2 - 2q^t}{2 + \bar{S}_2}.$$

3. Окружение в момент $t + 2$ совершает действие по своей истинной функции реакции:

$$q^{t+2} = \frac{\alpha_2 - Q_2^{t+1}}{3 + s^0}.$$

4. ОУ-игрок в момент $t + 3$ вычисляет СПВ по этому действию, которая равна S_2^0 , и возвращается к исходной функции реакции $Q_2^{t+3} = \frac{\alpha_2 - 2q}{2 + S_2^0}$.

5. В моменты $t + 4, t + 5$ и т.д. окружение и ОУ-игрок повторяют шаги 3, 4, в итоге равновесие в игре возвращается к исходному состоянию (q^0, Q_2^0) .

В результате этого процесса в моменты $t + 1, t + 2$ окружение достигает желаемого результата и получает максимальную прибыль. Известно [3–6], что описанный итерационный процесс достаточно быстро сходится к расчетному равновесию (8,б), поэтому будем считать, что с момента $t + 3$ состояние игры уже стабилизировалось. Следовательно, полагаем, что окружение получает дополнительный выигрыш в моменты $t + 1$ и $t + 2$, но может понести потери в момент t . Выигрыши, полученные в различные моменты времени, необходимо привести к сопоставимому уровню на основе метода дисконтирования $\pi_d(t) = \pi^t e^{-\rho t}$, использованного в динамических моделях олигополии [9]. Для оценки эффективности процесса управления сравним максимальные значения функции полезности окружения в моменты $0, t, t + 1$ и $t + 2$: управление эффективно, если

$$\pi^{*t} e^{-\rho t} + \pi^{*t+1} e^{-\rho(t+1)} + \pi^{*t+2} e^{-\rho(t+2)} \geq \pi^{*0} (e^{-\rho t} + e^{-\rho(t+1)} + e^{-\rho(t+2)}),$$

где ρ – коэффициент дисконтирования.

5. Экспериментальный анализ оптимального управления

Рассмотрим численные эксперименты, моделирующие процесс оптимального управления. Моделирование проведем на основе данных анализа тенденций спроса и издержек операторов телекоммуникационного рынка России [22], полученных для наиболее близкого к современности периода 2016–2021 гг. Линейная модель функции спроса на голосовой трафик мобильных операторов была сформирована в следующем виде:

$$P(Q) = a - bQ, \quad a = 1,6, \quad b = 0,000001;$$

функции издержек мобильных операторов рассмотрим без учета постоянных издержек, не влияющих на равновесие, в форме:

$$C_i(Q_i) = B_i Q_i, \quad B_1 = 0,0005, \quad B_2 = 0,0018, \quad B_3 = 0,0004,$$

где $i = 1$ соответствует МТС, $i = 2$ – Мегафон, $i = 3$ – Вымпелком.

Как и в разделе 4, будем считать второго игрока объектом управления (его действие обозначено через Q_2), а первого и третьего игроков – окружением ($j = 1, 3$), действия каждого из которых обозначим через q . На рис. 1 показаны линии оптимальных реакций игроков (обозначены через R) в случае равновесия Курно (значения СПВ равны $s = S_2 = 0$) и в случае трехстороннего лидерства по Штакельбергу (значения СПВ равны $s = S_2 = -0,5$); последний случай будем рассматривать как исходное положение в игре и обозначим точкой «0».

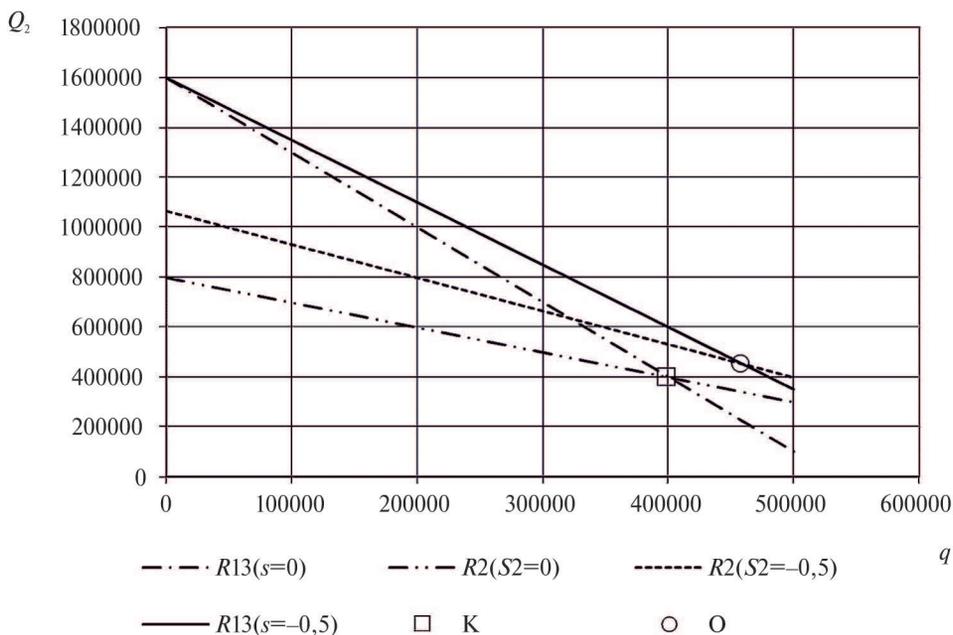


Рис. 1. Равновесие Курно (К) и трехстороннее лидерство по Штакельбергу (0).

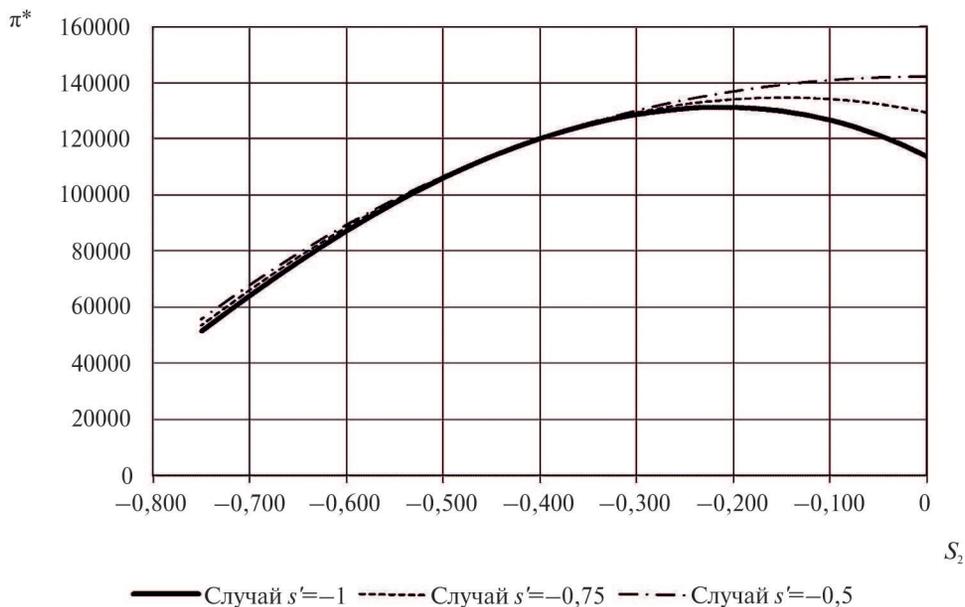


Рис. 2. Функции полезности окружения в зависимости от СПВ ОУ-игрока при различных значениях s' .

Анализ влияния характера изменения реагирования окружения на изменение реагирования ОУ-игрока, выражающегося величиной s' , показан в табл. 1 и на рис. 2. При этом использована следующая модельная зависимость:

$$s = -1 + s'S_2,$$

которая обеспечивает диапазон $s \in (-1, 0]$ при $S_2 \in (-1, 0]$ и $s' \in (-1, 0]$. Из данных табл. 1 следует, что снижение интенсивности изменения СПВ окружения, т.е. уменьшение модуля s' , обуславливает более близкое к нулю целевое значение СПВ ОУ-игрока \bar{S}_2 , а также увеличение максимальной полезности окружения π^* . Графики на рис. 2 демонстрируют наличие максимума функции полезности окружения при определенных значениях СПВ ОУ-игрока, причем условие второго порядка (8,г) также было проверено.

Таблица 1. Влияние изменения СПВ окружения при изменении СПВ ОУ-игрока на оптимум окружения

Случай	s'	\bar{S}_2	q^*	π^*
1	-1	-0,2338	5 079 027	128 969
2	-0,75	-0,1478	518 714	134 506
3	-0,5	-0,002	533 454	142 260

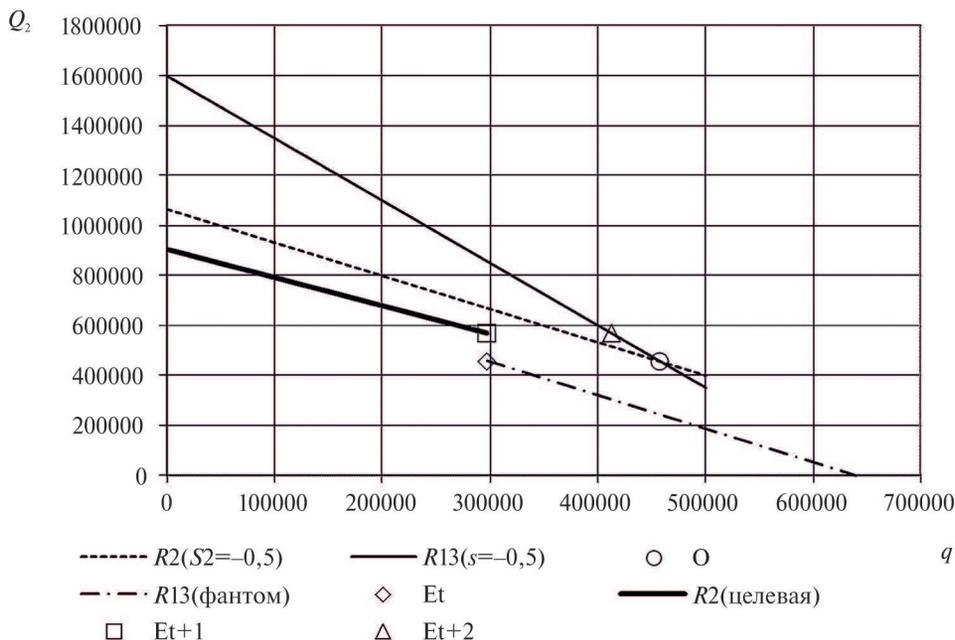


Рис. 3. Процесс информационного управления из лидерства по Штакельбергу (0).

На рис. 3 показан процесс информационного управления из состояния трехстороннего лидерства по Штакельбергу (точка 0) для случая $s' = -1$. На первом этапе действие окружения по его фантомной функции реакции $R13(\text{фантом})$ переводит игровую ситуацию в состояние E^t , причем ОУ-игрок совершает действие, соответствующее исходному положению в игре 0. Такое изменение действия окружения позволяет ОУ-игроку установить, что поведение окружения соответствует СПВ, равной \bar{S}_2 , поэтому он на втором шаге изменяет свою функцию реакции на желательную для окружения $R2(\text{целевая})$. Тогда игровая ситуация трансформируется в состояние $E^{(t+1)}$, а действие окружения остается на уровне E^t . В ответ на третьем шаге окружение возвращается на свою истинную функцию реакции $R13(s = -0,5)$, поскольку окружение вычисляло оптимальную для себя целевую СПВ ОУ-игрока именно при исходном значении своей СПВ. В результате игровая ситуация переходит в состояние $E^{(t+2)}$. Дальнейший процесс восстановления исходного равновесия в точке 0 не показан, поскольку в момент $t + 2$ действия игроков уже близки к этой точке.

Экономический анализ эффектов информационного управления представлен в табл. 2 при ставке дисконтирования 10%. Суммарная дисконтированная полезность окружения за три периода процесса управления превышает суммарное дисконтированное значение его полезности при исходном состоянии игры, что подтверждает целесообразность такого управления с позиций окружения. Отметим, что в период t окружение получило также повышенную полезность, что проявляется не всегда; в рассматриваемой ситуации первое

Таблица 2. Экономический эффект информационного управления, $\rho = 0,1$

Период времени	$e^{-\rho t}$	Номинальные значения		Дисконтированные значения	
		π^{*0}	π^{*t}	π^{*0}	π^{*t}
t	0,905	104 761	163 474	94 821	147 963
$t + 1$	0,819	104 761	129 758	85 823	106 301
$t + 2$	0,741	104 761	85 005	77 680	63 031
Сумма				258 324	317 295

действие окружения было меньше, чем в исходном равновесии, вследствие чего повысилась равновесная цена; если бы первое действие окружения было противоположным, то его полезность снизилась.

6. Заключение

В статье продемонстрирована принципиальная возможность осуществления информационного управления действиями игрока в игре триополии со стороны игроков окружения. Рассмотренный случай линейных моделей функций спроса и издержек позволил вывести выражения оптимальных реакций и равновесных действий игроков, разработать алгоритм и наглядно проиллюстрировать результаты информационного управления. Получены следующие ключевые результаты.

Выведена зависимость максимума функции полезности игрока от его равновесного действия при рефлексивном поведении игроков, показывающая, что выигрыш игрока зависит как от его СПВ, так и от СПВ других игроков через равновесное действие игрока. Из этого вытекает, что варьируя СПВ других игроков, рассматриваемый игрок может увеличить свой максимальный выигрыш, что является основанием для управления этими игроками, причем параметром управления будет их СПВ. Однако один игрок не может манипулировать остальными игроками, тогда как совокупность игроков, действуя согласованно, может изменить представление какого-то одного игрока.

Поэтому исследована иерархическая игра, в которой окружение управляет некоторым игроком путем нахождения оптимального для него значения СПВ. Получено уравнение, позволяющее вычислить оптимальное по функции полезности окружения значение СПВ этого игрока. Для побуждения управляемого игрока к выбору требуемого значения СПВ использован метод оптимальных реакций, для которых выведены формулы в случае триполии. Также в игре трех лиц, двое из которых имеют общую цель, определены равновесные действия и составлено уравнение для расчета оптимального СПВ управляемого игрока.

В рассматриваемой модели триполии разработан метод определения действия окружения на основе фантомной реакции, которое побуждает управляемого игрока использовать при реагировании оптимальную для окружения

СПВ. На основе этого разработан алгоритм, моделирующий последовательность действий окружения и игрока в процессе управления и определена эффективность этого процесса.

Проведены численные эксперименты, показавшие действенность и экономический эффект алгоритма управления. Несмотря на то, что повышенную прибыль субъект управления получает только в два момента игры, с практической точки зрения это достаточно существенно. В реальном бизнесе момент игры, как правило, соответствует периоду опубликования отчетности фирмы, которая в полном объеме формируется ежегодно. Поэтому один или два года, в которых игроки окружения получают преобладающие выигрыши, могут привести к большим финансовым потерям управляемого игрока и даже к его полному вытеснению с рынка. Следовательно, результаты исследования имеют практическую значимость.

СПИСОК ЛИТЕРАТУРЫ

1. Cournot A.A. *Researches into the Mathematical Principles of the Theory of Wealth*. London: Hafner. 1960. Original 1838.
2. Stackelberg H. *Market Structure and Equilibrium: 1st Edition*. Translation into English, Bazin, Urch & Hill, Springer, 2011. (Original 1934).
3. Алгазин Г.И., Алгазина Д.Г. Процессы рефлексии и равновесие в модели олигополии с лидером // *АиТ*. 2020. № 7. С. 113–128.
Algazin G.I., Algazina D.G. Reflexion Processes and Equilibrium in an Oligopoly Model with a Leader // Autom. Remote Control. 2020. V. 81(7). P. 1258–1270.
4. Алгазин Г.И., Алгазина Ю.Г. К аналитическому исследованию условий сходимости процессов рефлексивного коллективного поведения в моделях олигополии // *АиТ*. 2022. № 3. С. 84–109.
Algazin G.I., Algazina Y.G. To the Analytical Investigation of the Convergence Conditions of the Processes of Reflexive Collective Behavior in Oligopoly Models // Autom. Remote Control. 2022. V. 83(3). P. 367–388.
5. Алгазин Г.И., Алгазина Д.Г. Моделирование динамики коллективного поведения в рефлексивной игре с произвольным числом лидеров // *Информатика и автоматизация*. 2022. Т. 21. № 2. С. 339–375.
6. Алгазин Г.И., Алгазина Д.Г. Агрегативные игры олигополии при линейных функциях спроса и издержек агентов // *МАК: Математики – Алтайскому краю*. 2024. № 6. С. 99–101.
7. Xiao Y., Zhang S., Peng Y. Dynamic investigations in a Stackelberg model with differentiated products and bounded rationality // *J. Comput. Appl. Math*. 2022. V. 414. 114409.
8. Ougolnitsky G., Gorbaneva O. Sustainability of Intertwined Supply Networks // *A Game-Theoretic Approach Games*. 2022. V. 13(3). P. 35.
9. Угольницкий Г.А., Усов А.Б. Сравнительный анализ эффективности способов организации взаимодействия экономических агентов в моделях дуополии Курно с учётом экологических условий // *АиТ*. 2023. № 2. С. 150–168.
10. Al-Khedhairi A. Dynamical Study of Competition Cournot-like Duopoly Games Incorporating Fractional Order Derivatives and Seasonal Influences // *Int. J. Nonlinear Scie. Numer. Simulation*. 2020. V. 21(3–4). P. 339–359.

11. *Zouhar J., Zouharova M.* Stackelberg versus Cournot duopoly with asymmetric costs: primary markups, entry deterrence, and a comparison of social welfare and industry profits // *Econom. Theory Bulletin*. 2020. V. 8. P. 89–96.
12. *Collie D.R.* Taxation under oligopoly in a general equilibrium setting // *J. Public Economic Theory*. 2019. V. 21(4). P. 738–753.
13. *Ferrarese W.* When Multiple Merged Entities Lead in Stackelberg Oligopolies // *Rev. Industr. Organization*. 2020. V. 56(1). P. 131–142.
14. *Mukherjee A., Zeng C.* Social desirability of entry in a bilateral oligopoly-The implications of (non) sunk costs // *Mathem. Social Scie.* 2022. V. 118. P. 12–19.
15. *Zhou J.* Mixed bundling in oligopoly markets // *J. Econom. Theory*. 2021. V. 194. P. 105257.
16. *Shuai J., Yang H., Zhang L.* Dominant firm and competitive bundling in oligopoly markets // *Games Econom. Behav.* 2022. V. 132. P. 421–447.
17. *Zhang Y.* When should firms choose a risky new technology? An oligopolistic analysis // *Econom. Modelling*. 2020. V. 91. P. 687–693.
18. *Novikov D.A., Chkhartishvili A.G.* Reflexion and Control: Mathematical Models. London: CRC Press, 2014.
19. *Гераськин М.И.* Свойства предположительных вариаций в нелинейной модели олигополии Штакельберга // *АиТ*. 2020. № 6. С. 105–130.
Geraskin M.I. The Properties of Conjectural Variations in the Nonlinear Stackelberg Oligopoly Model // *Autom. Remote Control*. 2020. V. 81(6). P. 1051–1072.
20. *Гераськин М.И.* Информационное управление стратегией игрока в игре олигополии n лиц при рефлексивном поведении игроков // *АиТ*. 2025. № 6. С. 102–117.
21. *Гераськин М.И.* Объемные предположительные вариации в играх олигополии при различных функциях спроса и издержек и многоуровневом лидерстве // *АиТ*. 2024. № 7. С. 73–90.
Geraskin M.I. Quantity Conjectural Variations in Oligopoly Games under Different Demand and Cost Functions and Multilevel Leadership // *Autom. Remote Control*. 2024. V. 85 (7). P. 711–724.
22. *Гераськин М.И., Моисеева К.С.* Анализ тенденций спроса и издержек операторов телекоммуникационного рынка России // *Прикладная математика и вопросы управления*. 2023. № 3. С. 157–174.

Статья представлена к публикации членом редколлегии А.Г. Чхартшвили.

Поступила в редакцию 29.11.2024

После доработки 21.03.2025

Принята к публикации 25.03.2025

Интеллектуальные системы управления, анализ данных

© 2025 г. А.Ф. АНТИПИН, канд. техн. наук (andrejantipin@ya.ru)
(Стерлитамакский филиал Уфимского университета науки и технологий),
Е.В. АНТИПИНА, канд. физ.-мат. наук (stepashinaev@ya.ru),
С.А. МУСТАФИНА, д-р физ.-мат. наук (mustafina_sa@mail.ru)
(Уфимский университет науки и технологий)

МОДИФИКАЦИЯ МНОГОМЕРНОГО НЕЧЕТКОГО РЕГУЛЯТОРА С ЛОКАЛЬНЫМ ТЕРМ-ДЕЛЕНИЕМ ВХОДНЫХ ПЕРЕМЕННЫХ¹

Рассматривается модификация многомерного нечеткого регулятора с взаимосвязанными регулируемыми параметрами. Приводятся описание блок-схемы регулятора и выражения для расчета основных параметров. Описывается механизм перехода от общего терм-деления входных переменных к их локальному терм-делению относительно выходных переменных. Рассматривается разработанное авторами специальное программное обеспечение для визуальной разработки нечеткого регулятора, гибкой настройки термов, оценки и анализа системы продукционных правил. Теоретические результаты иллюстрируются примерами.

Ключевые слова: многомерный нечеткий регулятор, система продукционных правил, модификация, программное обеспечение, терм-деление.

DOI: 10.31857/S0005231025070053, **EDN:** JRZPLT

1. Введение

В настоящее время нечеткие регуляторы и их модификации, в том числе гибридные, все чаще находят свое применение при разработке автоматизированных систем управления (АСУ). Они обладают более гибкими настройками по сравнению с традиционными ПИД-регуляторами, что особенно актуально для многомерных и многосвязных систем со сложной математической моделью или отсутствием математической модели как таковой. С использованием нечеткой логики строятся экспертные и интеллектуальные системы управления [1–7].

Проблемам практической реализации нечетких регуляторов посвящен ряд работ отечественных и зарубежных авторов, в которых следует отметить [8–14]:

1) высокую сложность алгоритмов (в том числе фаззификации и дефаззификации) нечеткого регулирования;

¹ Исследование выполнено в рамках государственного задания Министерства науки и высшего образования Российской Федерации (код научной темы FZWU-2023-0002).

- 2) неоднозначность выбора функции принадлежности непрерывных величин;
- 3) значительный, а в отдельных случаях многократный рост числа продукционных правил при увеличении количества входных переменных;
- 4) отсутствие гибкости при реализации отличных от общепринятых подходов, связанных с нечеткими регуляторами, и др.

Кроме того, для синтеза и программной реализации нечетких регуляторов часто используется стороннее, по отношению к программируемым логическим контроллерам (ПЛК), программное обеспечение, в частности MATLAB [15–17], что ведет к дополнительным затратам времени и средств, а также затрудняет внесение изменений в процессе эксплуатации АСУ.

Для решения перечисленных выше проблем необходимо совершенствование характеристик нечетких регуляторов и алгоритмов их работы; появление новых модификаций, связанных с формированием системы продукционных правил; разработка специального программного обеспечения с генерацией программных блоков нечетких регуляторов для современных сред разработки программ для ПЛК.

Целью работы является разработка новой модификации многомерного нечеткого регулятора, ориентированной на устранение описанных выше недостатков и проблем, и программного обеспечения для настройки параметров и системы продукционных правил.

2. Описание модификации нечеткого регулятора

Рассмотрим разработанную авторами модификацию нечеткого регулятора, предназначенную для использования в АСУ с большим объемом продукционных правил, описывающих логику работы сложного технологического объекта при отсутствии адекватной математической модели. Цель модификации – решение проблем, связанных с составлением системы продукционных правил (СПП), устранение избыточности правил и снижение их максимального числа. Это достигается путем определения взаимосвязей переменных, а также перехода от общего к локальному терм-делению входных переменных.

На рис. 1 представлена блок-схема модификации многомерного нечеткого регулятора со взаимосвязанными регулируемыми параметрами, где Φ_1, \dots, Φ_n – фаззификаторы входных переменных x ; $\mathcal{D}_1, \dots, \mathcal{D}_m$ – дефаззификаторы выходных переменных y ; БФ и БД – блоки фаззификации и дефаззификации, управляющие работой фаззификаторов и дефаззификаторов; БЛВ – блок логического вывода регулятора с блоком продукционных правил БПП; T_t – термы переменных, которыми оперирует БЛВ в момент времени t ; БПр – блок предсказания термов входных переменных x ; БРФ – блоки расширения функций регулятора, в том числе для связи и обмена данными с внешними устройствами.

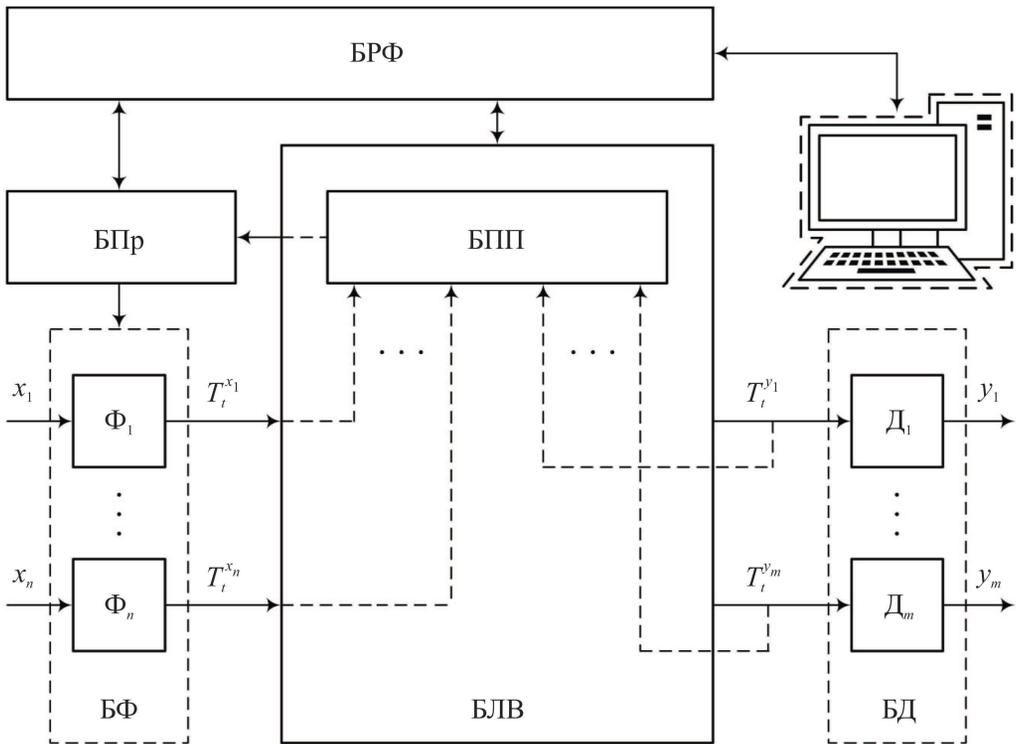


Рис. 1. Блок-схема модификации нечеткого регулятора.

Пусть терм-множество T^x непрерывной величины x имеет вид

$$T^x = \left\{ T_1^x \left(x_1^{(1)} \leq x_t < x_1^{(n)} \right), T_2^x \left(x_2^{(1)} \leq x_t < x_2^{(n)} \right), \dots, T_{k(x)}^x \left(x_{k(x)}^{(1)} \leq x_t \leq x_{k(x)}^{(n)} \right) \right\},$$

где $x^{(1)}, x^{(n)}$ – значения, которые определяют левую и правую границы интервалов термов T ; x_t – значение переменной x в момент времени t ; $k(x)$ – общее количество термов.

Для повышения иллюстративности был выбран трапецевидный (близкий к прямоугольному) тип функций принадлежности ч, для которого характерно отсутствие пересечения интервалов термов в границах общего терм-деления переменной.

Так как

$$x_k^{(1)} = x_{k-1}^{(n)}$$

по определению, то для описания $k(x) - 1$ термов T достаточно значения одной из границ интервалов. Границы первого или последнего терма описываются непосредственно двумя значениями. Например:

$$T^x = \left\{ T_1^x \left(x_1^{(1)} \right), T_2^x \left(x_2^{(n)} \right), \dots, T_{k(x)}^x \left(x_{k(x)}^{(1)} \leq x_t \leq x_{k(x)}^{(n)} \right) \right\}.$$

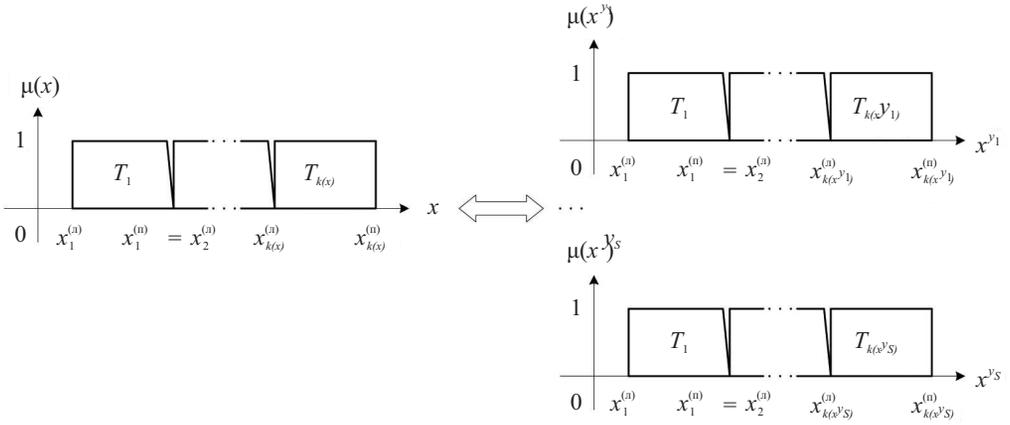


Рис. 2. Деление общего терм-множества на локальные терм-множества.

Известно, что максимальное количество продукционных правил R_{\max} регулятора зависит от суммарного числа термов входных переменных [18], т.е.

$$R_{\max} = \prod_{i=1}^n k(x_i),$$

где n – количество входных переменных регулятора.

При наличии выходных переменных с обратной связью (как показано на рис. 1) выражение для расчета R_{\max} примет вид

$$R_{\max} = \prod_{i=1}^n k(x_i) \prod_{j=1}^o k(y_j),$$

где o – количество выходных переменных регулятора с обратной связью.

Для многомерного нечеткого регулятора с m выходными переменными y общее терм-деление некоторых входных переменных x может оказаться избыточным для вычисления значений отдельных выходных переменных и приводит к ошибкам дефаззификации, связанным прежде всего с отсутствием продукционных правил, определенных техническим заданием.

Общее терм-деление может быть упрощено до локального терм-деления для конкретных переменных y . Так, локальное терм-множество T^{x^y} непрерывной величины x для взаимосвязанной переменной y аналитически можно представить следующим образом:

$$T^{x^y} = \left\{ T_1^{x^y} \left(x_1^{y(n)} \leq x_t < x_1^{y(n)} \right), \dots, T_{k(x^y)}^{x^y} \left(x_{k(x^y)}^{y(n)} \leq x_t \leq x_{k(x^y)}^{y(n)} \right) \right\},$$

при этом $k(x^y) \leq k(x)$.

Графически данный процесс можно представить в виде, показанном на рис. 2.

В этом случае выражение для расчета максимального количества продукционных правил R'_{\max} регулятора примет вид

$$R'_{\max} = \sum_{j=1}^m R_{\max}^{y_j} = \sum_{j=1}^m \prod_{i=1}^s k(x_i^{y_j}),$$

где R_{\max}^y – максимальное количество продукционных правил для конкретной выходной переменной y ; s – количество входных переменных x , связанных с выходной переменной y .

Для выполнения расчетов исходные данные по количеству термов входных переменных x для связанных с ними выходных переменных y удобно представить в виде табл. 1. При отсутствии взаимосвязи между переменными в соответствующую ячейку таблицы ставится символ «-».

Таблица 1. Данные для анализа терм-связей переменных

Выходная переменная	Количество термов входных переменных		
	x_1	...	x_n
y_1	$k(x_1^{y_1})$...	$k(x_n^{y_1})$
...
y_m	$k(x_1^{y_m})$...	$k(x_n^{y_m})$

При этом должно соблюдаться следующее условие по минимальному количеству термов выходных переменных:

$$k_{\min}(y) \leq R^y,$$

где R^y – количество продукционных правил для конкретной выходной переменной y .

При количестве продукционных правил R^y , меньшем суммарного числа термов y , некоторые термы останутся незадействованными, так как не участвуют в формировании системы продукционных правил. Такие термы следует удалить, объединив с соседними термами, чтобы не было разрывов внутри рабочих диапазонов значений выходных переменных. Если объединение термов невозможно, то такие переменные не могут в дальнейшем использоваться при реализации обратной связи во избежание ошибок фаззификации.

Подобный подход позволяет устранить так называемый конфликт локальных термов, вызванный наложением их интервалов при формировании общего терм-деления, как показано на рис. 3, что ведет к появлению новых термов и, как следствие, увеличению максимального числа продукционных правил регулятора.

Из рис. 3 видно, что для формирования общего терм-деления непрерывной величины x , которое имеет вид

$$T^x = \{T_1^x(0), T_2^x(20), T_3^x(40), T_4^x(60), T_5^x(80 \leq x_t \leq 100)\},$$

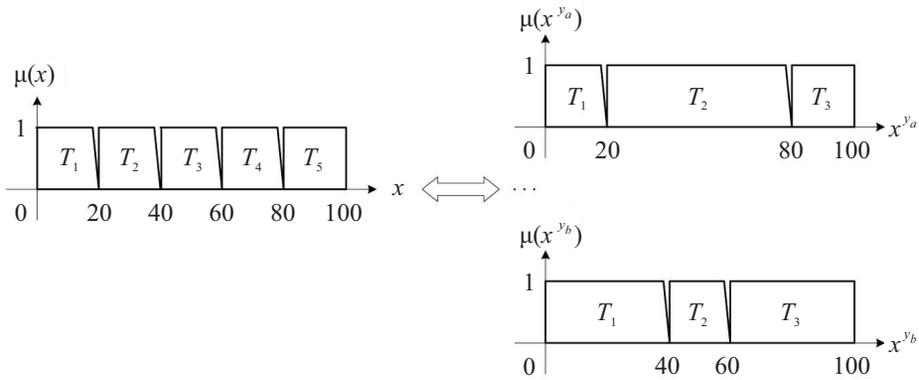


Рис. 3. Конфликт локальных термов при интерпретации переменных.

используются следующие локальные терм-деления x для выходных переменных y_a и y_b :

$$T^{x^{y_a}} = \{T_1^{x^{y_a}}(0), T_2^{x^{y_a}}(20), T_3^{x^{y_a}}(80 \leq x_t \leq 100)\},$$

$$T^{x^{y_b}} = \{T_1^{x^{y_b}}(0), T_2^{x^{y_b}}(40), T_3^{x^{y_b}}(60 \leq x_t \leq 100)\}.$$

Так, терм T_1^x соответствует терму $T_1^{x^{y_a}}$, терм T_3^x – терму $T_2^{x^{y_b}}$, а терм T_5^x – терму $T_3^{x^{y_a}}$. Термы T_2^x и T_4^x получены в результате объединения интервалов оставшихся термов из локальных терм-делений непрерывной величины x для выходных переменных y_a и y_b .

Как видно из примера, общее терм-деление непрерывной величины x приводит к появлению следующих проблем. Термы T_1^x и T_2^x , T_4^x и T_5^x дублируют продукционные правила для выходной переменной y_b , а термы T_2^x , T_3^x и T_4^x – для выходной переменной y_a , что сказывается на общем объеме продукционных правил, особенно при разработке многосвязных нечетких систем, где

$$y = f(x_1, \dots, x_s).$$

Кроме того, объединение термов может привести к некорректной работе дефазсификаторов при использовании некоторых типов функций принадлежности (например, треугольной) и к ряду сложностей при формировании системы продукционных правил.

Локальное терм-деление устраняет эти проблемы и позволяет вести разработку нечеткой АСУ в полном соответствии с техническим заданием.

При необходимости можно выделить критически важные продукционные правила, в которых используются термы непрерывных величин, помеченные как «критические» для указания их особой важности (например, при определении крайних диапазонов значений, которые сигнализируют аварийную ситуацию).

Количество критических правил для каждой группы взаимосвязанных переменных $R_{гр.кр}^g$ может быть рассчитано следующим образом:

$$R_{гр.кр}^g = R_{гр}^g - \prod_{i=1}^{n_g} (k(x_i) - k_{кр}(x_i)),$$

где $k_{кр}$ – количество критических термов, определенных для входных переменных x в объеме n_g , образующих группу с порядковым номером g ; $R_{гр}^g$ – максимальное количество продукционных правил группы:

$$R_{гр}^g = \prod_{i=1}^{n_g} k(x_i).$$

3. Описание программного обеспечения регулятора

Ввиду отсутствия в современных средах проектирования и разработки нечетких регуляторов возможностей для определения взаимосвязей переменных, перехода от общего терм-деления входных переменных к локальным терм-делениям применительно к выходным переменным, а также инструментов глубокого анализа системы продукционных правил авторами был разработан программный комплекс, предназначенный для визуальной разработки

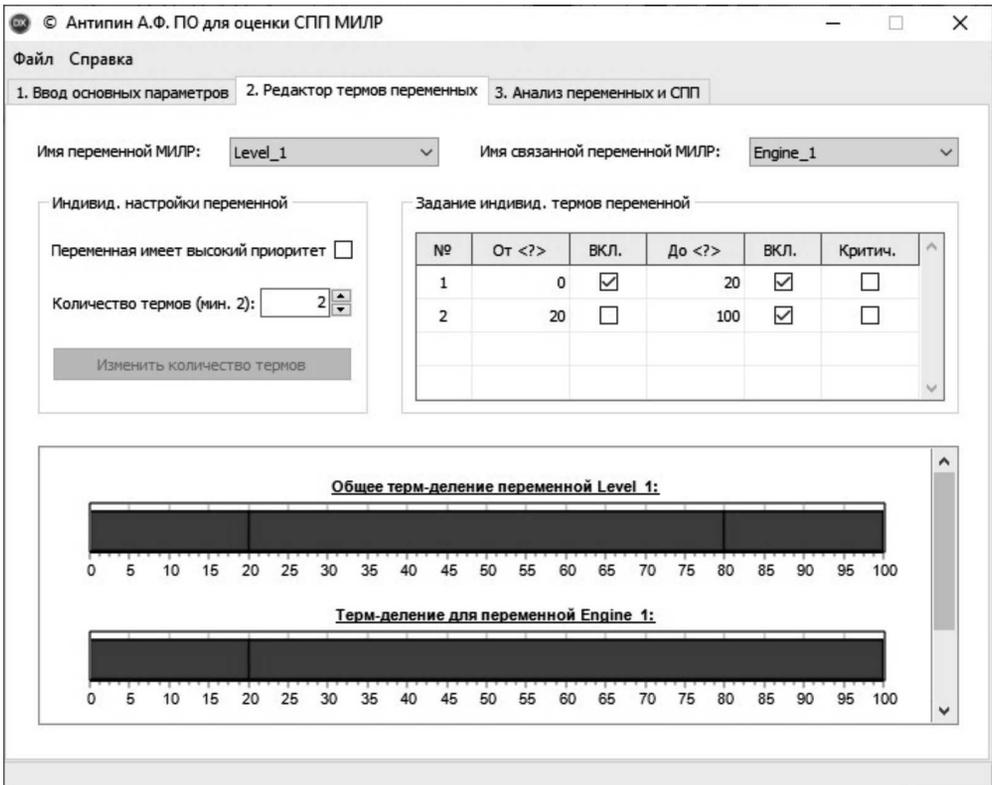


Рис. 4. Редактор термов непрерывных величин и их взаимосвязей.

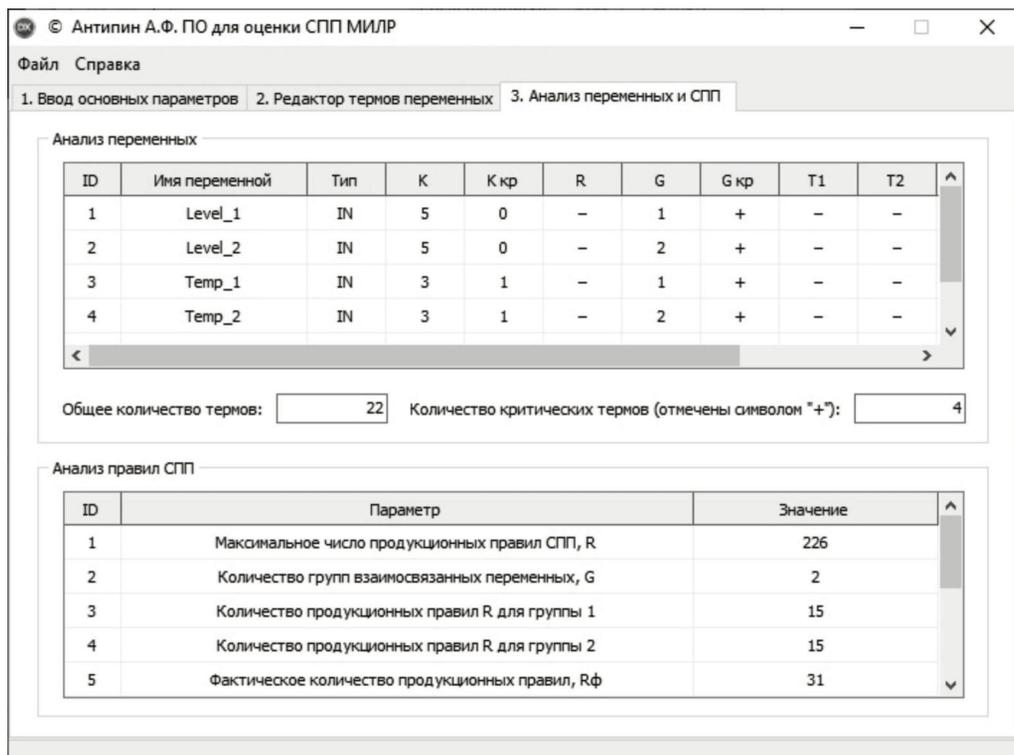


Рис. 5. Результаты анализа переменных и системы продукционных правил.

ПЛК-программы описанной в статье модификации многомерного нечеткого регулятора, расчета и анализа основных параметров, формирования продукционных правил и пр., что показано на рис. 4, 5.

При создании нового проекта нечеткого регулятора вводится следующая основная информация: количество входных и выходных переменных, их рабочие диапазоны значений и взаимные связи. Редактирование термов и ввод продукционных правил производится в специальном визуальном редакторе.

По окончании ввода данных имеется возможность выполнить анализ термделений переменных и системы продукционных правил. При этом оценивается общее количество термов и количество критических термов (для регулятора в целом и отдельно для каждой переменной), принадлежность переменных к конкретной группе (контур регулирования) и пр.

В отдельную таблицу выводятся значения параметров регулятора, которые зависят от содержимого системы продукционных правил (рис. 5):

- 1) максимальное R_{\max} и фактическое R'_{\max} количество продукционных правил, образующих систему продукционных правил регулятора;
- 2) количество групп взаимосвязанных переменных;
- 3) количество продукционных правил для каждой группы $R_{гр}^g$ и количество критических правил $R_{гр.кр}^g$ и пр.

```

VAR_INPUT
  NUM: INT;
  ...
  OUT_RM: ARRAY [0..R-1] OF BYTE;
END_VAR

VAR_OUTPUT
  ...
  OUT_M: REAL;
END_VAR

VAR
  ...
  OUT_MM: ARRAY [0..K(OUT_M)] OF REAL;
END_VAR
...

//ДЕФАЗЗИФИКАТОР M
IF BYTE_TO_INT(OUT_RM[NUM]) > 0
  THEN OUT_M := (OUT_MM[BYTE_TO_INT(OUT_RM[NUM]) - 1]
    + OUT_MM[BYTE_TO_INT(OUT_RM[NUM])]) / 2;
END_IF;

```

Рис. 6. Программный код функционального блока дефаззификации.

Настройки взаимосвязей и терм-деления переменных могут быть помещены в отдельный файл специального формата и загружены в систему автоматизированного проектирования, разработанную авторами и описанную в [19], которая содержит редактор системы продукционных правил, а также соответствующие инструменты для проверки ее полноты и правильности вводимых пользователем данных.

По окончании работы возможна генерация готовой программы регулятора и/или отдельных функциональных блоков (фаззификаторов, дефаззификаторов) в формате языков программирования SCL и AWL для программируемых логических контроллеров Simatic S7 с учетом требований стандарта ГОСТ Р МЭК 61131-7-2017 «Контроллеры программируемые. Ч. 7. Программирование нечеткого управления».

На рис. 6 представлен фрагмент сгенерированного программного кода функционального блока дефаззификации на языке программирования SCL для выходной непрерывной величины M , четкое значение которой находится путем вычисления середины интервала терма, активного в текущий момент времени [18].

4. Вычислительный эксперимент

Проведем вычислительный эксперимент, связанный с оценкой системы продукционных правил нечеткого регулятора при переходе от общего терм-деления входных переменных к локальному терм-делению.

Рассмотрим контур регулирования с четырьмя входными и двумя выходными непрерывными величинами x и y соответственно, схема взаимосвязей

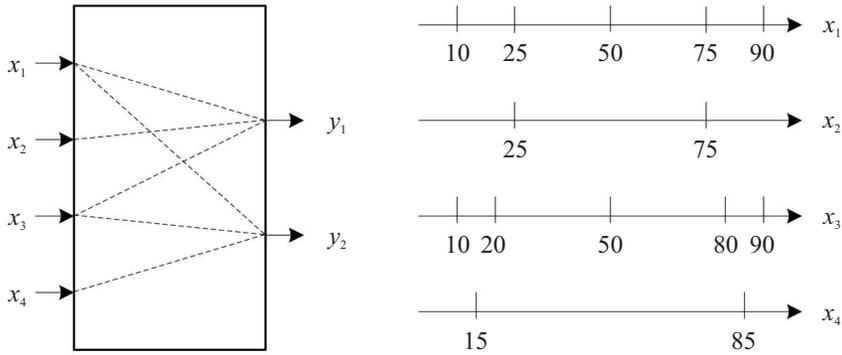


Рис. 7. Схема взаимосвязей и общее терм-деление входных переменных.

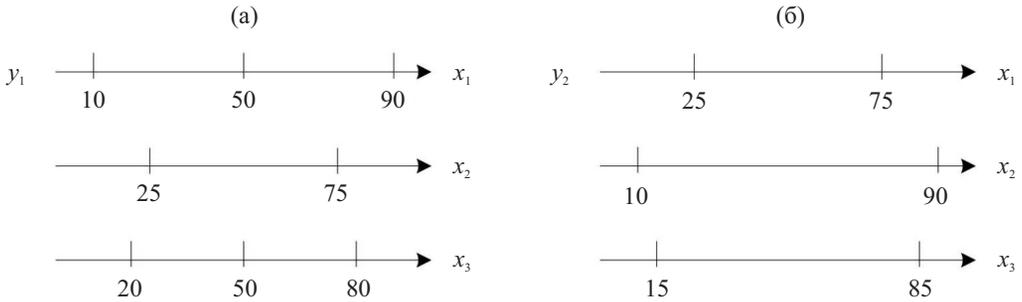


Рис. 8. Локальное терм-деление входных переменных x_1, x_2, x_3 и x_4 для выходных переменных (а) y_1 и (б) y_2 .

и общие терм-деления которых в рабочем диапазоне $[0; 100]$ (в упрощенном виде) приведены на рис. 7. Так как термы выходных величин не влияют на общий объем продукционных правил, то для выполнения расчетов они могут быть опущены.

Терм-множества входных непрерывных величин имеют вид:

$$\begin{aligned}
 T^{x_1} &= \{T_1(0), T_2(10), T_3(25), T_4(50), T_5(75), T_6(90; 100)\}, \\
 T^{x_2} &= \{T_1(0), T_2(25), T_3(75; 100)\}, \\
 T^{x_3} &= \{T_1(0), T_2(10), T_3(20), T_4(50), T_5(80), T_6(90; 100)\}, \\
 T^{x_4} &= \{T_1(0), T_2(15), T_3(85; 100)\}.
 \end{aligned}$$

Таким образом, $k(x_1) = 6, k(x_2) = 3, k(x_3) = 6, k(x_4) = 3$. Максимальное количество продукционных правил R_{\max} равно:

$$R_{\max} = \prod_{i=1}^4 k(x_i) = 324.$$

Данное количество правил будет предельно избыточным и сильно усложнит формирование системы продукционных правил регулятора. Однако определенные взаимосвязи позволяют на этапе разработки регулятора перейти от

общего к локальному терм-делению, а затем вводить продукционные правила отдельно для каждой выходной переменной.

Известно, что входные переменные x_1 , x_2 и x_3 связаны с выходной переменной y_1 , а x_1 , x_3 и x_4 – с y_2 . Их локальное терм-деление показано на рис. 8.

Терм-множества входных непрерывных величин x_1 , x_2 , x_3 при локальном терм-делении для выходной непрерывной величины y_1 будут иметь вид:

$$T^{x_1^{y_1}} = \{T_1(0), T_2(10), T_3(50), T_4(90; 100)\},$$

$$T^{x_2^{y_1}} = \{T_1(0), T_2(25), T_3(75; 100)\},$$

$$T^{x_3^{y_1}} = \{T_1(0), T_2(20), T_3(50), T_4(80; 100)\}.$$

Терм-множества входных непрерывных величин x_1 , x_3 , x_4 при локальном терм-делении для выходной непрерывной величины y_2 будут иметь вид:

$$T^{x_1^{y_2}} = \{T_1(0), T_2(25), T_3(75; 100)\},$$

$$T^{x_3^{y_2}} = \{T_1(0), T_2(10), T_3(90; 100)\},$$

$$T^{x_4^{y_2}} = \{T_1(0), T_2(15), T_3(85; 100)\}.$$

В табл. 2 приведены данные о количестве термов входных переменных x применительно к конкретным выходным переменным y .

Таблица 2. Данные для анализа терм-связей переменных

Выходная переменная	Количество термов входных переменных			
	x_1	x_2	x_3	x_4
y_1	4	3	4	–
y_2	3	–	3	3

Таким образом, с учетом известных взаимосвязей между входными и выходными величинами, максимальное количество продукционных правил R'_{\max} будет равно:

$$R'_{\max} = \sum_{j=1}^2 R_{\max}^{y_j} = k(x_1)k(x_2)k(x_3) + k(x_1)k(x_3)k(x_4) = 75.$$

Сокращение максимального числа продукционных правил ΔR_{\max} составит:

$$\Delta R_{\max} = \frac{R_{\max} - R'_{\max}}{R_{\max}} = 0,8,$$

или в 4,3 раз.

Данный процесс может быть автоматизирован при использовании описанного выше программного обеспечения. Результаты расчетов позволяют оценить сложность разработки, а также сохранить настройки взаимосвязей и

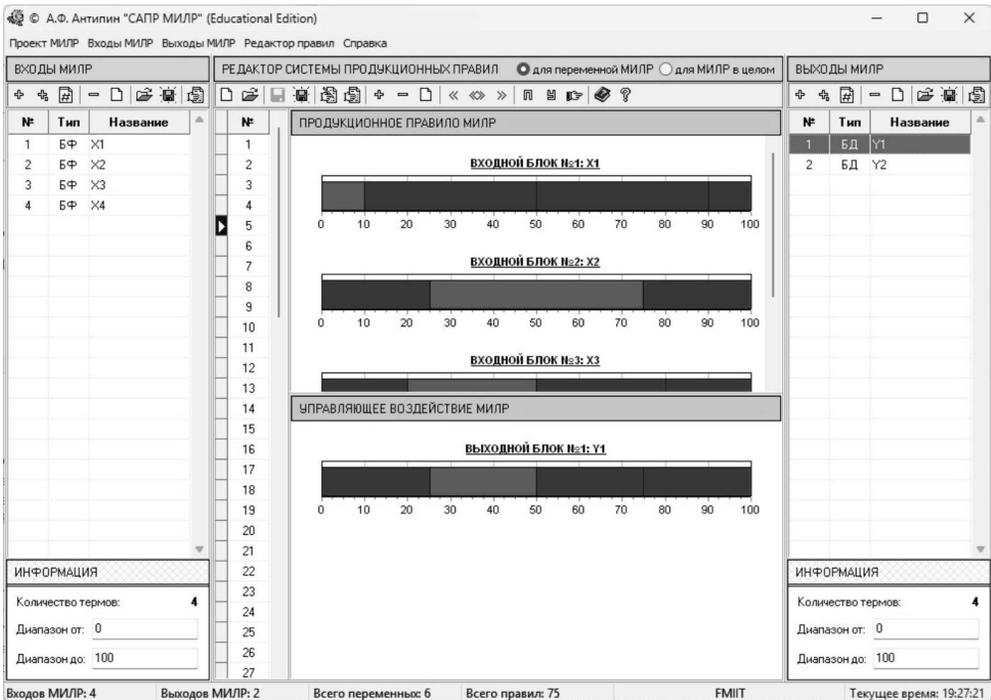


Рис. 9. Редактор системы автоматизированного проектирования.

терм-деления переменных для последующего формирования СПП в разработанной авторами системе автоматизированного проектирования, как показано на рис. 9.

Работа в единой среде визуального программирования упрощает разработку программы многомерного нечеткого регулятора со взаимосвязанными регулируемыми параметрами, а также расчет базовых характеристик.

5. Заключение

Использование при разработке АСУ описанной в статье модификации многомерного многосвязного нечеткого регулятора дает ряд преимуществ, связанных с гибкостью настройки термов переменных, сокращением максимального количества продукционных правил, заданием отдельной функции дефаззификации для каждой выходной переменной [20] и пр.

Разработанное авторами специальное программное обеспечение содержит инструменты для визуальной разработки нечеткого регулятора, оценки и анализа системы продукционных правил, генерации готовой программы для ПЛК Simatic S7. Данное программное обеспечение может применяться специалистами в сфере разработки АСУ сложными технологическими процессами, не имеющими адекватных математических моделей, и обеспечить создание надежных и эффективных ПЛК-программ нечетких регуляторов, сократив при этом срок их разработки.

СПИСОК ЛИТЕРАТУРЫ

1. *Dumitrescu C., Ciotirnae P., Vizitiu C.* Fuzzy Logic for Intelligent Control System Using Soft Computing Applications // *Sensors*. 2021. V. 21. No. 8. P. 2617. <https://doi.org/10.3390/s21082617>
2. *Васильев В.И., Ильясов Б.Г.* Интеллектуальные системы управления. Теория и практика. М.: Радиотехника, 2009.
3. *Kocian J., Koziorek J., Pokorný M.* Implementation of fuzzy logic control based on PLC // *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*. 2011. P. 1–8. <https://doi.org/10.1109/ETFA.2011.6059049>
4. *Nguyen L., Novák V.* Forecasting seasonal time series based on fuzzy techniques // *Fuzzy Sets and Systems*. 2019. V. 361. P. 114–129. <https://doi.org/10.1016/j.fss.2018.09.010>
5. *Пегат А.* Нечеткое моделирование и управление. М.: БИНОМ. Лаборатория знаний, 2009.
6. *Boukezzoula R., Coquin D., Nguyen T.L. et al.* Multi-sensor information fusion: Combination of fuzzy systems and evidence theory approaches in color recognition for the NAO humanoid robot // *Robot. Autonom. Syst.* 2018. V. 100. P. 302–316. <https://doi.org/10.1016/j.robot.2017.12.002>
7. *Melin P., Castillo O., Kacprzyk J. et al.* Fuzzy Logic in Intelligent System Design. Springer, 2018.
8. *Филимонов А.Б., Филимонов Н.Б.* Некоторые проблемные аспекты нечеткого ПИД регулирования // *Мехатроника, автоматизация, управление*. 2018. Т. 19. № 12. С. 762–769. <https://doi.org/10.17587/mau/19.762-769>
9. *Быковцев Ю.А., Лохин В.М.* Оценка точности системы управления с нечетким ПИД регулятором на основе аппроксимации статической характеристики регулятора // *Мехатроника, автоматизация, управление*. 2021. Т. 22. № 12. С. 619–624. <https://doi.org/10.17587/mau.22.619-624>
10. *Bobyry M.V., Emelyanov S.G.* A nonlinear method of learning neuro-fuzzy models for dynamic control systems // *Appl. Soft. Comput.* 2020. V. 8. P. 106030. <https://doi.org/10.1016/j.asoc.2019.106030>
11. *Karakuzu C., Karakaya F., Pavuslu M.A.* FPGA implementation of neuro-fuzzy system with improved PSO learning // *Neural Networks*. 2016. V. 76. P. 128–140. <https://doi.org/10.1016/j.neunet.2016.02.004>
12. *Zhirabok A.N., Suvorov A.Yu., Shumskii A.E.* Robust diagnosis of discrete systems with delay: Logic-dynamical approach // *J. Comput. Syst. Sci. Int.* 2014. V. 53. No. 1. P. 47–62. <https://doi.org/10.1134/S1064230714010134>
13. *Verma O.P., Jain V., Gumber R.* Simple fuzzy rule based edge detection // *J. Inform. Proc. Syst.* 2013. V. 9. No. 4. P. 575–591. <https://doi.org/10.3745/JIPS.2013.9.4.575>
14. *Yongming Li, Yali Li, Zhanyou M.* Computation tree logic model checking based on possibility measures // *Fuzzy Sets and Systems*. 2015. V. 262. P. 44–59. <https://doi.org/10.48550/arXiv.1401.4658>
15. *Игнатъев В.В., Соловьев В.В.* Метод синтеза нечетких регуляторов на основе кластеризации // *Программные продукты и системы*. 2021. Т. 34. № 4. С. 597–607. <https://doi.org/10.15827/0236-235X.136.597-607>
16. *Богачёв И.С., Коломыцев В.Г.* Методика настройки нечеткого регулятора с нелинейной характеристикой Р-типа средствами Matlab // *Современные наукоемкие технологии*. 2020. № 11-1. С. 9–14. <https://doi.org/10.17513/snt.38330>

17. *Buriboev A., Kang H.K., Ko M.-C. et al.* Application of Fuzzy Logic for Problems of Evaluating States of a Computing System // Appl. Sci. 2019. V. 9. No. 15. P. 3021. <https://doi.org/10.3390/app9153021>
18. *Антипин А.Ф., Антипина Е.В.* Многомерный многосвязный нечеткий интервально-логический регулятор // Мехатроника, автоматизация, управление. 2024. Т. 25. № 2. С. 72–78. <https://doi.org/10.17587/mau.25.72-78>
19. *Антипин А.Ф., Антипина Е.В.* Среда визуальной разработки программ многомерных нечетких интервально-логических регуляторов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23. № 5. С. 1041–1049. <https://doi.org/10.17586/2226-1494-2023-23-5-1041-1049>
20. *Антипин А.Ф., Антипина Е.В.* Моделирование технологических процессов и объектов на основе нечеткой логики с использованием многомерных интервально-логических регуляторов // Автометрия. 2020. Т. 56. № 1. С. 83–92. <https://doi.org/10.15372/AUT20200109>

Статья представлена к публикации членом редколлегии Р.А. Мунасыповым.

Поступила в редакцию 25.03.2024

После доработки 16.02.2025

Принята к публикации 26.02.2025

Интеллектуальные системы управления, анализ данных

© 2025 г. Ю.С. ПОПКОВ, д-р техн. наук. (popkov@isa.ru)

(Федеральный исследовательский центр

“Информатика и управление” РАН, Москва;

Институт проблем управления им. В.А. Трапезникова РАН, Москва),

А.Ю. ПОПКОВ, канд. техн. наук (apopkov@isa.ru),

Ю.А. ДУБНОВ (yury.dubnov@phystech.edu)

(Федеральный исследовательский центр

“Информатика и управление” РАН, Москва)

МЕТОДЫ РАНДОМИЗИРОВАННОГО МАШИННОГО ОБУЧЕНИЯ ДЛЯ ГЕНЕРАЦИИ АНСАМБЛЕЙ СЛУЧАЙНЫХ ДАННЫХ С ЗАДАННЫМИ ЧИСЛОВЫМИ ХАРАКТЕРИСТИКАМИ¹

Рассматривается задача генерации случайных ансамблей данных с заданными числовыми характеристиками. Развивается метод ее решения, использующий процедуры рандомизированного машинного обучения, которые строятся на последовательности задач функционального энтропийно-линейного программирования. В качестве ограничений в них рассматриваются нормированные моменты. Задача генерации сводится к системе нелинейных уравнений с интегральными компонентами. Адаптируется разработанный авторами асимптотический аналитический метод преобразования указанных уравнений к системе уравнений с полиномиальной левой частью. Развитые аналитические методы применены для генерации случайных ансамблей данных, прогнозирующих динамику стоимости финансовых активов.

Ключевые слова: энтропия, рандомизация, нормированные моменты, случайные ансамбли, машинное обучение, степенные ряды, аналитические функции, многомерные интегралы, многомерные полиномы.

DOI: 10.31857/S0005231025070065, EDN: JSAVQB

1. Введение

Методы машинного обучения, использующиеся для решения большого количества практических и научно-практических задач, основаны на использовании данных с требуемыми свойствами. В контексте методов машинного обучения, основанных на статистическом подходе, требования к данным реализуются в виде их вероятностных и числовых характеристик [1]. Помимо

¹ Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (проект № 075-15-2024-544).

традиционных задач классификации и прогнозирования [2–7], можно также отметить ряд таких областей, как, например, тестирование программного обеспечения [8–10] и контроля знаний [11–13], где данные либо служат для обучения моделей процессов, либо используются для статистического оценивания гипотез. Параллельно развивается так называемый «сценарный подход», в рамках которого для параметризованной модели процесса, чаще всего экспертно, составляются массивы параметров и генерируются соответствующие им ансамбли данных [14, 15]. В любом случае свойства данных должны учитываться как для корректного применения соответствующих теоретических методов и подходов, так и при практическом применении разработанных и обученных моделей.

Необходимо также отметить, что в современных условиях во многих областях недостатка в данных нет, так как они в избытке собираются и накапливаются в автоматическом режиме при функционировании информационных систем и различных технических устройств. Но в то же время по-прежнему существует проблема генерации необходимых данных (с требуемыми свойствами) в целях разработки, обучения и тестирования как методов, так и все тех же устройств. Конечно, в различных областях и задачах требуемые свойства могут быть формализованы различными способами. В данной работе такая формализация проводится в рамках вероятностной концепции, в частности, под данными понимаются выборки, обладающие подходящими функциями плотности распределения вероятностей (ПРВ). При этом предполагается, что подходящие распределения можно тем или иным способом сэмплировать, т.е. трансформировать в соответствующие случайные последовательности.

Таким образом, генерирование подходящих данных основано на восстановлении с учетом заданных требований, оптимальных в принятом смысле функций ПРВ. Поскольку при формулировке указанных требований имеет место довольно большая неопределенность, то естественным критерием оптимизации ПРВ является информационная энтропия [16–20]. Однако этого недостаточно, и часто требуется учитывать какие-то дополнительные свойства функций ПРВ. Некоторые из них можно сформулировать в терминах числовых характеристик, а именно, моментов, семиинвариантов и др. Следовательно, задача генерирования желательных функций ПРВ сводится к условной максимизации функционала информационной энтропии. По своему формальному представлению эта задача близка к некоторым математическим моделям рандомизированного машинного обучения (РМО), изучаемым в [21, 22]. Определенные отличия связаны с системой ограничений в рассматриваемой задаче.

Настоящая работа направлена на дальнейшее развитие методов рандомизированного машинного обучения в следующих направлениях:

- рандомизированное обучение при дополнительных ограничениях моментного типа;

- адаптация аналитического метода решения нелинейных уравнений с интегральными компонентами к рассматриваемой задаче;
- исследование эффективности развитых методов на примере задачи прогнозирования стоимости финансовых активов.

Отметим, что задачи РМО и задачи генерирования желаемых распределений сводятся к решению существенно нелинейных уравнений, содержащих так называемые интегральные компоненты. Они представляют собой многомерные интегралы на простых множествах (параллелепипедах) с экспоненциальными подынтегральными функциями, параметризованные множителями Лагранжа. С использованием аналитических свойств экспоненциальных функций указанные интегралы аппроксимируются параметризованными интегралами от многомерных полиномов, вычисление которых выполняется аналитически.

В результате указанных трансформаций многомерных параметризованных интегралов система нелинейных уравнений аппроксимируется системой уравнений с полиномиальной левой частью. Для их решения применяется аналитический метод, основанный на представлении решения в виде абстрактного степенного ряда [23, 24].

С учетом изложенного, статья организована следующим образом. В разделе 2 формулируется общая задача, на решение которой направлен предлагаемый метод, в разделе 3 приводится описание подхода к ее решению и необходимый теоретический инструментарий, раздел 4 посвящен результатам решения задачи прогнозирования стоимости финансового актива, в разделе 5 обсуждаются особенности полученных результатов и направления дальнейших исследований, раздел 6 посвящен концентрированному обзору результатов статьи.

2. Постановка задачи

Рассмотрим случайную последовательность $u[n]$, где $n \in \mathcal{N} = \overline{1, N}$, для которой задана матрица числовых характеристик $Y_{(s \times N)}$, элементы которой характеризуют значения *нормированных моментов*² порядка $k = \overline{1, s}$ в точках наблюдения n :

$$(1) \quad Y_{(s \times N)} = \left\{ \left(\mathcal{M}\{u^k[n]\} \right)^{1/k} \right\} = \left\{ y^k[n] \mid k = \overline{1, s}, n = \overline{1, N} \right\}.$$

В частности, такого рода информация часто встречается в задачах прогнозирования ценообразования торгуемых финансовых инструментов [26–28].

Задачу генерации данных с заданными свойствами можно сформулировать следующим образом.

² Здесь в качестве числовых характеристик рассматриваются нормированные моменты, но также могут быть избраны семинварианты или математические ожидания непрерывных функций от случайных последовательностей

В каждой точке n наблюдения требуется генерировать ансамбли (наборы) \mathcal{Z}_n случайных последовательностей $z[n]$ с s нормированными моментами

$$(2) \quad m^{(k)}[n] = \left(\mathcal{M}\{z^k[n]\} \right)^{1/k}, \quad k = \overline{1, s},$$

равными заданным нормированным моментам $y^k[n]$, определяемым (1).

Генератором ансамбля \mathcal{Z}_n является «вход–выход»-модель или «авто»-модель³. В том и другом варианте параметры $\mathbf{a} \in \mathcal{A} \subset R^r$ в моделях случайные, интервальные:

$$(3) \quad \mathbf{a} \in \mathcal{A} \subset R^r, \quad \mathcal{A} = [\mathbf{a}^-, \mathbf{a}^+],$$

и характеризуются для каждого момента времени n функцией ПРВ $P_{(n)}(\mathbf{a})$, которая предполагается непрерывно дифференцируемой.

В зависимости от наличия априорной информации о происхождении данных используются либо статическая, либо динамическая «вход–выход»-модель.

Статическая модель – генератор случайных последовательностей $z[n]$ – характеризуется нелинейной дифференцируемой функцией φ с параметрами \mathbf{a} :

$$(4) \quad z[n|\mathbf{a}] = \varphi(\mathbf{x}[n]|\mathbf{a}), \quad n = \overline{1, N},$$

где $\mathbf{x}[n] = \{x_1[n], \dots, x_m[n]\}$ является входом модели, а z – выходом.

В дальнейшем символ «|» означает, что для каждого момента времени n реализуются случайные параметры \mathbf{a} с ПРВ $P_{(n)}(\mathbf{a})$.

Динамическая модель характеризуется нелинейным непрерывным функционалом \mathcal{B} :

$$(5) \quad z[n|\mathbf{a}] = \mathcal{B}[\mathbf{x}[\tau], n - p \leq \tau \leq n | \mathbf{a}], \quad n = \overline{1, N},$$

где p – «память» модели, определяющая количество прошлых значений входа, влияющих на выход в данный момент времени. Параметры \mathbf{a} также случайные и интервальные, определяемые (3).

Сформулированная выше задача решается в два этапа. На первом этапе определяются оптимальные вероятностные характеристики случайных параметров, а именно, их функции ПРВ $P_{(n)}(\mathbf{a})$ для всех $n = \overline{1, N}$ при использовании статической модели (4) и для всех $n = \overline{1 - p, N}$ при использовании динамической модели (5). Второй этап состоит в сэмплировании этих ПРВ, т.е. в трансформации их в соответствующие ансамбли \mathcal{Z}_n .

³ Здесь будет использоваться «вход–выход»-модель, в примере – «авто»-модель, описываемая в виде разностных уравнений.

3. Материалы и методы

3.1. Оптимизация функций ПРВ параметров модели

Для решения задачи первого этапа воспользуемся методологией рандомизированного машинного обучения [22], в которой генерация указанных последовательностей осуществляется моделью типа «вход–выход» с оптимизированными по критерию информационной энтропии случайными параметрами.

Для удобства дальнейшего изложения введем следующие обозначения:

- вектор заданных нормированных моментов для каждого n

$$(6) \quad \mathbf{y}^{(n)} = \{y_1[n], \dots, y_s[n]\};$$

- вектор выхода модели для каждого n

$$(7) \quad \mathbf{z}^{(n)}(\mathbf{a}) = \{z_1[n | \mathbf{a}], \dots, z_s[n | \mathbf{a}]\};$$

- вектор нормированных моментов на выходе модели для каждого n

$$(8) \quad \mathbf{m}^{(n)} = \left\{ \int_{\mathcal{A}} P_{(n)}(\mathbf{a}) z[n | \mathbf{a}] d\mathbf{a}, \dots, \left(\int_{\mathcal{A}} P_{(n)}(\mathbf{a}) z^s[n | \mathbf{a}] d\mathbf{a} \right)^{1/s} \right\}.$$

Согласно [22], задача определения оптимальных ПРВ $P_{(n)}(\mathbf{a})$ может быть сформулирована для каждого n в виде максимизации функционала информационной энтропии:

$$(9) \quad \mathcal{H}_{(n)}[P_{(n)}(\mathbf{a})] = - \int_{\mathcal{A}} P_{(n)}(\mathbf{a}) \ln P_{(n)}(\mathbf{a}) d\mathbf{a} \Rightarrow \max$$

при ограничениях:

- нормировки

$$(10) \quad \int_{\mathcal{A}} P_{(n)}(\mathbf{a}) d\mathbf{a} = 1,$$

- балансов нормированных моментов выхода модели с данными

$$(11) \quad \mathbf{m}^{(n)} = \mathbf{y}^{(n)}.$$

Полагая, что функции ПРВ непрерывно-дифференцируемые, решение, параметризованное множителями Лагранжа $\boldsymbol{\lambda}^{(n)} = \{\lambda_1^{(n)}, \dots, \lambda_s^{(n)}\}$, имеет вид (см. [22])

$$(12) \quad P_{(n)}^*(\mathbf{a}) = \frac{\exp(-\langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a}) \rangle)}{\mathbb{P}_{(n)}(\boldsymbol{\lambda}^{(n)})},$$

где

$$(13) \quad \mathbb{P}_{(n)}(\boldsymbol{\lambda}^{(n)}) = \int_{\mathcal{A}} \exp\left(-\langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a}) \rangle\right) d\mathbf{a},$$

$\langle \bullet, \bullet \rangle$ обозначает скалярное произведение.

Множители Лагранжа $\boldsymbol{\lambda}^{(n)}$ определяются следующей системой из s уравнений, называемыми *балансовыми*:

$$(14) \quad \int_{\mathcal{A}} \exp\left(-\langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a}) \rangle\right) \left[\mathbf{z}^{(n)}(\mathbf{a}) - \mathbf{y}^{(n)}\right] d\mathbf{a} = \mathbf{0}.$$

Эта система уравнений определяет векторы множителей Лагранжа $\boldsymbol{\lambda}^{(n)}$ для каждого моменте n наблюдений из интервала $[1, N]$.

Задача второго этапа, а именно трансформация энтропийно-оптимальной функции ПРВ в соответствующую случайную последовательность, может быть реализована методами, изложенными в [25].

3.2. Аналитический метод решения балансовых уравнений

Из уравнения (14) следует, что для определения множителей Лагранжа требуется вычисление многомерных интегралов, а затем решение сформированных нелинейных уравнений. Разработанный аналитический метод позволяет объединить эти оба этапа.

В рассматриваемой задаче существуют некоторые полезные свойства, которыми можно воспользоваться для формирования приближенного аналитического метода. К ним относятся простая область определения интегралов, а именно, параллелепипед, и подынтегральные функции – экспоненты от непрерывных функций.

Заметим, что экспоненциальная функция аналитическая, а множители Лагранжа и выход модели ограничены. Поэтому имеет место полиномиальная аппроксимация степени q следующего вида:

$$(15) \quad \exp(-v_{(n)}) = \sum_{h=0}^q \frac{(-1)^h}{h!} v_{(n)}^h,$$

где

$$(16) \quad v_{(n)} = \langle \boldsymbol{\lambda}^{(n)}, \mathbf{z}^{(n)}(\mathbf{a}) \rangle < \pm M < \infty.$$

Здесь ⁴

$$(17) \quad (v_{(n)})^h = \sum_{i_j \geq 0; \sum_{j=1}^s i_j = h} A_{i_1, \dots, i_h}^{(h)} (\lambda_1^{(n)})^{i_1} \dots (\lambda_s^{(n)})^{i_h} \times \\ \times (z^1[n | \mathbf{a}])^{i_1} \dots (z^s[n | \mathbf{a}])^{i_h},$$

$$(18) \quad A_{i_1, \dots, i_h}^{(h)} = \frac{h!}{i_1! \dots i_h!}.$$

Тогда с использованием этих обозначений система (14) примет следующий вид:

$$(19) \quad \sum_{h=0}^q \frac{(-1)^h}{h!} \sum_{(i_1, \dots, i_h)=1}^s \lambda_{i_1}^{(n)} \dots \lambda_{i_h}^{(n)} u_{i_1, \dots, i_h; k}^{(n)} - v_n^k = 0,$$

где

$$(20) \quad u_{i_1, \dots, i_h; k}^{(n)} = \int_{\mathcal{A}} z^{i_1}[n | \mathbf{a}] \dots z^{i_h}[n | \mathbf{a}] \left(z^k[n | \mathbf{a}] - y^{(k)}[n] \right) d\mathbf{a}, \\ i_1, \dots, i_h, k = \overline{1, s}.$$

Здесь индексы i_1, \dots, i_h принимают значения в интервале $[1, s]$. Эта система содержит s переменных – множителей Лагранжа, и каждое уравнение представляет собой многомерный полином степени q .

Поскольку предполагается, что модель характеризуется непрерывной функцией (или функционалом), то она представима многомерным полиномом, и, следовательно, многомерные интегралы в (20) трансформируются в произведение одномерных интегралов, вычисляемых аналитически. Поэтому балансовая система в виде (19) имеет полиномиальную левую часть.

Рассмотрим семейство по параметру $\varepsilon \in [0, 1]$ систем следующего вида:

$$(21) \quad - \sum_{i_1=1}^s \lambda_{i_1}^{(n)} u_{i_1, k}^{(n)} + \varepsilon \sum_{h=1}^q \frac{(-1)^h}{h!} \sum_{\substack{(i_1, \dots, i_h)=1 \\ \sum_{j=1}^h i_j = h}}^s \lambda_{i_1}^{(n)} \dots \lambda_{i_h}^{(n)} u_{i_2, \dots, i_h; k}^{(n)} - v_n^k = 0.$$

При $\varepsilon = 0$ имеем так называемую *базовую* систему, линейную с квадратной матрицей $U^{(n)} = [u_{i_1, k}^{(n)}, | (i_1, k) = \overline{1, s}]$:

$$(22) \quad U^{(n)} \boldsymbol{\lambda}^{(n)} = -\mathbf{v}_n,$$

где $\mathbf{v}_n = \{v_n^1, \dots, v_n^s\}$.

⁴ Выражение в правой части равенства (17) представляет собой полиномы Ньютона степени h [31].

Если $\det U^{(n)} \neq 0$, то множители Лагранжа, образующие базовое решение, равны

$$(23) \quad \lambda_{(\bullet)}^{(n)} = -[U^{(n)}]^{-1} \mathbf{v}_n.$$

Представим решение системы (21) в виде абстрактного степенного ряда по параметру ε [23, 24]:

$$(24) \quad \lambda_{k,\star}^{(n)} = \lambda_{k,\bullet}^{(n)} + \varepsilon \lambda_{k,I}^{(n)} + \varepsilon^2 \lambda_{k,II}^{(n)} + \dots, \quad k = \overline{1, s},$$

где $\lambda_{k,I}^{(n)}, \lambda_{k,II}^{(n)}, \dots$ – I-ая, II-ая, ... коррекции базового решения.

Для последовательного определения коррекций, используя метод неопределенных коэффициентов [23], будем иметь вектор первой коррекции

$$(25) \quad \ast \lambda_I^{(n)} = -[U^{(n)}]^{-1} \mathbf{b}_{(1)}^{(n)}(\lambda_{(\bullet)}^{(n)}),$$

где

$$(26) \quad \mathbf{b}_{(I)}^{(n)}(\lambda_{(\bullet)}^{(n)}) = \left\{ \frac{1}{2} \sum_{(i_1, i_2)=1}^s \lambda_{i_1, \bullet}^{(n)} \lambda_{i_2, \bullet}^{(n)} u_{i_1, i_2, 1}^{(n)}, \dots, \frac{1}{2} \sum_{(i_1, i_2)=1}^s \lambda_{i_1, \bullet}^{(n)} \lambda_{i_2, \bullet}^{(n)} u_{i_1, i_2, s}^{(n)} \right\};$$

вектор второй коррекции

$$(27) \quad \ast \lambda_{II}^n = -[U^{(n)}]^{-1} \mathbf{b}_{(II)}^{(n)}(\lambda_{(\bullet)}^{(n)}, \lambda_{(I)}^n),$$

где

$$(28) \quad \left\{ \frac{1}{2} \sum_{(i_1, i_2)=1}^s \lambda_{i_1, I}^{(n)} \lambda_{i_2, \bullet}^{(n)} u_{i_1, i_2, k}^{(n)} + \frac{1}{3!} \sum_{(i_1, i_2, i_3)=1}^s \lambda_{i_1, \bullet}^{(n)} \lambda_{i_2, \bullet}^{(n)} \lambda_{i_3, \bullet}^{(n)} u_{i_1, i_2, i_3, k}^{(n)} \right\}.$$

Таким образом, решение балансовой системы (21) представимо в виде

$$(29) \quad \lambda_{\star}^{(n)} = \lambda_{(\bullet)}^{(n)} + \ast \lambda_I^{(n)} + \ast \lambda_{II}^{(n)} + \dots.$$

3.3. Модель ценообразования финансового актива

Рассмотрим процесс формирования стоимости финансового актива в процессе торговых сессий и применим для его прогнозирования развиваемый выше метод генерации случайных данных с заданными числовыми характеристиками.

Существует теоретический консенсус, что цена финансового инструмента в каждый момент времени является продуктом балансировки реального спроса и реального предложения. Однако трейдер ориентируется на ожидаемые спрос и предложение, которые могут существенно отличаться от реальных. В этих условиях важное значение приобретает прогноз стоимости при достаточно высоком уровне неопределенности.

Поэтому кажется естественной попытка максимизации информационной энтропии как меры неопределенности на обучающих ретроспективных данных, содержащих значения средних и дисперсий стоимости.

3.3.1. Модель динамики стоимости

Рассмотрим автономную модель в виде линейного разностного уравнения порядка p со случайными параметрами интервального типа (интервалы $[d, w]$ всех параметров одинаковые):

$$(30) \quad C[t] = \sum_{i=1}^p a_i C[t-i], \quad t \in \mathcal{T}, \quad a_i \in \mathcal{A}_i = [d, w], \quad \mathcal{A}^p = \prod_{i=1}^p \mathcal{A}_i.$$

Этой моделью будем пользоваться на этапах *обучения* $\mathcal{T}_{ln} = [t_0, t_0 + p]$ и *прогнозирования* $\mathcal{T}_{pr} = [t_0 + p + 1, t_0 + p + 1 + t_{pr}]$, где t_{pr} — одно- или двухдневные прогнозы (по аналогии могут быть рассмотрены и более длинные прогнозы).

Вероятностные свойства параметров характеризуются непрерывно дифференцированной функцией ПРВ $P_t(\mathbf{a})$.

Предполагается, что по результатам реальных торгов в каждой сессии t формируются два стоимостных индикаторов: средняя цена $m_C^*[t]$ и второй момент $D_C^*[t]$ цены в качестве характеристики ее волатильности:

$$(31) \quad D_C^*[t] = (V_C^*[t])^2 + (m_C^*[t])^2,$$

где $V_C^*[t]$ — оценка средне-квадратического отклонения цены, которая строится из величины максимального и минимального ее отклонения.

3.3.2. Данные

Пусть в реальных торгах на интервале $[t_0 - p, t_0 - 1]$ имеются данные об изменении средней стоимости:

$$m_C^*[t_0 - p], m_C^*[t_0 - p + 1], \dots, m_C^*[t_0 - 1]$$

и оценке второго момента:

$$D_C^*[t_0 - p], D_C^*[t_0 - p + 1], \dots, D_C^*[t_0 - 1].$$

Используя эти данные, получим значения стоимости на интервале обучения \mathcal{T}_{ln} , которые *генерируются моделью* (30) по данным на «прошлом» интервале $[t_0 - p, t_0 - 1]$:

$$(32) \quad \mathbb{C}[\mathbf{a} | t] = \sum_{i=1}^p a_i m_C^*[t-i], \quad t \in \mathcal{T}_{ln}.$$

3.3.3. Энтропийно-оптимальное оценивание функций ПРВ параметров модели

Для оптимизации функции ПРВ $P_t(\mathbf{a})$ воспользуемся методикой, развитой для рандомизированного машинного обучения [22] и изложенной в (9)–(11),

в рамках которой решаются следующие задачи для каждой торговой сессии $t \in \mathcal{T}_{ln}$:

$$(33) \quad \mathcal{H}_t[P_t(\mathbf{a})] = - \int_{\mathcal{A}^p} P_t(\mathbf{a}) \ln P_t(\mathbf{a}) d\mathbf{a} \Rightarrow \max_{P(\mathbf{a})}$$

при ограничениях:

$$(34) \quad \int_{\mathcal{A}^p} P_t(\mathbf{a}) d\mathbf{a} = 1,$$

$$(35) \quad \int_{\mathcal{A}^p} P_t(\mathbf{a}) \mathbb{C}[\mathbf{a} | t] d\mathbf{a} = m_C^*[t], \quad \int_{\mathcal{A}} P_t(\mathbf{a}) \mathbb{C}^2[\mathbf{a} | t] d\mathbf{a} = D_C^*[t], \quad t \in \mathcal{T}_{ln},$$

где $\mathbb{C}_t[\mathbf{a} | t]$ определяются равенствами (32).

Задача (33)–(35) согласно (12), (13) имеет аналитическое решение:

$$(36) \quad P_t^*(\mathbf{a}) = \frac{\exp\left(-\lambda_1^{(t)} \mathbb{C}[\mathbf{a} | t] - \lambda_2^{(t)} \mathbb{C}^2[\mathbf{a} | t]\right)}{\mathcal{P}_t(\lambda_1^{(t)}, \lambda_2^{(t)})}, \quad t \in \mathcal{T}_{ln}.$$

$$\mathcal{P}_t^*(\lambda_1^{(t)}, \lambda_2^{(t)}) = \int_{\mathcal{A}^p} \exp\left(-\lambda_1^{(t)} \mathbb{C}[\mathbf{a} | t] - \lambda_2^{(t)} \mathbb{C}^2[\mathbf{a} | t]\right) d\mathbf{a},$$

Множители Лагранжа $\lambda_1^{(t)}, \lambda_2^{(t)}$ определяются решением следующих двух балансовых уравнений:

$$(37) \quad \int_{\mathcal{A}^p} \exp\left(-\lambda_1^{(t)} \mathbb{C}[\mathbf{a} | t] - \lambda_2^{(t)} \mathbb{C}^2[\mathbf{a} | t]\right) (\mathbb{C}[\mathbf{a} | t - 1] - m_C^*[t]) d\mathbf{a} = 0,$$

$$\int_{\mathcal{A}^p} \exp\left(-\lambda_1^{(t)} \mathbb{C}[\mathbf{a} | t] - \lambda_2^{(t)} \mathbb{C}^2[\mathbf{a} | t]\right) (\mathbb{C}^2[\mathbf{a} | t - 1] - D_C^*[t]) d\mathbf{a} = 0, \quad t \in \mathcal{T}_{ln}.$$

Согласно развитому в [29] методу воспользуемся аппроксимациями экспоненты полиномом степени 2:

$$(38) \quad \exp(x) \approx \left(1 - \lambda_1^{(t)} \mathbb{C}[\mathbf{a} | t] + \frac{1}{2} \left(\lambda_1^{(t)} \mathbb{C}[\mathbf{a} | t]\right)^2\right) \mathbb{C}[\mathbf{a} | t],$$

$$\exp(y) \approx \left(1 - \lambda_2^{(t)} \mathbb{C}^2[\mathbf{a} | t] + \frac{1}{2} \left(\lambda_2^{(t)} \mathbb{C}^2[\mathbf{a} | t]\right)^2\right) \mathbb{C}^2[\mathbf{a} | t].$$

Используя аппроксимации (38), получим следующие формы балансовых уравнений:

$$(39) \quad \lambda_1 B_1(t) + \lambda_2 B_2(t) + \lambda_1^2 B_3(t) + \lambda_2^2 B_4(t) + \lambda_1 \lambda_2 B_5(t) = B_0(t),$$

$$\lambda_1 Z_1(t) + \lambda_2 Z_2(t) + \lambda_1^2 Z_3(t) + \lambda_2^2 Z_4(t) + \lambda_1 \lambda_2 Z_5(t) = Z_0(t),$$

$$t \in \mathcal{T} = [t_0, t_0 + p].$$

В первом из этих уравнений коэффициенты имеют вид:

$$(40) \quad \begin{aligned} B_0(t) &= A m_C^*[t] - I_{p,1}^{(t)}, & B_1(t) &= m_C^*[t] I_{p,1}^{(t)} - I_{p,2}^{(t)}, \\ B_2(t) &= m_C^*[t] I_{p,2}^{(t)} - I_{p,3}^{(t)}, & B_3(t) &= -\frac{1}{2} B_2^{(2,3)}(t), \\ B_4(t) &= -\frac{1}{2} \left(m_C^*[t] I_{p,4}^{(t)} - I_{p,5}^{(t)} \right) & B_5(t) &= -\left(m_C^*[t] I_{p,3}^{(t)} - I_{p,4}^{(t)} \right), \end{aligned}$$

Коэффициенты во втором уравнении имеют вид:

$$(41) \quad \begin{aligned} Z_0(t) &= A D_C^*[t] - I_{p,2}^{(t)}, & Z_1(t) &= D_C^*[t] I_{p,1}^{(t)} - I_{p,3}^{(t)}, \\ Z_2(t) &= D_C^*[t] I_{p,2}^{(t)} - I_{p,4}^{(t)}, & Z_3(t) &= -\frac{1}{2} Z_2^{(2,4)}(t), \\ Z_4(t) &= -\frac{1}{2} \left(D_C^*[t] I_{p,4}^{(t)} - I_{p,6}^{(t)} \right), & Z_5(t) &= -\left(D_C^*[t] I_{p,3}^{(t)} - I_{p,5}^{(t)} \right). \end{aligned}$$

В этих выражениях

$$(42) \quad \begin{aligned} A &= \int_{\mathcal{A}} d\mathbf{a} = (w - d)^p, \\ I_{p,n}^{(t)}(k_1, \dots, k_n) &= \underbrace{\int_d^w \dots \int_d^w}_{p} \mathbb{C}^n[\mathbf{a} | t] d\mathbf{a} = \\ &= \sum_{k_j \geq 0; \sum_{j=1}^n k_j = n} \frac{n!}{k_1! \dots k_n!} \left(\underbrace{\int_d^w \dots \int_d^w}_{p} a_1^{k_1} \dots a_p^{k_n} da_1 \dots da_p \right) \times \\ &\quad \times (m_C^*[t])^{k_1} \dots (m_C^*[t - p])^{k_n}, \quad n = \overline{0, \bar{p}}. \end{aligned}$$

4. Прогнозирование стоимости финансового актива

Экспериментальное прогнозирование предложенного метода будем проводить для задач *одно-* и *двухдневного прогнозирования* средней стоимости и дисперсии финансового актива – акций ПАО «Газпром» в течение 2020 г. на Московской бирже.

Рассматриваются 12 торговых сессий, каждая на начало месяца. Для удобства все вычисления проводятся в условных единицах, соответствующих 1/1000 исходной цены. Данные о стоимости приведены в табл. 1.

4.1. Обучение модели стоимости

Модель стоимости (30) имеет память $p = 2$, начальный момент обучения $t_0 = 3 \rightarrow$ март, интервалы для двух параметров, равные $d = -1$, $w = 2$.

Таблица 1. Стоимостные индикаторы акций ПАО «Газпром» в 2020 г.

Месяц	январь	февраль	март	апрель	май	июнь
Дата t	1	2	3	4	5	6
Цена m_C^*	0,259	0,223	0,208	0,178	0,188	0,200
Макс C_{\max}^*	0,262	0,240	0,212	0,196	0,202	0,208
Мин C_{\min}^*	0,227	0,201	0,158	0,177	0,182	0,190
V_C^*	0,017	0,020	0,027	0,009	0,011	0,009
D_C^*	0,084	0,069	0,070	0,041	0,046	0,049
Месяц	июль	август	сентябрь	октябрь	ноябрь	декабрь
Дата t	7	8	9	10	11	12
Цена m_C^*	0,195	0,182	0,181	0,171	0,154	0,183
Макс C_{\max}^*	0,202	0,195	0,186	0,173	0,189	0,215
Мин C_{\min}^*	0,179	0,180	0,170	0,154	0,152	0,182
V_C^*	0,011	0,007	0,009	0,010	0,019	0,017
D_C^*	0,206	0,189	0,190	0,181	0,173	0,200

Обучение модели стоимости будем проводить на интервале $\mathcal{T} = [t_0, t_0 + 2] = [3, 5]$. Исторический период $\mathcal{I}_p = [t_0 - 2, t_0 - 1] = [1, 2]$.

Поскольку модель (32) содержит два параметра, то приобретает следующий вид:

$$(43) \quad \mathbb{C}[\mathbf{a} | t] = a_1 m_C^*[t - 1] + a_2 m_C^*[t - 2].$$

Функции ПРВ в соответствующих торговых сессиях имеют вид (36) с моделью (43). Заметим, что энтропийно-оптимальные ПРВ параметров, генерируемые линейной моделью (43), отличны от нормального распределения.

Для определения значений множителей Лагранжа, т.е. решения балансовых уравнений, применялся приближенный аналитический метод, изложенный в разделе 2.2.

Энтропийно-оптимальные функции ПРВ (с приближенными до первой коррекции) для 3-й, 4-й и 5-й торговых сессий имеет вид:

$$(44) \quad P_3^*(\mathbf{a} | 1,068; -0,871) = 0,131 \exp(-0,238a_1 - 0,277a_2 + 0,043a_1^2 + 0,058a_2^2 + 0,100a_1a_2);$$

$$(45) \quad P_4^*(\mathbf{a} | 0,958; 0,102) = 0,133 \exp(-0,199a_1 - 0,214a_2 - 0,004a_1^2 - 0,005a_2^2 - 0,005a_1a_2);$$

$$(46) \quad P_5^*(\mathbf{a} | -1,994; 2,609) = 0,092 (0,355a_1 + 0,415a_2 - 0,083a_1^2 - 0,112a_2^2 - 0,193a_1a_2).$$

На рис. 1–3 показаны их графики.

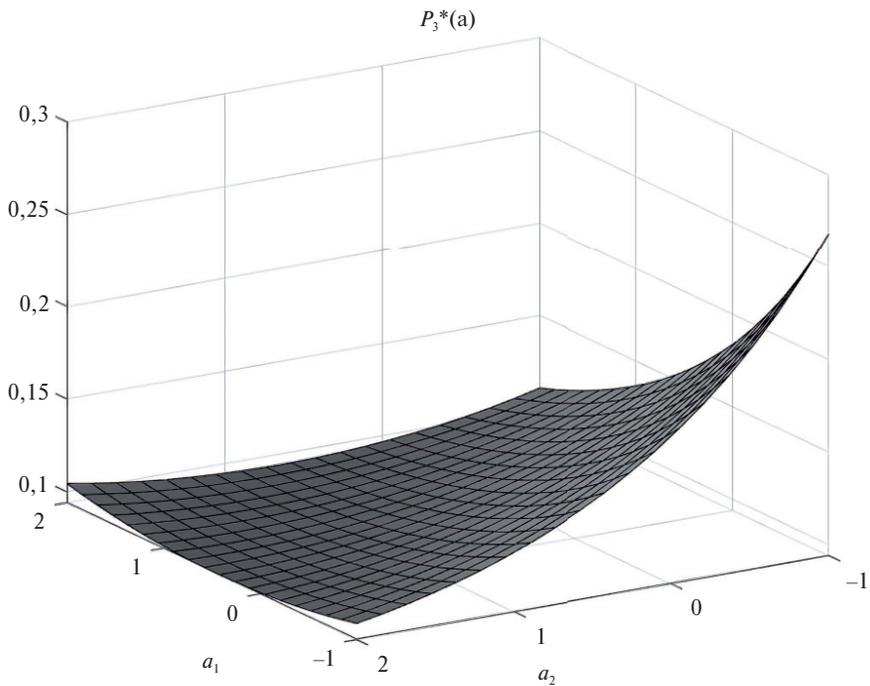


Рис. 1. Распределение $P_3^*(\mathbf{a})$.

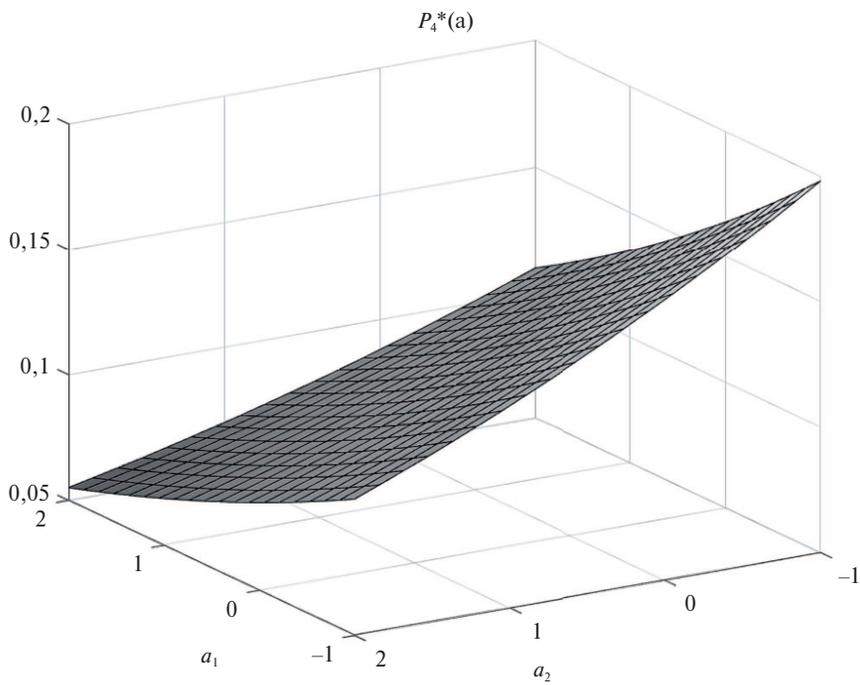


Рис. 2. Распределение $P_4^*(\mathbf{a})$.

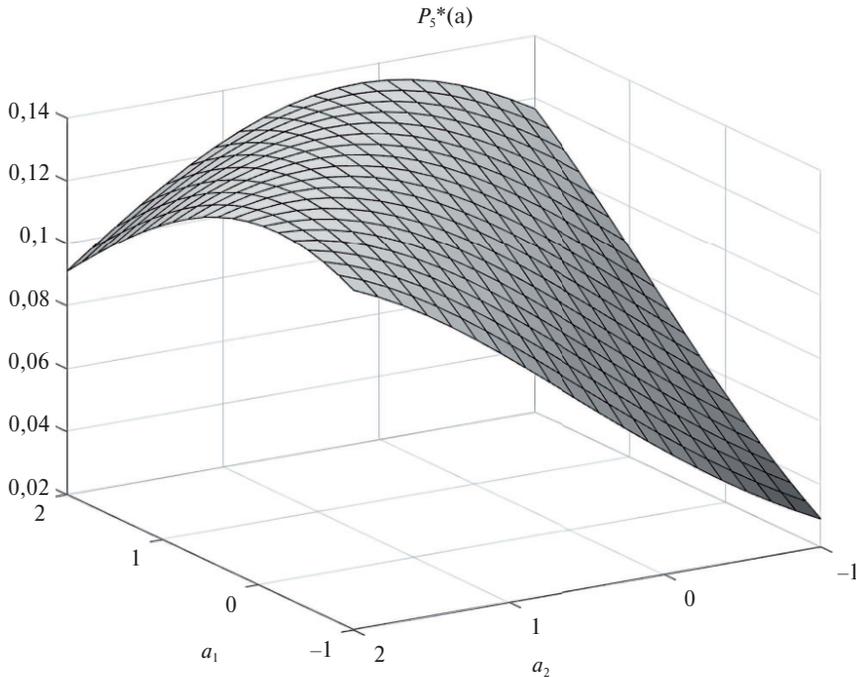


Рис. 3. Распределение $P_5^*(\mathbf{a})$.

4.2. Прогнозирование средней стоимости и среднего индикатора волатильности

Полученные выше энтропийно-оптимальные функции ПРВ $P_3^*(\mathbf{a})$, $P_4^*(\mathbf{a})$, $P_5^*(\mathbf{a})$ параметров модели (32) будем использовать для генерирования ансамблей данных, по которым вычисляются прогнозируемые значения $m_C[t]$ и $D_C[t]$ в торговых сессиях: апр. (4)–нояб. (11). Реализованные значения этих переменных известны (табл. 1), что позволяет оценить точность прогнозов при использовании различных стратегий прогнозирования.

4.2.1. Однодневные прогнозы $P_k^*(\mathbf{a}) \rightarrow (m_C[k+1], V_C[k+1])$

При однодневных прогнозах используется оптимальная ПРВ для k -ой торговой сессии и прогнозируются результаты $(k+1)$ -й сессии. Рассмотрим процедуру формирования прогноза $3 \rightarrow 4$. Для этого используется ПРВ $P_3^*(\mathbf{a})$ (44) и прогнозная модель (32), которая в данном примере приобретает следующий вид:

$$(47) \quad \mathbb{C}[\mathbf{a} | 4] = a_1 m_C^*[3] + a_2 m_C^*[2].$$

ПРВ $P_3^*(\mathbf{a})$ (44) трансформируем в случайную последовательность $\{a_1, a_2\}$. Генерируемый ансамбль содержит 1000 значений $\mathbb{C}[\mathbf{a} | 4]$. Вычисляем $\bar{m}_C[4] = \mathcal{M}(\mathbb{C}[\mathbf{a} | 4])$ и $\bar{\sigma}_C^2[4] = \mathcal{M}\{(\mathbb{C}[\mathbf{a} | 4] - \bar{m}_C[4])^2\}$, где под оператором $\mathcal{M}\{\bullet\}$ понимается эмпирическое среднее.

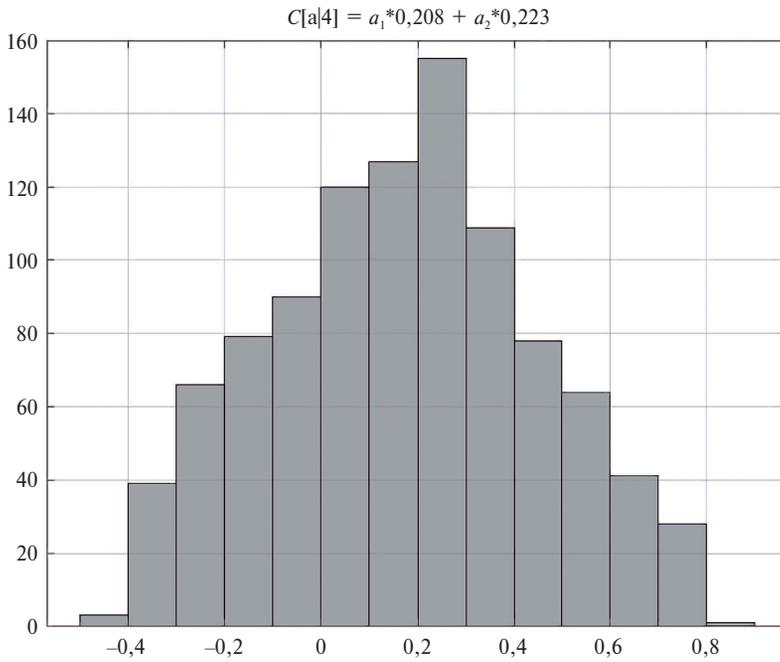


Рис. 4. Эмпирическая ПРВ $C[a|4]$ (однодневный прогноз 3 → 4).

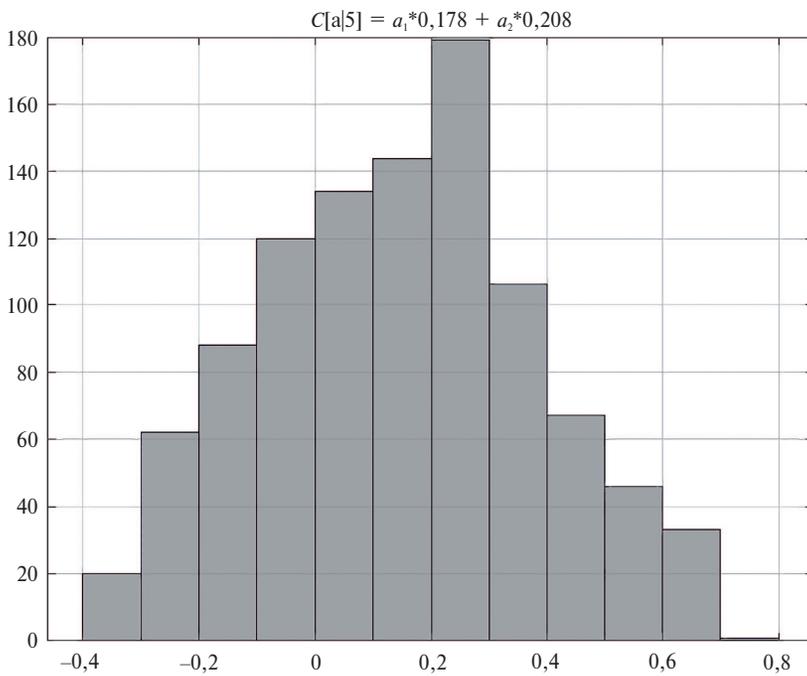


Рис. 5. Эмпирическая ПРВ $C[a|5]$ (однодневный прогноз 4 → 5).

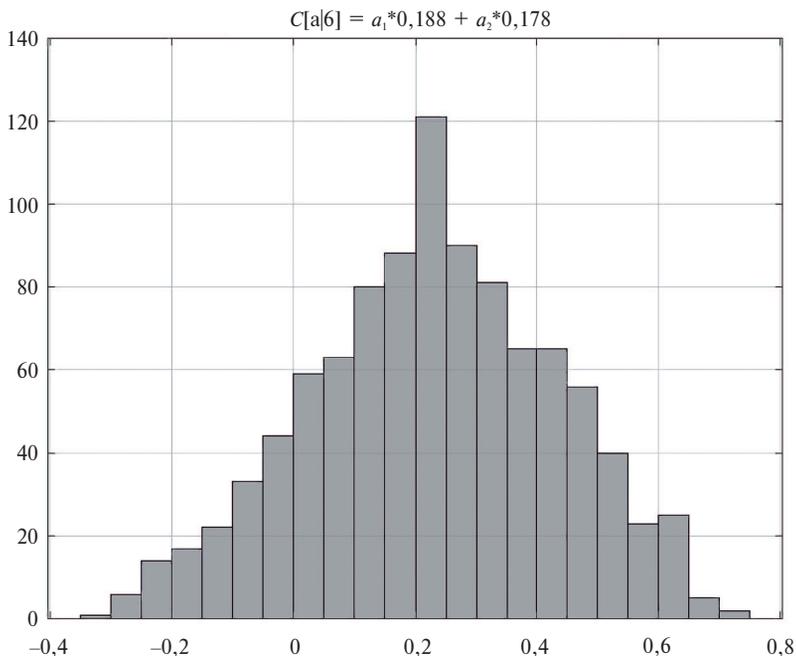


Рис. 6. Эмпирическая ПРВ $C[a|6]$ (однодневный прогноз $5 \rightarrow 6$).

Прогнозы $4 \rightarrow 5$ и $5 \rightarrow 6$ формируются аналогично. Результаты однодневных прогнозов и оценки их точности (в сравнении с реализованными значениями в торгах) приведены в табл. 2.

Таблица 2. Однодневные прогнозы

прогноз \bullet	$3 \rightarrow 4$	$4 \rightarrow 5$	$5 \rightarrow 6$
$\bar{m}_C[\bullet]$	0,175	0,145	0,229
$m_C^*[\bullet]$	0,178	0,188	0,200
$\bar{\sigma}_C^2[\bullet]$	0,076	0,056	0,041
$V_C^*[\bullet]$	0,041	0,046	0,049
$ \delta_m[\bullet] $	0,003	0,043	0,029
$ \delta_\sigma[\bullet] $	0,035	0,010	0,008

В этой таблице переменные

$$(48) \quad \begin{aligned} \delta_m[\bullet] &= \bar{m}_C[\bullet] - m_C^*[\bullet], \\ \delta_\sigma[\bullet] &= \bar{\sigma}_C^2[\bullet] - V_C^*[\bullet]. \end{aligned}$$

На рис. 4–6 показаны эмпирические ПРВ прогнозных значений стоимости в сессиях 4–6 при однодневных прогнозах.

«Интегральная» относительная ошибка прогноза средней стоимости при однодневных прогнозах имеет вид:

$$(49) \quad \Delta_m = \frac{\sqrt{\sum_{t=4}^6 \delta_m^2[t]}}{\sqrt{\sum_{t=4}^6 m_C^2[t] + \sum_{t=4}^6 (m_C^*[t])^2}} = 8\%.$$

«Интегральная» относительная ошибка прогноза средней дисперсии при однодневных прогнозах имеет вид:

$$(50) \quad \Delta_\sigma = \frac{\sqrt{\sum_{t=4}^6 \delta_\sigma^2[t]}}{\sqrt{\sum_{t=4}^6 \sigma_C^2[t] + \sum_{t=4}^6 (V_C^*[t])^2}} = 17\%.$$

4.2.2. Двухдневные прогнозы

$$P_k^*(\mathbf{a}) \rightarrow (m_C[k+1], V_C[k+1]), (m_C[k+2], V_C[k+2])$$

При двухдневных прогнозах используется оптимальная ПРВ для торговой сессии k и прогнозируется результат в сессии $k+2$.

Прогноз $3 \rightarrow 4, 5$, ПРВ $P_3^*(\mathbf{a})$ может быть реализован в варианте

$$(51) \quad \begin{aligned} \mathbb{C}[\mathbf{a} | 4] &= a_1 m^*[3] + a_2 m_C^*[2], \\ \mathbb{C}[\mathbf{a} | 5] &= a_1 m^*[4] + a_2 m_C^*[3]. \end{aligned}$$

и в варианте

$$(52) \quad \begin{aligned} \mathbb{C}[\mathbf{a} | 4] &= a_1 m^*[3] + a_2 m_C^*[2] = a_1 0,208 + a_2 0,223, \\ \mathbb{C}[\mathbf{a} | 5] &= a_1 \bar{m}^*[4] + a_2 m_C^*[3] = a_1 \bar{m}^*[4] + a_2 0,208, \\ \bar{m}^*[4] &= \tilde{\mathcal{M}}\{\mathbb{C}[\mathbf{a} | 4]\}. \end{aligned}$$

Двухдневные прогнозы $4 \rightarrow 5, 6$ и $5 \rightarrow 6, 7$ производятся аналогично и их результаты и оценки точности приведены в табл. 3.

Таблица 3. Двухдневные прогнозы

•	3 → 4	3 → 5	4 → 5	4 → 6	5 → 6	5 → 7
$\bar{m}_C[\bullet]$	0,175	0,185	0,145	0,206	0,229	0,200
$m_C^*[\bullet]$	0,178	0,188	0,188	0,200	0,200	0,195
$\bar{\sigma}_C^2[\bullet]$	0,056	0,060	0,056	0,012	0,041	0,024
$V_C^*[\bullet]$	0,041	0,051	0,046	0,009	0,049	0,011
$ \delta_m[\bullet] $	0,003	0,033	0,043	0,078	0,029	0,065
$ \delta_\sigma[\bullet] $	0,035	0,049	0,010	0,030	0,009	0,043

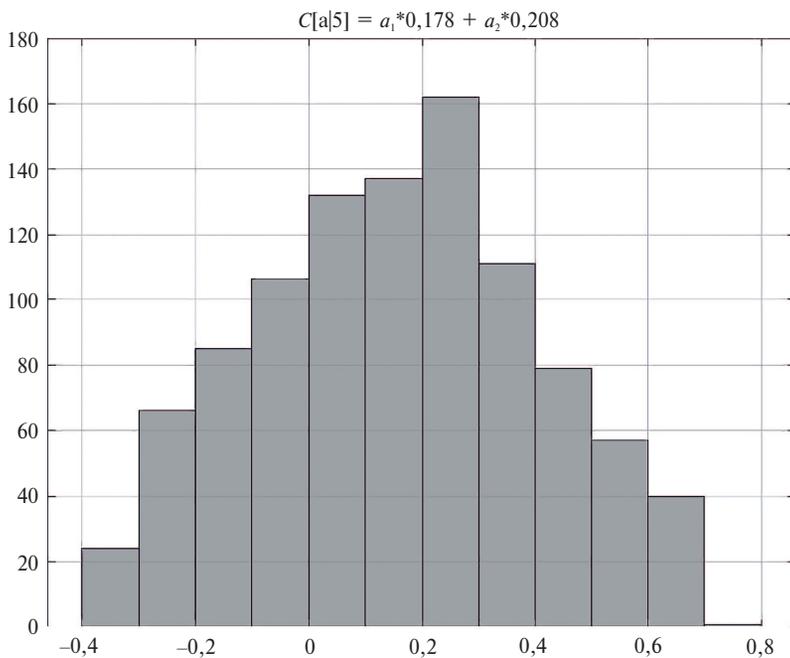


Рис. 7. Эмпирическая ПРВ $C[a | 5]$ (двухдневный прогноз 3 → 4, 5).

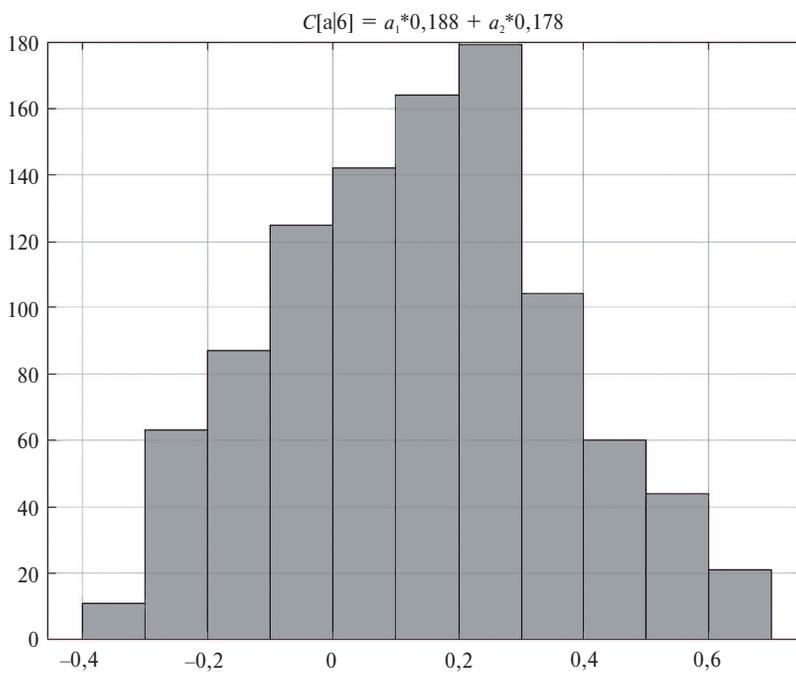


Рис. 8. Эмпирическая ПРВ $C[a | 6]$ (двухдневный прогноз 4 → 5, 6).

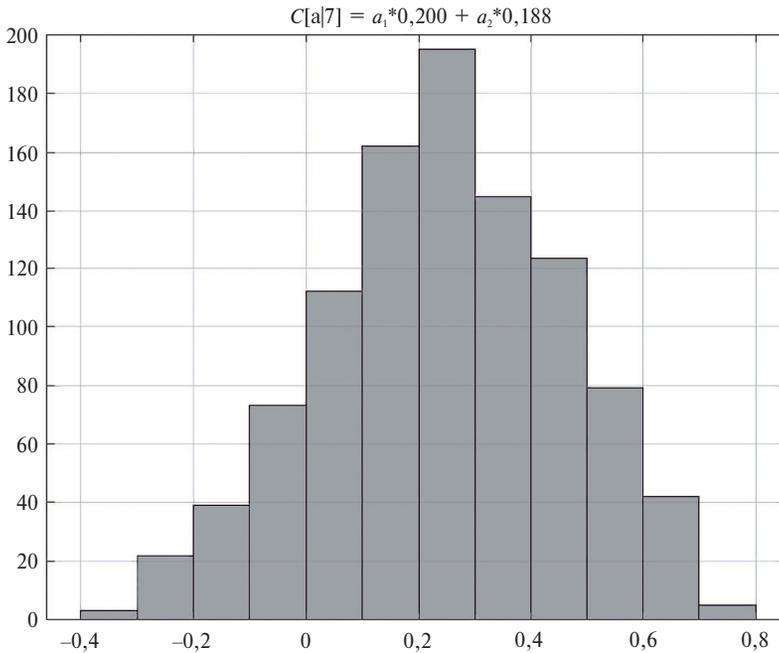


Рис. 9. Эмпирическая ПРВ $C[a | 7]$ (двухдневный прогноз 5 \rightarrow 6, 7)

Интегральная относительная ошибка прогноза средней стоимости для группы двухдневных прогнозов имеет вид

$$(53) \quad \Delta_m = \frac{\sqrt{\sum_{t=5}^7 \delta_m^2[t]}}{\sqrt{\sum_{t=5}^7 m_C^2[t] + \sum_{t=5}^7 (m_C^*[t])^2}} = 7,2\%.$$

Интегральная относительная ошибка прогноза средней волатильности для группы двухдневных прогнозов имеет вид

$$(54) \quad \Delta_\sigma = \frac{\sqrt{\sum_{t=5}^7 \delta_\sigma^2[t]}}{\sqrt{\sum_{t=5}^7 \sigma_C^2[t] + \sum_{t=5}^7 (V_C^*[t])^2}} = 25\%.$$

На рис. 7–9 показаны эмпирические ПРВ прогнозных значений стоимости в сессиях 5–7 при двухдневных прогнозах.

5. Обсуждение

Проблема генерации подходящих данных для тестирования и прогнозирования является весьма популярной в современной компьютерной науке. В статье предлагается адаптация и развитие технологии рандомизированного машинного обучения для генерации ансамблей данных с заданными числовыми характеристиками.

В отличие от существующей технологии предлагается расширение, учитывающее моментные характеристики от 1-го до s -го порядка. Показано, что это приводит к функциям ПРВ не гаусового класса даже в случае линейной модели данных. Предлагаемое расширение, так же как и существующее, сводится к решению соответствующих балансовых уравнений, содержащих интегральные компоненты. В статье развивается приближенный аналитический метод их решения, основанный на использовании степенных рядов и методе неопределенных коэффициентов.

Он применяется в задаче прогнозирования стоимости финансового инструмента, результаты которой сравнивались с реализованными данными по одно- и двухдневным прогнозам. В рамках проведенных исследований обнаружилась вполне приемлемая точность приближенного решения в составе двух коррекций. Однако необходимо более глубокое изучение приближенного метода, как его теоретических аспектов, так и численного исследования.

6. Заключение

В статье развита теория и алгоритм генерации ансамблей тестовых данных с заданными свойствами в виде числовых характеристик, основанная на модификации структуры процедуры рандомизированного машинного обучения [22]. Известно, что ядром указанной процедуры являются балансовые уравнения относительно множителей Лагранжа, содержащие так называемые интегральные компоненты, т.е. многомерные интегралы от экспоненциальных подынтегральных функций.

Для решения этих уравнений адаптирован метод асимптотического *аналитического* решения, развитый в [29], который позволяет свести задачу многомерного интегрирования к сумме произведений одномерных интегралов от степенных функций.

Разработан метод рандомизированного прогнозирования и применен к построению одно- и двухдневных прогнозов средних стоимости и дисперсии биржевого финансового актива.

СПИСОК ЛИТЕРАТУРЫ

1. *Rubinstein R.Y., Kroese D.P.* Simulation and the Monte Carlo Method, 2016, John Wiley & Sons.
2. *Vapnik V.N.* Statistical Learning Theory, Wiley, 1998.
3. *Bishop C.M.* Pattern Recognition and Machine Learning, Springer, 2006.
4. *Hastie T., Tibshirant R., Friedman J.* The Elements of Statistical Learning, Springer, 2009.
5. *Vovk V., Shafer G.* Good Randomized sequential probability forecasting is always possible // J. Royal. Stat. Soc. B. 2005. V. 67. No. 5. P. 747–763.
6. *Hong T., Prinson P., Fan S., Zareiypour H., Triccoli A., Hyndman R.J.* Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond // Int. J. Forecast. 2016. V. 32. No. 3. P. 896–913.

7. *Zhang et al.* Stock price prediction via discovering multy-frequency trading patterns // Proc. 23rd ACM SIGKDD international Conference on Knowledge Discovery and Data Mining. 2017. P. 2141–2149.
8. *Myers G.J.* The Art of Software Testing. John Wiley & Sons, 1979.
9. *Городецкий В.И., Грушицкий М.С., Хабалов А.В.* Многоагентные системы (обзор) // Новости искусственного интеллекта, 1998, № 2, С. 116.
10. *Patton R.* Software Testing, SAWS Publishers, 2005.
11. *Лысенков А.И., Бут Г.С., Диденко Д.А.* Система для разработки компьютерных тестов. <http://www.bytic.ru/cue99m/cf7pvke.html>, 2002.
12. *Мицель А.А., Погуда А.А.* Нейросетевой подход к задаче тестирования // Прикладная информатика, 2011. № 5 (35). С. 60–67.
13. *Заозерская Л.А., Платонова В.А.* Математические модели формирования оптимального комплекса структур тестов для контроля знаний // Омский научный вестник. 2012. № 3.
14. *Campi M.C., Garatti S., Prandini M.* The scenario approach for systems and control design // Ann. Rev. Control. 2009. V. 33. No. 2. P. 149–157.
15. *Chi Z., Liu Y., Turrini A., Zhang L., Jansen D.N.* A scenario approach for parametric Markov decision processes / In Principles of Verification: Cycling the Probabilistic Landscape: Essay Dedicated to Joost-Pieter Katoen on the Occasiion of His 60th Birthday, Part II. Cham, Springer Nature Switzerland. 2024. P. 234–266.
16. *Boltzmann L.* Vorlesungen uber Gastheory. Leipzig, 1896, V. 1, J.A.Barth; 1898, V. 2, J.A.Barth.
17. *Jaynes E.T.* Information theory and statistical Mechanics // Phys. Rev. 1957. V. 104. No. 4. P. 620–630.
18. *Jaynes E.T.* Gibbs vs Boltzmann entropy // Amer. J. Phys. 1965. V. 33. P. 391–398.
19. *Rosenkrantz R.D., Jaynes E.T.* Paper on Probability, Statistics, and Statistical Physics. Kluwer Academic Publishers, 1989.
20. *Jaynes E.T.* Probability theory: the logic of science. Cambridge Uni. Press, 2003.
21. *Попков Ю.С.* Асимптотическая эффективность оценок максимальной энтропии // Докл. АН. 2020. Т. 493. С. 100–103.
22. *Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A.* Entropy Randomization in Machine Learning. CRC Press, 2023.
23. *Красносельский М.А., Вайникко Г.М., Забрейко П.П., Рунтцкий Я.Б., Стеценко В.Я.* Приближенные решения операторных уравнений. М.: Наука, 1969.
24. *Малкин И.Г.* Некоторые задачи теории нелинейных колебаний. М.: УРСС, 2004.
25. *Darkhovsky, B.S., Popkov, Y.S., Popkov, A.Y., Aliev, A.S.* A Method of Generating Random Vectors with a Given Probability Density Function // Autom. Remote Control. 2018. V. 79. No. 9. P. 1569–1581. <https://doi.org/10.1134/S0005117918090035>
26. *Avellaneda M.* Minimum-relative-entropy calibration of asset-pricing models // Int. J. Theor. App. Finance. 1998. V. 1. No. 04. P. 447–472.
27. *Jackwerth J.C.* Recovering Risk Aversion from Option Prices and Realized Returns // Rev. Financ. Stud. 2000. V. 11. No. 2. P. 437.
28. *Ant-Sahalia Y., Lo A.W.* Nonparametric Risk Management and Implied Risk Avarision // J. Econom. 2000. V. 94. P. 4–5.

29. *Popkov Yu.S.* Analytic Method for Solving One Class of Nonlinear Equations // Doklady Mathematics. 2024. <https://doi.org/10.1134/S1064562424601392>
30. *Фихтенгольц Г.М.* Курс дифференциального и интегрального исчисления. М.: Физматгиз, 1962.
31. *Феллер В.* Введение в теорию вероятностей и ее приложения. М.: Мир, 1967.
32. *Соболев И.М.* Численные методы Монте-Карло. М.: Наука, 1973.
33. *Базвалов Н.С., Жидков Н.П., Кобельков Г.М.* Численные методы. М.: Бином, 2003.

Статья представлена к публикации членом редколлегии О.Н. Граничиным.

Поступила в редакцию 23.02.2025

После доработки 29.04.2025

Принята к публикации 29.05.2025

Оптимизация, системный анализ и исследование операций

© 2025 г. А.Н. КАРКИЩЕНКО, д-р физ.-мат. наук (karkishalex@gmail.com)
(Научно-конструкторское бюро робототехники и систем управления, Таганрог)

ОБ ЭМЕРДЖЕНТНОСТИ СИСТЕМ БИНАРНЫХ ОТОБРАЖЕНИЙ

Рассматривается возникновение эффекта эмерджентности в терминах отображений, реализуемых произвольными системами. При этом эмерджентность интерпретируется как «сверхаддитивное» расширение класса отображений при объединении систем или возникновении новых связей между элементами. В работе дается описание эмерджентных свойств в моделях систем, реализующих бинарные отображения. Структура взаимодействия элементов в таких системах описывается конечным ориентированным графом. Исследуются классы отображений при теоретико-множественных операциях над такими системами, и показано «сверхаддитивное» расширение отображений при объединении. Вводится и обосновывается коэффициент эмерджентности. Доказаны нижняя и верхняя оценки для этого коэффициента.

Ключевые слова: бинарное отображение, теоретико-множественные операции, эмерджентность, коэффициент эмерджентности.

DOI: 10.31857/S0005231025070076, **EDN:** JSCDTJ

1. Введение

В системном анализе понятие эмерджентности является одним из наиболее интересных свойств сложных систем. В стандартной формулировке *эмерджентность* – это появление новых свойств и качеств, не присущих элементам, входящим в состав системы. В качестве синонима эмерджентности используется также термин «системный эффект». История возникновения и развития понятия эмерджентности в рамках общей теории систем, а также базовые идеи, лежащие в основе этого явления, изложены в [1]. В частности, со ссылкой на [2] указывается, что эмерджентное поведение может быть понято, исходя из природы и поведения его частей плюс знание, как эти части взаимодействуют.

Большое количество работ посвящено философскому и общеметодологическому осмыслению этого явления в целом [3]. В [4] обосновываются категории сильной и слабой эмерджентности, эти понятия являются достаточно условными, и между ними трудно провести четкую границу. Более того, считается, что эмерджентность может проявляться на континууме от «слабого» до

«сильного» [5]. Эмерджентность считается *сильной*, если не существует приемлемой теории, которая могла бы объяснить или вывести поведение системы на основе свойств или поведения ее компонентов. Напротив, эмерджентность является *слабой*, если отчетливо наблюдается зависимость между поведением компонентов и поведением системы. В [6] делается попытка математически доказать возможность определить понятие сильной эмерджентности. Работа [7] посвящена формализации понятия эмерджентности, основанного на энтропии. Утверждается, что существует связь между энтропией и эмерджентностью в сложных системах в том смысле, что рост энтропии является индикатором возможного появления эмерджентных событий и состояний. В заключении авторы формулируют вопрос о том, как выразить энтропию системы через энтропию составляющих ее подсистем.

Попытки дать формальное описание таких эффектов в конкретных физических или инженерных системах вызывают значительные затруднения, поэтому примеров таких исследований значительно меньше. Здесь следует отметить работу [8], в которой рассматривается возникновение эмерджентных эффектов в искусственных системах, а также [9], посвященную исследованию принципов и подходов к построению инженерных систем, обладающих такими свойствами. Конкретные примеры проявлений системного эффекта в реальных физических системах можно найти, например, в работах [10–13].

Из-за отсутствия четкого понимания эмерджентность в разных исследованиях определяется по-своему. Например, в [14, 15] она понимается как способность взаимодействующих автономных агентов формировать согласованное поведение без внешнего управления, а только на основе внутреннего взаимодействия. Похожий взгляд отражен в работе [16], в которой рассматриваются процессы самоорганизации и эмерджентности при формировании общего словарного запаса на первичных стадиях возникновения человеческих сообществ. Основная идея также основана на моделировании взаимодействия агентов, действующих на основе небольшого количества простых правил, что и приводит к эмерджентным эффектам.

Исследование эмерджентности тесно связано с исследованием процессов, которые приводят к сложному поведению системы при относительно простом поведении ее элементов. Возможно, формальное обоснование подобных процессов даже в относительно узких конкретных случаях таких систем является ключевым вопросом понимания природы возникновения эмерджентности как таковой. Данные вопросы являются центральными в [17].

Системный эффект обнаруживается в различных по своей природе системах – полиметаллические соединения, химические процессы, коллективное принятие решений в социальных структурах и др. Многообразие его проявлений предопределяет формальные методы, которые наиболее приемлемы для исследования в каждом случае. Это могут быть уравнения, описывающие упругое взаимодействие твердых тел, законы протекания химических реакций или алгоритмы принятия компромиссных решений, в частности, ал-

горитмы группового управления в многоагентных системах. Исследование системных эффектов предполагает предварительное построение формальной модели, которая должна правильно предсказывать поведение исследуемого объекта или процесса при допустимых воздействиях на него.

Общепринято, что для существования любой системы необходимы системообразующие факторы, которые способствуют образованию системы, являются внешними по отношению к ее элементам, а также не вызываются необходимостью к объединению [18]. Во многих случаях роль таких факторов играют связи между элементами, природа которых определяется природой элементов, образующих систему.

Основное предположение, которое послужило мотивом для данной работы, состоит в том, что достаточно универсальным способом исследования эмерджентности является исследование отображений, осуществляемых сложной системой в зависимости от отображений, реализуемых элементами системы, и от состояния системообразующих факторов, т.е. структуры связей между элементами системы. Это почти очевидно в случае систем обработки информации, в том числе включающих в себя людей.

Ниже рассматривается класс систем, реализующих бинарные отображения. В определенном смысле их можно рассматривать как автоматные отображения. Несмотря на кажущуюся простоту, по широте их применения и выразительности автоматные отображения во многих случаях не уступают классическим алгебраическим структурам. Известно также, что полугруппы, группы и кольца автоматов, а также их функциональные системы способствовали решению ряда непростых абстрактно-математических проблем.

Модели систем бинарных отображений не позволяют полностью объяснить все процессы системной эмерджентности. Тем не менее, они демонстрируют важные свойства, которые присущи этому явлению. Например, так называемый «сверхаддитивный» эффект, который возникает при объединении систем, а также возникновение предельных циклов в множестве состояний. В физических системах предельным циклам соответствуют аттракторы, которые можно рассматривать как возможные макросостояния в многомасштабных системах. В случае систем бинарных отображений появляется возможность перейти к строго формальным моделям и получить ряд количественных характеристик процессов, отражающих эмерджентность. Следует отметить, что сверхаддитивные эффекты на классах бинарных отображений по уже почти общепринятой терминологии и классификации можно отнести к классу слабой эмерджентности [4–6], поскольку существует высокий уровень прослеживаемости между функциями элементов и функциями системы в целом.

Основная задача, которая ставится и решается в данной статье состоит в том, чтобы установить природу и строго объяснить причину возникновения эмерджентных эффектов в системах, реализующих бинарные отображения,

а также ввести и обосновать удобную количественную меру для измерения степени проявления таких эффектов.

Статья построена следующим образом. В разделе 2 приводятся основные определения, обозначения и термины, которые используются далее в тексте. Разделы 3 и 4 содержат вывод основных соотношений, описывающих реализуемые бинарные отображения в системах со структурой полного графа и произвольного ориентированного графа соответственно. В разделе 5 доказываются утверждения, описывающие классы бинарных отображений, которые получаются в результате теоретико-множественных операций над системами. Описывается также возникающий при этом системный эффект. Раздел 6 посвящен определению и обоснованию коэффициента эмерджентности для бинарных систем, а также доказательству двусторонних оценок для него.

2. Основные определения и обозначения

Будем использовать *кронекерово произведение* $A \otimes B$ матриц A и B , а именно, если $A = (a_{ij})$ – матрица размеров $t \times n$, а B – размеров $p \times q$, то по определению матрица $A \otimes B = (a_{ij}B)$ и имеет, очевидно, размеры $tp \times nq$. Определение и свойства кронекерова произведения можно найти, например, в [19]. В частности, далее потребуется равенство $(A \otimes B)(C \otimes D) = AC \otimes BD$, связывающее обычное и кронекерово произведения, причем матрицы A, B, C и D имеют размеры, при которых все произведения определены. Это свойство обобщается на произвольное число сомножителей. Кронекерово произведение матриц A_1, A_2, \dots, A_k будем для удобства обозначать $A_1 \otimes A_2 \otimes \dots \otimes A_k = \underset{i=1}{\overset{k}{\otimes}} A_i$. В частности, кронекерову k -ю степень матрицы A будем записывать в виде $\underbrace{A \otimes \dots \otimes A}_k = A^{[k]}$. Важно принимать во внимание некоммутативность кронекерова произведения.

Кронекеровым расширением вектора $a = (a_1 \ a_2 \ \dots \ a_n)^\top$ будем называть вектор

$$\begin{aligned} \bar{a} &= (1 \ a_1 \ a_2 \ (a_1 a_2) \ \dots \ (a_1 a_2 \dots a_n))^\top = \\ &= \begin{pmatrix} 1 \\ a_n \end{pmatrix} \otimes \begin{pmatrix} 1 \\ a_{n-1} \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ a_1 \end{pmatrix}. \end{aligned}$$

Под *кронекеровым расширением по строкам* матрицы A понимается матрица \bar{A} , каждая строка которой, рассматриваемая как вектор, является кронекеровым расширением соответствующей строки матрицы A . *Адамаровым произведением* матриц A и B называется матрица $A * B$, получающаяся в результате поэлементного перемножения этих матриц [19]. Это произведение определено для матриц, имеющих одинаковые размеры, и является коммутативной операцией. Матрицу будем называть *мономимальной* по столбцам, если каждый ее столбец состоит из нулей за исключением одного элемента, который равен единице.

Далее потребуются некоторые базовые определения из теории графов. Пусть задан ориентированный граф $H = (V, E)$, где $V = \{v_1, v_2, \dots, v_n\}$ – множество вершин, а $E = \{e_1, e_2, \dots, e_p\} \subseteq V \times V$ – множество дуг. В графе допускаются петли, т.е. дуги вида $e = (v, v)$. Если $v_i \in V$, то ρ_i будет обозначать *полу степень захода* вершины v_i , т.е. число дуг вида (v, v_i) , входящих в эту вершину, так что $\sum_{i=1}^n \rho_i = |E|$. Граф с n вершинами, который содержит все возможные дуги и петли, т.е. когда $E = V \times V$, $|E| = n^2$, называется *полным ориентированным графом с петлями* (или для краткости просто *полным графом*), будем обозначать его K_n . Если же $E = \emptyset$, то такой граф называется *нуль-графом* и обозначается O_n .

Матрицей смежности ориентированного графа H называется квадратная матрица $R = (r_{ij})$ размера $n \times n$, в которой $r_{ij} = 1$, если $(v_i, v_j) \in E$, и $r_{ij} = 0$ в противном случае. *Объединением* и *пересечением* графов $H_1 = (V, E_1)$ и $H_2 = (V, E_2)$ с одним и тем же множеством вершин называются графы $H_1 \cup H_2 = (V, E_1 \cup E_2)$ и $H_1 \cap H_2 = (V, E_1 \cap E_2)$ соответственно.

3. Системы со структурой полного графа

Под *системой* будем понимать совокупность n взаимодействующих функциональных элементов, которые функционируют в дискретном времени $t = 0, 1, 2, \dots$ и могут находиться в одном из двух *состояний* – 0 или 1. Состояние z_i i -го элемента в момент времени $t + 1$ определяется булевой функцией $z_i(t + 1) = f_i(x_1(t + 1), \dots, x_m(t + 1); z_1(t), \dots, z_n(t))$, зависящей от двух групп аргументов. Аргументы первой группы – это внешние аргументы $x_1(t + 1), \dots, x_m(t + 1)$, представляющие собой двоичные внешние сигналы, аргументы второй группы – это внутренние аргументы $z_1(t), \dots, z_n(t)$, которые являются состояниями соответствующих элементов системы в момент времени t . Функции z_i , $i = 1, \dots, n$, будем называть *функциями состояния* элементов, а упорядоченный набор состояний всех элементов – *состоянием системы*. Считаем, что система не зависит от случайных факторов.

Последовательность состояний системы можно рассматривать как бинарное отображение последовательности внешних аргументов и предыдущих состояний. В общем случае функции, реализуемые элементами, могут не зависеть от состояний некоторых других элементов, что определяется структурой связей между элементами. Поэтому последовательность состояний зависит от: 1) функций, реализуемых каждым элементом системы, и 2) структуры связей между элементами. Если система содержит n элементов с m входами, то каждый элемент может описываться одной из 2^{m+n} булевых функций.

Структуру системы удобно задавать с помощью ориентированного графа, вершины которого находятся во взаимно-однозначном соответствии с элементами системы, а дуги соответствуют связям между выходами одних элемен-

тов и входами других. Поскольку состояние каждого элемента может зависеть и от его собственного состояния, то граф системы может иметь петли. Описанную систему можно рассматривать как синхронную *последовательностную* логическую сеть, реализующую *автоматное* отображение [20].

Рассмотрим систему из n элементов со структурой K_n . Реализуемое ею отображение можно описать системой функций состояний, в которой правые части – некоторые булевы функции: $z_i(t+1) = f_i(x_1(t+1), \dots, x_m(t+1), z_1(t), \dots, z_n(t))$, $i = 1, \dots, n$. Далее будем рассматривать случай, когда отсутствуют внешние аргументы, т.е. $z_i(t+1) = f_i(z_1(t), \dots, z_n(t))$, $i = 1, \dots, n$. По аналогии с дифференциальными уравнениями такую систему можно назвать автономной, описывающей «собственное свободное движение» или, в терминологии системного анализа, закрытой системой, демонстрирующей свое поведение, зависящее от начального состояния $z_1(0), \dots, z_n(0)$. Опишем отображения, реализуемые такой системой при всех возможных функциях состояний элементов системы.

Как известно [21–23], любую булеву функцию от n переменных можно однозначно задать ее *арифметическим представлением*:

$$(1) \quad z_i(t+1) = f_i(z_1(t), \dots, z_n(t)) = a_0^i + a_1^i z_1(t) + a_2^i z_2(t) + a_{12}^i z_1(t) z_2(t) + \dots \\ \dots + a_{k_1 \dots k_p}^i z_{k_1}(t) \cdot \dots \cdot z_{k_p}(t) + \dots + a_{1 \dots n}^i z_1(t) \cdot \dots \cdot z_n(t), \quad i = 1, \dots, n,$$

где $a_0^i, a_1^i, a_2^i, a_{12}^i, \dots, a_{k_1 \dots k_p}^i, \dots, a_{1 \dots n}^i$ – целые числа, однозначно определяющие функцию f_i .

Рассмотрим кронекерово расширение вектора $z = (z_1 \ z_2 \ \dots \ z_n)^\top$:

$$\bar{z} = (1 \ z_1 \ z_2 \ (z_1 z_2) \ \dots \ (z_1 z_2 \dots z_n))^\top = \\ = \begin{pmatrix} 1 \\ z_n \end{pmatrix} \otimes \begin{pmatrix} 1 \\ z_{n-1} \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 \\ z_1 \end{pmatrix}.$$

Тогда систему соотношений (1) можно записать в матричном виде:

$$(2) \quad z(t+1) = A\bar{z}(t),$$

где A – матрица размера $n \times 2^n$ соответствующих коэффициентов.

Равенство (2) можно записать иначе. Вектор $\bar{z}(t)$ принимает 2^n различных значений, и каждому из них ставится в соответствие образ при отображении A . Поэтому можно записать 2^n равенств (2), соответствующих различным $\bar{z}(t)$. По индукции легко доказать, что все возможные 2^n значений вектора $\bar{z}(t)$, записанные в естественном порядке, образуют матрицу $Q_n = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{[n]}$. Обозначим через B матрицу, столбцы которой являются образами соответствующих векторов $\bar{z}(t)$. Очевидно, что эта матрица полностью и однозначно описывает реализуемое системой бинарное отображение. Тогда

все 2^n возможных равенств (2) можно записать компактно в виде $B = A Q_n$. Поскольку $Q_n^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^{[n]}$, то сразу получаем

$$(3) \quad A = B Q_n^{-1}.$$

Эта формула дает явное выражение коэффициентов арифметического представления функций состояния всех элементов системы, которая реализует бинарное отображение, заданное с помощью матрицы B . Однозначность решения (3) означает, что в системе со структурой K_n может быть осуществлено любое отображение. Иными словами, *класс отображений*, реализуемых в такой системе, совпадает с множеством всех возможных отображений. Если структура системы отличается от K_n , то это уже не будет выполняться.

Далее опишем классы реализуемых бинарных отображений в зависимости от структуры системы, задаваемой произвольным ориентированным графом (не обязательно графом K_n).

4. Системы с произвольной структурой и классы отображений

Рассмотрим общий случай, когда система имеет произвольную структуру, задаваемую ориентированным графом $H = (V, E)$. Функция состояния i -го элемента зависит от состояния j -го элемента в том случае, если одновременно выполняются два условия: 1) имеется связь между входом i -го элемента и выходом j -го и 2) значение состояния j -го элемента является для i -го существенной переменной. Первое условие определяется структурой системы, а второе – функцией состояния. В этом отражается взаимосвязь внутриэлементной и надэлементной организации системы. Система со структурой K_n является максимально «информированной», поскольку каждый элемент имеет информацию о состояниях всех остальных элементов (включая его самого) и может быть настроен на выполнение любой функции.

Охарактеризуем класс отображений, реализуемых системой с произвольной структурой H . Пусть $R = (r_{ij})$ – матрица смежности графа H . Воспользовавшись арифметическим представлением (1), получим аналогичное (2) матричное соотношение $z(t+1) = A \bar{z}(t)$, где, как и раньше, A – матрица, каждая строка которой образована коэффициентами соответствующих арифметических представлений.

Однако в общем случае $z_i(t+1)$ зависит не от всех значений $z_j(t)$, $j = 1, \dots, n$, а лишь от некоторых, определяемых структурой системы. Говоря точнее, $z_i(t+1)$ существенно зависит от $z_j(t)$, если $r_{ji} = 1$, и, соответственно, не зависит при $r_{ji} = 0$. Таким образом, вся информация о структурной зависимости i -го элемента от остальных элементов содержится в i -м столбце матрицы R . Если $z_i(t+1)$ структурно не зависит от $z_j(t)$, то можно считать, что коэффициент a_j^i в разложении равен нулю или, в общем случае, можно записать $a_j^i = r_{ji} u_j^i$. Ясно, что в остальных членах разложения, на-

пример, $a_{k_1 \dots k_2}^i z_{k_1}(t) \cdot \dots \cdot z_{k_2}(t)$, коэффициент $a_{k_1 \dots k_2}^i$ также равен нулю, если отсутствует хотя бы одна из связей k_1 -го, \dots , k_p -го элементов со входом i -го элемента. Это равносильно равенству: $a_{k_1 \dots k_p}^i = r_{k_1 i} \cdot \dots \cdot r_{k_p i} u_{k_1 \dots k_p}^i$. Следовательно, матрицу A можно записать в виде адамарова произведения двух матриц:

$$A = \begin{pmatrix} a_0^1 & a_1^1 & a_2^1 & a_{12}^1 & \dots & a_{1 \dots n}^1 \\ a_0^2 & a_1^2 & a_2^2 & a_{12}^2 & \dots & a_{1 \dots n}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_0^n & a_1^n & a_2^n & a_{12}^n & \dots & a_{1 \dots n}^n \end{pmatrix} = \\ = \begin{pmatrix} u_0^1 & u_1^1 & u_2^1 & u_{12}^1 & \dots & u_{1 \dots n}^1 \\ u_0^2 & u_1^2 & u_2^2 & u_{12}^2 & \dots & u_{1 \dots n}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ u_0^n & u_1^n & u_2^n & u_{12}^n & \dots & u_{1 \dots n}^n \end{pmatrix} * \begin{pmatrix} 1 & r_{11} & r_{21} & (r_{11} \cdot r_{21}) & \dots & (r_{11} \cdot \dots \cdot r_{n1}) \\ 1 & r_{12} & r_{22} & (r_{12} \cdot r_{22}) & \dots & (r_{12} \cdot \dots \cdot r_{n2}) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & r_{1n} & r_{2n} & (r_{1n} \cdot r_{2n}) & \dots & (r_{1n} \cdot \dots \cdot r_{nn}) \end{pmatrix}.$$

Первая матрица в произведении, обозначим ее через U , – собственно матрица коэффициентов представления. Вторая матрица отражает структурную зависимость элементов системы. Она, очевидно, является кронекеровым расширением по строкам $\overline{R^\top}$ транспонированной матрицы смежности R . Матрица $\overline{R^\top}$, очевидно, состоит из нулей и единиц, как и матрица R . Ввиду этого, матричное соотношение можно записать следующим образом:

$$(4) \quad z(t+1) = \left(U * \overline{R^\top} \right) \bar{z}(t).$$

Это выражение определяет отображения, реализуемые системой с произвольной структурой H , определяемой матрицей смежности R . В частности, если $H = K_n$, все элементы матрицы R , следовательно, и матрицы $\overline{R^\top}$, равны единице, и формула (4) совпадает с формулой (2).

Подобно тому, как из (2) было получено выражение (3) для коэффициентов арифметического представления, определяющее матрицу A , из соотношения (4) можно выразить $U * \overline{R^\top}$:

$$(5) \quad U * \overline{R^\top} = BQ_n^{-1}.$$

Однако в отличие от предыдущего случая в данном соотношении матрица B не может быть произвольной; она должна быть двоичной матрицей, но такой чтобы в правой части (5) были гарантированно равны нулю элементы, местоположения которых определяются положением нулевых элементов матрицы $\overline{R^\top}$.

Каждое состояние системы можно отождествить с двоичным вектором. Если записать все возможные состояния системы в естественном порядке, то они образуют двоичную матрицу G размера $n \times 2^n$. Для $n = 2, 3$ эти матрицы показаны ниже:

$$G = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Каждый столбец матрицы B является одним из возможных состояний системы, поэтому можно записать $B = GP$, где P – мономиальная по столбцам матрица размера $2^n \times 2^n$. Иными словами, GP – это та же самая матрица B , что и в выражении для коэффициентов арифметического представления (3), но записанная явно с помощью мономиальной матрицы P , т.е. множества матриц B и всех мономиальных матриц P находятся во взаимно однозначном соответствии. Это позволяет в дальнейшем оперировать не всеми возможными образами двоичных векторов при отображении, а всеми мономиальными по столбцам матрицами, что с формальной точки намного удобнее. Подставляя это равенство в формулу (5) для матрицы коэффициентов арифметического представления в случае произвольной структуры, получаем

$$(6) \quad U * \overline{R^T} = GPQ_n^{-1}.$$

В дальнейшем будем называть это равенство *соотношением для коэффициентов*. Если матрица P удовлетворяет данному условию при некоторой матрице U , то она полностью описывает бинарное отображение, реализуемое системой с заданной структурой H . Таким образом, существует взаимно однозначное соответствие между отображениями и мономиальными матрицами, удовлетворяющими соотношению для коэффициентов (6) при некоторой матрице U . Поэтому каждое отображение можно отождествлять с соответствующей мономиальной матрицей P .

Обозначим через \mathfrak{P} множество всех мономиальных матриц размера $2^n \times 2^n$. Очевидно, что \mathfrak{P} – это полугруппа с единицей, которая содержит, в частности, подполугруппу изоморфную симметрической группе S_{2^n} . Множество $\mathcal{P}(H) \subseteq \mathfrak{P}$ преобразований P , удовлетворяющих соотношению для коэффициентов (6) при некоторой матрице U , определяет *класс отображений*, реализуемых системой со структурой графа H . Это можно пояснить следующим образом.

Отображение всех возможных состояний системы полностью определяется совокупностью конкретных отображений, реализуемых каждым элементом этой системы, т.е. булевыми функциями, описывающими их функционирование. В свою очередь, эти функции полностью определяются соответствующими коэффициентами их арифметических представлений. Но эти коэффициенты однозначно определяются матрицей B , а в силу сказанного выше, соответствующей мономиальной матрицей P . В случае структуры полного графа для *любого возможного отображения* состояний системы можно вычислить коэффициенты арифметических представлений функций элементов, при которых система будет реализовывать это отображение.

Для системы с произвольной структурой это не так, поскольку функции, реализуемые элементами системы, а значит, и системой в целом, ограничены имеющейся структурой системы, т.е. связями между элементами. Поэтому не каждое отображение может быть реализовано в такой системе. Поскольку отображение состояний системы полностью определяется мономиальной матрицей P , то это означает, что не для каждой матрицы P могут

быть найдены функции элементов, а значит, и коэффициенты их арифметических представлений так, чтобы выполнялось соотношение для коэффициентов (или равенство) (6). Все возможные отображения, которые могут быть реализованы в системе при фиксированной структуре за счет всевозможных выборов функций элементов, и образуют класс отображений при заданной структуре.

Охарактеризуем матрицы P , входящие в множество $\mathcal{P}(H)$. Для этого потребуется вспомогательное «индексное» множество. Пусть для удобства матрицы $\overline{R^T}$ и U заданы с помощью обычной двухиндексной нумерации, т.е. $\overline{R^T} = (\bar{r}_{ij})$, $U = (u_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, 2^n$. Введем множество $\text{Ind}\overline{R^T} = \{(i, j) \mid i = 1, \dots, n, j = 1, \dots, 2^n, \bar{r}_{ij} = 0\}$. Иначе говоря, $\text{Ind}\overline{R^T}$ образовано всевозможными парами индексов, которые определяют позиции нулевых элементов матрицы $\overline{R^T}$. Следующее утверждение дает описание множества $\mathcal{P}(H)$, не зависящее от U .

Утверждение 1. $\mathcal{P}(H) = \left\{ P \mid P \in \mathfrak{F}, g_i P q_j^{(-1)} = 0, (i, j) \in \text{Ind}\overline{R^T} \right\}$.

Доказательство. Пусть g_i и $q_j^{(-1)}$ обозначают i -ю строку в G и j -й столбец в Q_n^{-1} . Тогда (i, j) -элемент матрицы, стоящей в соотношении для коэффициентов (6) справа, очевидно, равен $g_i P q_j^{(-1)}$. Если соответствующий элемент матрицы $\overline{R^T}$ равен нулю, т.е. $\bar{r}_{ij} = 0$, то равен нулю также (i, j) -элемент матрицы $U * \overline{R^T}$. Поэтому для того, чтобы выполнялось матричное равенство для коэффициентов (6), необходимо должно быть $g_i P q_j^{(-1)} = 0$. В противном случае, т.е. при $\bar{r}_{ij} = 1$, (i, j) -элемент матрицы $U * \overline{R^T}$ равен (i, j) -элементу матрицы U . Так как этот элемент является одним из коэффициентов, то, полагая его равным $g_i P q_j^{(-1)}$, получаем необходимое равенство элементов слева и справа. Поэтому мономиальная матрица $P \in \mathfrak{F}$ удовлетворяет соотношению для коэффициентов (6), если она удовлетворяет системе соотношений $g_i P q_j^{(-1)} = 0, (i, j) \in \text{Ind}\overline{R^T}$, что доказывает утверждение.

Каждая структура H порождает класс $\mathcal{P}(H)$ бинарных отображений, образующихся при всевозможных функциональных настройках элементов системы. В частности, если $H = K_n$, то R – матрица из единиц. В этом случае, очевидно, $\mathcal{P}(K_n) = \mathfrak{F}$, и значит, в такой системе может быть реализовано любое отображение. Если структура системы – нуль-граф, т.е. $H = O_n$, то $R = \mathbf{0}$, и $\mathcal{P}(O_n)$ содержит 2^n матриц P , характеризующихся тем, что у них элементы какой-то из строк равны единице, а остальные – нулю.

Следующее утверждение дает явное выражение для мощности $|\mathcal{P}(H)|$ класса отображений $\mathcal{P}(H)$, реализуемых системой со структурой H .

Утверждение 2. Пусть структура системы задается ориентированным графом $H = (V, E)$, $|V| = n$, ρ_i – полустепень захода вершины i в графе H , $i = 1, \dots, n$. Тогда

$$(7) \quad |\mathcal{P}(H)| = 2^{\sum_{i=1}^n \rho_i}.$$

Доказательство. Число различных бинарных отображений, реализуемых i -м элементом с ρ_i входами при всех возможных функциональных настройках, равно $2^{2^{\rho_i}}$. Поэтому $|\mathcal{P}(H)| = 2^{2^{\rho_1}} \cdot 2^{2^{\rho_2}} \cdot \dots \cdot 2^{2^{\rho_n}} = 2^{\sum_{i=1}^n 2^{\rho_i}}$.

Следствие 1. 1) $|\mathcal{P}(K_n)| = 2^{n2^n}$; 2) $|\mathcal{P}(O_n)| = 2^n$.

В следующем разделе дается описание классов бинарных отображений, которые получаются в результате теоретико-множественных операций над системами.

5. Операции над системами и классы реализуемых отображений

Пусть имеются две системы со структурами H_1 и H_2 . Возникает вопрос: существуют ли классы отображений, которые при подходящих функциях состояния элементов могут быть реализованы как в первой, так и во второй системе? Если существуют, то как их охарактеризовать? Исследуем этот вопрос.

Напомним еще раз, что, как было описано выше, каждое отображение возможных состояний системы с заданной структурой и фиксированными функциями элементов, однозначно описывается мономиальной по столбцам матрицей P . При всех возможных функциях элементов получается класс отображений, реализуемых системой с данной структурой. Поэтому описание классов отображений сводится к описанию всех возможных и допустимых мономиальных матриц.

Если некоторое отображение, определяемое матрицей P , реализуется как первой, так и второй системой, то $P \in \mathcal{P}(H_1)$ и $P \in \mathcal{P}(H_2)$, следовательно, $P \in \mathcal{P}(H_1) \cap \mathcal{P}(H_2)$. Справедливо и обратное, любая матрица P , лежащая в пересечении множеств $\mathcal{P}(H_1)$ и $\mathcal{P}(H_2)$, определяет отображение, реализуемое каждой системой.

Утверждение 3. $\mathcal{P}(H_1 \cap H_2) = \mathcal{P}(H_1) \cap \mathcal{P}(H_2)$ для любых H_1 и H_2 .

Доказательство. Пусть $R_1 = (r_{ij}^{(1)})$ и $R_2 = (r_{ij}^{(2)})$ – матрицы смежности графов H_1 и H_2 . Тогда $\mathcal{P}(H_1)$ определяется системой соотношений $g_i P q_j^{(-1)} = 0$, $(i, j) \in \text{Ind} \overline{R_1}$, а $\mathcal{P}(H_2)$ – системой соотношений $g_i P q_j^{(-1)} = 0$, $(i, j) \in \text{Ind} \overline{R_2}$. Если $P \in \mathcal{P}(H_1) \cap \mathcal{P}(H_2)$, то это значит, что P должно удовлетворять каждой из этих систем или, иначе, удовлетворять системе соотношений $g_i P q_j^{(-1)} = 0$, $(i, j) \in \text{Ind} \overline{R_1} \cup \text{Ind} \overline{R_2}$.

Введем матрицу \overline{R}^T , которая определяется следующим образом: (i, j) -элемент этой матрицы равен нулю, если равен нулю соответствующий элемент хотя бы одной из матриц $\overline{R_1}$ или $\overline{R_2}$, и единице, когда оба элемента равны единице. Тогда, очевидно, $\text{Ind} \overline{R}^T = \text{Ind} \overline{R_1} \cup \text{Ind} \overline{R_2}$. Непосредственно из определения матрицы \overline{R}^T следует, что $\overline{R}^T = \overline{R_1} * \overline{R_2}$. Покажем теперь, что \overline{R}^T является кронекеровым расширением по строкам матрицы $R^T = (R_1 * R_2)^T =$

$= R_1^\top * R_2^\top$, т.е. $\overline{R^\top} = \overline{R_1^\top * R_2^\top}$. Для этого достаточно доказать соотношение

$$(8) \quad \overline{R_1^\top * R_2^\top} = \overline{R_1^\top} * \overline{R_2^\top}.$$

Прежде всего, заметим, что первые столбцы этих матриц равны, поскольку содержат только единицы. Покажем, что равны и остальные элементы. Произвольный элемент матрицы, стоящей слева, имеет вид $\left(r_{k_1 i}^{(1)} r_{k_1 i}^{(2)}\right) \cdot \dots \cdot \left(r_{k_p i}^{(1)} r_{k_p i}^{(2)}\right)$, $p = 1, \dots, n$, а соответствующий элемент правой матрицы — $\left(r_{k_1 i}^{(1)} \cdot \dots \cdot r_{k_p i}^{(1)}\right) \cdot \left(r_{k_1 i}^{(2)} \cdot \dots \cdot r_{k_p i}^{(2)}\right)$. Оба выражения отличаются лишь порядком сомножителей, следовательно, равны. Поэтому равенство (8) справедливо. Но $R = R_1 * R_2$ является матрицей смежности графа $H = H_1 \cap H_2$, из чего следует справедливость утверждения.

Следствие 2. $\mathcal{P}(H_1 \cap \dots \cap H_k) = \mathcal{P}(H_1) \cap \dots \cap \mathcal{P}(H_k)$ для любых H_1, \dots, H_k .

Данное следствие доказывается по индукции. Утверждение 3 устанавливает точную связь между классами бинарных отображений и системами при их пересечении. Аналогичное соотношение для объединения систем в общем случае неверно, однако имеет место следующее включение.

Утверждение 4. $\mathcal{P}(H_1 \cup H_2) \supseteq \mathcal{P}(H_1) \cup \mathcal{P}(H_2)$ для любых H_1 и H_2 .

Доказательство. Пользуясь утверждением 3, имеем

$$\mathcal{P}(H_1 \cup H_2) \cap \mathcal{P}(H_1) = \mathcal{P}((H_1 \cup H_2) \cap H_1) = \mathcal{P}(H_1).$$

Отсюда следует, что $\mathcal{P}(H_1) \subseteq \mathcal{P}(H_1 \cup H_2)$. Аналогично получаем $\mathcal{P}(H_2) \subseteq \mathcal{P}(H_1 \cup H_2)$. Из этих двух включений следует справедливость утверждения.

Содержательный смысл утверждения 4 состоит в том, что множество классов отображений, реализуемых системой со структурой $H = H_1 \cup H_2$, больше, чем простое объединение классов отображений объединяемых систем. Этот факт в известном смысле отражает свойственный системам «сверхаддитивный эффект», соответствующий тезису «целое больше суммы своих частей». Классы отображений, реализуемые в системе со структурой графа H и не реализуемые ни в одной из объединяемых систем, демонстрируют эффект эмерджентности, понимаемой в узком смысле. Данное утверждение легко обобщается на любое число систем.

Следствие 3. $\mathcal{P}(H_1 \cup \dots \cup H_k) \supseteq \mathcal{P}(H_1) \cup \dots \cup \mathcal{P}(H_k)$ для любых H_1, \dots, H_k .

Из утверждения 4 вытекает свойство монотонности по включению систем.

Следствие 4. Из $H_1 \subseteq H_2$ следует $\mathcal{P}(H_1) \subseteq \mathcal{P}(H_2)$.

Охарактеризуем более подробно левую и правую части утверждения 4.

Утверждение 5. Множество $\mathcal{P}(H_1 \cup H_2)$ состоит из тех и только тех матриц $P \in \mathfrak{P}$, которые удовлетворяют системе соотношений $g_i P q_j^{(-1)} = 0$, $(i, j) \in \text{Ind}(R_1 + R_2 - R_1 * R_2)^\top$.

Доказательство. Пусть $H = H_1 \cup H_2$. Очевидно, что матрица смежности R структуры H связана с матрицами смежности R_1 и R_2 структур H_1 и H_2 равенством $R = R_1 + R_2 - R_1 * R_2$. Ввиду этого из утверждения 1 следует справедливость доказываемого утверждения.

Утверждение 6. Множество $\mathcal{P}(H_1) \cup \mathcal{P}(H_2)$ состоит из тех и только тех матриц $P \in \mathfrak{P}$, которые удовлетворяют системе соотношений

$$(9) \quad g_i P q_j^{(-1)} g_k P q_l^{(-1)} = 0, \quad (i, j) \in \text{Ind} \overline{R_1^\top}, \quad (k, l) \in \text{Ind} \overline{R_2^\top}.$$

Доказательство. Если $P \in \mathcal{P}(H_1)$ или $P \in \mathcal{P}(H_2)$, то P удовлетворяет системе соотношений (9). Обратное, если P не принадлежит, например, $\mathcal{P}(H_1)$, но удовлетворяет (9), то при некотором $(i, j) \in \text{Ind} \overline{R_1^\top}$ имеет место $g_i P q_j^{(-1)} \neq 0$. Если рассмотреть уравнения системы соотношений (9), в которых индексы i и j совпадают с выбранными, т.е. $g_i P q_j^{(-1)} g_k P q_l^{(-1)} = 0$, $(k, l) \in \text{Ind} \overline{R_2^\top}$, то получим соотношения $g_k P q_l^{(-1)} = 0$, $(k, l) \in \text{Ind} \overline{R_2^\top}$. Следовательно, $P \in \mathcal{P}(H_2)$. Подобным образом, матрица P , не принадлежащая $\mathcal{P}(H_2)$ и удовлетворяющая (9), принадлежит $\mathcal{P}(H_1)$, что и доказывает утверждение.

6. Коэффициент эмерджентности

Как показано выше, при объединении систем проявляется «сверхаддитивное» свойство, которое выражается в том, что новая система способна реализовать отображения, которые не реализуемы ни одной из исходных систем. Чтобы количественно охарактеризовать этот эффект необходимо ввести подходящий скалярный показатель, который можно интерпретировать как степень несводимости свойств системы к свойствам ее элементов.

Наиболее простой и прямолинейный подход состоит в использовании отношения $\frac{|\mathcal{P}(H_1 \cup H_2)|}{|\mathcal{P}(H_1) \cup \mathcal{P}(H_2)|}$, показывающего кратность превышения количества реализуемых отображений при объединении систем. Однако несмотря на простоту, такой показатель оказывается трудно интерпретируемым. Это связано с тем, что он принимает очень большие значения из-за дважды экспоненциального роста мощности классов этих отображений, как это установлено в утверждении 2. Следующие соображения предполагают несколько иной подход к построению такого показателя. Рассмотрим простые примеры.

Пусть граф структуры трехэлементной системы имеет соответственно $n = 3$ вершины и $\rho_1 = \rho_2 = 1$, $\rho_3 = 2$ – полустепени захода его вершин. Тогда в силу утверждения 2 мощность класса реализуемых отображений равна $2^{2^1+2^1+2^2} = 2^8 = 256$. С другой стороны, в точности такую мощность будет иметь класс отображений, реализуемых системой со структурой K_2 с двумя элементами, т.е. $2^{2 \cdot 2^2} = 2^8 = 256$.

Еще один аналогичный пример. В графе H имеется $n = 6$ вершин с полустепенями захода $2, 1, 1, 2, 3, 2$. Тогда $|\mathcal{P}(H)| = 2^{2^2+2^1+2^1+2^2+2^3+2^2} = 2^{24}$. Но такой же по мощности класс отображений реализует структура K_3 , т.е. $|\mathcal{P}(K_3)| = 2^{3 \cdot 2^3} = 2^{24} = |\mathcal{P}(H)|$. Иными словами, в каждом из этих случаев для некоторого μ выполняется равенство $2^{\mu 2^\mu} = 2^{2^{\rho_1} + 2^{\rho_2} + \dots + 2^{\rho_n}}$ или

$$(10) \quad \mu 2^\mu = 2^{\rho_1} + 2^{\rho_2} + \dots + 2^{\rho_n}.$$

В этих примерах μ – наименьшее число вершин полного графа с петлями, который реализует в точности такое же число отображений, что и некоторый граф с заданными полустепенями захода вершин. В рассмотренных случаях μ было целым числом и могло быть интерпретировано как число вершин полного ориентированного графа. В общем случае решение μ уравнения (10) при произвольных $\rho_1, \rho_2, \dots, \rho_n$ не будет целым. Тем не менее, спекулятивно рассуждая, можно считать, что решение уравнения (10) является скалярной характеристикой рассматриваемого класса отображений как минимального, но «размытого» (не целого) числа вершин полного ориентированного графа с петлями, реализующего такое же число различных отображений. Поскольку функция $\mu 2^\mu$ является положительной и монотонно возрастающей на $[0, \infty)$, то при любой правой части уравнение (10) имеет единственный корень.

Величину μ можно рассматривать как меру $\mu = \mu[\mathcal{P}(H)]$ множества $\mathcal{P}(H)$ на булеане $2^{\mathfrak{B}}$, где \mathfrak{B} , как было определено выше, – множество всех мономатриц соответствующего размера. Данная мера, очевидно, не является аддитивной, но является *монотонной* по включению классов, поскольку $\mu[\mathcal{P}(H_1)] \leq \mu[\mathcal{P}(H_2)]$ при $\mathcal{P}(H_1) \subseteq \mathcal{P}(H_2)$.

Коэффициентом эмерджентности $\kappa(H_1, H_2)$ при объединении систем H_1 и H_2 будем называть величину $\kappa(H_1, H_2) = \frac{\mu[\mathcal{P}(H_1 \cup H_2)]}{\mu[\mathcal{P}(H_1) \cup \mathcal{P}(H_2)]}$. Из этого определения следует, что $\mu[\mathcal{P}(H_1 \cup H_2)] = \mu_1$ и $\mu[\mathcal{P}(H_1) \cup \mathcal{P}(H_2)] = \mu_2$ – корни соответствующих уравнений

$$(11) \quad \mu_1 2^{\mu_1} = \log_2 |\mathcal{P}(H_1 \cup H_2)|,$$

$$(12) \quad \mu_2 2^{\mu_2} = \log_2 |\mathcal{P}(H_1) \cup \mathcal{P}(H_2)| = \log_2 (|\mathcal{P}(H_1)| + |\mathcal{P}(H_2)| - |\mathcal{P}(H_1 \cap H_2)|).$$

Введенное понятие коэффициента эмерджентности легко может быть обобщено на произвольное число объединяемых систем. Следующая теорема дает двусторонние оценки введенного коэффициента.

Теорема 1. Коэффициент эмерджентности $\kappa(H_1, H_2)$ при объединении любых двух n -элементных систем H_1 и H_2 удовлетворяет неравенствам

$$1 \leq \kappa(H_1, H_2) < 1 + \sqrt{1 + \frac{1}{e \ln 2}} \approx 2,2372.$$

Полученная верхняя оценка является достаточно точной, что особенно заметно при возрастании n . В нижеследующей таблице приведены значения

максимального коэффициента эмерджентности для $n = 2, \dots, 20$ (с точностью до четырех знаков после запятой). Сравнение приведенных данных с найденной верхней оценкой позволяет высказать предположение о ее асимптотической точности.

Значения максимального коэффициента эмерджентности

n	Максимальный коэффициент эмерджентности	n	Максимальный коэффициент эмерджентности
		11	2,0822
2	1,2553	12	2,1245
3	1,3729	13	2,1567
4	1,4979	14	2,1807
5	1,5981	15	2,1980
6	1,6946	16	2,2102
7	1,7869	17	2,2184
8	1,8803	18	2,2236
9	1,9610	19	2,2265
10	2,0281	20	2,2278

7. Заключение

В работе предложен аппарат для исследования системных эффектов, возникающих в произвольных системах бинарного отображения. Несмотря на двоичную природу реализуемых отображений, для исследования используется стандартная алгебраическая техника, что упрощает описание и понимание эффектов эмерджентности в узком смысле. Проведенный анализ позволяет сделать вывод, что наблюдаемые в таких системах классы отображений и их преобразования при теоретико-множественных операциях над структурами этих систем схожи с эффектом «расширения» при объединении классических алгебраических структур. Полученные результаты, в частности, утверждения 3, 4 и следствия из них, а также оценки для коэффициента эмерджентности, даваемые теоремой 1, позволяют установить в каких случаях и какой в количественном измерении эмерджентный эффект может быть достигнут при агрегировании бинарных систем. Это можно использовать, например, для оптимизации коллективного поведения отдельных групп функционирующих объектов или субъектов.

Рассмотренная в работе модель двоичных отображений – достаточно узкий класс отображений. Однако из-за своей относительной простоты она позволяет выявить некоторые закономерности, присущие системным взаимодействиям и даже количественно их охарактеризовать. В этом, собственно, и состояла мотивация данной работы. Обобщение данных исследований на более общий класс системных взаимодействий (даже в автономном случае) по-

требует привлечения более сложного математического аппарата, связанного с морфизмами на дифференциальных многообразиях и построение подходящих мер на таких объектах. Это тема отдельного непростого исследования. Но подобная теория может дать более естественное описание эффектов эмерджентности в физических средах и в технических системах.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1.

Левое неравенство сразу следует из утверждения 4 и монотонности функции $x2^x$. Докажем правое неравенство.

Пусть $H_1 = (V, E_1)$ и $H_2 = (V, E_2)$ – произвольные ориентированные графы и $|V| = n$. Рассмотрим множества дуг $E_C = E_1 \cap E_2$, $E_A = E_1 \setminus E_C$ и $E_B = E_2 \setminus E_C$. Эти множества, очевидно, попарно не пересекаются. Построим три графа $H_A = (V, E_A)$, $H_B = (V, E_B)$ и $H_C = (V, E_C)$; обозначим через $\rho_i^A, \rho_i^B, \rho_i^C, i = 1, \dots, n$, – соответственно полустепени захода их вершин. Тогда $H_1 = H_A \cup H_C = (V, E_A \cup E_C)$, $H_2 = H_B \cup H_C = (V, E_B \cup E_C)$ и в силу утверждения 2 можно записать уравнения (11) и (12) следующим образом:

$$\mu_1 2^{\mu_1} = \sum_{i=1}^n 2^{\rho_i^A + \rho_i^B + \rho_i^C}, \quad \mu_2 2^{\mu_2} = \log_2 \left(\sum_{i=1}^n 2^{\rho_i^A + \rho_i^C} + \sum_{i=1}^n 2^{\rho_i^B + \rho_i^C} - \sum_{i=1}^n 2^{\rho_i^C} \right).$$

Пусть $\rho_i^A + \rho_i^B + \rho_i^C = m_i$ для некоторых фиксированных $m_i \leq n, i = 1, \dots, n$. В этом случае правая часть в первом уравнении является постоянной величиной, и, значит, в силу монотонности функции $x2^x$ коэффициент эмерджентности κ будет достигать максимального значения при наименьшем значении правой части второго уравнения. Для отыскания такого значения необходимо решить задачу

$$\varphi = \sum_{i=1}^n 2^{m_i - \rho_i^A} + \sum_{i=1}^n 2^{m_i - \rho_i^B} - \sum_{i=1}^n 2^{m_i - \rho_i^A - \rho_i^B} \xrightarrow{\rho_i^A, \rho_i^B} \min.$$

Найдем стационарные точки функции φ . Вычислив частные производные по $\rho_j^A, \rho_j^B, j = 1, \dots, n$, и приравняв их к нулю, после эквивалентных преобразований получаем:

$$(II.1) \quad \begin{aligned} 2^{\rho_j^B} \cdot \sum_{i=1}^n 2^{m_i - \rho_i^A} &= \sum_{i=1}^n 2^{m_i - \rho_i^A - \rho_i^B}, \\ 2^{\rho_j^A} \cdot \sum_{i=1}^n 2^{m_i - \rho_i^B} &= \sum_{i=1}^n 2^{m_i - \rho_i^A - \rho_i^B}, \end{aligned} \quad j = 1, \dots, n.$$

Поскольку правые части равны, то $2^{-\rho_j^A} \cdot \sum_{i=1}^n 2^{m_i - \rho_i^A} = 2^{-\rho_j^B} \cdot \sum_{i=1}^n 2^{m_i - \rho_i^B}$, $j = 1, \dots, n$. Обозначим $X = \sum_{i=1}^n 2^{m_i - \rho_i^A}$, $Y = \sum_{i=1}^n 2^{m_i - \rho_i^B}$ и перепишем выражение в виде:

$$(II.2) \quad 2^{m_j - \rho_j^A} \cdot 2^X = 2^{m_j - \rho_j^B} \cdot 2^Y.$$

Просуммируем последнее равенство по j , получим: $X2^X = Y2^Y$. Функция $x2^x$ строго монотонно возрастает при $x \geq 0$, поэтому последнее уравнение имеет единственное решение $X = Y$, а значит, из (П.2) получаем, что $\rho_j^A = \rho_j^B = \rho_j$, $j = 1, \dots, n$. Подставим эти условия в любое из уравнений (П.1), после логарифмирования и перегруппировки слагаемых получим: $\rho_j + \sum_{i=1}^n (2^{m_i - \rho_i} - 2^{m_i - 2\rho_i}) = 0$. При $\rho_i \geq 0$ все слагаемые в левой части неотрицательны, поэтому последнее равенство возможно только тогда, когда все эти слагаемые равны нулю, откуда следует, что все $\rho_i = 0$. В допустимой области это единственная стационарная точка, в которой может достигаться экстремум. Эта точка лежит на границе области, и значение функции в ней равно $\varphi = 2^{\sum_{i=1}^n m_i}$. Однако с помощью стандартной техники исследования на экстремум с помощью вторых производных можно показать, что квадратичная форма, описывающая второй дифференциал функции φ в найденной точке, является знаконеопределенной, т.е. в этой точке экстремума нет. Это значит, что экстремальное значение функция φ принимает на границе области $\rho_i^A, \rho_i^B \geq 0$, $\rho_i^A + \rho_i^B \leq m_i$, т.е. при условии $\rho_i^A, \rho_i^B \geq 0$, $\rho_i^A + \rho_i^B = m_i$, $i = 1, \dots, n$.

Для отыскания стационарной точки на границе подставим $\rho_i^B = m_i - \rho_i^A$ в выражение для функции f : $f = 2^{\sum_{i=1}^n m_i - \rho_i^A} + 2^{\sum_{i=1}^n \rho_i^A} - 2^n$. Вычислив производные по ρ_j^A и приравняв их к нулю, после преобразований получим систему уравнений для отыскания стационарных точек:

$$(П.3) \quad 2^{\sum_{i=1}^n m_i - \rho_i^A} \cdot 2^{m_j - \rho_j^A} = 2^{\sum_{i=1}^n \rho_i^A} \cdot 2^{\rho_j^A}, \quad j = 1, \dots, n.$$

Просуммируем все уравнения по j и обозначим $X = \sum_{i=1}^n 2^{m_i - \rho_i^A}$ и $Y = \sum_{i=1}^n 2^{\rho_i^A}$, в результате получим $X2^X = Y2^Y$. Данное равенство возможно только при $X = Y$. Поэтому из (П.3) получаем $2^{m_j - \rho_j^A} = 2^{\rho_j^A}$, откуда $\rho_j^A = \frac{m_j}{2} = \rho_j^B$, $j = 1, \dots, n$. Рутинное исследование с помощью вторых производных показывает, что в этой точке действительно достигается минимум функции $\varphi_{\min} = 2^{1 + \sum_{i=1}^n \frac{m_i}{2}} - 2^n$. Таким образом, для нахождения наибольшего значения коэффициента эмерджентности κ при условии $\rho_i^A + \rho_i^B + \rho_i^C = m_i$ получаем уравнения:

$$(П.4) \quad \mu_1 2^{\mu_1} = \sum_{i=1}^n 2^{m_i},$$

$$(П.5) \quad \mu_2 2^{\mu_2} = \log_2 \left(2 \cdot 2^{\sum_{i=1}^n \frac{m_i}{2}} - 2^n \right).$$

Оценим снизу правую часть (П.5). Поскольку $2^{\sum_{i=1}^n 2^{\frac{m_i}{2}}} - 2^n \geq 0$ при любых m_i , $i = 1, \dots, n$, то

$$\log_2 \left(2 \cdot 2^{\sum_{i=1}^n 2^{\frac{m_i}{2}}} - 2^n \right) \geq \log_2 2^{\sum_{i=1}^n 2^{\frac{m_i}{2}}} = \sum_{i=1}^n 2^{\frac{m_i}{2}} > \left(\sum_{i=1}^n 2^{m_i} \right)^{\frac{1}{2}}.$$

Наряду с уравнением (П.5) рассмотрим уравнение $\bar{\mu}_2 2^{\bar{\mu}_2} = \left(\sum_{i=1}^n 2^{m_i} \right)^{\frac{1}{2}}$. В силу монотонности функции $x2^x$ заключаем, что $\bar{\mu}_2 < \mu_2$. Следовательно, $\kappa = \frac{\mu_1}{\mu_2} < \frac{\mu_1}{\bar{\mu}_2} = \bar{\kappa}$, поэтому $\bar{\kappa}$ можно рассматривать как верхнюю оценку для κ .

Поделим теперь уравнение (П.4) на $(\bar{\mu}_2 2^{\bar{\mu}_2})^2$: $\frac{\mu_1 2^{\mu_1}}{\bar{\mu}_2^2 2^{2\bar{\mu}_2}} = \frac{\bar{\kappa}}{\bar{\mu}_2} 2^{\bar{\mu}_2(\bar{\kappa}-2)} = 1$. Рассмотрим последнее уравнение как неявно заданную функцию $\bar{\kappa}$, зависящую от $\bar{\mu}_2$, т.е. $F(\bar{\mu}_2, \bar{\kappa}) = \frac{\bar{\kappa}}{\bar{\mu}_2} 2^{\bar{\mu}_2(\bar{\kappa}-2)} - 1 = 0$. Найдем экстремум неявной функции $\bar{\kappa}$ от $\bar{\mu}_2$. Для того, чтобы было $\bar{\kappa}'_{\bar{\mu}_2} = 0$, должно выполняться равенство $F'_{\bar{\mu}_2} = -\frac{\bar{\kappa}}{\bar{\mu}_2^2} 2^{\bar{\mu}_2(\bar{\kappa}-2)} + \frac{\bar{\kappa}}{\bar{\mu}_2} 2^{\bar{\mu}_2(\bar{\kappa}-2)} (\bar{\kappa} - 2) \ln 2 = 0$. Решая совместно уравнения $F'_{\bar{\mu}_2} = 0$ и $F = 0$, находим два решения этих уравнений: $\bar{\kappa}_{1,2} = 1 \pm \sqrt{1 + \frac{1}{e \ln 2}}$. Поскольку коэффициент эмерджентности заведомо положительная величина, то получаем окончательно, $\kappa < \bar{\kappa} = 1 + \sqrt{1 + \frac{1}{e \ln 2}} \approx 2,2372$. С помощью стандартных методов можно показать, что найденное значение является единственным максимумом исследуемой функции, что доказывает верхнюю оценку для κ .

Теорема 1 доказана.

СПИСОК ЛИТЕРАТУРЫ

1. *Dampier R.I.* Editorial for the Special Issue on “Emergent Properties of Complex Systems”: Emergence and levels of abstraction // *Int. J. Syst. Sci.* 2000. V. 31. No. 7. P. 811–818. <https://doi.org/10.1080/002077200406543>
2. *Crick F.* The Astonishing Hypothesis. The Scientific Search for the Soul. London: Simon and Schuster, 1994.
3. *Goldstein J.A.* Emergence and Radical Novelty: From Theory to Methods // *Handbook of Research Methods in Complexity Science.* Edward Elgar Publishing, 2018. Chapter 23. P. 507–524.
4. *Chalmers D.J.* Strong and Weak Emergence / In: *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion.* Oxford: Oxford University Press, 2006. P. 244–254.
5. *Bedau M.A.* Weak Emergence // *JSTOR. Philosophical Perspectives: Mind, Causation, and World.* 1997. V. 11. P. 375–399.
6. *Bar-Yam Y.* A Mathematical Theory of Strong Emergence Using Multiscale Variety // *Complexity.* 2004. V. 9. No. 6. P. 15–24.
7. *Johnson J.J. IV, Tolk A., Sousa-Poza A.* A Theory of Emergence and Entropy in Systems of Systems // *Procedia Comput. Sci.* 2013. No. 20. P. 283–289.

8. *Stepney S., Polack F.A.C., Turner H.R.* Engineering Emergence // 11th International Conference on Engineering of Complex Computer Systems (ICECCS 2006). Stanford. USA. 2006.
9. *Fromm J.* On Engineering and Emergence // arXiv preprint nlin/0601002. 2006.
10. *Alexander J.R. Jr, Challef S.* Control: An Emergent Property of Biological Neurons // Int. J. Syst. Sci. 2000. V. 31. No. 7. P. 895–909.
11. *Chen Y., Reggia J.A.* The Temporal Correlation Hypothesis for Self-Organizing Feature Maps // Int. J. Syst. Sci. 2000. V. 31. No. 7. P. 911–921.
12. *Watters P.A.* Time-Invariant Long-Range Correlations in Electroencephalogram Dynamics // Int. J. Syst. Sci. 2000. V. 31. No. 7. P. 819–825.
13. *Channon A.D., Damper R.I.* Towards the Evolutionary Emergence of Increasingly Complex Advantageous Behaviours // Int. J. Syst. Sci. 2000. V. 31. No. 7. P. 843–860.
14. *Cucker F., Smale S.* Emergent Behavior in Flocks // IEEE Trans. Autom. Control. 2007. V. 52. No. 5. P. 852–862.
15. *Cucker F., Smale S.* On the Mathematics of Emergence // Jpn. J. Math. 2007. V. 2. P. 197–227.
16. *Ke J., Minett J., Au C.-P., Wang W.S.-Y.* Self-Organization and Selection in the Emergence of Vocabulary // Complexity. 2002. V. 7. No. 3. P. 41–54.
17. *Benjamin D.P.* On the Emergence of Intelligent Global Behaviours from Simple Local Actions // Int. J. Syst. Sci. 2000. V. 31. No. 7. P. 861–872.
<https://doi.org/10.1080/002077200406589>
18. *Сурмин Ю.П.* Теория систем и системный анализ: уч. пособие. Киев: МАУП, 2003.
19. *Marcus M., Mink H.* A Survey of Matrix Theory and Matrix Inequalities. Boston. Allys and Bacon, 1964.
20. *Cavanagh J.* Sequential Logic. Analysis and Synthesis. CRC Press, Taylor&Francis Group. 2007.
21. *Малюгин В.Д.* Реализация булевых функций арифметическими полиномами // АиТ. 1982. № 4. С. 84–93.
22. *Rubio A.J.* Algebraic Representation of Boolean Functions // Int. J. Electron. 1984. V. 56. No. 5. P. 735–739. <https://doi.org/10.1080/00207218408938868>
23. *Jain J.* Arithmetic Transform of Boolean Functions / In: Representations of Discrete Functions. Springer. Boston. MA. 1996. P. 133–161.
<https://doi.org/10.1007/978-1-4613-1385-4-6>

Статья представлена к публикации членом редколлегии Д.В. Виноградовым.

Поступила в редакцию 02.02.2025

После доработки 28.03.2025

Принята к публикации 17.04.2025

СОДЕРЖАНИЕ

Управление в технических системах

- Булычев Ю.Г.** Гибридный метод комплексирования разнесенных датчиков для адаптивной идентификации стационарных и подвижных объектов в условиях априорной неопределенности 3
- Хуторцев В.В.** Программное управление поиском объектов в условиях ограничений на количество линий обработки в независимых каналах поисковой системы..... 21
- Штакин Д.В., Плотников А.А., Снегирев О.Ю., Торгашов А.Ю.** Построение виртуального анализатора с несколькими выходами на основе предиктора ошибок для управления качеством целевого продукта сложной ректификационной колонны..... 41

Управление в социально-экономических системах

- Гераськин М.И.** Алгоритм управления действиями игрока в игре триполии при линейных функциях спроса и издержек 61

Интеллектуальные системы управления, анализ данных

- Антипин А.Ф., Антипина Е.В., Мустафина С.А.** Модификация многомерного нечеткого регулятора с локальным терм-делением входных переменных 76
- Попков Ю.С., Попков А.Ю., Дубнов Ю.А.** Методы рандомизированного машинного обучения для генерации ансамблей случайных данных с заданными числовыми характеристиками 90

Оптимизация, системный анализ и исследование операций

- Каркищенко А.Н.** Об эмерджентности систем бинарных отображений..... 112

C O N T E N T S

Control in Technical Systems

Bulychev Y.G. Hybrid Method of Integrating Spaced Sensors for Adaptive Identification of Stationary and Mobile Objects under Conditions of a Priori Uncertainty	3
Khutortsev V.V. Software Control of Object Search in Conditions of Restrictions on the Number of Processing Lines in Independent Channels of the Search Engine	21
Shtakin D.V., Plotnikov A.A., Snegirev O.Yu., Torgashov A.Yu. Development of Soft Sensor With Multiple Outputs Based on an Error Predictor for Control of the Target Product Quality of a Complex Distillation Column.....	41

Control in Social Economic Systems

Geraskin M.I. Algorithm for Control of Player's Actions in Tripolia Game with Linear Functions of Demand and Costs.....	61
--	----

Intellectual Control Systems, Data Analysis

Antipin A.F., Antipina E.V., Mustafina S.A. Modification of Multidimensional Fuzzy Regulator with Local Term-Division of Input Variables	76
Popkov Y.S., Popkov A.Y., Dubnov Y.A. Methods of Randomized Machine Learning for Generation of Random Data with Given Numerical Properties.....	90

Optimization, System Analysis, and Operations Research

Karkishchenko A.N. On the Emergence of Binary Mapping Systems.....	112
---	-----