

УДК 621.513.8

В.С. Степашко

Достижения и перспективы индуктивного моделирования

Описаны типичные задачи, решаемые средствами индуктивного моделирования. Приведены сведения о развитии данного научного направления в Украине и за рубежом. Дана характеристика основных фундаментальных, прикладных и технологических достижений. Сформулированы наиболее перспективные пути дальнейших исследований.

Ключевые слова: индуктивное моделирование, МГУА, самоорганизация, структурно-параметрическая идентификация, интеллектуальный анализ данных, принятие решений.

Описано типові задачі, що розв'язуються засобами індуктивного моделювання. Подано відомості про розвиток цього наукового напрямку в Україні і за кордоном, а також характеристику основних фундаментальних, прикладних та технологічних досягнень. Сформульовано найперспективніші шляхи подальших досліджень.

Ключові слова: індуктивне моделювання, МГУА, самоорганізація, структурно-параметрична ідентифікація, інтелектуальний аналіз даних, прийняття рішень.

Введение. Плодотворное научное направления под названием *индуктивное моделирование сложных процессов и систем* сформировано академиком НАН Украины Алексеем Григорьевичем Ивахненко и продолжает активно развиваться усилиями ученых его научной школы в разных странах мира, и прежде всего в профильном отделе Международного научно-учебного центра информационных технологий и систем. Центральное место в этой области занимает *метод группового учета аргументов* (МГУА) как средство самоорганизации моделей сложных процессов и систем различной природы для задач принятия решений.

Статья А.Г. Ивахненко [1], опубликованная в 1968 г. в журнале «Автоматика», положила начало главному этапу в его научном творчестве. Возникло новое научное направление, которое сам автор называл *эвристическая самоорганизация моделей* [2], *самоорганизация моделей по экспериментальным данным* [3], *индуктивный метод моделирования сложных систем* [4]. Наконец, в 1998 году, при образовании отдела Информационных технологий индуктивного моделирования в Международном Центре, это направление получило краткое название *индуктивное моделирование*, активно воспринятое на международном уровне и став-

шее ныне общепринятым для научных форумов, статей и книг в Украине и за рубежом.

Понятие *индуктивное моделирование* можно определить как самоорганизующийся процесс эволюционного перехода от первичных данных к математическим моделям, отражающим закономерности функционирования моделируемых объектов и систем, неявно содержащиеся в имеющихся экспериментальных, опытных, статистических данных. В статье дана постановка и описание типичных задач, решаемых средствами индуктивного моделирования, приведены сведения о развитии данного научного направления в Украине и за рубежом, характеризуются основные фундаментальные, прикладные и технологические достижения, а также формулируются наиболее перспективные пути развития дальнейших исследований.

Общая характеристика индуктивного подхода к построению моделей

Общепринятый подход к решению задач моделирования означает, что исследователь проходит весь традиционный *дедуктивный* путь «от общей теории – к конкретной модели»: наблюдение за объектом; изучение его внутренней структуры; познание принципов его функционирования; поиск адекватного математического аппарата для их описания; создание теории и апро-

бация собственно математической модели исследуемого объекта.

Вместо этого А.Г. Ивахненко сформулировал принципиально новый *индуктивный* подход, основанный на интенсивном использовании компьютеров, как путь «от конкретных данных – к общей модели»: исследователь после получения данных наблюдений выдвигает гипотезу о классе функций, формирует процедуру автоматической генерации многих тысяч вариантов моделей в этом классе и задает критерий выбора наилучшей модели из всех генерируемых. При этом наиболее трудоемкая рутинная работа перекладывается на компьютер, благодаря чему появляется возможность минимизировать влияние субъективных факторов и получить объективный результат при максимальной активизации творческих возможностей исследователя.

МГУА как основной метод индуктивного моделирования – это оригинальное средство автоматического решения задач синтеза моделей по экспериментальным данным в условиях неопределенности. Такая задача заключается в построении математической модели, приближающей неизвестную закономерность функционирования исследуемого объекта или процесса, информация о которой неявно содержится в выборке (таблице) имеющихся данных.

МГУА отличается от других методов активным применением принципов

- автоматической генерации усложняемых структур моделей;
- неокончательных решений (множественного выбора моделей на каждом этапе);
- последовательной селекции по внешним критериям для построения моделей оптимальной сложности.

Он имеет оригинальную многорядную процедуру автоматической генерации структур моделей, имитирующую процесс биологической селекции с попарным учетом последовательных признаков. Для сравнения и выбора лучших моделей применяются внешние критерии, основанные на делении выборки на две и более частей, причем оценивание параметров и проверка качества моделей выполняется на разных подвыборках. Это дает возможность обходиться без обре-

зительных априорных предположений, поскольку деление выборки на части с близкими статистическими характеристиками позволяет неявно, автоматически учесть разные виды априорной неопределенности при построении модели.

Краткие сведения о развитии индуктивно-го моделирования как научного направления

Индуктивное моделирование, начавшее свое развитие с публикации в [1], состоит в построении моделей на основе эмпирических данных, результатов измерений, наблюдений, экспериментов, когда этот статистический материал используется для формирования математических моделей, с той или иной степенью правдоподобности описывающих закономерности, отраженные в данных, т.е. сюда относятся задачи, где структура модели, отражающей неизвестную закономерность, заранее не задана, и ее следует также определить на основе данных вместе с оценками параметров.

Индуктивный процесс моделирования его автор трактовал как акт самоорганизации: из «хаоса» начальных данных формируется «организованная структура» в виде математической модели. Это было отражено в первых названиях этого направления, имеющего целью выявление неявных причинно-следственных связей и закономерностей, скрытых в имеющихся данных, и представление их в явной форме математических моделей. Это позволяет отнести МГУА к эффективным методам интеллектуального анализа данных и средствам вычислительного интеллекта, поскольку он направлен на автоматический поиск и описание закономерностей с выбором структуры и параметров линейных, нелинейных, разностных и других моделей на основе выборки данных в условиях неопределенности и неполноты исходной информации.

В 1970-е годы внимание автора и его учеников уделялось повышению эффективности МГУА и изучению его прикладных возможностей без строгого обоснования. Эффективность метода многократно подтверждалась решением конкретных задач моделирования сложных объектов и процессов в областях экологии, экономики, техники и гидрометеорологии [4]. Наглядность метода привела к быстрому росту

его популярности среди прикладников и в нашей стране, и за рубежом. Параллельно развивались исследования по разработке алгоритмов самоорганизации в смежной к моделированию области – в задачах классификации и распознавания образов [3, 4].

В 1980-е годы продолжались интенсивные исследования по применению МГУА к решению практических задач: моделирование экологических процессов и систем [4], водных и нефтяных полей [5], экономических процессов [6]. В частности, некоторые результаты решения таких задач зарубежными авторами изложены в [7, 8].

В начале 1980-х годов А.Г. Ивахненко установил органичную аналогию между задачей построения модели по зашумленным экспериментальным данным и задачей прохождения сигнала через канал с шумом [9]. Это позволило, используя идеи теории помехоустойчивого приема сигналов Котельникова, построить теорию помехоустойчивого моделирования, основы которой изложены в [10], где приведен детальный анализ алгоритмов переборного (комбинаторного) типа.

Основной результат теории: сложность оптимальной прогнозирующей модели зависит от уровня шума в данных – чем выше уровень, тем проще должна быть оптимальная модель. Это составляет основу теории МГУА как метода автоматической адаптации сложности оптимальной модели к уровню информативности зашумленной выборки данных, что и составляет суть самоорганизации моделей. В те же годы были получены существенные теоретические результаты в исследовании многорядных (итерационных) алгоритмов МГУА и условий их сходимости [11].

В 1994 г. в США вышла книга А.Г. Ивахненко об индуктивных обучающихся алгоритмах моделирования [12], содержащая основные результаты его школы 1980-х – начала 1990-х гг. Эти результаты содержат алгоритмы, основанные на таких идеях, как *объективный системный анализ* (ОСА) [13] и *объективная компьютерная кластеризация* (ОКК) [14], а также непараметрическое прогнози-

вание (корреляционные модели [15] и ком-плексообразование групповых аналогов [16]).

Однако главным результатом 1990-х годов стала идея А.Г. Ивахненко о дальнейшем развитии возможностей МГУА как нейронной сети: структура МГУА является специфической вычислительной сетью с нейронами, имеющими полиномиальную функцию активации, для которой принято название *полиномиальная нейронная сеть* (например, [17]). Идея развития состоит в том, что если создать сеть из нейронов в виде алгоритмов МГУА, полученная структура превратится в сеть с «активными» нейронами, архитектура каждого из которых настраивается на конкретную выборку данных. Это направление стало весьма популярным среди исследователей и продолжает развиваться как «нейросети с активными нейронами» [18]. В эти же годы МГУА начали относить к наиболее успешным средствам интеллектуального анализа данных [19].

О развитии теории МГУА. В теоретических исследованиях МГУА анализируются прежде всего алгоритмы генерации (перебора) вариантов структур моделей, а также критерии выбора (селекции) лучших из генерируемых моделей. Основные свойства *переборных* алгоритмов МГУА изложены в [10], а *итерационных* – в [11] (включая вопросы их сходимости).

Задача анализа эффективности критериев селекции моделей рассматривается в рамках теории помехоустойчивого моделирования на основе МГУА [10], опирающейся на аналогию между задачами теорий связи и моделирования. Продуктивные результаты связаны с анализом свойств так называемого «идеального» *J*-критерия, впервые введенного в [20] и исследованного в [10, 21, 22]. Асимптотические свойства внешних критериев селекции моделей изучались в [23–25]. Сформированы также основы теории МГУА для объектов с многомерным выходом [26, 27].

Основным аналитическим аппаратом теории индуктивного моделирования является *метод критических дисперсий*, который можно эффективно применять как для ограниченной выборки, так и в асимптотике. Этот аппарат, впервые предложенный в [20], был применен в [10] и по-

следовательно развивался в [28–30]. В систематизированном виде метод представлен в [31].

Общая постановка задачи моделирования по экспериментальным данным

Задача построения моделей по экспериментальным данным может быть сведена к поиску экстремума некоторого критерия CR на дискретном множестве различных моделей Φ :

$$f^* = \operatorname{argmin}_{f \in \Phi} CR(f). \quad (1)$$

Для конкретизации этой постановки предположим, что задана выборка $W = [Xy]$, содержащая n наблюдений за m входными (независимыми) переменными и одной выходной, и эти данные представлены в виде матрицы $X = [x_{ij}, i = 1, \dots, n; j = 1, \dots, m]$ и вектора $y = (y_1 \dots y_m)^T$, причем $n \geq m$.

В общем случае процесс решения задачи структурно-параметрической идентификации (1) включает в себя прежде всего следующие основные этапы:

1. задание выборки данных и априорной информации;
2. выбор или задание класса базисных функций и преобразование данных;
3. генерация различных структур моделей в этом классе;
4. оценивание параметров генерируемых структур и формирование множества Φ ;
5. минимизация заданного критерия $CR(f)$ и выбор оптимальной модели f ;
6. проверка адекватности полученной модели;
7. принятие решения о завершении процесса и/или применение модели.

Перечисленные этапы описывают произвольный процесс построения моделей, причем в зависимости от априорной информации и цели моделирования те или иные этапы могут отсутствовать. Например, в случае задачи параметрической идентификации множество Φ состоит из одной модели (задается одна структура), т.е. исключаются этапы 3 и 5 процесса.

Задача идентификации состоит в формировании некоторого множества Φ моделей вида

$$\hat{y}_f = f(X, \hat{\theta}_f), \quad f \in \Phi, \quad (2)$$

и поиске оптимальной модели по условию

$$f^* = \operatorname{argmin}_{f \in \Phi} CR(y, f(X, \hat{\theta}_f)), \quad (3)$$

причем оценки $\hat{\theta}_f$ для каждой $f \in \Phi$ являются решением еще одной экстремальной задачи:

$$\hat{\theta}_f = \operatorname{argmin}_{\theta_f \in R^{s_f}} QR(y, X, \theta_f), \quad (4)$$

где s_f – сложность модели f , равная числу оцениваемых параметров в модели (2), а $QR \neq CR$ – критерий качества решения задачи *параметрической* идентификации каждой частной модели, генерируемой в задаче *структурной* идентификации.

Для оценивания параметров в алгоритмах МГУА применяется прежде всего МНК, когда в качестве QR в (4) выступает функция потерь (сумма квадратов ошибок):

$$QR(\theta_f) = \|y - \hat{y}_f\|^2 = \|y - X\hat{\theta}_f\|^2. \quad (5)$$

В предположении, что в измерениях y присутствует независимый шум с нулевым средним и конечной дисперсией, наилучшим методом является МНК, тогда решение задачи (4), (5) имеет вид

$$\hat{\theta}_f = (X_f^T X_f)^{-1} X_f^T y, \quad (6)$$

где X_f – подматрица матрицы X со столбцами, соответствующими частной модели $f \in \Phi$.

Экстремальные задачи (3), (4) содержат ряд неопределенностей, свойственных задаче моделирования по данным наблюдений и так или иначе влияющих на качество ее решения. Виды неопределенностей можно разделить на две основные группы: относящиеся к данным, т.е. к физической информации о моделируемом объекте (*структурная, стохастическая* и *информационная* неопределенность), а также к средствам обработки данных, т.е. к применяемой технологии моделирования (*функциональная, параметрическая* и *критериальная* неопределенность).

Методы моделирования можно анализировать с точки зрения степени учета всех указанных видов неопределенности. Здесь крайние позиции занимают регрессионный анализ и МГУА: первый требует снятия всех неопределенностей, а второй обеспечивает их автоматический учет.

Из приведенной постановки следует, что каждый метод моделирования можно описать четырьмя основными компонентами: класс моделей, генератор структур, метод оценивания параметров и критерий селекции моделей. Все эти компоненты присутствуют в алгоритмах МГУА, и это целесообразно использовать для характеристики особенностей метода.

Классы моделей, применяемые в МГУА

В практике моделирования встречаются следующие основные *типы задач*:

- построение регрессионных моделей *статических* объектов;
- моделирование *временных* рядов или процессов;
- моделирование *динамических* объектов, процессов и систем.

Модели статических объектов. Термин *статический объект* объединяет все случаи, когда необходимо построить модель вида

$$y_k = b^T g(x_k) = \sum_{j=1}^m b_j g_j(x_k) + \xi_k, \quad (11)$$

где $k=1, \dots, n$ – номер точки наблюдения, b_j – параметры модели, $g(x)$ – известная m -мерная вектор-функция от MX входных переменных, ξ_k – независимый некоррелированный случайный вектор с нулевыми средним и конечной дисперсией.

Базисными функциями $g(x)$ в алгоритмах МГУА служат прежде всего полиномы:

$$b^T g(x) = \sum_{j=1}^m b_j \prod_{v=1}^{MX} x_v^{c_{jv}}, \quad (12)$$

тогда каждая компонента $g(x)$ является одночленом этого полинома. При этом степени c_{jv} могут принимать целые значения от нуля до ST , причем сумма всех m степеней не превышает ST .

Модели временных рядов. Для моделирования процессов как временных рядов в алгоритмах МГУА применяются полиномиальные, тригонометрические, экспоненциальные, логистические и логарифмические функции времени. В случае тригонометрических функций модели имеют вид

$$b^T g(t) = a_0 + \sum_{j=1}^m (\alpha_j \sin \omega_j t + \beta_j \cos \omega_j t) =$$

$$= a_0 + \sum_{j=1}^m b_j \sin(\omega_j t + \varphi_j), \quad (13)$$

т.е. $g_j(t) = \sin(\omega_j t + \varphi_j)$ при заданных ω_j, φ_j , причем частоты предполагаются некрратными.

Для моделирования временных рядов применяются также авторегрессионные модели:

$$y_k = \sum_{i=1}^{LY} a_i y_{k-i} + \xi_k, \quad (14)$$

где LY – порядок авторегрессии (число запаздываний).

Модели динамических объектов. Задачи моделирования динамики решаются, прежде всего, в классе линейных динамических моделей при наличии вектора внешних воздействий x . Тогда общий вид модели динамики многомерной системы представляется разностным уравнением

$$y_k = \sum_{\alpha=1}^{LY} a_{\alpha}^T y_{k-\alpha} + \sum_{v=1}^{LX} b_v^T x_{k-v+1} + \xi_k, \quad (15)$$

где LY, LX – число учитываемых прошлых значений (запаздываний, лагов) для выходных и входных переменных, a_{α}, b_v – векторы неизвестных параметров, ξ_k – входной процесс типа белого шума. Из (15) можно получить предыдущие классы моделей (11) и (14) как частные случаи.

В каждом алгоритме МГУА имеется блок преобразования данных в соответствии с заданным классом моделей, после чего формируется и решается задача структурно-параметрической идентификации (3), где специфика типа исходной задачи уже не учитывается. Например, в одномерном случае ($MY=1$) выражения (12) и (15) приводятся к виду обычной регрессионной модели

$$y_i = f(x_i, \theta) + \xi_i = x_i^T \theta + \xi_i, \quad (16)$$

если в (12) обозначить $\theta = \overset{\Delta}{b}$, $x = g(x)$, а в (15) – соответственно $\theta = (\overset{\Delta}{a_1}, \dots, \overset{\Delta}{a_{LY}}, \overset{\Delta}{b_1}, \dots, \overset{\Delta}{b_{LX}})^T$, $x = (\overset{\Delta}{y_{k-1}}, \dots, \overset{\Delta}{y_{k-LY}}, \overset{\Delta}{x_k}, \dots, \overset{\Delta}{x_{k-LX+1}})^T$. Тогда задача (3) состоит в определении оптимальной структуры и значений параметра θ , или построении модели с оптимальным подмножеством аргументов.

Далее рассмотрим обобщенное представление моделей вида (16), линейных по параметрам

рам, а в разработанных алгоритмах имеется большинство упомянутых классов моделей.

Сравнительный анализ методов генерации структур моделей

Большинство известных генераторов структур МГУА естественным образом разделяются на две группы – переборные и итерационные, по аналогии с методами оптимизации. Рассмотрим характерные особенности этих методов, отличающихся способами генерации структур моделей и организацией поиска минимума заданного критерия.

Переборные алгоритмы [10]. Они предназначены для решения задачи перебором моделей из конечного множества Φ , элементы которого могут быть вычислены независимо путем оценки параметров всех различных моделей. При этом задача (1) оказывается аналогичной задаче дискретного программирования и может быть решена полным или направленным перебором.

Наиболее известна реализация *полного перебора* в виде комбинаторного алгоритма *COMBI* [32]. В случае линейного объекта с m входами генерируются все возможные модели вида

$$y_v = X_v \hat{\theta}_v, \quad v = 1, \dots, 2^m, \quad (17)$$

где десятичному числу v ставится в соответствие двоичный структурный вектор d_v , в котором единицы указывают на присутствие аргументов, нули – на их отсутствие. Здесь число сравниваемых моделей равно 2^m , и полный перебор эффективен только примерно до $m = 25$.

Цель методов *направленного перебора* – отыскать глобальный минимум $CR(s)$, т.е. результат полного перебора при существенно меньших вычислениях. Они работоспособны при достаточно больших m – свыше 100, например, алгоритм *MULTI* [33], в котором применяется процедура вида

$$\hat{y}_s^i = (X_{s-1}^i | x_s^i) \hat{\theta}_s, \quad s = \overline{1, m}, i, l = \overline{1, F_{s-1}}, \quad (18)$$

где s – номер этапа (и сложность структуры); F_s – число лучших структур (свобода выбора); j – индекс регрессора, отсутствующего в матрице X_{s-1}^i . При этом анализируется подмножество $\Phi_s \subseteq \Phi$, с большой вероятностью содержащее

результат полного перебора. Если $F_s = C_m^s$, то Φ_s и Φ совпадают – перебор становится полным. Алгоритм имеет полиномиальную сложность m^3 .

Вопросам исследования и применения переборных алгоритмов посвящена монография [34], более подробное описание всех существующих алгоритмов этого типа дано в [35].

Итерационные алгоритмы [11]. Они известны только в теории самоорганизации моделей, и все их варианты (как и переборные алгоритмы) относятся к группе алгоритмов МГУА. По принципу действия итерационные методы подобны методам оптимизации посредством последовательных приближений, однако используют принцип неокончательных решений (свободу выбора F).

В зависимости от способа последовательных приближений итерационные алгоритмы МГУА можно разделить на две группы: *многорядные* и *релаксационные*. *Многорядные* алгоритмы построены на аналогии с биологической селекцией живых организмов: усложнение моделей от r -го ряда селекции к $(r + 1)$ -му происходит путем попарного «скрещивания» лучших моделей предыдущего ряда по линейному или нелинейному частному описанию, например:

$$y_i^{r+1} = f_i(y_i^r, y_j^r) = a_{1l} + a_{2l} y_i^r + a_{3l} y_j^r + a_{4l} y_i^r y_j^r, \quad (19)$$

$$r = 0, 1, \dots; \quad i, j = \overline{1, F}; \quad l = \overline{1, C_F^2},$$

где F – заданная «свобода выбора» – количество отобранных лучших моделей предыдущего ряда, а при $r = 0$ справа учитываются исходные регрессоры. В таких алгоритмах процедура усложнения прекращается после начала возрастания значений критерия CR .

В алгоритмах МГУА *релаксационного* типа на каждом ряде (шаге приближения) модели усложняются «скрещиванием» лучших моделей предыдущего ряда с исходными аргументами:

$$y_i^{r+1} = f_i(y_i^r, x_k) = a_{1l} + a_{2l} y_i^r + a_{3l} x_k + a_{4l} y_i^r x_k, \quad (20)$$

$$i = \overline{1, F}; \quad k = \overline{1, m}.$$

Вопросам описания, исследования и применения современных высокоэффективных алгоритмов МГУА релаксационного типа посвящена монография [36].

Отметим, что в 1970–80 гг. были сконструированы варианты алгоритмов итерационного типа [11]. Там же приведена оригинальная структура итерационного алгоритма СМЛ, основанная на переборе не пар, а троек аргументов:

$$y_i^r = a_i^{r-1} y_i^{r-1} + b_i^{r-1} y_j^{r-1} y_k^{r-1}. \quad (21)$$

Следует указать, что для итерационных алгоритмов, в отличие от переборных, множество моделей можно описать, но нельзя просмотреть ввиду его бесконечности, а в процессе решения формируется некоторое его (бесконечное) подмножество, зависящее от параметров алгоритма. При этом увеличивается разнообразие генерируемых структур моделей, но возникает необходимость доказательства сходимости итерационного процесса решения задачи (1), которое пока получено только для случая так называемой «внутренней сходимости» [11, 36].

Однако высокую эффективность этих алгоритмов демонстрируют многочисленные примеры решения практических задач моделирования. В целом итерационные методы работоспособны при весьма больших m – порядка 1000. Они позволяют строить удовлетворительные модели даже в вырожденных задачах, когда длина выборки $n < m$. Современное структурированное описание архитектур различных итерационных алгоритмов МГУА представлено в [37].

Выражения (17)–(21) упрощенно представляют четыре основных типа генераторов структур в алгоритмах МГУА. Все они строят модели, линейные по параметрам, но в общем случае нелинейные по входам. Отметим, что в прикладном регрессионном анализе имеются простые варианты перебора (методы всех регрессий, включения–исключения, шаговой регрессии), в которых не используется принцип неокончательных решений и внешние критерии, т.е. алгоритмы МГУА предоставляют более широкие возможности генерации моделей. Итерационные генераторы МГУА – оригинальны, и само название МГУА ассоциируется прежде всего с ними.

Внешние критерии качества моделей

Такие критерии предназначены для применения в задаче (3) и основаны на разделении выборки на две или более частей [10] – напри-

мер, в простейшем случае на две непересекающиеся подвыборки (подмножества точек) A и B , $A \cup B = W$.

Приведем вычислительные формулы для основных внешних критериев, приняв следующие стандартизованные обозначения: МНК-оценка параметров на некоторой подвыборке G равна

$$\hat{\theta}_G = (X_G^T X_G)^{-1} X_G^T y_G, \quad G = A, B, W, \quad (22)$$

а значение ошибки на некоторой подвыборке Q по модели, оценки параметров которой вычислены на G , равно

$$\Delta(Q|G) = \|y_Q - X_Q \hat{\theta}_G\|^2, \quad Q = A, B, W, C. \quad (23)$$

Отметим, что разбиение выборки данных на две подвыборки соответствует следующему:

$$W = [X : y] = \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} X_A : y_A \\ X_B : y_B \end{bmatrix}, \quad n = n_A + n_B. \quad (24)$$

Критерии точности. При $Q = G$ величина $\varepsilon_G = \Delta(G|G)$ – остаточная сумма квадратов RSS , а при $G \neq Q$ она является так называемым *критерием регулярности*:

$$AR_B = \Delta(B|A), \quad AR_A = \Delta(A|B). \quad (25)$$

Применяется также симметричный критерий регулярности: $AD = \Delta(B|A) + \Delta(A|B)$.

Критерии согласованности. Основным в этой группе является критерий *непротиворечивости* (несмещенности), определяемый выражениями

$$CB = \left\| X_W \hat{\theta}_A - X_W \hat{\theta}_B \right\|^2 = \\ = (\hat{\theta}_A - \hat{\theta}_B)^T X_W^T X_W (\hat{\theta}_A - \hat{\theta}_B). \quad (26)$$

К данной группе относится также критерий *вариативности*

$$CV = (\hat{\theta}_A - \hat{\theta}_W)^T X_W^T X_W (\hat{\theta}_W - \hat{\theta}_B) = \\ = (\varepsilon_W - (\varepsilon_A + \varepsilon_B)). \quad (27)$$

Имеются также критерии с многократным разбиением выборки – например, JN (джек-найф или «усредненный критерий регулярности» (УКР)). Подробнее о критериях МГУА в [10, 11].

Современные тенденции в развитии алгоритмов самоорганизации моделей

МГУА как полиномиальная нейросеть. В 1990-х годах, когда искусственные нейронные сети получили широкую известность, А.Г. Ивах-

ненко и многочисленные пользователи его метода начали называть типичную структуру МГУА также *нейросетью*. Теперь за границей многорядный алгоритм МГУА чаще всего называют *Polynomial Neural Network (PNN)*, т.е. Полиномиальная Нейронная Сеть (ПНС) [17, 19]. При этом основной элемент итерационных алгоритмов МГУА, а именно частное описание, можно также рассматривать как элементарный нейрон ПНС МГУА. Оригинальность и эффективность нейросети, состоящей из таких нейронов, заключается в скорости процесса локальной настройки их весов и автоматической глобальной оптимизации (*самоорганизации*) структуры сети (числа узлов и количества рядов или *скрытых слоев*). Более детально полиномиальные нейросети МГУА и вопросы их использования рассмотрены в [19].

Нейросеть с активными нейронами. Типичный *нейрон МГУА* в виде квадратичного полинома от двух аргументов или его упрощенного варианта (19) можно назвать *пассивным*, поскольку он имеет фиксированную структуру, и все другие нейроны (частные описания) имеют такую же структуру, т.е. классическая полиномиальная нейросеть МГУА – *гомогенна*. В 1990-х годах А.Г. Ивахненко предложил новый тип сети МГУА – с активными нейронами [18], или *гетерогенную* сеть, в которой каждый из нейронов является, в свою очередь, также алгоритмом МГУА, благодаря чему структура нейрона оптимизируется. В результате все нейроны могут получить разную структуру, что повышает гибкость настройки сети на конкретную задачу. Такие сети называются также *дважды многорядными* [19].

МГУА-подобная нейросеть с обратной связью. В [38] была предложена МГУА-подобная нейронная сеть, имеющая цикл обратной связи: выходы нейронов каждого слоя «скрещиваются» с исходными переменными. В этом алгоритме сеть автоматически выбирает одну из трех архитектур для каждого активного нейрона: сигмоидальную, радиальную или полиномиальную переходную функцию. При этом структурные параметры, например, количество слоев, нейронов

в скрытых слоях и входных переменных выбираются автоматически, т.е. самоорганизуются.

Эволюция групп адаптивных моделей. Эволюционный алгоритм под названием *GAME (group of adaptive models evolution)*, разработанный в Чешском техническом университете в Праге, базируется на общей архитектуре МГУА [39]. В *GAME* коэффициенты различных передаточных функций каждого узла сети оцениваются по данным выборки.

Главные модификации этой МГУА-подобной системы с активными нейронами:

- передаточная функция узла может иметь несколько типов – линейная, полиномиальная, логистическая, РБФ и другие, а также в виде перцептрона, причем выбор типа узлов, образующих сеть, определяется значением критерия – т.е. структура сети – *гетерогенна* (неоднородна);
- количество входов узла растет с ростом глубины узла в сети, передаточные функции узлов отражают растущее количество входов, причем в сети существуют междурядные связи;
- при построении сети находятся не все размещения узлов, а их случайные подмножества;
- оригинальный МГУА создает одну оптимальную модель, а данный алгоритм – группу моделей, которые локально оптимальны для своего подмножества данных.

В такой модифицированной архитектуре МГУА начальные значения весов и коэффициентов узлов назначаются случайно, их входы и передаточные функции также отбираются случайно. Поэтому топология моделей, построенных по одной и той же выборке, может отличаться.

Другие пути развития идей самоорганизации моделей. В настоящее время имеется много различных тенденций в развитии методов моделирования на основе МГУА: создание гибридных архитектур в сочетании с алгоритмами вычислительного интеллекта [40]; построение нейросетей с нечеткими функциями активации [41]; разработка алгоритмов моделирования на принципах действия иммунных систем [42]; решение задач моделирования при нестохастических предположениях – так называемые интервальные модели [43]; применение распараллелива-

ния операций [44]; использование онтологических подходов [45] и многое другое.

Обзоры исторических аспектов развития индуктивного моделирования представлены в [46, 47]. Более подробную информацию о разработанных алгоритмах, средствах и технологиях, их исследовании и применении можно найти на базовом сайте [48] и на сайте [49] отдела ИТИМ.

Специализация отдела Информационных технологий индуктивного моделирования

Исследовательская стратегия отдела ИТИМ охватывает полный жизненный цикл научного поиска: *методологию* моделирования по выборкам данных; *теорию* индуктивного построения моделей оптимальной сложности; *алгоритмизацию* высокопродуктивных средств моделирования; *интеллектуализацию* технологий построения моделей; *компьютерные эксперименты* для оценивания эффективности разработанных технологий; *решение реальных задач* моделирования и прогнозирования; апробацию и прикладное *внедрение* разработанных инструментальных средств в системах мониторинга, управления и поддержки принятия решений.

Приоритетные *научные направления* деятельности отдела таковы:

- развитие методологии структурной идентификации оптимальных прогнозных моделей сложных процессов в условиях неполноты и неопределенности априорной информации;
- развитие теории индуктивного моделирования сложных процессов и систем по зашумленным статистическим и экспериментальным данным на основе МГУА;
- создание высокопродуктивных компьютерных средств индуктивного моделирования на основе рекуррентных процедур и кластерных технологий распараллеливания вычислений;
- создание интеллектуальных информационных технологий и инструментальных средств индуктивного моделирования и прогнозирования сложных процессов и систем;
- решение прикладных задач анализа, моделирования и прогнозирования экономических, экологических и технологических процессов с

целью выявления их закономерностей, оптимизации, управления и принятия решений;

- разработка интеллектуальных технологий информационной поддержки решений в процессах управления разного уровня на основе методов и средств индуктивного моделирования.

Профильные научно-прикладные направления исследований отдела связаны с решением задач моделирования и прогнозирования в областях:

- социально-экономические процессы, прежде всего в сфере экономической безопасности государства, с целью информационной поддержки управленческих решений;
- экологические процессы (в водной, воздушной и грунтовой средах);
- технологические процессы (например, задачи мониторинга литейного производства);
- медико-биологические процессы (оценивание индивидуальной эффективности лекарственных средств и построение классификаторов для задач дифференциальной диагностики).

Основные научные результаты отдела ИТИМ

Разработана *теория индуктивного моделирования* на основе МГУА как эффективного метода автоматического построения оптимальных прогнозирующих моделей сложных объектов и процессов на основе статистических и экспериментальных данных в условиях неполноты и неопределенности информации. Аналитическим аппаратом теории служит *метод критических дисперсий* [31], позволивший объяснить природу эффективности МГУА как метода построения оптимальных моделей с минимальной дисперсией ошибки прогнозирования в условиях неопределенности. В рамках теории решена задача оптимизации разбиения выборки данных на две части по условию близости ковариационных (информационных) матриц этих частей [36].

Получила дальнейшее развитие теория МГУА в результате разработки и теоретического обоснования двухкритериального *метода доопределения* выбора модели с использованием нового критерия несмещенности ошибок [50], позволяющего устранить неоднозначность выбора мо-

дели в реальных задачах моделирования. В этом методе последовательно применяются критерий регулярности для отбора наиболее точных моделей и несмещенности ошибок для окончательного выбора модели оптимальной сложности [34].

Разработаны теоретические и прикладные принципы конструирования и реализации *высокоэффективных переборных алгоритмов* МГУА на основе рекуррентных вычислений [34], процедур интеллектуального распараллеливания операций [44] и последовательной селекции информативных переменных [51], что повышает размерность решаемых задач анализа, моделирования и прогнозирования сложных процессов и систем.

Разработан метод комплексного *интегрального оценивания* и прогнозирования состояния сложных иерархически организованных систем, представленных статистическими данными о динамике изменения групп взаимосвязанных первичных показателей [52]. Построение интегральных индексов базируется на оригинальном методе нелинейной нормализации статистических показателей.

Разработаны теоретические и прикладные принципы конструирования гибридных архитектур итерационных алгоритмов МГУА как обобщение структур алгоритмов многорядного, релаксационного и комбинаторного типов, на основе чего создан *обобщенный итерационный алгоритм* (ОИА) МГУА [37], существенно повышающий эффективность решения задач индуктивного моделирования сложных систем. В архитектуре ОИА как нейросети реализована идея *активных нейронов* в виде комбинаторного алгоритма МГУА для автоматической настройки сложности (степени и структуры полинома) нейронов.

Разработан, теоретически обоснован и исследован *обобщенный релаксационный итерационный алгоритм* (ОРИА) МГУА, базирующийся на использовании быстродействующих рекуррентных вычислений и матриц нормальных уравнений, что позволяет эффективно решать задачи индуктивного моделирования по выборкам сверхбольшой размерности [36].

Разработана концепция, теоретические основы и математическое обеспечение принципиально нового класса переборных и итерационных алгоритмов МГУА с применением *рекуррентно-параллельных вычислений* на кластерных системах, что способствует созданию высокопродуктивных интеллектуальных программных средств и технологии индуктивного моделирования, эффективность которых пропорциональна количеству процессоров при обеспечении равномерной их загрузки [53].

Разработаны концептуальные и теоретические принципы конструирования технологий *интеллектуального моделирования* сложных систем на основе использования баз знаний, средств индуктивного анализа данных и интеллектуального интерфейса пользователя [54]. Такие технологии должны иметь три основных инструментальных уровня: *автономное* интеллектуальное моделирование поведения системы по имеющейся базе данных; *встроенное* интеллектуальное моделирование в составе системы управления или принятия решений в реальном времени; *комплексное* интеллектуальное моделирование сложной системы для автоматического выявления как оптимальных режимов ее работы, так и критических сценариев.

Разработаны теоретические принципы и инструментальные средства интеллектуального моделирования и прогнозирования многомерных взаимосвязанных социально-экономических процессов по статистическим данным в классе дискретных динамических *моделей векторной авторегрессии* [55]. Эти средства служат основой конструирования систем информационной поддержки оперативных управленческих решений для органов государственного управления разных уровней, в том числе в сфере экономической безопасности Украины.

Технологии и инструментальные средства

Компьютерная технология АСТРИД [56] предназначена для построения математических моделей сложных объектов и процессов по статистическим (экспериментальными) данным с целью выявления закономерностей, идентификации, прогнозирования для задач информационной поддержки принятия решений. Предо-

ставляет возможность строить модели статических объектов, временных рядов и динамических объектов и процессов в таких классах структур: линейных, полиномиальных, авторегрессионных, разностных (динамических), нелинейных сетевого типа и др.

Комплекс инструментальных средств [34] для конструирования, исследования и применения методов моделирования и их составляющих элементов, проведения экспериментов по компьютерному тестированию методов моделирования и их компонентов (классов моделей, генераторов структур моделей, методов оценивания параметров и критериев селекции моделей) средствами разработанного инструментального комплекса.

Кроссплатформенный *программный комплекс с развитым интерфейсом* [57] на языке *Java* для индуктивного моделирования и прогнозирования сложных объектов и процессов большой размерности на основе быстродействующего переборного алгоритма МГУА с последовательной селекцией информативных и/или отсевом неинформативных аргументов [51].

Программный комплекс АСТРИД-ОИА [58] для индуктивного моделирования сложных систем на основе различных итерационных алгоритмов МГУА позволяет использовать обобщенный алгоритм ОИА МГУА и все его частные случаи [37] в режиме доступа онлайн как через Интернет, так и в локальной сети. Интерфейс комплекса дает возможность строить модели в трех основных режимах: двух автоматических – стандартном и планируемом, а также интерактивном, когда процесс самоорганизации модели можно корректировать или останавливать на любом этапе с сохранением и дальнейшим восстановлением промежуточных результатов.

Компьютерная система АСПИС [59] для построения прогнозирующих моделей сложных систем по данным наблюдений на основе быстродействующего обобщенного релаксационного итерационного алгоритма ОРИА МГУА [36]. Реализована на языке программирования *C++* в объектно-ориентированном стиле с использованием современных паттернов проектирования, позволяет решать задачи с тысячами аргументов и сотнями тысяч наблюдений.

Система информационной поддержки управленческих решений (ИПУР) позволяет в автоматизированном режиме решать задачи комплексного оценивания, анализа и прогнозирования состояния сложных систем взаимосвязанных социально-экономических процессов с целью принятия обоснованных и эффективных решений в разных сферах государственной и предпринимательской деятельности [60]. Опытный образец системы прошел апробацию как инструментальной информационной поддержки управленческих решений в сфере экономической безопасности на основе выявления скрытых закономерностей развития социально-экономических процессов и устранения неэффективных решений.

Основные прикладные результаты исследований

Совместно с Институтом микробиологии и вирусологии им. Д.К. Заболотного НАН Украины на основе опытных данных многолетних натуральных наблюдений построены математические модели динамики численности микроорганизмов в почве зависимо от экологических факторов и дозы загрязнения тяжелыми металлами [61] с целью экологического мониторинга и оценки эффективности реабилитации загрязненных почв.

По статистическим данным Минэкономразвития Украины построены многомерные разностные модели динамики реального экономического процесса в виде 11 взаимозависимых показателей энергетической сферы Украины в классе моделей множественной авторегрессии, позволяющие строить краткосрочный прогноз сложного векторного процесса [34]. Этот подход успешно применен также для моделирования и прогнозирования взаимозависимых показателей инвестиционной сферы [55].

Совместно с Московским авиационным институтом построена нелинейная модель зависимости коэффициента распыления (разрушения) поверхности космического аппарата под действием струй ионизированных газов в зависимости от физических свойств разных материалов покрытия поверхности [34]. Модель описывает этот объект точнее, чем традиционные модели с физической параметризацией, позволяет рассчи-

тывать значения коэффициентов распыления в зависимости от свойств потенциально применяемых материалов и подбирать материал с наименьшим коэффициентом распыляемости.

Совместно с Институтом биологии южных морей НАН Украины построены и внедрены модели для количественного оценивания экологических последствий влияния загрязнений морской воды битумоидными веществами на общее число видов донных организмов в бухтах Севастопольская, Южная и Карантинная, на основе чего разработана методика зонирования загрязнений в бухтах города Севастополя [62].

Совместно с Национальной медицинской академией последипломного образования МОЗ Украины им. П.Л. Шупика разработана и апробирована технология индуктивного моделирования и прогнозирования результатов тестирования образцов крови разными медицинскими препаратами. Разработана методика, позволяющая сократить временные и финансовые затраты на проведение предварительных клинических исследований с целью определения наиболее эффективных препаратов для лечения конкретного пациента [36].

Совместно с Институтом гематологии и трансфузиологии АМН Украины разработана технология автоматического построения классификаторов для дифференциальной диагностики заболеваний крови с целью повышения точности диагностики и снижения рисков постановки ошибочного диагноза при сокращении временных и финансовых затрат [36].

Совместно с Физико-технологическим институтом металлов и сплавов НАН Украины разработан проект комплексной компьютерной технологии поддержки решений в процессе охлаждения металлической отливки [63]. Решена задача построения моделей зависимости промежуточной температуры отливки от параметров, влияющих на процесс охлаждения, с целью оптимизации режима литья для обеспечения надлежащего качества изделия.

Перспективы исследований в области индуктивного моделирования

Развитие методологии и теории интеллектуального моделирования на основе МГУА:

- концептуальные вопросы развития методологии индуктивного моделирования на новых парадигмах нейросетей с активными нейронами и гибридных архитектур;

- теоретические исследования эффективности и точности решения задач индуктивного моделирования в условиях неопределенности и неполноты информации в статистических выборках данных, типичных для экономических, экологических и технологических задач;

- развитие теории МГУА как эффективного метода построения помехоустойчивых моделей для прогнозирования, классификации, распознавания и кластеризации сложных процессов и систем по экспериментальным данным.

Разработка теории и оригинальных архитектур итерационных (нейросетевых) алгоритмов МГУА с принципиально новыми свойствами:

- повышение эффективности и точности решения задач структурно-параметрической идентификации моделей сложных систем на основе обобщения структур известных алгоритмов;

- разработка новых гибридных архитектур итерационных алгоритмов на основе идеи активных нейронов и применения эволюционных и мультиагентных подходов;

- анализ и доказательство сходимости новых итерационных алгоритмов к истинной модели по структуре и параметрам.

Развитие теории и вычислительных структур переборных алгоритмов МГУА:

- радикальное повышение эффективности переборных алгоритмов на основе распараллеливания и применения рекуррентных процедур;

- разработка новых процедур индуктивного поиска оптимальной модели в переборных алгоритмах на основе эффективных схем последовательной селекции лучших вариантов моделей;

- конструирование гибридных переборных алгоритмов на основе процедур решения задач дискретной оптимизации с применением идей вычислительного интеллекта.

Развитие методологии, теории и алгоритмов индуктивного решения задач классификации, распознавания и кластеризации:

- разработка теории и алгоритмов индуктивного построения оптимальных правил классификации для задач со сверхбольшими объемами данных;

- разработка высокоэффективных методов индуктивного решения задач многоальтернативного распознавания на основе итерационных алгоритмов МГУА;

- разработка интеллектуальных алгоритмов анализа и автоматической кластеризации данных для задач выявления закономерностей.

Разработка интеллектуальных технологий информационной поддержки решений на основе методов и средств индуктивного моделирования:

- разработка инструментальных средств в форме конструктора алгоритмов как оболочки для интерактивного синтеза и автоматического программирования новых средств моделирования;

- разработка средств интеллектуального интерфейса для эффективной поддержки принятия решений пользователя в процессе решения задач индуктивного моделирования на основе МГУА;

- разработка комплекса инструментальных средств для поддержки конструирования интеллектуальных технологий информационной поддержки решений по управлению сложными системами различной природы.

Применение разработанных методов и средств в прикладных задачах интеллектуальной поддержки решений на основе технологий индуктивного моделирования:

- поддержка оперативных решений по управлению социально-экономическими процессами разного уровня, в том числе в сфере экономической безопасности государства;

- поддержка решения задач медицинской диагностики и прогнозирования эффективности лекарственных средств и процедур;

- интеллектуальное моделирование и управление техническими и технологическими процессами и робототехническими системами;

- поддержка решений по управлению экологическим состоянием процессов разного уровня в водохозяйственных системах, орошаемом зем-

леделии, морских прибрежных зонах, других задачах, решаемых на основе данных мониторинга окружающей среды;

- поддержка решения задач автоматического поиска и анализа категорий, жанров и структур текстовой цифровой информации;

- выявление знаний на основе автоматизированного анализа контента сообщений в социальных сетях с целью поддержки принятия управленческих решений разного уровня;

- повышение эффективности поиска целевой информации в сети Интернет путем индуктивного моделирования и содержательной селекции результатов поисковой выдачи.

Заключение. Благодаря обзору состояния исследований в области индуктивного моделирования можно утверждать, что МГУА – это перспективная основа современных информационных технологий получения знаний из данных наблюдений, или один из оригинальных и наиболее эффективных методов интеллектуального анализа данных.

Развитие идей самоорганизации моделей предполагает совершенствование теории индуктивного моделирования, разработку и реализацию новых высокопродуктивных алгоритмов в современных компьютерных технологиях и применение этих технологий для решения широкого спектра реальных прикладных задач моделирования, прогнозирования, управления и принятия решений в системах различной природы – экономических, экологических, технологических и многих других.

1. *Ивахненко О.Г.* Метод группового учета аргументов – конкурент методу стохастичної апроксимації // Автоматика. – 1968. – № 3. – С. 58–72.
2. *Ивахненко А.Г.* Системы эвристической самоорганизации в технической кибернетике. – Киев: Техніка, 1971. – 392 с.
3. *Ивахненко А.Г.* Долгосрочное прогнозирование и управление сложными системами. – Там же, 1975. – 311 с.
4. *Ивахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наук. думка, 1982. – 296 с.
5. *Ивахненко А.Г., Пека Ю.П., Востров Н.П.* Комбинированный метод моделирования водных и нефтяных полей. – Там же, 1984. – 185 с.

6. *Ивахненко А.Г., Мюллер Й.А.* Самоорганизация прогнозирующих моделей. – Киев: Техніка, 1985. – 223 с.
7. *Ivachnenko A.G., Müller J.A.* Selbstorganisation von Vorhersagemodellen. – Berlin: VEB Verlag Technik, 1984. – 223 с.
8. *Self-organizing methods in modeling: GMDH type algorithms* / Ed. S.J. Farlow. – New York, Basel: Marcel Dekker Inc., 1984. – 350 p.
9. *Ивахненко А.Г., Карпинский А.М.* Самоорганизация моделей на ЭВМ в терминах общей теории связи (теории информации) // Автоматика. – 1982. – № 4. – С. 7–26.
10. *Ивахненко А.Г., Степашко В.С.* Помехоустойчивость моделирования. – Киев: Наук. думка, 1985. – 216 с.
11. *Ивахненко А.Г., Юрачковский Ю.П.* Моделирование сложных систем по экспериментальным данным. – М.: Радио и связь, 1987. – 120 с.
12. *Madala H.R., Ivakhnenko A.G.* Inductive learning algorithms for complex systems modeling. – New York: Boca Raton, CRC Press, 1994. – 384 p.
13. *Ивахненко А.Г., Костенко Ю.В., Голуусов И.В.* Системный анализ и долгосрочное количественное прогнозирование квазистатических систем на основе самоорганизации моделей. Ч. 2. Объективный системный анализ без априорного указания внешних воздействий // Автоматика. – 1983. – № 3. – С. 3–11.
14. *Ивахненко А.Г.* Объективная компьютерная кластеризация на основе теории самоорганизации моделей // Там же. – 1987. – № 5. – С. 1–9.
15. *Ивахненко А.Г., Кротов Г.И., Строчкова Т.И.* Самоорганизация безразмерных гармонико-экспоненциальных и корреляционных прогнозирующих моделей стандартной структуры // Там же. – 1984. – № 4. – С. 18–29.
16. *Ivakhnenko A.G.* Inductive Sorting Method for the Forecasting of Multidimensional Random Processes and Events with the Help of Analogs Forecast Complexing // Pattern Recognition and Image Analysis. – 1991. – 1, N 1. – P. 99–108.
17. *Kondo T.* GMDH Neural Network Algorithm Using the Heuristic Self-Organization Method and its Application to the Pattern Identification Problem // Proc. of the 37th SICE Annual Conf. – SICE'98. – Tokyo: IEEE, 1998. – P. 1143–1148.
18. *Ivakhnenko A.G., Ivakhnenko G.A., Mueller J.A.* Self-Organization of Neuronets with Active Neurons // Pattern Recognition and Image Analysis. – 1994. – 4, N 4. – P. 177–188.
19. *Muller J.-A., Lemke F.* Self-Organizing Data Mining. An Intelligent Approach to Extract Knowledge from Data. – Berlin, Dresden, 1999. – 225 p.
20. *Степашко В.С.* Потенциальная помехоустойчивость моделирования по комбинаторному алгоритму МГУА без использования информации о помехах // Автоматика. – 1983. – № 3. – С. 18–27.
21. *Кочерга Ю.Л.* J-оптимальная редукция структуры модели в схеме Гаусса–Маркова // Автоматика. – 1988. – № 4. – С. 34–36.
22. *Аксенова Т.И., Юрачковский Ю.П.* Характеризация J-оптимальной модели в задаче моделирования по выборкам наблюдений // Там же. – № 4. – С. 37–43.
23. *Степашко В.С.* Асимптотические свойства внешних критериев селекции моделей // Там же. – № 6. – С. 75–82.
24. *Дышин О.А.* Асимптотические свойства помехоустойчивости критериев точности моделей // Там же. – 1989. – № 5. – С. 53–56.
25. *Аксенова Т.И.* Достаточные условия сходимости внешних критериев выбора моделей // Там же. – С. 53–56.
26. *Сарычев А.П.* Системный критерий регулярности в методе группового учета аргументов // Проблемы управления и информатики. – 2006. – № 6. – С. 25–37.
27. *Сарычев А.П.* Идентификация состояний структурно-неопределенных систем. – Днепропетровск: Ин-т техн. механики НАНУ и НК АУ, 2008. – 268 с.
28. *Степашко В.С.* Исследование прогнозирующих свойств рекуррентного структурно-параметрического идентификатора // Автоматика. – 1991. – № 3. – С. 33–41.
29. *Степашко В.С.* Структурная идентификация прогнозирующих моделей в условиях планируемого эксперимента // Там же. – 1992. – № 1. – С. 26–35.
30. *Степашко В.С.* Анализ эффективности критериев структурной идентификации прогнозирующих моделей // Проблемы управления и информатики. – 1994. – № 3–4. – С. 13–21.
31. *Степашко В.С.* Метод критических дисперсий как аналитический аппарат теории индуктивного моделирования // Там же. – 2008. – № 2. – С. 8–26.
32. *Степашко В.С.* Комбинаторный алгоритм МГУА с оптимальной схемой перебора моделей // Автоматика. – 1981. – № 3. – С. 31–36.
33. *Степашко В.С.* Конечная селекционная процедура сокращения полного перебора моделей // Автоматика. – 1983. – № 4. – С. 84–88.
34. *Степашко В.С., Єфіменко С.М., Савченко Є.А.* Компьютерный эксперимент в индуктивном моделировании. – К.: Наук. думка, 2014. – 222 с.
35. *Мороз О.Г., Степашко В.С.* Порівняльний аналіз генераторів структур моделей у перебірних алгоритмах МГУА // Індуктивне моделювання складних систем. – 8. – К.: МННЦ ІТтаС НАНУ, 2016. – С. 133–148.
36. *Павлов А.В., Степашко В.С., Кондрашова Н.В.* Эффективные методы самоорганизации моделей. – К.: Академперіодика, 2014. – 200 с.
37. *Степашко В.С., Булгакова О.С.* Обобщенный итерационный алгоритм метода группового учета аргументов // УСиМ. – 2013. – № 2. – С. 5–17.
38. *Kondo T., Ueno J.* Feedback GMDH-Type Neural Network Self-Selecting Optimum Neural Network Architecture and Its Application to 3-Dimensional Medical

- Image Recognition of the Lungs / Proc. of the II Int. Workshop on Inductive Modeling IWIM-2007, 19–23 Sept. 2007. – Prague: Czech Techn. Univ., 2007. – P. 63–70. – ISBN 978-80-01-03881-9.
39. *Kordik P.* Fully automated knowledge extraction using group of adaptive model evolution: Ph.D. thesis. / Electrical Engineering and Inform. Techn. – Prague: CTU, 2006. – 150 p.
 40. *Hybrid Particle Swarm Optimization and Group Method of Data Handling for Inductive Modeling* / G. Onwubolu, A. Sharma, A. Dayal et al. // Proc. of 2nd Int. Conf. on Inductive Modeling (ICIM'2008). – Kyiv: IRTC ITS NASU, 2008. – P. 95–103.
 41. *Bodyanskiy Ye., Zaychenko Yu., Pavlikovskaya Ye.* The Neo-Fuzzy Neural Network Structure Optimization Using the GMDH for the Solving Forecasting and Classification Problems / Proc. of the 3rd Int. Workshop on Inductive Modeling IWIM-2009, 14–19 Sept. 2009, Krynica, Poland. – Prague: Czech Techn. Univ., 2009. – P. 100–107.
 42. *Lytvynenko V.* Hybrid GMDH Cooperative Immune Network For Time Series Forecasting // Proc. of the 4th Int. Conf. on Inductive Modeling ICIM-2013, Sept. 16–20, 2013, Kyiv. – Kyiv: IRTC ITS NASU, 2013. – P. 179–187.
 43. *Дивак М.П.* Задачі моделювання статичних систем з інтервальними даними. – Тернопіль: Економічна думка, 2011. – 216 с.
 44. *Stepashko V., Yefimenko S.* Parallel algorithms for solving combinatorial macromodeling problems // *Przegląd Elektrotechniczny (Electrical Review)*. – 2009. – **85**, N 4. – P. 98–99.
 45. *Валькман Ю.Р., Степашко П.В.* На пути построения онтологии интеллектуального моделирования // *Индуктивное моделирование складных систем*. – 7. – К.: МННЦ ІТтаС НАНУ, 2015. – С. 101–115.
 46. *Степашко В.С.* Элементы теории индуктивного моделирования / Стан та перспективи розвитку інформатики в Україні / Кол. авторів. Гол. ред. акад. Сергієнко І В. – К.: Наук. думка, 2010. – 1008 с. / С. 481–496.
 47. *Степашко В.С.* Самоорганизация прогнозирующих моделей сложных процессов и систем // ХУ Всерос. науч.-техн. конф. «Нейроинформатика-2013»: Лекции по нейроинформатике / Отв. ред. Ю.В. Тюменцев – М.: НИЯУ МИФИ, 2013. – 320 с. / С. 150–170.
 48. www.gmdh.net
 49. www.mgua.irtc.org.ua
 50. *Ивахненко А.Г., Савченко Е.А.* Исследование эффективности метода доопределения выбора модели в задачах моделирования с применением МГУА // *Проблемы управления и информатики*. – 2008. – № 2. – С. 65–76.
 51. *Самойленко А.А.* Весовой критерий определения информативности аргументов в методах построения моделей с последовательной селекцией переменных // *УСиМ*. – 2013. – № 2. – С. 33–39.
 52. *Степашко В.С., Мельник І.М., Волощук Р.В.* Моделі синтезу інтегральної оцінки стану складної системи взаємозв'язаних первинних показників // *Моделювання та керування станом еколого-економічних систем регіону: Зб. наук. праць*. – 3. – К.: МННЦ ІТтаС НАНУ, 2006. – С. 275–284.
 53. *Yefimenko S., Stepashko V.* Intelligent Recurrent-and-Parallel Computing for Solving Inductive Modeling Problems // Proc. of 16th Int. Conf. on Computational Problems of Electrical Engineering CPEE'2015, Lviv, Ukraine, Sept. 2–5, 2015. – Львів: Львівська політехніка, 2015. – 274 с. / P. 236–238.
 54. *Степашко В.С.* Концептуальные основы интеллектуального моделирования // *УСиМ*. – 2016. – № 4. – С. 3–15.
 55. *Ефименко С.Н.* Построение систем прогнозных моделей многомерных взаимосвязанных процессов // Там же. – С. 80–86.
 56. *Степашко В.С., Кона Ю.В.* Опыт применения системы АСТРИД для моделирования экономических процессов по статистическим данным // *КВТ*. – 1998. – **117**. – С. 24–31.
 57. *Самойленко О.А.* Проектування перебірних алгоритмів МГУА як основних компонентів підсистеми моделювання // *Индуктивное моделирование складных систем*. – 3. – К.: МННЦ ІТтаС НАНУ, 2011. – С. 191–208.
 58. *Булгакова О.С., Зосімов В.В., Степашко В.С.* Програмний комплекс моделювання складних систем на основі ітераційних алгоритмів МГУА з можливістю мережевого доступу // *Системні дослідження та інформаційні технології*. – 2014. – № 1. – С. 43–55.
 59. *Павлов А.В.* Проектирование системы автоматизированной структурно-параметрической идентификации // *Индуктивное моделирование складных систем*. – 7. – К.: МННЦ ІТтаС НАНУ, 2015. – С. 202–219.
 60. *Stepashko V., Samoilenko O., Voloschuk R.* Informational Support of Managerial Decisions as a New Kind of Business Intelligence Systems. – *Computational Models for Business and Engineering Domains* / G. Setlak, K. Markov (Eds.). – Rzeszow, Poland; Sofia, Bulgaria: ITNEA, 2014. – 298 p. / – P. 269–279.
 61. *Іутиньська Г.О., Кона Ю.В., Степашко В.С.* Моделювання динаміки чисельності мікроорганізмів у ґрунті, забрудненому важкими металами // *Мікробіологічний журнал*. – 2002. – **64**. – № 3. – С. 59–67.
 62. *Альмов С.В., Булгакова О.С., Степашко В.С.* Моделювання впливу забруднення Чорного моря на загальне число видів донних організмів: Зб. наук. праць СНУАЕтаП. – 3 (39). – 2011. – С. 54–62.
 63. *Токова О.В., Савченко Є.А.* Підхід до розроблення системи інформаційної підтримки рішень у ливар-

UDC 621.513.8

V.S. Stepashko

The Achievements and Prospects of Inductive Modeling

Keywords: Inductive modeling, GMDH, self-organization, structural-parametric identification, data mining, decision-making.

Introduction. Inductive modeling as the scientific school was formed by academician Alexey Ivakhnenko and is still actively being developed. The worldwide known Group method of data handling (GMDH) as a means of self-organizing models keeps the central place in this studying. The concept “inductive modeling” can be defined as a self-organizing process of evolutionary transition from primary data to mathematical models reflecting patterns of the simulated systems functioning, implicitly contained in the available experimental or statistical data.

Purpose. The typical problems solved by means of inductive modeling are described, information on this scientific school in Ukraine and abroad is presented, the fundamental basic and the technological achievements are defined, the most promising ways of further research are formulated.

Methods. The goal of this article is being achieved by presenting a comprehensive survey of the main publications in this area and structuring the material in accordance with the historical, fundamental, technological and applied aspects of the inductive modeling.

Results. As a result of the references survey in the field of inductive modeling, it can be stated that the GMDH is a promising basis of modern information technology for discovering knowledge from data, or one of original and efficient methods of data mining.

Conclusion. The further development of the model assumes the improvement of the theory of inductive modeling, the development and implementation of new high-performance algorithms in up-to-date computer technologies and the application of these technologies to solve a wide range of real-life problems of modeling, forecasting, control and decision-making in systems of different fields – economic, ecological, technological and others.



Для соответствия научно-метрическим базам при подаче статей к рассмотрению, авторы должны подать метаданные на английском языке:

- ФИО
- место и адрес работы каждого автора
- расширенную аннотацию (до 2000 знаков с пробелами и рубриками:
Introduction, Purpose, Methods, Results, Conclusion)
- список пристатейной литературы в переводе или транслитерации.

При оформлении списков литературы к расширенной аннотации на английском языке, можно пользоваться сайтом

<http://translit.net> для русских ссылок

<http://ukrlit.org/transliteratsiia> для украинских.