

15. Ker A. General Framework for Structural Steganalysis of LSB Replacement. *VII International Workshop on Information Hiding*. Berlin, Springer-Verlag, 2005. Vol. 3727, pp. 296–311.
16. Barsukov V.S. Steganograficheskie tekhnologii zashchity dokumentov, avtorskikh prav i informatsii [Steganographic security technology for documents, copyrights and information]. *Obzor spetsialnoy tekhniki*, 2000, no. 2, pp. 31–40.
17. Craver S. On public-key steganography in the presence of an active warden. *Information Hiding, II International Workshop*. Portland, Oregon, 1998, vol. 1525, pp. 355–368.
18. Gribunin V.G., Okov I.N., Turintsev I.V. *Tsifrovaya steganografiya* [Digital steganography]. Moscow, Solon-Press Publ., 2002. 272 p.
19. Agranovskiy A.V., Devyanin P.N., Khadi R.A., Cheremushkin A.V. *Osnovy kompyuternoy steganografii* [Fundamentals of computer steganography]. Moscow, Radio i svyaz Publ., 2003. 152 p.
20. Kokhanovich G.F., Puzyrenko A.Yu. *Kompyuternaya steganografiya. Teoriya i praktika* [Computer steganography. Theory and practice]. Kiev, MK-Press Publ., 2006. 288 p.

УДК 004.4::004.85[3+5]

## АВТОМАТИЧЕСКАЯ СИСТЕМА МЕТА-ОБУЧЕНИЯ С ПОДДЕРЖКОЙ ВЫБОРА ОПТИМАЛЬНОГО АЛГОРИТМА РЕШЕНИЯ ЗАДАЧИ И ВЫЧИСЛЕНИЯ ОПТИМАЛЬНЫХ ПАРАМЕТРОВ ЕГО ФУНКЦИОНИРОВАНИЯ

Орлов Андрей Александрович,

ассистент каф. промышленной электроники Томского Государственного  
Университета Систем Управления и Радиоэлектроники,  
634050, Россия, Томск, пр. Ленина, д. 40. Email: d1scnc@gmail.com

**Актуальность исследования** обусловлена необходимостью повышения эффективности работы автоматических систем интеллектуального анализа данных, основанных на мета-обучении.

**Цель исследования** состоит в разработке автоматической системы мета-обучения с поддержкой выбора оптимального алгоритма решения задачи и вычисления оптимальных параметров его функционирования.

**Методы исследования:** индуктивное моделирование, методы статистической обработки результатов.

**В результате исследования** проведена систематизация известных систем мета-обучения на основании выработанных классификационных признаков, учитывающих внутреннюю организацию систем. Сформулированы требования к реализации автоматической системы мета-обучения. Предложен способ построения системы мета-обучения, удовлетворяющей всем сформулированным требованиям и производящей накопление мета-знаний, построение на их основе мета-моделей, выбор оптимального алгоритма из набора доступных и вычисление оптимальных параметров его функционирования. Разработана объектно-ориентированная архитектура программной платформы для реализации любой из систем мета-обучения, представленных в систематизации. Эффективность реализованной автоматической системы мета-обучения с использованием алгоритмов методов группового учета аргументов проверена экспериментально при решении набора задач, относящихся к классу задач прогнозирования временных последовательностей (1428 временных последовательностей из тестового набора, известного под названием «M3 Competition»).

**Ключевые слова:**

Мета-обучение, мета-характеристики данных, мета-модель, программная платформа, объектно-ориентированный анализ и проектирование, прогнозирование временных последовательностей, метод группового учета аргументов.

**Введение**

Задачей Интеллектуального Анализа Данных (ИАД, в англоязычной литературе используется термин «Data Mining») является обнаружение (извлечение) в доступных исследователю исходных данных ранее неизвестных, неочевидных, но практически полезных знаний [1]. В настоящее время существует большое количество алгоритмов искусственного интеллекта (включая машинное обучение), математической статистики, оптимизации и прогнозирования и пр., применяемых для решения задачи ИАД: искусственные нейронные сети, генетические алгоритмы, деревья решений, алгоритмы нечеткой логики, корреляционный и регрессионный анализ и т. д. Каждый из существующих алгоритмов показал свою эффективность при решении разнообразных практических задач. Од-

нако в работе [2] было показано, что не существует единственного алгоритма, способного максимально эффективно решать задачу ИАД во всех возможных практических применениях, поэтому решение каждой новой практической задачи требует привлечения некоторых экспертных знаний для выбора наиболее подходящего алгоритма из числа доступных.

В работе [3] был формализован подход к проблеме выбора алгоритма (рис. 1): на основании набора мета-характеристик (meta-features, MF)  $f(x) \in F$  ( $F$  – пространство мета-характеристик) для выборки данных  $x$  из пространства проблем (задач)  $X$  функция  $S(f(x))$  («selection mapping») производит выбор такого алгоритма  $a$  из пространства доступных алгоритмов  $A$  таким образом, что его эффективность  $p(a, x)$  («performance mapping») на

выборке данных  $x$  максимальна ( $p(a, x)$  – показатель эффективности, ПЭ).

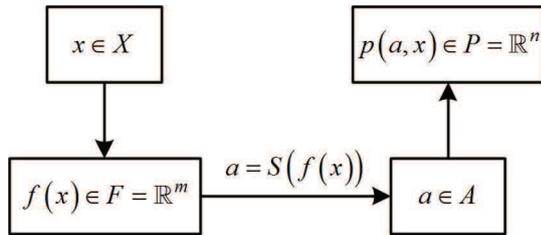


Рис. 1. Выбор алгоритма на основании мета-характеристик данных

Подход, предложенный в работе [3], предполагал, что выполнение функции «selection mapping» возложено на эксперта в области ИАД, либо в качестве  $S(f(x))$  используется составленный экспертом набор правил (эвристик) выбора алгоритма. Очевидным недостатком применения подхода являлась необходимость привлечения экспертов на стадии разработки системы ИАД либо на стадии ее использования. С одной стороны это требовало дополнительных материальных и временных затрат, а с другой стороны ограничивало применение данного подхода в автоматических системах ИАД. Более того, возникала проблема эффективности накопления экспертных знаний и улучшения имеющихся эвристик выбора алгоритма, поскольку данные процессы основывались на увеличении уровня квалификации самого эксперта.

Решение описанных проблем производилось исследователями в рамках направления «мета-обучение», области ИАД, изучающей способы увеличения эффективности «обучающихся систем» ИАД путем накопления информации об эффективности обучения в таких системах (показатели эффективности обучения, их взаимосвязь с характеристиками задачи, особенности функционирования алгоритмов решения задачи и т. д.), в том числе и в автоматическом режиме [4].

#### Обзор существующих систем мета-обучения

До настоящего времени было предложено множество вариантов построения систем мета-обучения – например, в аналитических обзорах каждой из 10 глав только одной книги [5], посвященной мета-обучению, упомянуты более 40 систем, разрабатываемых различными научными коллективами. Несмотря на то, что в публикациях по системам мета-обучения приводится аналитический обзор и сравнение с некоторыми существующими системами, выделяются лишь некоторые внешние критерии оценки систем, как то результирующая точность, скорость принятия системой решений и т. д. Известна работа [5. С. 117–156], в которой сделана попытка описания строения систем мета-обучения, однако рассматривается лишь состав внутренних модулей без выделения общности в их организации.

В связи с этим автором было принято решение о разработке классификационных признаков, учи-

тывающих внутреннюю организацию систем, с целью систематизации существующих систем мета-обучения. Ниже описаны сформулированные классификационные признаки, согласно которым в табл. 1, 2 приводится информация для каждой из систем.

1. **Класс** разделяет приведенные системы на 4 группы по их предназначению:

- a. **Algorithm Selection** – для заданной выборки данных произвести выбор оптимального алгоритма из набора доступных (согласно задаваемому ПЭ);
- b. **Parameter Tuning** – для заданной выборки данных определить оптимальные значения параметров функционирования заданного алгоритма;
- c. **Data Mining Assistant** – выработать экспертные рекомендации, содействующие специалисту по ИАД в выборе оптимального алгоритма для заданной выборки данных;
- d. **Multi-Level Data Mining** – организовать процесс автоматического многоуровневого ИАД, при котором осуществляется не только перебор алгоритмов ИАД и их параметров, но и сравнение и выбор алгоритмов поиска оптимальных алгоритмов ИАД и оптимизации их параметров.

2. **Режим** распределяет ответственность за осуществление той или иной функции, системы мета-обучения: «Р» – ручной, когда принятие решений полностью возложено на эксперта; «А» – автоматический, когда эксперт уведомляется о решениях, принимаемых системой мета-обучения. Знаком «-» отмечены случаи, когда функция не поддерживается системой. В табл. 1, 2 рассматриваются следующие функции систем мета-обучения:

- a. **обучение** – способность системы производить накопление мета-знаний об эффективности доступных алгоритмов ИАД, мета-характеристиках данных и их взаимосвязи (в терминах работы [3]: определение вида и/или построение функции «selection mapping»  $S$ );
- b. **использование** – осуществление действий по выбору/рекомендации оптимального алгоритма ИАД и/или параметров его функционирования на основе анализа поставленной задачи – анализа заданной выборки данных (в терминах работы [3] для выбора алгоритма: вычисление мета-характеристик данных  $f$  и применение  $S$  для выбора алгоритма  $a$ );
- c. **выбор мета-характеристик** (значимых) – определение подмножества заданных мета-характеристик, отражающих лишь значимые свойства выборок данных, определяющих корректный выбор алгоритма для решения поставленной задачи. Иными словами, это исключение нерепрезентативных мета-характеристик;

- d. **настройка параметров алгоритмов** – определение таких значений параметров функционирования алгоритмов, при которых достигается максимизация ПЭ при решении поставленной задачи;
- e. **проверка эффективности** (лучшего алгоритма) – оценка допустимости применения выбранного системой алгоритма для решения поставленной задачи. Это необходимо для детектирования случаев, когда все доступные системе мета-обучения алгоритмы неудовлетворительно решают поставленную задачу, поэтому выбор лучшего из них не гарантирует ожидаемое высокое значение ПЭ.
3. **Вид мета-характеристик** определяет набор групп мета-характеристик, используемых системой для анализа различных свойств заданной выборки данных:
- a. **КОЛ** – количественные: длина выборки, число переменных, число классов в задаче классификации и т. п.;
- b. **ОБЛ** – область происхождения выборки данных (финансы, промышленность, информационные технологии и т. д.);
- c. **СТАТ** – статистические: среднее, минимальное и максимальное значения, дисперсия, корреляция и т. п.;
- d. **ГЕОМ** – (для задач классификации) – характеристики пересечений классов по входным данным, делимость, диаметры классов и расстояния между классами в пространстве параметров и т. д.;
- e. **ПОВ** – поведение временной последовательности: характеристики тренда, периодичность, автокорреляция, число точек изменения направления и т. п.;
- f. **ЭКСП** – экспертные знания: особенности области применения, описание входных и выходных данных и т. д.;
- g. **ХАОС** – показатель Ляпунова, размерность фрактального пространства и т. д.;
- h. **ОРИ** – ориентирование, landmarking – запуск относительно быстрых и простых алгоритмов ИАД, значение ПЭ которых есть мета-характеристика данных;
- i. **ГРАФ** – характеристики графов: число вершин, ребер, показатели связности, оценка решения сверху и снизу, статистические показатели и др.;
- j. **ШАБ** – (для алгоритмов) – содержание в алгоритме ИАД определенных структурных шаблонов обработки данных.
4. **Показатель эффективности алгоритма и системы** указывает способ расчета ПЭ отдельных доступных алгоритмов и системы в целом, включая известные критерии качества решения задач ИАД.
5. **Содержание базы мета-данных** иллюстрирует способ представления мета-знаний в виде хранимой информации для последующего использования.
6. **Мета-модель** содержит информацию о входных и выходных величинах мета-модели (разделены знаком «->»).
7. **Способы получения оптимального алгоритма и настройки оптимальных параметров алгоритма** определяет процедуру поиска оптимального алгоритма из доступных и/или вычисления оптимальных параметров его функционирования: **ВЫБ** – выбор из фиксированного множества, **ПЕР** – перебор элементов вычисляемого набора значений, **ОПТ** – решение задачи оптимизации.
8. **Задача** – краткая характеристика практического применения, на котором разработчиками показана эффективность системы мета-обучения. Помимо количественных характеристик указан вид задачи: **КС** – классификация, **КЗ** – кластеризация, **РЕ** – регрессия, **ПР** – прогнозирование, **РГ** – раскраска графов, **УК** – задача об упаковке в контейнеры (Bin Packing Problem, BPP).
- Следует отметить, что в табл. 1, 2 представлены системы мета-обучения, действующие с применением возможностей вычислительной техники. Аналитический обзор других систем мета-обучения (например, экспертных систем мета-обучения, систем построения онтологий и пр.) был выполнен в работе [5. С. 117–156]. Знаком «?» и наклонной штриховкой отмечены ячейки, информации по которым недостаточно.
- На основании анализа табл. 1, 2 можно сделать следующие выводы.
1. Большинство систем мета-обучения (17 из 18) поддерживают **автоматический режим обучения**, что избавляет от необходимости трудоемкого ручного формирования базы мета-знаний. Также большинство систем, относящихся к классам Algorithm Selection и Multi-Level Data Mining, поддерживают **автоматический режим использования** (11 из 12). Это позволяет включать данные системы в состав автоматических систем ИАД.
  2. Используются 10 видов мета-характеристик данных, причем наиболее часто в задаче классификации используются виды КОЛ и СТАТ (соответственно, 7 и 5 из 11), а в задаче прогнозирования – ПОВ и СТАТ (3 и 2 из 3). Это подтверждает, что репрезентативность различных мета-характеристик данных определяется видом задачи, решаемой системой мета-обучения, и подчеркивает актуальность выбора значимых мета-характеристик. Однако **автоматический режим выбора мета-характеристик** поддерживается лишь в 4 системах из 18.
  3. 11 систем из 18 осуществляют формирование мета-модели на основе накопленных при обучении мета-знаний, причем только 3 системы строят модели, позволяющие **прогнозировать значения ПЭ** на основании мета-характеристик заданной выборки данных, а значит, и производить **оценку качества решения**.

4. 8 систем из 18 позволяют производить **вычисление оптимальных параметров** функционирования алгоритмов, причем 6 из них – **в автоматическом режиме**. При этом только 3 системы из 6 основываются на **анализе свойств заданной выборки** данных (на расчете мета-характеристик). Остальные системы опираются на значение внешнего ПЭ, что увеличивает вероятность получения смещенных оценок оптимальных параметров функционирования алгоритмов (в том числе вследствие явления «переобучения» – «overfitting»).
5. Только одна система оценивает эффективность полученного оптимального алгоритма и допустимость его применения для решения поставленной задачи, остальные системы осуществляют выбор и/или генерируют рекомендации безотносительно предполагаемой эффективности работы выбранного алгоритма (качества решения задачи). Это ограничивает применение данных систем в отсутствии контроля со стороны эксперта.

На основании данных выводов сформулируем **требования к реализации автоматической системы мета-обучения**, лишенной упомянутых недостатков.

**Требование А:** Поддержка автоматических режимов обучения, использования, выбора значимых мета-характеристик, настройки параметров алгоритмов и проверки эффективности.

**Требование В:** Накопление мета-знаний и построение на их основе мета-моделей, отражающих закономерности в процессах обучения системы и алгоритмов.

**Требование С:** Расчет оптимальных параметров функционирования и выбор оптимального алгоритма на основании анализа свойств заданной выборки данных (мета-характеристик) и предсказания производительности алгоритмов.

**Требование Д:** Поддержка принятия решений о допустимости применения найденного оптимального алгоритма и его оптимальных параметров для решения поставленной задачи.

В табл. 1, 2 для каждой существующей системы проведен анализ ее соответствия сформулированным требованиям, причем степень соответствия с точки зрения возможности и трудоемкости доработки каждой системы отмечена соответствующим цветом ячейки (примечание к табл. 2). Из анализа следует, что доработка каждой существующей системы является либо трудоемкой, либо приводящей к существенному изменению ее структуры. Таким образом, разработка системы мета-обучения, удовлетворяющей всем сформулированным требованиям, является актуальной.

#### Организация системы мета-обучения

С целью построения системы мета-обучения, удовлетворяющей сформулированным выше требованиям, в работе [21] автором был предложен способ, представленный на рис. 2.

Выбор алгоритма  $a$  из пространства доступных алгоритмов  $A$  и набора его оптимальных параметров функционирования  $q^a$  из пространства параметров  $Q^a$  производится функцией  $L(f(x))$  на основании набора мета-характеристик  $f(x) \in F$  ( $F$  – пространство мета-характеристик) для выборки данных  $x$  из пространства проблем  $X$  таким образом, что эффективность  $p(a(q^a), x)$  алгоритма  $a$  на выборке данных  $x$  максимальна ( $p(a(q^a), x)$  – ПЭ).

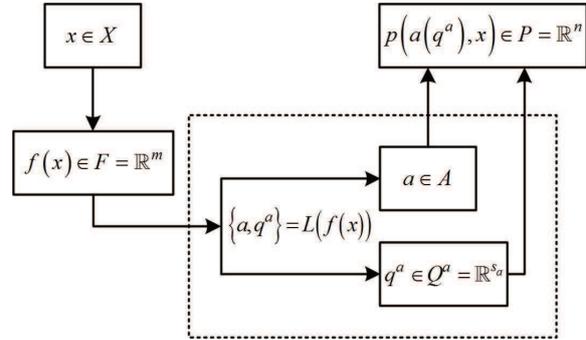


Рис. 2. Выбор алгоритма и параметров его функционирования на основании мета-характеристик данных

Рассмотрим подробнее процесс обучения предлагаемой системы мета-обучения.

Пусть система располагает информацией о некотором наборе из  $K$  выборок данных  $x_k$ , где  $k=1, \bar{K}$ . Для каждой выборки данных значение мета-характеристик рассчитывается как  $f_k = f(x_k)$ , причем  $f_k = [f_k^1, f_k^2, \dots, f_k^m]$  – вектор из  $m$  мета-характеристик данных ( $m$  – размерность  $F$ ).

Системе доступен набор из  $G$  алгоритмов  $a_g$ , где  $g=1, \bar{G}$ .

Система также обладает информацией об эффективности каждого из доступных ей алгоритмов  $a_g: q_{i,j}^g$ , где  $i \in I_g \subseteq I = \{1, 2, \dots, G\}$  – информация о производительности алгоритма  $a_g$  может быть доступна не для всех выборок  $x_k; j=1, J_i^g, J_i^g$  – количество наборов параметров  $q_{i,j}^g$  алгоритма  $a_g$ , для которых доступно значение показателя эффективности на выборке  $x_i$ . При этом каждый набор параметров  $q_{i,j}^g = [q_{i,j}^{g,1}, q_{i,j}^{g,2}, \dots, q_{i,j}^{g,s_g}]$  – вектор из  $s_g$  параметров алгоритма  $a_g$  (разные алгоритмы могут иметь различное число параметров  $s_g$ ). Составим матрицу мета-характеристик  $M^F$ , матрицы параметров алгоритма  $M_{g,i}^Q$  и матрицы показателей эффективности  $M_{g,i}^P$ :

$$M^F = \begin{bmatrix} f_1^1 & \dots & f_1^m \\ \dots & \dots & \dots \\ f_K^1 & \dots & f_K^m \end{bmatrix}, \quad M_{g,i}^Q = \begin{bmatrix} q_1^1 & \dots & q_1^{s_g} \\ \dots & \dots & \dots \\ q_{J_i^g}^1 & \dots & q_{J_i^g}^{s_g} \end{bmatrix},$$

$$M_{g,i}^P = \begin{bmatrix} p_1^1 & \dots & p_1^n \\ \dots & \dots & \dots \\ p_{J_i^g}^1 & \dots & p_{J_i^g}^n \end{bmatrix}.$$

С учетом введенных обозначений для каждого алгоритма  $a_g$  и выборки данных  $x_i$  составим матрицы доступной мета-информации для выборки данных  $x_i$ :

Таблица 1. Обзор систем мета-обучения. Класс «Algorithm Selection»

Класс системы	Система мета-обучения	Мета-обучение											Задача			
		Режим обучения	Режим использования	Вид мета-характеристик	Режим выбора MF	Показатель эффективности алгоритма	Содержание базы мета-данных (содержание записей)	Мета-модель	Способ получения оптимального алгоритма	Режим настройки параметров алгоритмов	Способ настройки параметров алгоритмов	Показатель эффективности системы	Режим проверки эффективности	Количество выборок данных	Количество MF	Количество алгоритмов
Algorithm Selection	NOEMON Prudencio, Ludermir [6]	A	A	ПОВ	P	MAE	MF, Algo -> MAE	MF -> Binary Compare (MLP)	ВЫБ Ранжирование	-	- (def)	SRC	P	ПП		
	Prudencio, Ludermir [7]	A	A	СТАТ	P	?	MF -> вектор рангов алгоритмов	MF -> Binary Compare (MLP)	ВЫБ Ранжирование	-	-	Average Error Rate	P	PE		
	Romero, Olmo, Ventura [8]	A	A	КОЛ, сложности, ОБЛ	P	5 ПЭ: Sen, Prec, F-M, Кар, AUC	MF -> вектор рангов алгоритмов	-	Nearest Neighbour, Euclidean distance on MF	-	- (def)	F-Measure (precision, recall)	P	KC		
	Matijas, Suykens, Krajcar [9]	A	A	СТАТ, ПОВ	A	MASE + NRMSE	MF -> вектор рангов алгоритмов	MF->rank (объединение 7 алгоритмов классификации)	ВЫБ Ранжирование	-	- (def)	MASE NRMSE MAPE	P	ПП		
	Cacoveanu, Vidrighin, Potolea [10]	A	A	КОЛ СТАТ ОРИ	P	CA	MF, Algo -> CA	-	ВЫБ наиболее похожих случаев (kNN)	-	-	CA	P	KC		
	Cruz-reyes et. al. [11]	A	A	СТАТ	P	Число контейнеров	MF, Algo -> Group/algo (kNN)	MF->Group (Discriminant Analysis, C4.5, SOM)	ВЫБ Применение мета-модели – классификатора	-	-	Число контейнеров	P	УК		
	Wang, Smith-Miles, Hyndman [12]	A P	A	СТАТ ПОВ ХАОС	P	?	MF, Algo -> ranking (FA)	MF->ranking (SOM)	ВЫБ (DT)	-	-	?	A	ПП		
	Abdelmessih, Shafait, Reif, Goldstein [13]	A	A	КОЛ СТАТ ОРИ	P	CA	MF -> CA	MF -> CA (SVM, RMSE)	ВЫБ (на основе прогноза ПЭ)	A	ПЕР (GS)	(RMSE для мета-моделей)	P	KC		
	Molina, Luna, Romero, Ventura [14]	A	P	КОЛ	P	CA	MF, Params -> Variation Accuracy	MF+Params -> Variation Accuracy (DT)	-	P	ПЕР	Variation of Accuracy	P	KC		
	Musliu, Schwengerer [15]	A	A	ГРАФ	P	Минимальное число цветов	MF -> Число цветов	MF -> ranking (6 АК)	ВЫБ (применение АК)	-	- (def)	SSR	P	ПГ		
													645	5	3	
													50	10	?	
													32	16	19	
													65 + 4	13	7	
													13 + 2	11	7	
													243 0	5	7	
													315	13	4	
													90	7	7	
													14	3	1	
													859	8 + 82	6	

Таблица 2. Обзор систем мета-обучения. Классы «Parameter Tuning», «Data Mining Assistant» и «Multi-level Data Mining»

Класс системы	Система мета-обучения	Мета-обучение											Задача			
		Режим обучения	Режим использования	Вид мета-характеристик	Режим выбора MF	Показатель эффективности алгоритма	Содержание базы мета-данных (содержание записей)	Мета-модель	Способ получения оптимального алгоритма	Режим настройки параметров алгоритмов	Способ настройки параметров алгоритмов	Показатель эффективности системы	Режим проверки эффективности	Количество выборок данных	Количество MF	Количество алгоритмов
Parameter Tuning	Konen et al. [16]	A	P	-	A	WCE	-	-	-	A	ОПТ	WCE	P	KC		
	Soares, Bradzil [17]	A	P	КОЛ СТАТ	P	NMSE	MF, Params -> NMSE	MF -> ranking (SOM)	-	A	ПЕР (kNN)	SRC	P	PE		
	Reif, Shafait, Dengel [18]	A	A	КОЛ СТАТ ОРИ	A	CA	MF -> Optimal Params (GS)	-	-	A	ПЕР (CBR) + ОПТ (GA)	CA	P	KC		
Data Mining Assistant	Sleeman, Rissakis et al. [19]	P	P	КОЛ ЭКСП	P	CA; Calcula- tion time; complexi- ty etc.	MF -> ПЭ	набор пра- вил выбора алгоритма	ВЫБ согласно набору правил	P	ОПТ (диа- лого- вый ре- жим)	ПЕР	P	1	7	10
	Kalousis, Hilario [20]	A	A	СТАТ	P	CE	MF, Algo -> Class.Error	-	ВЫБ наиболее похожих (CBR)	-	- (def)	Strict Error Loose Error	P	KC/[K3]/ [PE]		
	Hilario, Nguyen, Do, Woznica, Kalousis [5. С. 179–224]	A	P	Данные: СТАТ ГЕОМ Алго- ритм: ШАБ	A	CA	Data MF, Algo MF -> CA	Data MF, Algo MF -> CA	ВЫБ (на осно- ве про- гноза ПЭ)	-	-	CA	P	KC		
Multi-level Data Mining	Kordik, Cerny [5. С. 179–224]	A	A	-	-	CA RMSE	-	-	ВЫБ Лучший по ПЭ	A	ОПТ (GA)	CA RMSE	P	KC / PE		
	Jankowski, Grabczewski [5. С. 1–76]	A	A	КОЛ	-	CA	-	-	ВЫБ Лучший по ПЭ	A	ПЕР ОПТ	CA	P	KC		

Цвет	Соответствие сформулированным требованиям
X	Соответствует. Доработка не требуется
X	Не соответствует. Доработка возможна и не приведет к изменению структуры системы
X	Не соответствует. Доработка либо чрезвычайно трудоемкая либо приведет к существенному изменению структуры системы

MAE – Mean Absolute Error, RMSE – Root Mean Square Error, NMSE – Normalized Mean Square Error, CA – Classification Accuracy (процент верных классификаций), CE – Classification Error (процент ошибок классификации), WCE – Weighted Classification Error (взвешенная ошибка классификации), FA – Forecasting Accuracy (точность прогноза), SRC – Spearman Rank Correlation (мера совпадения с идеальным вариантом ранжирования), SSR – Selection Success Rate (процент случаев верного выбора лучшего алгоритма), Algo – алгоритм, Params – параметры функционирования алгоритма, АК – алгоритм классификации, GS – Grid Search (перебор всех значений в заданных пределах с заданным шагом), MLP – Multi-layer perceptron, kNN – k-nearest-neighbor algorithm (метод k ближайших соседей), SOM – Self-Organizing Map (самоорганизующаяся карта), SVM – Support Vector Machine (метод опорных векторов), CBR – Case-Based Reasoning (основанный на правилах выбор наиболее похожего случая), GA – Genetic Algorithm (генетический алгоритм), DT – Decision Tree (дерево решений).

$$M_{g,j}^{left} = [M_i^F | M_{g,i}^Q] = \begin{bmatrix} f_i^1 & \dots & f_i^m & q_1^1 & \dots & q_1^{s_g} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ f_i^1 & \dots & f_i^m & q_{j_i^g}^1 & \dots & q_{j_i^g}^{s_g} \end{bmatrix}$$

$$M_{g,j}^{right} = M_{g,j}^P = \begin{bmatrix} P_1^1 & \dots & P_1^n \\ \dots & \dots & \dots \\ P_{j_i^g}^1 & \dots & P_{j_i^g}^n \end{bmatrix}$$

Объединение данных матриц для всех  $i \in I_g$  для каждого алгоритма  $a_g$  позволит получить матрицы мета-информации  $M_{g,i}^{left}$  и  $M_{g,i}^{right}$ , составляющие выборку данных для обучения мета-моделей вида  $p_g = \mu^g(f, q)$ , способных прогнозировать значение ПЭ для алгоритма  $a_g$  при заданных значениях мета-характеристик  $f$  и параметрах алгоритма  $q$ . Мета-модели  $\mu^g(f, q)$  являются решением задачи идентификации вида

$$\mu^g = \arg \min_{\mu \in \Omega} Z(M_{g,i}^P, \mu(M^F, M_{g,i}^Q)) =$$

$$= \arg \min_{\mu \in \Omega} Z(M_{g,i}^{right}, \mu(M_{g,i}^{left})),$$

где  $\Omega$  – множество всех мета-моделей;  $Z(M^{true}, M^\mu)$  – критерий оценки качества мета-моделей.

Таким образом, в общем виде функция  $L(f)$  может быть записана следующим образом:

$$\{a, q\} = L(f) =$$

$$= \left\{ \begin{array}{l} a^{g_{opt}}, q^{g_{opt}} \mid g_{opt} = \arg \max_{g=1..G} \mu^g(f, q_g^{opt}), \\ q_g^{opt} = \arg \max_{q_g \in R^{s_g}} \mu^g(f, q_g) \end{array} \right\}.$$

В работах [18] и [22] было указано, что большой размер пространства поиска  $R^{s_g}$  параметров  $q_g^{opt}$  при наличии обычно небольшого объема доступной мета-информации ( $J_i^g$  записей для каждой пары алгоритма  $a_g$  и выборки данных  $x_i$ ) приводит к ухудшению качества прогнозирующих мета-моделей. Одним из возможных решений данной проблемы является обучение дополнительных мета-моделей вида  $q_g^{opt} = \eta^g(f)$ , производящих расчет оптимальных параметров для каждого алгоритма  $a_g$  на основании значений мета-характеристик данных  $f$ :

$$\eta^g = \arg \min_{\eta \in \Theta} Y(M_{g,i}^{Q_{opt}}, \eta(M^F)),$$

где  $Y(M^{true}, M^\eta)$  – критерий оценки качества мета-модели, причем

$$M_{g,i}^{Q_{opt}} = \left\{ M_{g,i,j_{opt}}^{Q_{opt}} \mid j_{opt} = \arg \max_{j=1..J_i^g} P_{i,j}^g \right\}.$$

Таким образом, функция примет вид:

$$\{a, q\} = L^n(f) =$$

$$= \left\{ a^{g_{opt}}, q^{g_{opt}} \mid g_{opt} = \arg \min_{g=1..G} \mu^g(f, q_g^{opt}), q_g^{opt} = \eta^g(f) \right\}.$$

Схема, соответствующая данному модифицированному способу (в сравнении с предложенным на рис. 2), представлена на рис. 3.

### Программная платформа для реализации системы мета-обучения

Архитектура программной платформы для реализации систем мета-обучения, полученная в результате объектно-ориентированного анализа и проектирования, представлена на рис. 4.

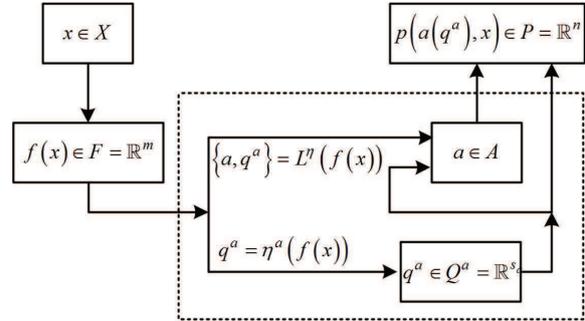


Рис. 3. Выбор алгоритма и параметров его функционирования на основании мета-характеристик данных на основе двух видов мета-моделей

Алгоритм (Algorithm) – объекты данного класса осуществляют некоторую последовательность действий по решению поставленной задачи (метод solve ()) с целью получения результата, выраженного абстракцией Result. Исходная информация к решению задачи передается в виде выборки данных Data, а семантика данных определяется условием конкретной решаемой задачи. Работа алгоритма регулируется набором параметров функционирования params ( $q^a$ ), специфичных для каждого алгоритма ( $a$ ).

Ответственностью класса Criterion является оценка качества решения задачи на основании имеющихся результатов и исходных данных. В соответствии с введенными ранее обозначениями, оценка представляет собой вектор числовых значений  $p$ . MetaFeature – интерфейс расчета мета-характеристик ( $f^m$ ) для задаваемой выборки данных ( $x$ ).

Модель (Model) – это сущность, отражающая некоторую взаимосвязь или закономерность в данных (Data). Структура и параметры модели определяют ее уникальность среди других моделей. Модель может являться результатом работы некоторого алгоритма (ModelAlgorithm на рис. 4) по решению им поставленной задачи (например, задачи структурной и параметрической идентификации системы, классификации, прогнозирования и пр.).

Ответственностью класса MetaLearner является организация процесса мета-обучения. Объекты класса MetaLearner имеют доступ к некоторому задаваемому набору алгоритмов (algorithms), которые могут быть использованы для решения (solve()) поставленной задачи, а также к набору критериев оценки качества решения задачи (criteria). В процессе работы производится накопление мета-данных (metadata), которые используются для построения мета-моделей (MetaModel, наследованный от Model) с использованием задаваемой реализации интерфейса алгоритмов построения моделей (ModelAlgorithm). Полученные мета-мо-



существующими аналогами: на сегодняшний день в базе данных ИФ (International Institute of Forecasters) содержатся прогнозы, полученные с использованием 24 современных методов прогнозирования, начиная от «наивного прогнозировщика» (Naive) и заканчивая искусственными нейронными сетями (ANN) [24].

Значения каждой временной последовательности  $x_t$  следуют с фиксированным шагом по времени  $\Delta t$ , равным одному году (645 последовательностей), кварталу (756 последовательностей) или месяцу (1428 последовательностей) [25]. Для каждой последовательности задается число контрольных точек в конце, которые **не могут участвовать в обучении** прогнозирующих моделей, но **должны быть использованы для оценки** точности прогнозирования.

Для экспериментов были выбраны месячные данные: 1428 последовательностей. Для удобства сравнительного анализа результатов прогнозирования все используемые последовательности были нормализованы к интервалу  $[-1; 1]$ .

Были определены 8 мета-характеристик временных последовательностей, подробная информация по которым приведена в табл. 3. Для каждой из мета-характеристик приведен вид согласно классификации, введенной ранее в аналитическом обзоре систем мета-обучения, а также способ их вычисления. На рис. 5 представлены гистограммы

распределений значений мета-характеристик, рассчитанные для всех 1428 выборок данных ( $N$  – число выборок, значение мета-характеристики которых попадает в соответствующий интервал гистограммы), а на рис. 6 визуализированы значения линейного коэффициента их взаимной корреляции (более темный цвет означает более коррелированные значения, знаком «+» и «-» отмечены соответственно, положительные и отрицательные значения корреляции). Гистограммы распределений для каждой мета-характеристики имеют различный вид, а абсолютное значение корреляции в среднем составляет 0,29, что свидетельствует о том, что мета-характеристики отражают независимые свойства временных последовательностей.

В качестве алгоритма прогнозирования применялся один из «прогнозирующих модулей», описанных автором в работе [21], соответствующий типу с названием «Sliding Window» и характеризующийся параметром  $T$  – размером скользящего окна, определяющим количество точек предыстории, используемых прогнозирующей моделью для построения прогноза на следующий шаг. Данный параметр был выбран в качестве единственного параметра функционирования алгоритма  $q^a$ . В экспериментах значение  $T$  полагалось равным 1, 2, 3, 4, 5, 6 и 7. Построение прогнозирующих моделей производилось с использованием следующих алгоритмов МГУА:

Таблица 3. Описание применяемых мета-характеристик временных последовательностей

Метка	Вид	Описание и способ вычисления	
length	КОЛ	Длина временной последовательности (число точек)	
noise	СТАТ	Средняя величина абсолютного изменения за шаг: $f^{noise} = \frac{1}{T} \sum_{t=1}^{T-1}  x_{t+1} - x_t $	
trend	ПОВ	Коэффициент наклона прямой линейного тренда $f^{trend} = a$ . Расчет коэффициента $a$ – из уравнения $x_t = at + b$ по методу наименьших квадратов	
period	ПОВ	Период гармонического сигнала, рассчитываемый согласно критерию баланса ординат из предположения наличия единственной гармоники во входной временной последовательности [26. С. 76–78]	
land1in	ОРИ	Значение CR на обучающей выборке данных: $f^{land1in} = CR_{0, \lfloor T/2 \rfloor}^{combi-1}$	Осуществлялось построение прогнозирующей модели с использованием комбинаторного алгоритма метода МГУА [27. С. 32–38] с ограничением на 1-ю (land1, линейная модель) и 7-ю (land2, нелинейная модель) степень полинома. Производился поиск модели вида $x_t = \phi(t)$ . Обучающая выборка составляла первую половину точек временной последовательности, проверочная – вторую. В качестве внешнего критерия использовался критерий регулярности [27. С. 84]: $CR_{T1, T2} = \frac{1}{T2 - T1} \sum_{t=T1}^{T2-1} (x_t - \hat{x}_t)^2,$ где $\hat{x}_t$ – выход модели, $T1$ и $T2$ – начальный и конечный индексы части временной последовательности
land1out	ОРИ	Значение CR на проверочной выборке данных $f^{land1out} = CR_{\lfloor T/2 \rfloor + 1, T}^{combi-1}$	
land2in	ОРИ	Значение CR на обучающей выборке данных: $f^{land2in} = CR_{0, \lfloor T/2 \rfloor}^{combi-7}$	
land2out	ОРИ	Значение CR на проверочной выборке данных $f^{land2out} = CR_{\lfloor T/2 \rfloor + 1, T}^{combi-7}$	

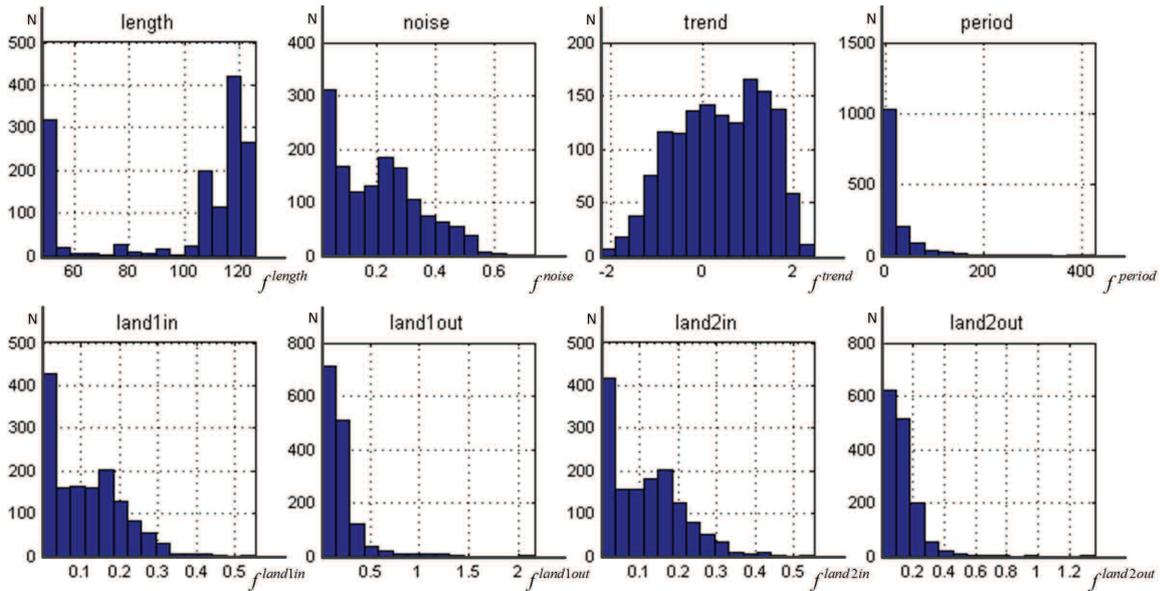


Рис. 5. Гистограммы распределений значений мета-характеристик для месячных временных последовательностей «МЗ Competition»

- Комбинаторный алгоритм; ограничение на степень полинома 1; обучающая выборка – первые 2/3 точек временной последовательности, проверочная – последующие 1/3 точек; внешний критерий селекции моделей – CR. Далее используется метка «COMBI-row1».
- Алгоритм дважды многорядной полиномиальной сети; ограничение на размер сети 2×2, 3×3, 4×4, 5×5 (число слоев×число нейронов в слое); двухвходовые нейроны; комбинаторный алгоритм обучения нейронов, максимальная степень полиномиального частного описания нейрона 2; минимальное изменение внешнего критерия от слоя к слою 10<sup>-5</sup>; обучающая выборка – первые 2/3 точек временной последовательности, проверочная – последующие 1/3 точек; внешний критерий селекции моделей – CR. Далее используются метки «PNN-2×2», «PNN-3×3» и т. п.

Таким образом, для каждой временной последовательности осуществлялся запуск 7×(1+4)=35 прогнозирующих модулей для получения прогноза на контрольной выборке данных. В качестве ПЭ алгоритмов  $(p(a(q^n), x))$  использовался критерий CR, и среднее его значение на контрольных выборках составило 3,52.

Согласно предложенному ранее модифицированному способу (рис. 3), в системе мета-обучения производилось построение двух видов мета-моделей.

- $q_g^{opt} = \eta^g(f)$ , где размер вектора мета-характеристик  $f$  равнялся 8,  $q_g^{opt} - 1$ , а число строк матрицы  $M_g^{q, opt}$  составляло 1428.
- $\mu^g(f, q_g^{opt})$ , где размеры векторов  $f$  и  $q_g^{opt}$  составляли 8 и 1 соответственно, а число строк матрицы  $M^F - 1428$ .

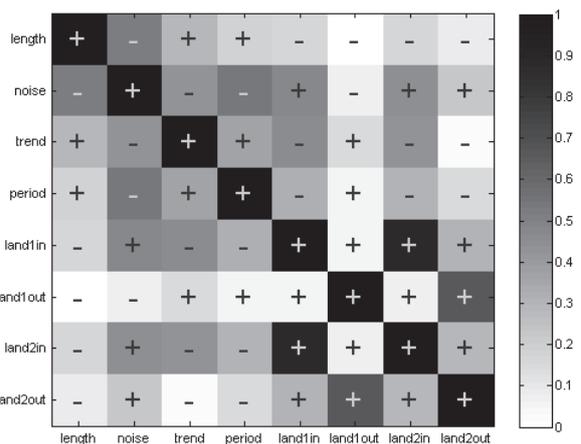


Рис. 6. Величины линейного значения взаимной корреляции значений мета-характеристик для месячных временных последовательностей «МЗ Competition»

Построение мета-моделей осуществлялось с помощью алгоритма дважды многорядной полиномиальной сети МГУА с ограничением на размер 10×10 и внешним критерием CR (разделение выборки на обучающую и проверочную – пополам; двухвходовые нейроны, комбинаторный алгоритм обучения нейронов, максимальная степень полиномиального частного описания 2).

Результатом работы модуля самодиагностики (реализация требования D), применявшегося в том числе для контроля качества мета-знаний, явилось исключение из обучающих выборок 0,1...2,7 % информации по причине несоответствия ее минимальным ожидаемым значениям ПЭ для поставленной задачи прогнозирования (CR>10<sup>3</sup>).

Также стоит отметить, что в процессе построения моделей вида  $q_g^{opt} = \eta^g(f)$  алгоритмом МГУА бы-

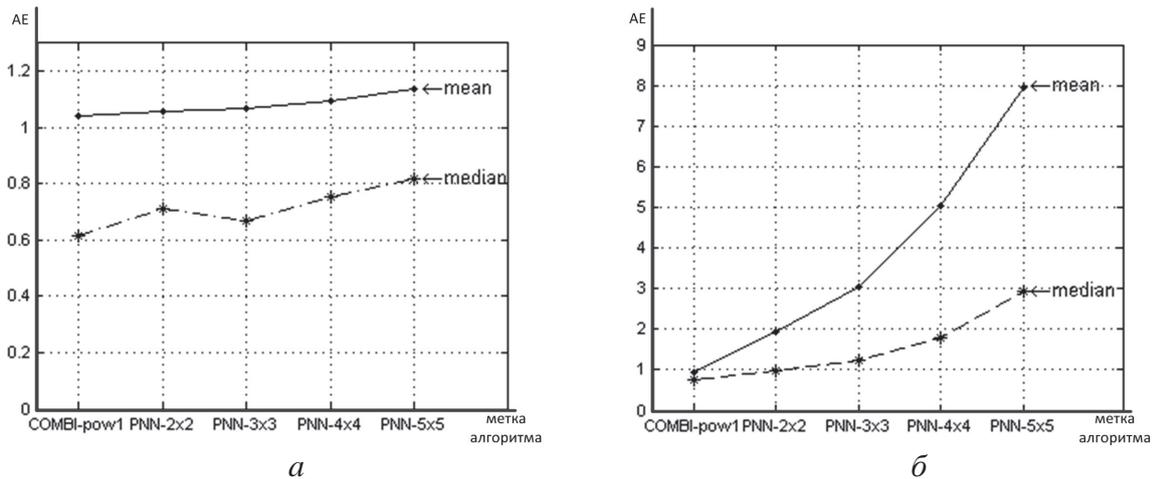


Рис. 7. Средние и медианные значения абсолютных ошибок мета-моделей вида: а)  $q_g^{opt} = \eta^g(f)$ ; б)  $\mu^g(f, q_g^{opt})$

ли автоматически исключены мета-характеристики «length» (для 3 мета-моделей из 5), «period» (3 из 5), «land1out» (5 из 5) и «land2out» (4 из 5) как нерепрезентативные. Для моделей вида  $\mu^g(f, q_g^{opt})$  – только «land2in» (3 из 5).

На рис. 7, а приведены среднее (mean) и медианное (median) значения абсолютной погрешности (АЕ) мета-моделей вида  $q_g^{opt} = \eta^g(f)$  для доступных системе алгоритмов прогнозирования. Можно видеть, что медианное значение лежит в пределах [0,6; 0,8], что является удовлетворительным показателем для модели, вычисляющей прогноз оптимального значения параметра, лежащего в диапазоне [1; 7]. На рис. 7, б показаны средние (mean) и медианные (median) значения абсолютной погрешности мета-моделей вида  $\mu^g(f, q_g^{opt})$ . На графике наблюдается возрастание ошибки мета-модели с увеличением сложности прогнозирующих моделей, генерируемых соответствующим алгоритмом. Тем не менее, погрешность находится на приемлемом уровне [1; 8] для показателя, принимающего значение [0; 110].

Проверка эффективности системы мета-обучения проводилась путем передачи очередной временной последовательности системе с целью решения задачи построения прогноза на заданное число шагов вперед. Система производила выбор оптимального алгоритма путем применения обученных мета-моделей обоих видов согласно модифицированному способу, описанному ранее. Среднее значение CR на контрольной выборке для системы составило 1,95, что на 44,7 % меньше, чем среднее значение CR отдельных алгоритмов. Стоит отметить, что среднее значение CR для системы мета-обучения, использующей идеально точные мета-

модели (мета-модели со средним значением абсолютного отклонения, равным нулю, т. е. имеющие нулевую ошибку прогнозирования), составило 1,55, что лишь на 20,4 % лучше полученного результата.

#### Заключение

Известные на сегодняшний день аналитические обзоры систем мета-обучения выделяют лишь некоторые внешние критерии оценки систем либо рассматривают состав внутренних модулей без выделения общности в организации систем. В данной работе сформулированы классификационные признаки, учитывающие внутреннюю организацию систем, и проведена систематизация существующих систем мета-обучения. Разработана объектно-ориентированная архитектура программной платформы, позволяющая реализовать любую из систем мета-обучения, представленных в систематизации. Предложен способ построения системы мета-обучения, реализованный в рамках программной платформы и удовлетворяющий всем сформулированным требованиям. Проведенные эксперименты по решению задачи прогнозирования с использованием набора алгоритмов МГУА показали эффективность разработанной системы мета-обучения по выбору оптимального алгоритма прогнозирования и расчету оптимальных параметров его функционирования (улучшение значения критерия качества на 44,7 %). Направлением дальнейших работ является анализ поведения алгоритмов выбора оптимального алгоритма решения задачи с целью дальнейшего повышения эффективности системы.

## СПИСОК ЛИТЕРАТУРЫ

1. U. Fayyad, Piatetsky-Shapiro G., Smyth P. From Data Mining to Knowledge Discovery in Databases // *AI Magazine: International Journal*. – 1996. – V. 17. – № 3. – P. 37–54.
2. Wolpert D.H. The Lack of a Priori Distinctions between Learning Algorithms // *Neural Computation: International Journal*. – 1996. – V. 8. – Iss. 7. – P. 1341–1390.
3. Rice J. The Algorithm Selection Problem // *Advances in Computers*. – 1976. – V. 15. – P. 65–118.
4. Vilalta R., Drissi Y. A Perspective View and Survey of Meta-Learning // *Artificial Intelligence Review: International Journal*. – 2002. – V. 18. – Iss. 2. – P. 77–95.
5. Meta-Learning in Computational Intelligence. *Studies in Computational Intelligence Series / Ed. N. Jankowski, W. Duch, K. Grabczewski*. – Berlin: Springer, 2011. – V. 358. – 359 p.
6. Prudencio R.B.C., Ludermir T.B. Selecting and ranking time series models using the NOEMON approach // *Proceedings of the Joint International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP 2003)*. – Istanbul, Turkey, 2003. – P. 654–661.
7. Prudencio R.B.C., Ludermir T.B. Active Meta-Learning with Uncertainty Sampling and Outlier Detection // *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008)*. – Hong Kong, 2008. – P. 346–351.
8. Romero C., Olmo J.L., Ventura S. A meta-learning approach for recommending a subset of white-box classification algorithms for Moodle datasets // *Proceeding of the VI International Conference on Educational Data Mining (EDM 2013)*. – Memphis, Tennessee, USA, 2013. – P. 268–271.
9. Matijas M., Suykens J.A.K., Krajcar S. Load forecasting using a multivariate meta-learning system // *Expert Systems with Applications: International Journal*. – 2013. – V. 40. – Iss. 11. – P. 4427–4437.
10. Căcovăanu S., Vidrighin C., Potolea R. Evolutional meta-learning framework for automatic classifier selection // *Proceedings of the V International Conference on Intelligent Computer Communication and Processing (IEEE ICCP 2009)*. – Cluj-Napoca, Romania, 2009. – P. 27–30.
11. Algorithm Selection: From Meta-Learning to Hyper-Heuristics / L. Cruz-Reyes, C. Gomez-Santillan, J. Perez-Ortega, V. Landero, M. Quiroz, A. Ochoa // *Intelligent Systems*. – InTech, 2012. – Ch. 6. – 27 p. URL: <http://www.intechopen.com/books/intelligent-systems/algorithm-selection-from-meta-learning-to-hyper-heuristics> (дата обращения: 07.02.2014).
12. Wang X., Smith-Miles K., Hyndman R. Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series // *Neurocomputing: International Journal*. – 2009. – V. 72. – Iss. 10–12. – P. 2581–2594.
13. Landmarking for Meta-Learning using RapidMiner / S.D. Abdelmessih, F. Shafait, M. Reif, M. Goldstein // *Proceedings of the RapidMiner Community Meeting and Conference*. – Dortmund, Germany, 2010. – 6 p. URL: [http://www.dfki.de/web/forschung/iwi/publikationen/renameFileForDownload?filename=Sarah-Landmarking-RCOMM10.pdf&file\\_id=uploads\\_781](http://www.dfki.de/web/forschung/iwi/publikationen/renameFileForDownload?filename=Sarah-Landmarking-RCOMM10.pdf&file_id=uploads_781) (дата обращения: 13.12.2013).
14. Meta-Learning Approach for Automatic Parameter Tuning: A Case Study with Educational Datasets / M.M. Molina, J.M. Luna, C. Romero, S. Ventura // *Proceeding of the V International Conference on Educational Data Mining (EDM 2012)*. – Chania, Greece, 2012. – P. 180–183.
15. Musliu N., Schwengerer M. Algorithm Selection for the Graph Coloring Problem // *Proceedings of the VII International Conference on Learning and Intelligent Optimization (LION 7)*. – Catania, Italy, 2013. – P. 389–403.
16. Tuned data mining: A benchmark study on different tuners / W. Konen, T. Bartz-beielstein, P. Koch, M. Friese, O. Flasch, B. Naujoks // *Proceedings of the XIII Annual Conference on Genetic and Evolutionary Computation (GECCO'11)*. – ACM New York, USA, 2011. – P. 1995–2002.
17. Soares C., Brazdil P.B. Selecting parameters of SVM using meta-learning and kernel matrix-based meta-features // *Proceedings of the ACM symposium on Applied computing (SAC'06)*. – New York, USA, 2006. – P. 564–568.
18. Reif M., Shafait F., Dengel A. Meta-learning for evolutionary parameter optimization of classifiers // *Machine Learning: International Journal*. – 2012. – V. 87. – Iss. 3. – P. 357–380.
19. Consultant-2: pre- and post-processing of machine learning applications / D. Sleeman, M. Rissakis, S. Craw, N. Graner, S. Sharma // *International Journal of Human-Computer Studies*. – 1995. – V. 43. – Iss. 1. – P. 43–63.
20. Kalousis A., Hilario M. Representational issues in meta-learning // *Proceedings of the XX International Conference on Machine Learning (ICML-03)*. – Washington, DC, USA, 2003. – P. 313–320.
21. Orlov A.A. Non-Stationary Time Series Forecasting on Basis of Analysis and Prediction of Forecasting Models Efficiency // *Proceedings of the IV International Conference on Inductive Modeling (ICIM-2013)*. – Kyiv, 2013. – P. 192–199.
22. Gomes T.A.F., Prudencio R.B.C., Soares C. Combining meta-learning and search techniques to select parameters for support vector machines // *Neurocomputing: International Journal*. – 2012. – V. 75. – Iss. 1. – P. 3–13.
23. Орлов А.А. Принципы построения архитектуры программной платформы для реализации алгоритмов метода группового учета аргументов (МГУА) // *Управляющие системы и машины: Международный журнал*. – 2013. – № 2. – С. 65–71.
24. Makridakis S., Hibon M. M3-Competition // *INSEAD*. – Fontainebleau, France, 1999. – 41 p. URL: <http://www.insead.edu/facultyresearch/research/doc.cfm?did=1094> (дата обращения: 23.03.2014).
25. Makridakis S., Hibon M. The M3-Competition: results, conclusions and implications // *International Journal of Forecasting*. – 2000. – V. 16. – Iss. 4. – P. 451–476.
26. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем. – Киев: Наукова Думка, 1982. – 296 с.
27. Madala H.R., Ivakhnenko A.G. *Inductive Learning Algorithms for Complex System Modeling*. – Boca Raton; Ann Arbor; London; Tokyo: CRC Press, 1994. – 368 p.

Поступила 31.03.2014 г.

UDC 004.4::004.85[3+5]

## AUTOMATIC META-LEARNING SYSTEM SUPPORTING SELECTION OF OPTIMAL ALGORITHM FOR PROBLEM SOLVING AND CALCULATION OF OPTIMAL PARAMETERS OF ITS FUNCTIONING

**Andrey A. Orlov,**

 Tomsk State University of Control Systems and Radioelectronics,  
 40, Lenin avenue, Tomsk, 634050, Russia. E-mail: d1scnc@gmail.com

**The relevance of the work** is caused by necessity of increasing efficiency of automatic data mining systems based on meta-learning. **The main aim of the study** is to design an automatic meta-learning system supporting selection of optimal algorithm for problem solving and calculation of optimal parameters of its functioning.

**The methods used in the study:** inductive modeling, methods of statistical analysis of results.

**Results:** The known meta-learning systems were integrated based on produced classification features taking into account internal structure of systems. The author has stated the requirements for implementation of the automatic meta-learning system and has offered the way to build a meta-learning system satisfying all stated requirements and accumulating meta-knowledge, building meta-models on its basis, selecting optimal algorithm from a set of available ones and calculating optimal parameters of its functioning. The object-oriented architecture of a software framework for implementation of any meta-learning system presented in the systematization was developed. The efficiency of the implemented automatic meta-learning system using algorithms of group method of data handling was experimentally examined being applied to solution of problems related to the short-term time series forecasting (1428 time series from the testing set known as «M3 Competition»).

### Key words:

Meta-learning, meta-features of data, meta-model, software framework, object-oriented analysis and design, forecasting of time series, group method of data handling.

### REFERENCES

1. Fayyad U., Piatetsky-Shapiro G., Smyth P. From Data Mining to Knowledge Discovery in Databases. *AI Magazine: International Journal*, 1996, vol. 17, no. 3, pp. 37–54.
2. Wolpert D.H. The Lack of A Priori Distinctions Between Learning Algorithms. *Neural Computation: International Journal*, 1996, vol. 8, Iss. 7, pp. 1341–1390.
3. Rice J. The Algorithm Selection Problem. *Advances in Computers*, 1976, vol. 15, pp. 65–118.
4. Vilalta R., Drissi Y. A Perspective View and Survey of Meta-Learning. *Artificial Intelligence Review: International Journal*, 2002, vol. 18, Iss. 2, pp. 77–95.
5. *Meta-Learning in Computational Intelligence. Studies in Computational Intelligence Series*. Ed. N. Jankowski, W. Duch, K. Grabczewski. Berlin, Springer, 2011. Vol. 358. 359 p.
6. Prudencio R.B.C., Ludermir T.B. Selecting and ranking time series models using the NOEMON approach. *Proceedings of the Joint International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP 2003)*. Istanbul, Turkey, 2003. pp. 654–661.
7. Prudencio R.B.C., Ludermir T.B. Active Meta-Learning with Uncertainty Sampling and Outlier Detection. *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2008)*. Hong Kong, 2008. pp. 346–351.
8. Romero C., Olmo J.L., Ventura S. A meta-learning approach for recommending a subset of white-box classification algorithms for Moodle datasets. *Proceeding of the VI International Conference on Educational Data Mining (EDM 2013)*. Memphis, Tennessee, USA, 2013. pp. 268–271.
9. Matijas M., Suykens J.A.K., Krajcar S. Load forecasting using a multivariate meta-learning system. *Expert Systems with Applications: International Journal*, 2013. vol. 40, Iss. 11, pp. 4427–4437.
10. Cacoveanu S., Vidrighin C., Potolea R. Evolutional meta-learning framework for automatic classifier selection. *Proceedings of the V International Conference on Intelligent Computer Communication and Processing (IEEE ICCP 2009)*, 2009. pp. 27–30.
11. Cruz-Reyes L., Gomez-Santillan C., Perez-Ortega J., Landero V., Quiroz M., Ochoa A. Algorithm Selection: From Meta-Learning to Hyper-Heuristics. *Intelligent Systems. InTech*, 2012. Ch. 6, 27 p. Available at: <http://www.intechopen.com/books/intelligent-systems/algorithm-selection-from-meta-learning-to-hyper-heuristics> (accessed 07 February 2014).
12. Wang X., Smith-Miles K., Hyndman R. Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series. *Neurocomputing: International Journal*, 2009, vol. 72, Iss. 10–12, pp. 2581–2594.
13. Abdelmessih S.D., Shafait F., Reif M., Goldstein M. Landmarking for Meta-Learning using RapidMiner. *Proceedings of the RapidMiner Community Meeting and Conference*. Dortmund, Germany, 2010. 6 p. Available at: [http://www.dfki.de/web/forschung/iwi/publikationen/renamefilefordownload?filename=Sarah-Landmarking-RCOMM10.pdf&file\\_id=uploads\\_781](http://www.dfki.de/web/forschung/iwi/publikationen/renamefilefordownload?filename=Sarah-Landmarking-RCOMM10.pdf&file_id=uploads_781) (accessed 13 December 2013).
14. Molina M.M., Luna J.M., Romero C., Ventura S. Meta-Learning Approach for Automatic Parameter Tuning: A Case Study with Educational Datasets. *Proceeding of the V International Conference on Educational Data Mining (EDM 2012)*. Chania, Greece, 2012. pp. 180–183.
15. Musliu N., Schwengerer M. Algorithm Selection for the Graph Coloring Problem. *Proceedings of the VII International Conference on Learning and Intelligent Optimization (LION 7)*. Catania, Italy, 2013. pp. 389–403.
16. Konen W., Bartz-beielstein T., Koch P., Friese M., Flasch O., Naujoks B. Tuned data mining: A benchmark study on different tuners. *Proceedings of the XIII Annual Conference on Genetic and Evolutionary Computation (GECCO '11)*. ACM New York, USA, 2011. pp. 1995–2002.
17. Soares C., Brazdil P.B. Selecting parameters of SVM using meta-learning and kernel matrix-based meta-features. *Proceedings of the ACM symposium on Applied computing (SAC '06)*. New York, USA, 2006. pp. 564–568.
18. Reif M., Shafait F., Dengel A. Meta-learning for evolutionary parameter optimization of classifiers. *Machine Learning: International Journal*, 2012, vol. 87, Iss. 3, pp. 357–380.
19. Sleeman D., Rissakis M., Craw S., Graner N., Sharma S. Consultant-2: pre- and post-processing of machine learning applications. *International Journal of Human-Computer Studies*, 1995, vol. 43, Iss. 1, pp. 43–63.

20. Kalousis A., Hilario M. Representational issues in meta-learning. *Proceedings of the XX International Conference on Machine Learning (ICML-03)*. Washington, DC, USA, 2003. pp. 313–320.
21. Orlov A.A. Non-Stationary Time Series Forecasting on Basis of Analysis and Prediction of Forecasting Models Efficiency. *Proc. the IV International Conference on Inductive Modelling (ICIM-2013)*. Kyiv, 2013. pp. 192–199.
22. Gomes T.A.F., Prudencio R.B.C., Soares C. Combining meta-learning and search techniques to select parameters for support vector machines. *Neurocomputing: International Journal*. – Elsevier Science Publishers, Amsterdam, 2012, vol. 75. Iss. 1. pp. 3–13.
23. Orlov A.A. Printsipy postroeniya arkhitektury programnoy platformy dlya realizatsii algoritmov metoda gruppovogo ucheta argumentov (MGUA) [Principles of constructing software framework architecture to implement the algorithms of group method of data handling (GMDH)]. *Upravlyayushchie Sistemy i Mashiny: Mezhdunarodnyi Zhurnal – Control Systems and Computers: International Journal*, 2013, no. 2. pp. 65–71.
24. Makridakis S., Hibon M. M3-Competition. *INSEAD*. Fontainebleau, France, 1999. 41 p. Available at: <http://www.insead.edu/facultyresearch/research/doc.cfm?did=1094> (accessed 23 March 2014).
25. Makridakis S., Hibon M. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 2000, vol. 16, Iss. 4, pp. 451–476.
26. Ivakhnenko A.G. *Induktivnyy metod samoorganizatsii modeley slozhnykh system* [Inductive method of self-organization of models of complex systems]. Kiev, Naukova Dumka, 1982. 296 p.
27. Madala H.R., Ivakhnenko A.G. *Inductive Learning Algorithms for Complex System Modeling*. Boca Raton; Ann Arbor; London; Tokyo, CRC Press, 1994. 368 p.

УДК 004.822, 004.942; 620.9:001.891.57

## ИНТЕГРАЦИЯ СЕМАНТИЧЕСКИХ МОДЕЛЕЙ В ИССЛЕДОВАНИЯХ ПРОБЛЕМ ЭНЕРГЕТИЧЕСКОЙ БЕЗОПАСНОСТИ

**Массель Алексей Геннадьевич,**

канд. техн. наук, ст. науч. сотр. Института систем энергетики  
им. Л.А. Мелентьева (ИСЭМ) СО РАН,  
Россия, 664033, г. Иркутск, ул. Лермонтова, 130. E-mail: amassel@gmail.com

**Тюрюмин Вадим Олегович,**

аспирант Института систем энергетики им. Л.А. Мелентьева (ИСЭМ)  
СО РАН Россия, 664033, г. Иркутск, ул. Лермонтова, 130.  
E-mail: vadim.tyuryumin@gmail.com

Актуальность работы определяется, с одной стороны, важностью проблем энергетической безопасности и их исследований, с другой – необходимостью развития интеллектуальных инструментальных средств исследований и их интеграции с традиционными математическими моделями.

**Цель работы:** усовершенствование двухуровневой технологии исследований проблем энергетической безопасности за счет интеграции семантических моделей и расширения онтологического пространства знаний о предметной области онтологиями событий.

**Методы исследования:** В предложенной ранее двухуровневой технологии исследований проблем энергетической безопасности на верхнем (первом) уровне выполняется качественный анализ (экспресс-анализ) с применением методов и средств семантического (онтологического, когнитивного и событийного) моделирования, на нижнем (втором) – количественный анализ на основе численных расчетов с использованием традиционных программных комплексов. Когнитивные модели используются для моделирования угроз энергетической безопасности, событийные модели – для моделирования вариантов развития чрезвычайных ситуаций в энергетике. Совместное использование когнитивного и событийного моделирования позволяет получить более объективную оценку ситуации. Вводятся онтологии событий, используемые для перехода от когнитивных к событийным моделям. Рассматривается использование Joiner-сетей для последующего анализа и обработки событийных моделей.

**Результаты:** Предложены правила генерирования онтологий событий. Описан алгоритм автоматизированного перехода от когнитивных карт к событийным моделям на основе онтологий (с использованием онтологий событий) как один из этапов усовершенствования двухуровневой технологии.

Приведены разработанные онтологии, когнитивные и событийные карты и построенная на их основе событийная модель с использованием Joiner-сетей.

### **Ключевые слова:**

Энергетическая безопасность, семантическое моделирование, онтологическое, когнитивное и событийное моделирование, онтологии событий, Joiner-сети.

### **Введение**

В Институте систем энергетики им. Л.А. Мелентьева СО РАН проводятся комплексные исследования, важную роль в которых играют исследования проблем энергетической безопасности.

Энергетическая безопасность (ЭБ) рассматривается как часть национальной безопасности, а именно как защищенность граждан, общества, государства, экономики от угроз дефицита в обеспечении их обоснованных потребностей топливно-энерге-