

по технологическому критерию, поэтому ее можно рекомендовать лишь как инструмент управления локальными проектами, сосредоточенными внутри отдельных небольших компаний. Системе Project Kaiser не хватает функциональности для решения многих задач проектного управления, однако она обладает неоспоримым ценовым преимуществом и поэтому может быть рекомендована для компаний, не имеющих достаточных финансовых возможностей для приобретения других систем. Удобным является наличие в системе online-решения, позволяющего легко организовать работу над проектом территориально удаленных пользователей. Система Deltek Open Plan также может использоваться в рассматриваемой сфере для проектов средней сложности по решаемым задачам и требованиям к технологиям доступа к данным.

Литература

1. Управление проектами в России: Информационные системы. URL: <http://www.project-open.ru/category/Информационные-системы/> (дата обращения: 10.06.2013).
2. Продукты MicrosoftProject – программное обеспечение для управления портфелями проектов. Официальный сайт. URL: <http://office.microsoft.com/ru-ru/project/> (дата обращения: 10.06.2013).
3. Primavera Enterprise Project Portfolio Management. Официальный сайт. URL: <http://www.oracle.com/ru/products/applications/primavera/> (дата обращения: 10.06.2013).
4. Enterprise Program Management Software. Deltek Open

Plan. Официальный сайт. URL: <http://www.deltek.com/products/ipm/> (дата обращения: 10.06.2013).

5. Спайдер Проджект: Управление проектами. Официальный сайт. URL: <http://spiderproject.ru/> (дата обращения: 10.06.2013).

6. Project Kaiser: Система управления проектами. Официальный сайт. URL: <http://www.projectkaiser.com/ru/> (дата обращения: 10.06.2013).

7. Сводный перечень российских технологических платформ. Официальный сайт НИУ ВШЭ. URL: <http://www.hse.ru/org/hse/tp/catalogue> (дата обращения: 10.06.2013).

References

1. *Upravlenie proektami v Rossii: Informatsionnye sistemy* [Project management in Russia: information systems], available at: <http://www.project-open.ru/category/Информационные-системы/> (accessed 10 June 2013).

2. *Produkty MicrosoftProject – programmnoe obespechenie dlya upravleniya portfelyami proektov* [MicrosoftProject products – software for project portfolio management], available at: <http://office.microsoft.com/ru-ru/project/> (accessed 10 June 2013).

3. *Primavera Enterprise Project Portfolio Management*, available at: <http://www.oracle.com/ru/products/applications/primavera/> (accessed 10 June 2013).

4. *Enterprise Program Management Software. Deltek Open Plan*, available at: <http://www.deltek.com/products/ipm/> (accessed 10 June 2013).

5. *Spider Project*, available at: <http://spiderproject.ru/> (accessed 10 June 2013).

6. *Project Kaiser*, available at: <http://www.projectkaiser.com/ru/> (accessed 10 June 2013).

7. *Svodnyy perechen rossyskikh tekhnologicheskikh platform* [Consolidated listing of Russian technological platforms], available at: <http://www.hse.ru/org/hse/tp/catalogue> (accessed 10 June 2013).

УДК 004.942

ПРОГРАММНАЯ СИСТЕМА АНАЛИЗА ИНДИКАТОРОВ АКТИВНОСТИ РЕГИОНОВ РОССИИ НА БАЗЕ ОНТОЛОГИЧЕСКИХ МОДЕЛЕЙ И ПАТТЕРНОВ ДАННЫХ

(Работа выполнена при финансовой поддержке Минобрнауки России по ГК от 14.06.2012 г. № 07.514.11.4144 в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 гг.»)

А.Д. Московский, бакалавр; М.А. Ровбо, бакалавр; В.Ф. Хорошевский, д.т.н., профессор
(Национальный исследовательский университет Московского физико-технического университета,
Институтский пер., 9, г. Долгопрудный, 141700, Россия,
strohkopf@fnbic.ru, rovboma@gmail.com, vkhoroshevsky@hse.ru)

Обсуждаются вопросы разработки и реализации автоматизированного рабочего места аналитика, обеспечивающего поддержку процессов интеллектуального анализа больших объемов статистических данных науки, образования и инновационной деятельности. Фиксируются научно-техническая проблема, решение которой реализуется в системе, цели и задачи разработки, основными из которых являются формирование системы индикаторов науки, образования и инновационной деятельности, построение системы агрегатов статистических данных с целью формирования опорных точек для анализа ситуации в регионах России в различных аналитических разрезах, формирование векторов индикаторов для наборов статистических данных и их агрегатов, а также определение семантической близости статистических данных и их агрегатов, анализ поведения индикаторов аналогичных регионов РФ, формирование динамических групп статистических данных и их агрегатов, анализ поведения индикаторов науки, образования и инновационной деятельности в регионах РФ. Дается краткий обзор состояния исследований и разработок в данной области. Математическое обеспечение представленной в работе системы базируется на использовании гибридного подхода, в рамках которого обеспечивается интеграция классических математических методов корреляционного анализа, анализа паттернов данных и временных рядов с методами семантической интерпретации получаемых результатов. При разра-

ботке и реализации программного обеспечения системы особое внимание уделено поддержке бизнес-процессов выявления трендов изменения индикаторов и нетипичной динамики индикаторов, определению характерных векторов направленности индикаторов Best Performance, а также дружественным интерфейсам пользователя.

Ключевые слова: индикаторы науки, образования, инновационной деятельности, паттерны данных, корреляционный анализ, кластеризация и классификация, анализ временных рядов, онтологическое моделирование, система анализа индикаторов активности, бизнес-сценарии анализа статистических данных.

RUSSIAN REGIONS ACTIVITY INDICATORS ANALYSIS SOFTWARE SYSTEM BASED ON ONTOLOGICAL MODELS AND DATA PATTERNS

Moskovsky A.D., bachelor; Rovbo M.A., bachelor; Khoroshevsky V.F., Ph.D., professor

(National Research University Moscow Institute of Physics and Technology,

Institutskiy lane, 9, Dolgoprudny, 141700, Russia,

strohkopf@fnbic.ru, rovboma@gmail.com, vkhoroshevsky@hse.ru)

Abstract. An analyst workbench development and implementation based on intelligent mining of large amounts of statistical data in the domains of science, education and innovation is discussed in the paper. Scientific and technical challenges are pointed out. The R&D objectives are carried out: a specification of an indicators system of science, education and innovation, as well as aggregation of statistical data for providing analysis of scientific, educational and innovation activity of Russian regions in various aspects. Methods to form vectors of specific indicators and aggregates for statistical datasets are developed. Business tasks are solved including identification of semantically similar vectors for building clusters of “similar” regions, and dynamic analysis of statistical data for monitoring performance of Russian regions in science, education and innovation. A brief overview of the art state in domain is provided. A hybrid approach proposed in the paper integrates classical mathematical methods for correlation analysis, pattern and time series analysis with methods of their semantic interpretation.

The developed software provides analytical support for trends identification, atypical dynamics of indicators and “best performance” patterns specification. Implementation of user-friendly interfaces is also discussed.

Keywords: indicators of science, education and innovation; data patterns; correlation analysis; clustering and classification; time-series analysis; ontological modeling; activity indicators analysis system; business scenario of analysis of statistical data.

Инновационное развитие регионов России является одной из проблем, решение которой определяет положение страны в рамках мирового постиндустриального общества. В связи с этим особую важность приобретает мониторинг состояния науки, образования и инноваций. Уже сегодня в этом направлении ведутся активные исследования и разработки, но сложность задачи настолько велика, что требуются новые подходы и методы ее решения. Одним из таких подходов является использование паттернов, выделенных в результате применения классических методов обработки статистических данных, для поиска взаимосвязей исследуемых объектов, их классификации и исследования процессов развития объектов во времени.

В настоящей работе обсуждаются вопросы создания системы *анализа индикаторов данных науки, образования и инновационной деятельности (АИДА)*.

Постановка задачи и состояние исследований и разработок

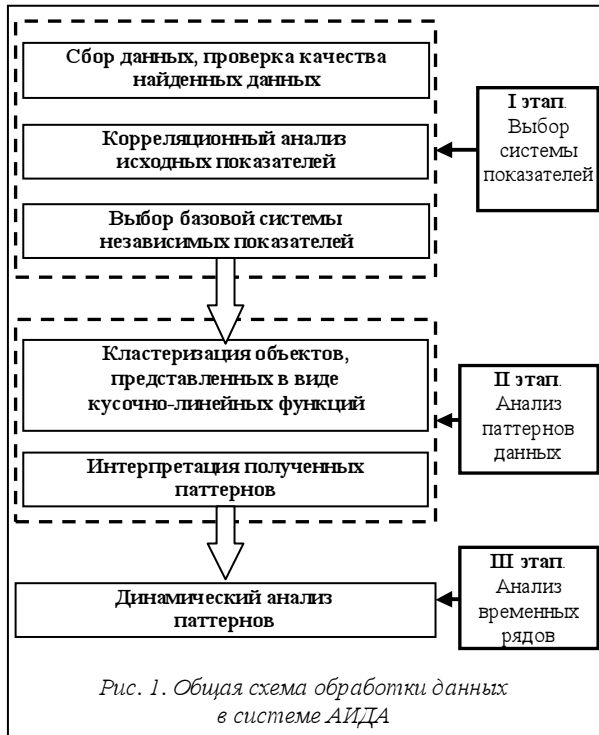
Научно-технической проблемой, решение которой обсуждается в настоящей работе, являются разработка и реализация рабочего места аналитика, поддерживающего процессы интеллектуального анализа больших объемов статистических данных науки, образования и инновационной деятельности с использованием гибридного подхода, где обеспечивается интеграция классических ма-

тематических методов корреляционного анализа, анализа временных рядов и методов интерпретации получаемых результатов.

С учетом сказанного основными задачами обсуждаемой в работе НИР была разработка системы, обеспечивающей

- формирование независимой системы индикаторов науки, образования и инновационной деятельности;
- построение системы агрегатов статистических данных с целью формирования опорных точек для анализа ситуации в регионах России в различных аналитических разрезах;
- формирование векторов индикаторов для наборов статистических данных и их агрегатов;
- определение семантической близости векторов и формирование кластеров индикаторов похожих регионов РФ;
- формирование динамических групп статистических данных и их агрегатов;
- анализ поведения индикаторов науки, образования и инновационной деятельности в регионах РФ.

Следует отметить, что математическая статистика имеет давнюю историю и серьезные научно-технические результаты [1]. В последнее время активно развивается и теория анализа паттернов данных [2, 3], а полученные в этих областях результаты могут использоваться для решения вышперечисленных задач [4, 5]. Не менее активно ведутся разработки соответствующих програм-

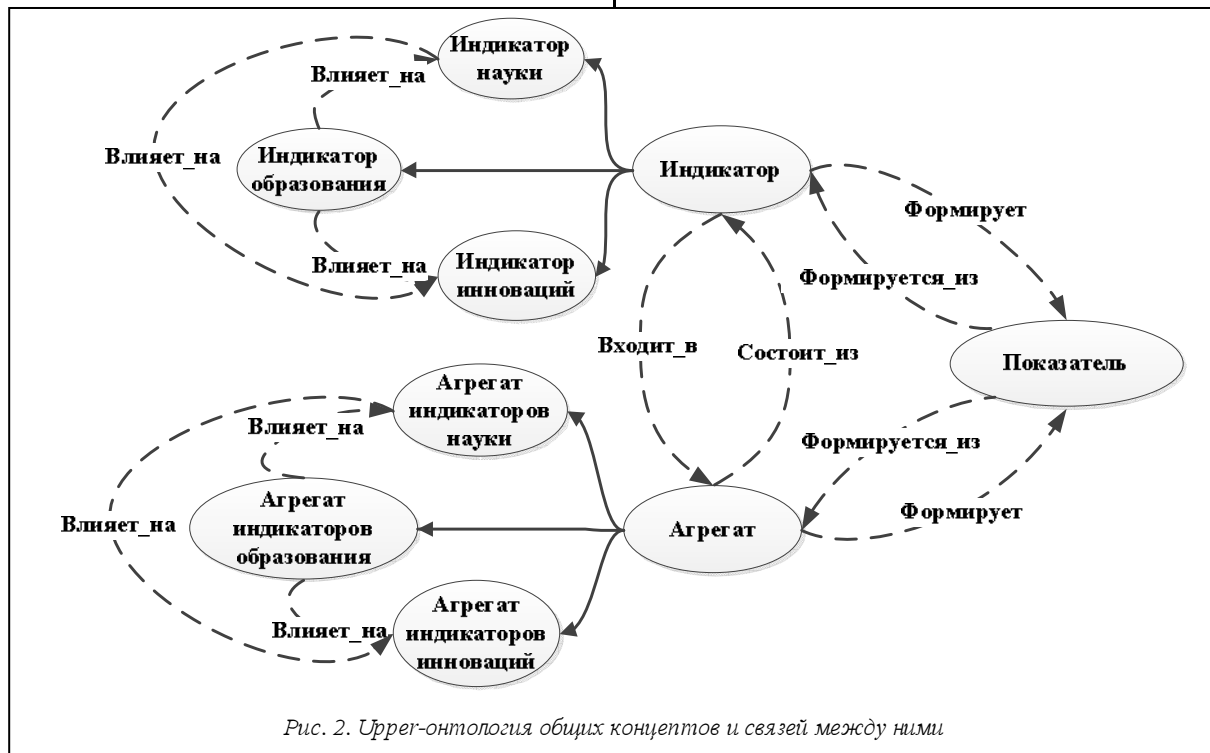


мных средств [6, 7]. В качестве примеров наиболее мощных и распространенных статистических пакетов можно назвать SAS, SPSS, STATGRAPHICS, STATISTICA, STADIA и другие. Препятствиями на пути активного использования таких па-

пользователя серьезной подготовки в области статистики и смежных областях, а также предъявляют достаточно высокие требования к вычислительной мощности компьютеров для развертывания и работы. К тому же такие системы весьма дороги (от \$1 000 до \$15 000 [7]). Поэтому разработка и реализация АРМ аналитика, с одной стороны, использующего достаточно мощные методы анализа данных, а с другой – обеспечивающего пользователей гибкими средствами варьирования и обработки исходных данных при щадящих требованиях к памяти и быстродействию, представляется серьезной научно-технической и практически значимой задачей.

Математическое обеспечение системы АИДА

Для изучения статистических связей в триаде «образование–наука–инновации» необходимы данные о процессах на уровне тех объектов, в которых эти процессы совершаются. Так, для анализа инновационных процессов нужны статистические данные на уровне отдельных компаний и предприятий, науки – на уровне научных подразделений университетов, фирм и министерств, включая институты РАН и других академий, образования – на уровне отдельных университетов. Однако, как показывает анализ доступных данных, в настоящее время подобной статистики не существует. В обсуждаемой НИР было принято



кетов и систем при решении прикладных задач является то, что они предполагают наличие у

решение за единицы наблюдения взять отдельные регионы РФ, поскольку благодаря их относитель-

но большому числу (83 региона) применение методов анализа данных является возможным и целесообразным. Стандартные проблемы при таком выборе связаны с тем, что исходные измерения носят здесь непрямой, агрегированный, характер, что предопределяет непрямой, косвенный, характер проявления статистических связей, а исследуемые процессы нестабильны в силу незавершенности системы функционирования производственных, экономических и социальных процессов в РФ. Дополнительные сложности создает несовершенство инструментария статистического анализа социально-экономических процессов. С учетом вышесказанного предлагаемое решение концентрируется вокруг подходов, разработанных для анализа слабо структурированных процессов, прежде всего вокруг факторного и кластерного анализа, а также майнинга данных, машинного обнаружения закономерностей и онтологического моделирования.

Таким образом, в основе математического обеспечения обсуждаемой системы лежат отказ от построения всеобъемлющей модели изучаемого процесса и переход на уровень изучения и моделирования структуры данных. При этом происходят, по крайней мере на уровне статистико-математической обработки, отказ от понятия причинной связи и замена его на понятие корреляции, а поддержка процессов анализа данных осуществляется с использованием системы онтологических моделей индикаторов и агрегатов науки, образования и инновационной деятельности. Общая схема обработки данных в рассматриваемой системе представлена на рисунке 1.

В рамках разработки онтологических моделей

индикаторов науки, образования и инновационной деятельности используется многоуровневая система взаимосвязанных онтологий. С учетом этого все базовые онтологические модели опираются на Uppreg-онтологию общих концептов и связей между ними, схема которой приведена на рисунке 2.

В процессе разработки математического обеспечения системы АИДА были спроектированы онтологические модели индикаторов науки, образования и инноваций. Для примера на рисунке 3 представлен фрагмент интеллект-карты индикаторов «Кадры науки», а на рисунке 4 – фрагмент результатов онтологического моделирования индикаторов инновационной деятельности в системе Protégé [8].

В соответствии с общей схемой обработки данных в системе АИДА (рис. 1) на первом этапе осуществляются корреляционный анализ исходных показателей и выбор базовой системы независимых показателей, а также агрегация. Методы такого анализа хорошо известны [1–4]. Заметим только, что формирование агрегатов осуществляется на базе факторного анализа исходных данных и метода главных компонент, нормирование каждого показателя – методом деления на размах, а при агрегировании в блоки показатели принимаются равнозначными. Для примера в таблице представлены результаты корреляционного анализа агрегатов (блоков), сформированных для обработки на исходных данных индикаторов.

Результаты корреляционного анализа агрегатов (блоков) индикаторов за 2007 г.

Блок	Блок					
	1	2	3	4	5	6

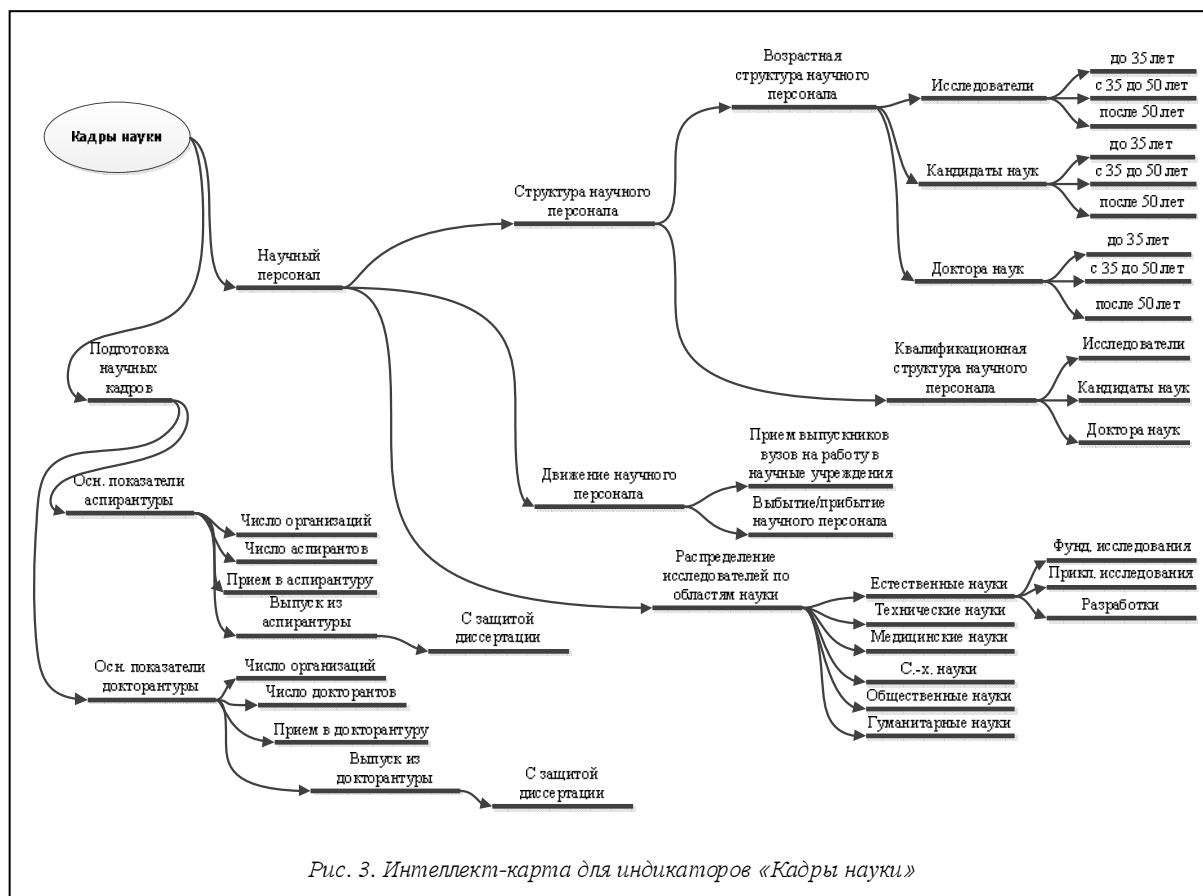


Рис. 3. Интеллект-карта для индикаторов «Кадры науки»

1	1					
2	-0,2	1				
3	0,16	-0,1	1			
4	0,11	0,11	0,7	1		
5	-0,2	0,14	0,22	0,33	1	
6	-0,1	0,05	0,05	0,13	0,21	1

Примечание: блок 1 – социально-экономические условия, блок 2 – образовательный потенциал, блок 3 – потенциал научно-технической деятельности, блок 4 – результативность исследований и разработок, блок 5 – потенциал инновационной деятельности, блок 6 – результативность инновационной деятельности.

В соответствии с общей схемой обработки данных в системе АИДА дальнейшая обработка исходных данных предполагает формирование паттернов данных и их анализ.

Как известно [2–4], анализ паттернов связан с поиском взаимосвязей исследуемых объектов, построением их классификации и исследованием развития объектов во времени. При этом суть ме-

вие последовательность номеров кластеров (траектория) $\{y_i^1, y_i^2, \dots, y_i^t, \dots, y_i^k\}$, к которым данный объект x_t принадлежал в момент времени t . На основании полученных траекторий развития выделяются динамические группы объектов, демонстрирующих одинаковое поведение и сходные характеристики во времени.

Спецификой использованного в системе АИДА метода анализа паттернов является то, что для выделения однородных групп среди всех объектов они оцениваются не по абсолютным значениям их признаков, а по углам наклона кривых, обозначающих объекты в системе параллельных координат [9].

Таким образом, в рассматриваемой системе АИДА базовыми функционалами статистического анализа являются факторный анализ исходных данных, модифицированный метод К-средних и методика отыскания главных компонент, а также методы структурного распознавания образов и по-

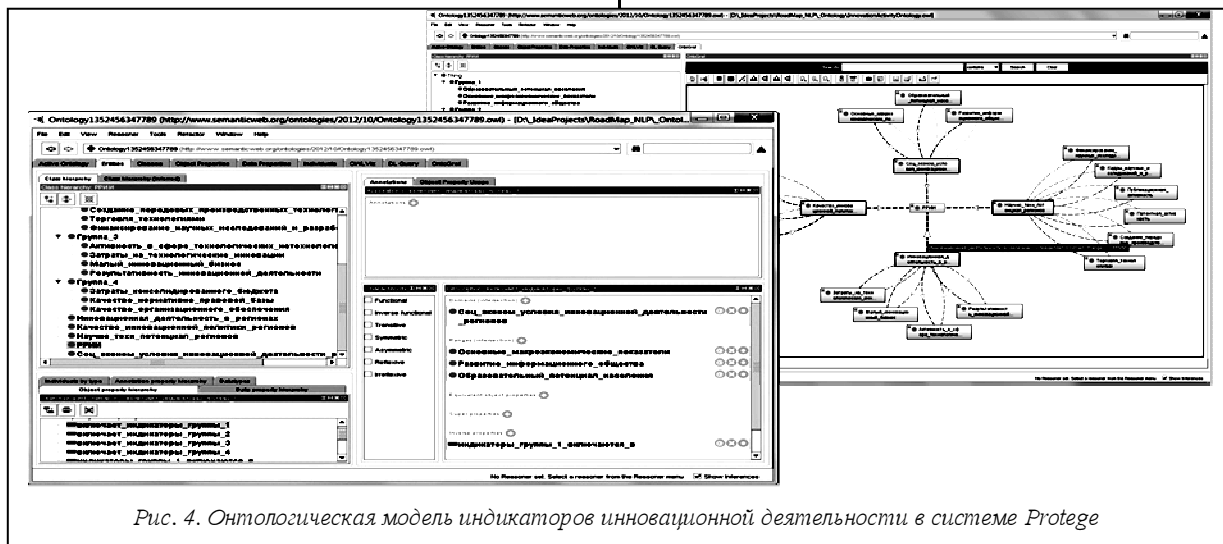


Рис. 4. Онтологическая модель индикаторов инновационной деятельности в системе Protege

тода анализа паттернов в следующем:

- заданы множество объектов X и множество меток (номеров, имен) кластеров Y , а также мера близости (функция расстояния) $p(x, x')$ между объектами $x, x' \in X$;
- требуется разбить выборку $X^m = \{x_1, \dots, x_m\} \in X$ на непересекающиеся подмножества, называемые паттернами, так, чтобы каждый паттерн состоял из объектов, близких по метрике p , а объекты разных паттернов существенно отличались.

Для этого каждому объекту $x \in X$ ставится в соответствие n -мерный вектор $z_x = \{z_1^x, \dots, z_n^x\}$, характеризующий объект x по n признакам. По каждому вектору z_x строится кривая, проходящая через точки z_1^x, \dots, z_n^x , то есть функция $f^x(i) = z_i^x$ для всех $i=1-n$. Как правило, в качестве функции f^x выбирается кусочно-линейная функция.

Для проведения динамического анализа каждому объекту выборки $x_t \in X$ ставится в соответ-

ствии ассоциаций.

Проектирование и реализация рабочего места аналитика АИДА

Основные принципы, положенные в основу разработки и реализации системы АИДА, следующие:

- поддержка всех этапов обработки статистических данных науки, образования и инновационной деятельности;
- возможности варьирования исходных данных и пересчета получаемых результатов;
- получение отчетов по результатам сессии в форматах Excel, Word, PDF и др.;
- дружелюбный интерфейс, не требующий опыта работы со сложными программными комплексами;
- возможность расширения ПО системы но-

выми функциональными компонентами.

Для реализации системы выбрано трехуровневое решение в программной архитектуре клиент-сервер. При этом приложение размещается на рабочей станции, а основными составными частями реализации являются компоненты с графическим интерфейсом, обменивающиеся данными через слой бизнес-логики с сервером БД. Бизнес-процессы обеспечиваются взаимодействием между клиентской частью, сервером БД, а также внешним обработчиком статистических данных.

В результате выбора трехуровневой программной архитектуры клиент-сервер достигаются встроенная безопасность данных, обеспечение централизации хранения и доступа к данным, а также простота сопровождения и развития системы.

Статистические данные науки, образования и инновационной деятельности хранятся на серверной компоненте, реализованной на основе сервера БД Microsoft SQL Server версии не ниже 2008, которая обеспечивает полноценный контроль безопасности с использованием встроенного сервиса разграничения доступа Windows Authentication. Управление доступом к данным определяется политиками, настраиваемыми в одной точке, что обеспечивает удобство администрирования и переноса данных. Бизнес-процессы распределены между компонентами, взаимодействующими друг

однозначно определенных протоколов, вследствие чего любой компонент не зависит от изменений в смежных компонентах и может дорабатываться без внесения изменений в систему в целом.

Бизнес-процессы анализа статистических данных науки, образования и инновационной деятельности декомпозированы в системе АИДА на следующие логические модули: модуль хранения статистических данных и их агрегатов, паттернов данных и результатов их анализа; собственно модуль анализа паттернов данных; модуль визуализации и модуль формирования отчетности.

Каждый модуль предоставляет смежным модулям определенные интерфейсы, содержащие методы, события и свойства, и обладает свойствами обеспечения пригодности для повторного использования, стандартизации относительно бизнес-задачи, контекстной независимости, расширяемости и инкапсуляции.

Для реализации компонентной архитектуры выбрана платформа Microsoft .NET Framework [10], обеспечивающая системную поддержку построения решений с применением компонентного подхода. Компоненты пользовательского интерфейса слоя представления, бизнес-компоненты слоя бизнес-логики и компоненты интерфейса данных формируются в виде сборок .NET Framework и выполняются под управлением среды выполнения .NET Framework.

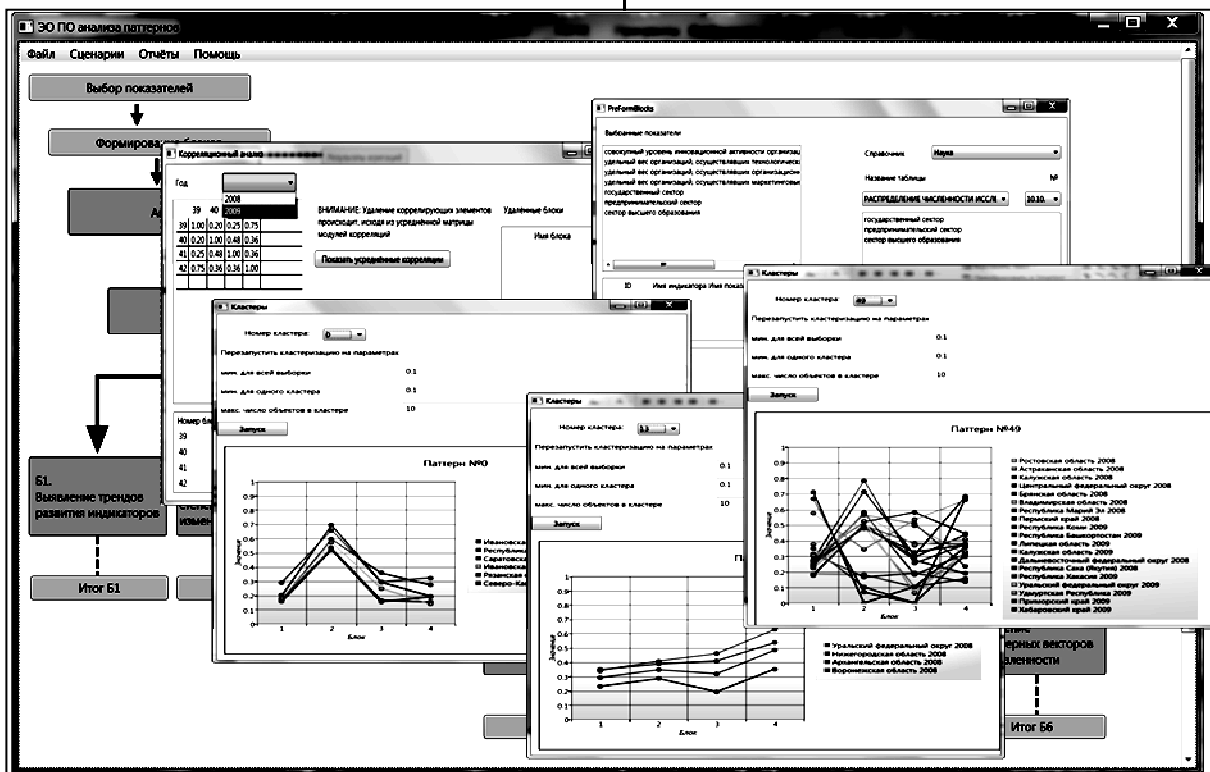


Рис. 5. Результаты первого этапа обработки данных в системе АИДА

с другом посредством открытых интерфейсов и

В процессе реализации вышеописанной много-

слоистой архитектуры применен комплексный шаблон проектирования «Отделение представления» (Separated Presentation), в рамках которого используется перечень шаблонов, определяющих взаимодействие пользователя с визуальным интерфейсом, представление, бизнес-логику и модель данных ПО.

Для обеспечения полного функционального охвата в системе АИДА реализованы

- методы корреляционного анализа для подготовки исходных данных для корректной дальнейшей обработки;
- методы формирования агрегатов признаков и/или объектов, включая факторный анализ, и метод главных компонент (количественный анализ), формирование отдельных кластеров или кластерных разбиений (номинальный анализ) и формирование ранжирований (порядковый анализ), а также определение связи концепций на основе методов регрессионного или лог-линейного анализа (количественный анализ), методов распознавания образов и построения ассоциаций (качественный анализ);
- модифицированный метод К-средних и методика отыскания главной компоненты для получения адекватной системы сформированных кластеров.

В рамках реализации бизнес-процессов анализа данных в системе АИДА обеспечивается выполнение следующих функций:

- выявление трендов изменения индикаторов (показателей, рассчитываемых в ходе агрегации и преобразования статистических данных);
- выявление неявно выраженного взаимного влияния индикаторов;
- выявление нетипичной динамики индикаторов;
- оценка степени однородности изменения индикаторов (их агрегатов);
- идентификация взаимного влияния траекторий развития индикаторов (их агрегатов);
- определение характерных векторов направленности индикаторов и их качественная оценка для различных аналитических измерений, включая разные виды экономической деятельности, высоко-, средне- и низкотехнологичные отрасли, федеральные округа, регионы-доноры и дотационные регионы, формы собственности предприятий;
- определение характерных векторов направленности индикаторов Best Performance и их качественная оценка.

Для обеспечения комфортной работы пользователя в рамках решения задач анализа данных науки, образования и инновационной деятельности в системе АИДА поддерживается дружелюбный интерфейс.

Так, все функционалы системы отражаются на технологической схеме, где отдельные этапы обработки информации представлены в виде кнопок,

активация которых обеспечивает выполнение соответствующих программных модулей.

Анализ индикаторов активности регионов России в системе АИДА

В качестве исходных для такого анализа были выбраны статистические данные из сборников ВШЭ, а также данные Федеральной службы государственной статистики. В результате обработки исходных данных проведен анализ активности регионов России по показателям науки, образования и инновационной деятельности в 2007–2010 гг., получена классификация регионов по схожести внутренней структуры показателей, также построены траектории развития регионов с течением времени и выявлены группы регионов, придерживающихся одинаковой стратегии развития показателей.

Детальное обсуждение результатов проведенного анализа представлено в работе [4], в данном случае остановимся на нескольких функционалах системы АИДА, демонстрирующих ее возможности.

В соответствии с общей схемой (рис. 1) первый этап обработки данных в системе АИДА заключается в выборе системы показателей. Его результаты для случая анализа индикаторов развития науки, образования и инновационной деятельности в регионах РФ показаны в системе экранных форм, представленной на рисунке 5.

Полученные результаты открывают пользователю доступ к бизнес-процессам анализа индикаторов науки, образования и инноваций, среди которых одним из самых сложных является функционал выявления регионов в соответствии со сценарием Best Performance.

Решение задачи такого анализа в системе АИДА осуществляется в две стадии. На первой из них пользователь из множества паттернов данных регионов выбирает те, которые, по его мнению, являются примерами Best Performance. Так, на-



Рис. 6. Формирование класса Best Performance

пример, на рисунке 6 представлена экранная форма результатов формирования такого класса, где в

качестве базиса выбрана Республика Татарстан как один из регионов-доноров в РФ.

На второй стадии, выполнение которой запускается нажатием кнопки «Старт», происходит разбиение множества паттернов данных индикаторов регионов на два класса: Best Performance и «Остальные регионы» (рис. 7).

Аналогично в системе АИДА выполняются и остальные бизнес-процессы формирования и анализа паттернов данных индикаторов науки, образования и инновационной деятельности регионов РФ.

По результатам сессии пользователь может сформировать отчет в одном из поддерживаемых системой форматов: Excel, Word, PDF и др.

Таким образом, представленная в настоящей работе система автоматизированной обработки индикаторов данных обеспечивает решение достаточно сложного спектра задач анализа статистиче-

просов поддержки функционалов выявления трендов изменения индикаторов, неявно выраженного взаимного влияния индикаторов и нетипичной динамики индикаторов, оценке степени однородности изменения индикаторов (их агрегатов) и идентификации взаимного влияния траекторий развития индикаторов (их агрегатов), а также определению характерных векторов направленности индикаторов Best Performance и их оценке.

Направления дальнейших исследований и разработок по данной теме предполагают разработку методов и средств прогнозирования динамики изменения индикаторов науки, образования и инновационной деятельности, методов и средств автоматизированного построения структурных описаний паттернов данных, экспертных систем поддержки принятия решений на основе использования паттернов данных.

Кроме того, планируется доведение экспери-

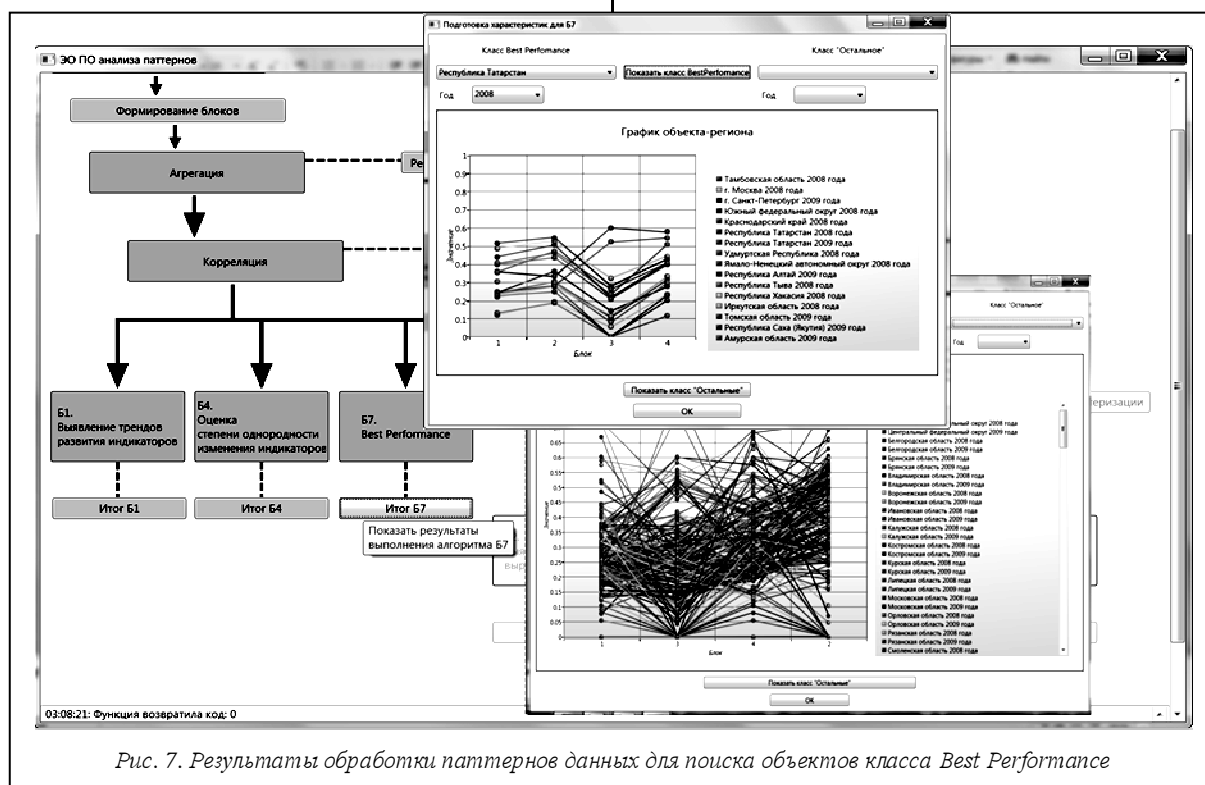


Рис. 7. Результаты обработки паттернов данных для поиска объектов класса Best Performance

ской информации о развитии регионов РФ.

В заключение можно отметить следующее. В статье представлены результаты разработки и реализации рабочего места аналитика, поддерживающего процессы интеллектуального анализа больших объемов статистических данных науки, образования и инновационной деятельности с использованием гибридного подхода, где обеспечивается интеграция классических математических методов корреляционного анализа, анализа паттернов данных и временных рядов, а также методов интерпретации получаемых результатов. При этом особое внимание уделено обсуждению во-

ментального образца системы АИДА до программного продукта и выведение его на рынок.

Литература

1. Ким Дж-О., Мюллер Ч.У., Клекка У.Р. [и др.]. Факторный, дискриминантный и кластерный анализ; [пер. с англ.]. М.: Финансы и статистика, 1989. 215 с.
2. Wang W., Yang J., Mining Sequential Patterns from Large Data Sets, Series: Advances in Database Systems, Springer, 2005, Vol. 28, 160 p.
3. Mahdi E., Fazekas G., Finding Sequential Patterns from Large Sequence Data // Intern. Journ. of Comp. Sc. Is., 2010, Vol. 7, Is. 1, no. 1.
4. Алескерев Ф.Т., Гохберг Л.М., Егорова Л.Г., Мячин А., Сагиева Г.С. Анализ данных науки, образования и иннова-

ционной деятельности с использованием методов анализа паттернов // Препринт WP7/2012/07. Нац. исслед. ун-т ВШЭ. М.: Изд. дом ВШЭ, 2012.

5. Хорошевский В.Ф. Семантическая интерпретация паттернов данных на основе структурного подхода // Искусственный интеллект и принятие решений. 2013. № 2.

6. Gualtieri M., Powers S., Brown V., The Forrester Wave™: Big Data Predictive Analytics Solutions, Q1 2013. Forrester research, Inc., 2013, January 13.

7. Ганюшкин М. Вам TUDA // Открытые системы. 2013. № 2.

8. Protege Homepage, URL: <http://protege.stanford.edu/> (дата обращения: 26.05.2013).

9. Few S., Multivariate Analysis Using Parallel Coordinates, URL: http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf (дата обращения: 26.05.2013).

10. Microsoft .NET Framework, URL: <http://www.microsoft.com/net> (дата обращения: 26.05.2013).

References

1. Kim Jae-On, Mueller Ch W., Klecka W.R., *Factor, Discriminant, and Cluster Analysis*, Beverly Hills, CA, Sage Publications, 1989.

2. Wang W., Yang J., *Mining Sequential Patterns from Large Data Sets*, Advances in Database Systems series, Vol. 28, Springer, 2005.

3. Mahdi E., Fazekas G., *IJCSI*, Vol. 7, iss. 1, no. 1, 2010.

4. Aleskerov F.T., Gokhberg L.M., Egorova L.G., Myachin A., Sagieva G.S., *Analiz dannykh nauki, obrazovaniya i innovatsionnoy deyatel'nosti s ispol'zovaniem metodov analiza patternov* [Data analysis of science, education and innovation using patterns analysis methods], Moscow, HSE Publ., 2012.

5. Khoroshevsky V.F., *Iskusstvennyy intellekt i prinyatie resheniy* [Artificial intelligence and decision making], no. 2, 2013.

6. Gualtieri M., Powers S., Brown V., *The Forrester Wave™: Big Data Predictive Analytics Solutions*, Q1 2013, Forrester research, Inc., 2013.

7. Ganushkin M., *Otkrytye sistemy* [Open systems], no. 2, 2013.

8. Protege Homepage, available at: <http://protege.stanford.edu/> (accessed 26 May 2013).

9. Few S., *Multivariate Analysis Using Parallel Coordinates*, available at: http://www.perceptualedge.com/articles/b-eye/parallel_coordinates.pdf (accessed 15 May 2013).

10. Microsoft .NET Framework, available at: <http://www.microsoft.com/net> (accessed 25 May 2013).

УДК 621.396

РЕШЕНИЕ ЗАДАЧИ ОПТИМАЛЬНОЙ МАРШРУТИЗАЦИИ ПО КРИТЕРИЮ СРЕДНЕЙ ЗАДЕРЖКИ

Б.И. Марголис, д.т.н., профессор; М.М. Музанна, аспирант
(Тверской государственной технической университет,
наб. Аф. Никитина, 22, г. Тверь, 170026, Россия,
borismargolis@yandex.ru, mohamed1984a@yahoo.com)

Рассмотрена потоковая модель информационной сети с альтернативной маршрутизацией сообщений для произвольного количества коммутационных узлов и каналов связи. Сформулированы критерий оптимизации по времени задержки сообщений, ограничения на пропускные способности каналов и условия сохранения потоков в сети. С использованием метода неопределенных множителей Лагранжа получены аналитические соотношения для оптимального распределения информационных потоков в сети из трех линий связи. Полученные соотношения распространены на произвольное количество линий связи. Показана справедливость формул для расчета оптимального распределения потоков для частного случая сети, состоящей из двух линий связи. Разработана программа расчета в среде MatLab, позволяющая по топологической структуре сети, матрице входных потоков и пропускной способности каналов найти оптимальное распределение потоков в сети, состоящей из произвольного количества узлов и линий связи. Рассмотрен пример оптимального разделения входных потоков для сети из четырех узлов и трех линий связи, обеспечивающего минимум среднего времени задержки. Продемонстрировано изменение оптимального распределения потоков для рассмотренного примера при уменьшении числа линий связи до двух. Намечены подходы к решению задачи оптимального распределения потоков при наличии в топологии моделируемой сети промежуточных узлов.

Ключевые слова: информационная сеть, коммутационный узел, канал связи, средняя задержка сообщений, пропускная способность, входной поток, топология сети, оптимальная маршрутизация.

PROBLEM COST ROUTING CRITERION AVERAGE DELAY

Margolis B.I., Ph.D., professor; **Muzanna M.M.**, postgraduate
(Tver State Technical University, 22, Quay Nikitin, Tver, 170026, Russia,
borismargolis@yandex.ru, mohamed1984a@yahoo.com)

Abstract. The flow model of an information network with alternative routing of messages to any number of switching nodes and links. We formulate the optimization criterion for the delay posts, restrictions on the channel capacity and the conservation of flows on the network. Using the method of Lagrange multipliers, analytical relations for the optimal distribution of information flows in a network of three lines. The relations obtained extended to any number of lines. The validity of the formulas for calculating the optimal flow distribution for the special case of a network consisting of two lines. The program of calculation among MatLab, which allows for the topological structure of the network, the matrix of the input streams and bandwidth to find the optimal flow distribution in the network, consisting of an arbitrary number of nodes and links. An ex-