

ТЕХНОЛОГИИ ВИЗУАЛИЗАЦИИ ДЛЯ ПРИКЛАДНЫХ ЗАДАЧ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ¹

А.А. ЛЕЖЕБОКОВ, Э.В. КУЛИЕВ

Институт компьютерных технологий и информационной безопасности
Южного федерального университета
347928, Ростовская область, г. Таганрог, пер. Некрасовский, 44.
E-mail: info@ictis.sfedu.ru

В работе рассматриваются современные подходы к разработке новых и модификации существующих технологий визуализации данных и знаний. Классические подходы к визуализации утрачивают актуальность в связи с переходом на кроссплатформенные технические решения и необходимостью визуализировать большие наборы данных, в том числе относящихся к категории «big data». Решение прикладных задач интеллектуального анализа данных, таких как классификация, кластеризация, поиск зависимостей и закономерностей в данных, построение многомерных кубов, является, несомненно, важной и актуальной задачей, качество решения которой напрямую связано с используемой технологией и способом визуализации. В работе отмечается необходимость своевременного и рационального использования таких технологий, как виртуальная и дополненная реальность, которые позволили совершить скачок в развитии технических средств визуализации, в том числе многомерных данных. Наличие расширенных интерактивных возможностей и средств манипулирования данными позволяет сделать вывод о зарождении нового этапа в развитии технологий визуализации.

Ключевые слова: визуализация, интеллектуальный анализ данных, прикладная задача, большие наборы данных, классификация.

ВВЕДЕНИЕ

В настоящее время наблюдается взрывной рост накопленных данных и информации в разных областях, в основном вызванный ростом технологических возможностей хранения и обработки данных. Широкое использование Интернета и связанных с ним технологий увеличило потребность и необходимость разработки новых эффективных методов интеллектуального анализа данных и визуализации результатов. Интеллектуальный анализ данных, который касается обнаружения и извлечения фрагментов знаний из больших хранилищ данных, направлен на решение важнейших прикладных задач в области автоматизации обнаружения скрытых шаблонов и отношений (связей) между данными, которые не всегда могут быть очевидными [1].

Инструменты интеллектуального анализа данных включают методы классификации (такие как деревья решений, правила индукционных программ и нейронные сети), алгоритмы кластеризации и подходы к построению правил ассоциации. Интеллектуальный анализ данных плодотворно используется во многих областях, включая маркетинг, медицину, финансы, инженерию и биоинформатику. Однако существует еще ряд факторов, которые препятствуют широкому внедрению и использованию этих технологических решений. Это в основном связано с тем, что результаты многих методов интеллектуального анализа данных часто трудно понять. Например, трудно проанализировать работу метода по извлечению данных, в результате которой будут созданы 300 страниц правил. Визуальное представление знаний, заложенных в такие правила, поможет повысить практическую значимость и применимость результатов. Визуализация данных и информации дает пользователям возможность получить новое представление о знаниях, которые извлека-

¹ Работа выполнена при финансовой поддержке гранта РФФИ № 18-29-22019

ются из больших хранилищ. В данной работе описан ряд важных проблем интеллектуального анализа данных и представлены новые модели и методы, используемые для улучшения понимания результатов анализа сложных и больших данных.

Необходимо выполнять визуализацию исходных данных до и во время выполнения интеллектуального анализа данных. Посредством визуализации может быть оценено качество данных в процессе обнаружения знаний, который включает в себя следующие шаги: предварительную обработку данных, интеллектуальный анализ данных и отчетность [1]. Важной частью работы является то, как знания, обнаруженные инструментами интеллектуального анализа данных, могут быть визуализированы в процессе обучения. Отдельным направлением исследований являются технологии дополненной, виртуальной и смешанной реальности, которые могут быть использованы для представления интерактивных перспектив многомерных данных и процессов интеллектуального анализа данных.

1. ПОСТАНОВКА ЗАДАЧИ

Под визуализацией данных мы понимаем такой способ представления многомерного распределения данных на двумерной/многомерной плоскости, при котором по крайней мере качественно отражены основные закономерности, присущие исходному распределению, – его кластерная структура, топологические особенности, внутренние зависимости между признаками, информация о расположении данных в исходном пространстве и т.д. В качестве основных применений методов визуализации можно указать следующие [2]:

- а) наглядное представление геометрической метафоры данных;
- б) лаконичное описание внутренних закономерностей, заключенных в наборе данных;
- в) сжатие информации, заключенной в данных;
- г) восстановление пробелов в данных;
- д) решение задач прогноза и построения регрессионных зависимостей между признаками.

Один из способов целенаправленного проецирования в пространства малой размерности (в зарубежной литературе – *projecting pursuit*) заключается в следующем: найти такое отображение U (способ проецирования) из исходного пространства на двумерную плоскость, которое бы оптимизировало заданный критерий качества Q – некоторый функционал от координат точек данных до и после процедуры проецирования: $Q(U, X)$. Здесь под X понимается весь набор многомерных данных, а Q зависит от параметров отображения.

Например, если каждой точке многопараметрических данных приписать две координаты (например, с использованием квазилинейной модели), то это позволит построить в пространстве данных гладкое многообразие, которое обладает свойством обобщать заключенную в данных информацию и служит для лаконичного описания, сжатия информации или для восстановления пробелов в данных. Тогда проецирование данных в пространство меньшей размерности заключается в настройке процедуры построения моделирующей двумерной поверхности, вложенной в многомерное пространство признаков.

Рассмотрим все этапы анализа данных [2-5].

Формулирование цели. Каждое исследование должно отвечать на ряд поставленных вопросов – не нужно плодить исследования для исследований.

Сбор данных. На этом этапе аналитик или работает с уже собранными данными, или участвует в процессе постановки задания на сбор данных (фактически решает, какая информация ему необходима и в каком виде).

В первом случае особое внимание стоит уделить правильной интерпретации данных, которые записаны в базу, и зачастую смириться с существующим форматом данных, дизайном таблиц и т. д. Во втором случае аналитик сталкивается с проблемой построения грамотного сценария сбора данных – он может особенно перестараться в планировании А/В-тестов, логировании событий и т. п. Здесь важна коммуникация с программистами, которые могут помочь в понимании процессов и оценке масштабов планируемых записей.

Подготовка данных. «Мусор на входе – мусор на выходе» – правило, о котором всегда нужно помнить. Структурирование, устранение ошибок, изменение форматов содержимо-

го, разбор аномальных результатов, очистка от выбросов, устранение дубликатов, интеграция данных из разных источников – одни из важнейших пунктов в анализе данных.

Иногда требуется расширение метрик, например, добавление вычислительной информации (прирост, ранг, номер и т. п.). Иногда следует сократить количество признаков (переменных) или перейти к вспомогательным переменным, принимающим одно из двух значений: true (1) / false (0). На этом этапе сырые данные превращаются в полезную входную информацию для моделирования и анализа.

Исследование данных. Для правильной интерпретации многомерных данных необходимо посмотреть на них в разрезе как конкретного признака, так и группы признаков. Также следует представить ключевые показатели в динамике с планами и фактическими результатами. Именно на этом этапе подбирается формат будущей визуализации.

Визуализация и построение выводов. Каждое исследование должно заканчиваться результатами и выводами. Даже если они негативные, их стоит проговорить и обсудить. При этом правильная постановка задачи, методика проведения сбора данных, правильная интерпретация результатов, выявленные ошибки и многое другое должны послужить базой для последующих исследований. Время, затрачиваемое аналитиком на каждую фазу, зависит от многих переменных: начиная от опыта работы и уровня знания данных, заканчивая перечнем используемых инструментов и технических характеристик ПК.

Необходимо также понимать, что процесс анализа данных имеет итерационный характер и может быть представлен циклом (рис. 1) [6].

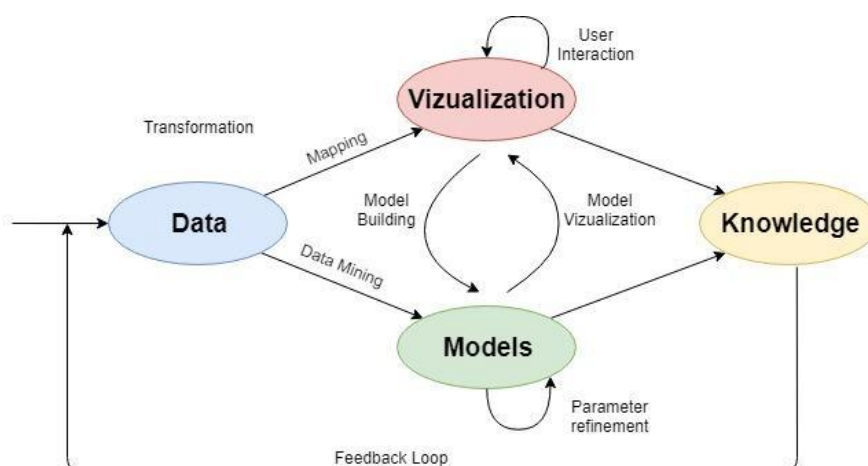


Рис. 1. Процесс анализа данных

2. АНАЛИЗ ДАННЫХ

Люди интуитивно ищут новые особенности, модели, тенденции, связи и отношения в различных данных. Посредством визуализации данных и полученных концептуальных описаний (например, в форме правил) может быть получен качественный обзор больших и сложных наборов данных. Кроме того, визуализация данных и правил может помочь в определении областей интереса и соответствующих параметров для более сфокусированного количественного анализа. Пользователь может таким образом получить «грубое представление» о качестве данных с точки зрения их правильности, адекватности, полноты, актуальности и т. д. Таким образом, использование визуализации данных и правил значительно расширяет диапазон ситуационных и имитационных моделей, которые могут быть понятны широкому кругу пользователей, что облегчает управление компромиссом «точность в сравнении с понятностью» [1].

Методы интеллектуального анализа данных строят модель данных посредством повторяющихся вычислений, чтобы найти статистически значимые отношения в данных. Тем не менее система зрительного восприятия человека может обнаружить шаблоны в данных, которые неизвестны инструменту интеллектуального анализа данных. Эта комбинация различных сильных сторон визуальной системы человека и инструментов интеллектуального анализа данных может впоследствии привести к открытию новых идей и улучшений в понимании человеком проблемы (анализируемого набора данных). Визуальный анализ данных использует возможности системы человеческого зрения, что делает ее эффективным инструментом для понимания распределения данных, обнаружения паттернов, кластеров и «выбросов» (всплесков) в данных. Визуальный анализ данных в настоящее время является активной областью прикладных исследований. Примеры коммерческих пакетов интеллектуального анализа включают [1, 3, 7]:

- систему интеллектуального анализа данных DBMiner, которая является частью RuleQuest,
- набор инструментов для интеллектуального анализа данных, Clementine, разработанный Integral Solutions Ltd (ISL),
- Enterprise Miner, разработанный SAS Institute,
- Intelligent Miner производства IBM и различные другие инструменты.

Прикладные инструменты на основе нейронных сетей, такие как NeuroSolutions и SNNS и байесовская сеть, включая Hugin, TETRAD и Bayesware Discoverer, также включают средства визуализации. Примеры связанных исследовательских проектов и подходов к визуализации включают MLC ++, WEKA, AlgorithmMatrix и C4.5 / See5.

Визуальный интеллектуальный анализ данных объединяет визуализацию данных и интеллектуальный анализ данных и тесно связан с компьютерной трехмерной и интерактивной графикой, мультимедийными системами, человеко-машинными интерфейсами, с задачами распознавания образов и с задачами повышения производительности вычислений.

3. ВИЗУАЛИЗАЦИЯ ДАННЫХ

Визуализация данных и визуализация информации в настоящее время дополнительно исследуется с точки зрения определения преимуществ для извлечения знаний.

Визуализация данных предоставляет мощный механизм, помогающий пользователю решать прикладные задачи во время предварительной обработки данных и фактического анализа данных. Благодаря визуализации исходных данных пользователь может просматривать визуальные образы, чтобы «прочувствовать» свойства этих данных. Например, большие образцы могут визуализироваться и анализироваться. В частности, визуализация может быть использована для обнаружения «всплесков», которые выделяют «неожиданности» в данных, то есть экземпляры данных, которые не соответствуют общему поведению или общей модели данных. Кроме того, пользователь помогает в выборе соответствующих данных через визуальный интерфейс. Преобразование данных является важным этапом предварительной обработки данных. Во время преобразования данных визуализация данных может помочь пользователю обеспечить правильность этого преобразования. То есть пользователь может определить, являются ли два представления (оригинальные и преобразованные) данных эквивалентными. Визуализация также может быть использована для помощи пользователям при интеграции источников данных, помогая им видеть отношения в различных форматах данных [7].

Методы визуализации данных классифицируются по трем аспектам.

Во-первых, фокус (символичность); во-вторых, их стимул (2D против 3D); и, наконец, их отображение (статическое или динамическое). Кроме того, данные в хранилище данных могут рассматриваться как различные уровни детализации или абстракции или как различные комбинации атрибутов или размеров. Данные могут быть представлены в различных визуальных форматах, включая коробочные графики, точечные диаграммы, 3D-кубы, диаграммы распределения данных, кривые, визуализацию объема, поверхности или графики-ссылки.

Точечные диаграммы относятся к визуализации элементов данных по двум осям, а именно значениям X и Y. Согласно последним исследованиям график рассеяния является наиболее популярным инструментом визуализации, так как это может помочь найти кластеры, выбросы, тенденции и корреляции. На рис. 2 показан пример графика рассеяния, используемый в системе ViziMine, которая визуализирует как набор данных, так и подмножество данных, которые охватываются определенным правилом, обнаруженным инструментом интеллектуального анализа данных.

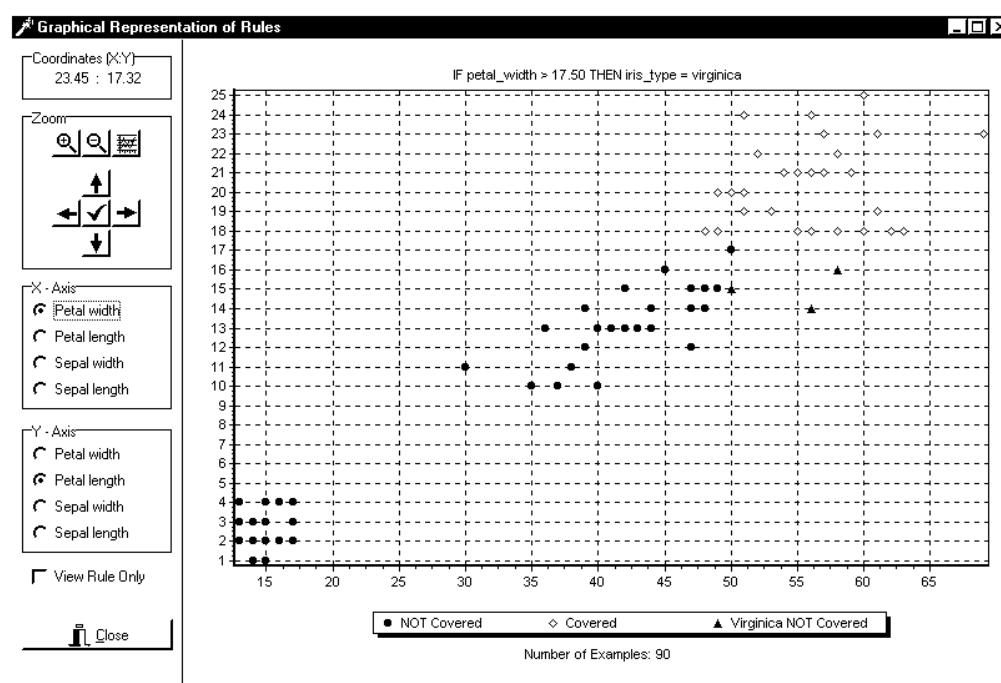


Рис. 2. Пример графика рассеивания

На рис. 2 показан график рассеяния для правила IF (ширина лепестка > 17,50), затем Iris-type = Virginica [8].

Другие методы визуализации данных включают в себя 3D-кубы, используемые в диаграммах отношений, где данные сравниваются как итоги разных категорий. В поверхностных диаграммах точки данных визуализируются с помощью линий между ними. Площадь, определяемая линией, вместе с нижней частью графика впоследствии заполняется. Связи или линейные графики отображают отношения между точками данных путем установки соединительной линии. Они обычно используются для 2D данных, где значение X не повторяется.

Продвинутые методы визуализации могут значительно расширить диапазон ситуационных и имитационных моделей, которые могут быть использованы экспертами в своей прикладной области, тем самым облегчая так называемый компромисс между точностью и понятностью.

Однако из-за так называемого «проклятия размерности», которое относится к проблеме, связанной с работой с многочисленными измерениями, обычно высокоточные модели наименее понятны. В системе интеллектуального анализа данных целью визуализации данных является получение первоначального понимания данных и их качества. Фактическая точная оценка данных и открытие новых знаний являются следующими задачами инструментов интеллектуального анализа данных. Следовательно, визуальное отображение должно быть в высшей степени понятным, в том числе за счет уменьшения точности.

Использование одного или нескольких из вышеупомянутых методов визуализации данных, таким образом, помогает пользователю получить исходную модель данных, чтобы обнаружить возможные «всплески» и получить интуитивно понятную оценку качества данных, используемую для интеллектуального анализа данных.

4. ВИЗУАЛИЗАЦИЯ ИНФОРМАЦИИ

Развитие средств визуализации ситуационных и имитационных моделей интеллектуального анализа данных может быть четко определено в двух ключевых областях, а именно: понимание и доверие. Понимание означает больше, чем просто понимание, обычно оно также включает контекст. Если пользователь может понять, что было обнаружено в контексте бизнес-проблемы, он будет доверять данным и базовой модели и, следовательно, использовать ее. Визуализация модели также позволяет пользователю обсуждать и объяснять логику модели другим пользователям. Таким образом, общее доверие к модели увеличивается, и последующие действия, предпринятые в результате анализа, являются оправданными [9].

Искусство визуализации информации можно рассматривать как сочетание трех четко определенных понятных категорий: когнитивные науки, графика и информационная графика. Подход к визуализации должен обеспечить легкое понимание и знание предметной области, изучение визуальных параметров и получение полезных результатов. Характерные особенности должны быть закодированы графически, и интерактивный процесс должен оказаться полезным для пользователя. Формат знаний, извлекаемых в процессе майнинга (анализа), зависит от типа интеллектуального анализа данных, решаемой задачи и ее сложности. Примеры включают в себя правила классификации, правила ассоциации, временные последовательности и случайные графики. Визуализация этих результатов анализа данных включает в себя представление результатов или знаний, полученных из интеллектуального анализа данных, в визуальных формах, таких как дерево решений, правила ассоциации, кластеры, выбросы и обобщенные правила. Например, кремниевая графика (SGI) Набор инструментов MineSet 3.0 использует диаграммы подключения для визуализации деревьев решений и упрощает байесовские классификаторы и классификаторы таблиц решений. Другие примеры включают в себя визуализатор доказательств, который используется для визуализации байесовских классификаторов; систему DB-Discover, которая использует многоатрибутное обобщение для обобщения данных; и усовершенствованную систему обнаружения NASD, которая использует деревья решений и визуализацию правил ассоциации для наблюдения за фондовым рынком NASDAQ.

В качестве альтернативы визуализация конструкций, созданных инструментом интеллектуального анализа данных (например, правила, дерево решений, ветвей и т.д.), может быть получена с помощью точечных диаграмм и коробочных блоков. Например, точечные диаграммы могут использоваться для указания точек данных, охватываемых одним цветом, а точки, не покрытые анализом, – другим цветом. Этот метод визуализации позволяет пользователям задавать простые вопросы, интуитивно понятные вопросы в интерактивном режиме. То есть пользователь может выполнить некоторые формы анализа «что, если».

5. БУДУЩИЕ ТЕНДЕНЦИИ

Трехмерная визуализация может показать гораздо больше информации, чем двумерная визуализация, сохраняя при этом ее простоту. Эта техника визуализации быстро показывает количество и относительную силу отношений между элементами, помогая сосредоточить внимание на важных данных объектах и правилах. Поэтому она помогает как при предварительной обработке данных, так и в процессе «добычи (извлечения)» знаний.

В двух измерениях представление данных ограничено двумерными графическими элементами. В трехмерном пространстве размеры могут быть использованы как для двух-, так и для трехмерных графических элементов. Эти элементы гораздо более многочисленны и разнообразны при визуализации в трех измерениях по сравнению с двумерным пространством. Кроме того, трехмерные представления (или дескрипторы) могут быть объемными или поверхностными в зависимости от того, представляет интерес внутренняя структура данных или нет. Поверхностное представление привлекает во внимание только внешний вид или оболочку объекта, в то время как объемный подход присваивает значение каждому элементу объема. Последний подход довольно распространен в биомедицинских изображениях, таких как САТ-сканирования. Многие методы доступны для визуализации данных в трех измерениях. Например, это представление данных с помощью глифов. Глиф может быть определен как трехмерный объект, подходящий для представления данных или подмножеств данных. Объект выбран для облегчения как визуализации, так и процесса извлечения данных. Глиф должен быть самоочевидным и однозначным. Глифы могут иметь различные атрибуты, такие как цвет и масштаб. При использовании этих атрибутов для построения глифа применяется контент на основе дескриптора. Даже если большинство символов представляют собой жесткие объекты, нежесткие и сочлененные объекты могут быть также повторно использованы при построении. Затем можно использовать деформацию и положение глифа, чтобы представлять определенное поведение набора данных. Кроме того, глифы могут быть анимированы, чтобы смоделировать некоторый динамический процесс [10].

Трехмерная визуализация может быть сделана более эффективной с помощью виртуальной реальности (VR). Виртуальная среда (VE) представляет собой трехмерную среду, характеризующуюся тем, что она интерактивна, иллюстративна и интуитивно понятна. Окружающая среда имеет большое значение в интеллектуальном анализе данных, поскольку в традиционной визуализации человеческий субъект смотрит на данные извне, в то время как в среде виртуальной реальности пользователь является частью мира данных. Это означает, что пользователь может использовать все свои чувства, чтобы ориентироваться и понимать данные. Это также подразумевает, что виртуальное представление является более интуитивным. VR особенно хорошо приспособлен для представления масштаба и топологий различных наборов данных. Это становится еще более очевидным, когда используется стерео-визуализация, так как стереозрение позволяет аналитику иметь реальное восприятие глубины. Эта глубина восприятия важна для оценки относительных расстояний и масштабов между глифами. Такая оценка может быть затруднена без стереозрения, если сцена не соответствует парадигме, которую наш мозг привык обрабатывать. В некоторых случаях восприятие глубины может быть улучшено с помощью метафор.

Совместные виртуальные среды (CVE) можно рассматривать как серьезный прорыв в анализе данных для извлечения знаний. По аналогии их можно рассматривать как эквивалент совместных агентов в визуализации. Традиционно один или несколько аналитиков выполняют визуализацию на уникальном ресурсе. Эта операционная модель отражает тот

факт, что многие предприятия распределены по всему миру так же, как их данные и специалисты по анализу этих данных. Следовательно, эти предприятия не могут централизовать все свои операции интеллектуального анализа данных в одном центре. Они не только должны сотрудничать в процессе интеллектуального анализа данных, который в определенной степени может выполняться автоматически совместными агентами, но они также должны сотрудничать в области аспектов визуализации и визуального анализа данных.

ЗАКЛЮЧЕНИЕ

Способность визуализировать результаты интеллектуального анализа данных помогает пользователю понять и доверять знаниям, заложенным в них. Визуализация данных и информации дает пользователю возможность получить интуитивно понятное представление о данных и результатах, например, в форме правил. Эта способность может быть плодотворно использована во многих сферах бизнеса, например, для выявления мошенничества, диагностики в медицинских системах, построения сложных доменных моделей и при проверке кредитоспособности.

Виртуальная реальность и совместные виртуальные среды открывают новые возможности для сбора данных. VR идеально приспособлена для анализа буквенно-цифровых данных и отображения их на бесконечном количестве измерений. Совместные виртуальные среды обеспечивают основу для совместного и распределенного анализа данных путем проведения иммерсивного и синергетического анализа данных. Использование виртуальной реальности для эффективного управления и визуализации мультимедийных данных открывает новые интересные направления исследований.

ЛИТЕРАТУРА

1. Data Mining Community Top Resource. URL: <http://www.kdnuggets.com/> (дата обращения: 03.06.2019).
2. *Deng H., Runger G., Tuv E.* Bias of importance measures for multi-valued attributes and solutions // *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN)*. 2011. P. 293–300. 13.
3. *Feindt M.A.* Neural Bayesian Estimator for Conditional Probability Densities, 2004. URL: <http://arxiv.org/abs/physics/0402093> (дата обращения: 03.06.2019).
4. *Кравченко Ю.А., Лежебоков А.А., Запорожец Д.Ю.* Способы интеллектуального анализа данных в сложных системах // *Известия Кабардино-Балкарского научного центра РАН*. 2012. № 3 (47). С. 52-57.
5. *Кравченко Ю.А.* Интегрированные интеллектуальные обучающие системы управления знаниями // *Информатика, вычислительная техника и инженерное образование*. 2012. № 3 (10). С. 18-24.
6. *Толстых А.М., Шкаленко Б.И., Кравченко Ю.А., Нагоева О.В.* Перспективы развития методов интеллектуального анализа данных на основе модели виртуальной реальности // *Известия Кабардино-Балкарского научного центра РАН*. 2018. № 6-3 (86). С. 146-151.
7. *Бова В.В., Лежебоков А.А.* Разработка когнитивной архитектуры bdi интеллектуального агента // *Известия Кабардино-Балкарского научного центра РАН*. 2017. № 6-2 (80). С. 33-42.
8. *Лежебоков А.А., Ивакин В.П., Шкаленко Б.И.* Развитие средств взаимодействия и управления объектами виртуальной реальности // *Виртуальная и дополненная реальность-2016: состояние и перспективы. Материалы конференции*. 2016. С. 167-171.

9. Лежебоков А.А., Шкаленко Б.И., Ивакин В.П. Дополненная реальность в проектировании // *Виртуальная и дополненная реальность-2016: состояние и перспективы*. Материалы конференции. 2016. С. 359-362.

10. Лежебоков А.А. Разработка архитектуры и алгоритма работы модуля визуализации результатов экспериментальных исследований и научных данных // *Информатика, вычислительная техника и инженерное образование*. 2016. № 4 (28). С. 43-48.

REFERENCES

1. Data Mining Community Top Resource. URL: <http://www.kdnuggets.com/> (date of contact: 06/03/2019).

2. Deng H., Runger G., Tuv E. Bias of importance measures for multi-valued attributes and solutions // *Proceedings of the 21st International Conference on Artificial Neural Networks (ICANN)*. 2011.P. 293-300. 13.

3. Feindt M.A. Neural Bayesian Estimator for Conditional Probability Densities, 2004. URL: <http://arxiv.org/abs/physics/0402093> (accessed June 3, 2019).

4. Kravchenko Yu.A., Lezhebokov A.A., Zaporozhets D.Yu. *Sposoby intellektual'nogo analiza dannykh v slozhnykh sistemakh* [Methods of data mining in complex systems] // *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [News of the Kabardin-Balkar Scientific Center of the Russian Academy of Sciences]. 2012. No. 3 (47). S. 52-57.

5. Kravchenko Yu.A. *Integrirovannyye intellektual'nyye obuchayushchiye sistemy upravleniya znaniyami* [Integrated intelligent learning management systems] // *Informatika, vychislitel'naya tekhnika i inzhenernoye obrazovaniye* [Computer Science, Computing, and Engineering Education]. 2012. No. 3 (10). S. 18-24.

6. Tolstykh A.M., Shkalenko B.I., Kravchenko Yu.A., Nagoeva O.V. *Perspektivy razvitiya metodov intellektual'nogo analiza dannykh na osnove modeli virtual'noy real'nosti* [Prospects for the development of data mining methods based on the virtual reality model] // *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [News of the Kabardin-Balkar Scientific Center of the Russian Academy of Sciences]. 2018. No. 6-3 (86). S. 146-151.

7. Bova V.V., Lezhebokov A.A. *Razrabotka kognitivnoy arkhitektury bdi intellektual'nogo agenta* [Development of bdi cognitive architecture of intellectual agent] // *Izvestiya Kabardino-Balkarskogo nauchnogo tsentra RAN* [News of the Kabardin-Balkar Scientific Center of the Russian Academy of Sciences]. 2017. No. 6-2 (80). S. 33-42.

8. Lezhebokov A.A., Ivakin V.P., Shkalenko B.I. *Razvitiye sredstv vzaimodeystviya i upravleniya ob'yektami virtual'noy real'nosti* [Development of means of interaction and management of virtual reality objects] // *Virtual'naya i dopolnennaya real'nost'-2016: sostoyaniye i perspektivy. Materialy konferentsii* [Virtual and Augmented Reality-2016: status and prospects. Conference proceedings]. 2016. S. 167-171.

9. Lezhebokov A.A., Shkalenko B.I., Ivakin V.P. *Dopolnennaya real'nost' v proyektirovani* [Augmented reality in designing] // *Virtual'naya i dopolnennaya real'nost'-2016: sostoyaniye i perspektivy. Materialy konferentsii* [Virtual and augmented reality-2016: state and prospects. Conference proceedings]. 2016. S. 359-362.

10. Lezhebokov A.A. *Razrabotka arkhitektury i algoritma raboty modulya vizualizatsii rezul'tatov eksperimental'nykh issledovaniy i nauchnykh dannykh* [Development of the architecture and algorithm of the module for visualizing the results of experimental studies and scientific data] // *Informatika, vychislitel'naya tekhnika i inzhenernoye obrazovaniye* [Computer Science, Computing and Engineering Education]. 2016. No. 4 (28). S. 43-48.

VISUALIZATION TECHNOLOGIES FOR APPLIED PROBLEMS OF INTELLECTUAL DATA ANALYSIS

A.A. LEZHEBOKOV, E.V. KULIEV

Southern Federal University
of the Russian Academy of Sciences
347928, Rostov region, Taganrog, Nekrasovsky lane, 44
E-mail: info@ictis.sfedu.ru

Modern methods of data and knowledge visualization are studied in the article. Classical approaches to visualization lose relevance in connection with the transition to cross-platform technical solutions and the need to visualize large data sets, including those belonging to the “big data” categories. Intelligent analytical data such as classification, clustering, search for dependencies and patterns in data, construction of multidimensional cubes, which are undoubtedly necessary, and the quality of solutions are directly related to the technological and visible visualization used. The paper notes the need for timely and rational use of such technologies as virtual and augmented reality, which made it possible to make a jump in the development of technical visualization tools, including multidimensional data. The presence of enhanced interactive capabilities and tools for data manipulation allow us to conclude about the emergence of a new stage in the development of visualization technologies.

Keywords: *visualization, data mining, application, large data sets, classification.*

Работа поступила 15.07.2019 г.