

РЕЗУЛЬТАТИВНОСТЬ ВИЗУАЛЬНОГО АНАЛИЗА В ЗАДАЧАХ ПРИНЯТИЯ РЕШЕНИЙ

Шкляр Алексей Викторович

Институт кибернетики Томского политехнического университета, Россия

634034, г. Томск, ул. Советская 84/3, e-mail: shklyarav@tpu.ru

Аннотация. При изучении результатов экспериментальных измерений, относящихся к большинству областей научных исследований, одной из проблем, обладающих возрастающей актуальностью, является создание средств анализа, применимых для изучения больших объемов исходных данных. Существующие подходы к анализу экспериментальных данных обладают рядом недостатков, к которым следует отнести значительную ресурсоемкость и высокие требования к подготовке специалистов, участвующих в анализе. В работе показаны некоторые результаты решения задачи создания средств визуального анализа экспериментальных данных, направленного на увеличение результативности исследования эмпирических данных.

Ключевые слова: визуальная аналитика, результативность визуального анализа, восприятие, комплексная характеристика визуализации.

Цитирование: Шкляр А.В. Результативность визуального анализа в задачах принятия решений // Информационные и математические технологии в науке и управлении. 2018. № 3 (11). С. 147–155. DOI:10.25729/2413-0133-2018-3-16

1. Визуальный анализ. Современные средства получения цифровых данных усложняют задачу обработки эмпирических данных, дополняя ее, в том числе, необходимостью исследования значительных объемов изменяющихся данных. Широкое прикладное значение имеет решение задач, связанных с обработкой данных и принятием управляющих решений в режиме реального времени. Одним из перспективных направлений в разработке инструментов оперативного анализа данных следует считать создание средств визуальной аналитики, целенаправленно использующих преимущества визуального восприятия.

В случае традиционного подхода к привлечению возможностей визуализации, задача анализа больших объемов экспериментальных данных может возникать лишь в исключительных случаях, отличающихся высоким уровнем формализованности и опытом исследования аналогичных вопросов. В качестве возможного направления развития визуальной аналитики выступает создание средств визуализации, обладающих свойством представления наблюдателю значений динамических переменных [1]. В данном случае, под динамическими понимаются не только информативные характеристики, содержащие сведения, зависящие от времени, но и величины, интерпретация смысла которых возможна при сопоставлении результатов различных измерений. Существующие подходы к созданию средств визуальной аналитики практически не используют особенности визуального восприятия изменяющихся зрительных образов, которое существенно отличается от

наблюдения за статичными изображениями, и может стать основой увеличения результативности решений задач анализа больших объемов данных.

Практическое исследование экспериментальных данных, во многих случаях, связано с необходимостью соответствия средств визуальной аналитики требованию масштабируемости. Основанием для этого является распространенная необходимость переформулирования уже существующей задачи исследования данных в частный вариант изучения лишь фрагмента данных или, что еще сложнее, в вариант задачи с накоплением данных. Применение визуализации в качестве аналитического инструмента предлагает естественное решение этой проблемы, например, благодаря целенаправленному использованию особенностей визуального восприятия, обеспечивающего одновременное целостное восприятие сведений, относящихся к разным объектам, но объединенным в поле зрения в обобщенный информативный образ [6].

2. Результативность в визуальном анализе. Сравнение возможностей визуальной аналитики с другими подходами подразумевает существование количественной меры, позволяющей делать выводы в пользу одного из сопоставляемых вариантов. В некоторых исследованиях возможностей визуализации определены различные подходы к измерению сложности производимых исследователем действий или трудоемкости совершаемых им операций [7], но применительно к визуальной аналитике и ее практическому использованию эти вопросы остаются открытыми. В роли простейшей численной оценки может выступать продолжительность процедур, выполняемых при разработке и использовании средств визуального анализа. Однако, препятствием, затрудняющим развитие этого подхода, является необходимость в высоком уровне формализации процесса визуального анализа данных.

Опираясь на общие положения теории сложности, возможно определение характеристики визуальной аналитики, аналогичной определению сложности алгоритмов. С этой точки зрения, сложность процесса визуального анализа определяется объемом ресурсов, необходимым для решения задачи визуальной аналитики. На этом основании, с учетом разнообразия привлекаемых к анализу данных усилий, особое значение приобретает вопрос о значимости различных видов используемых ресурсов, а также об их взаимозаменяемости [8]. Наряду с основными видами ресурсов — временными и вычислительными, для определения ресурсоемкости визуализации становится актуальным учет усилий, совершаемых человеком или группой людей для достижения цели визуального исследования, в том числе интеллектуальных, физических или эмоциональных.

3. Визуальная модель данных. Определим модель M данных D (визуальная модель данных) в визуальной аналитике как объект зрительного восприятия (образ), сопоставленный этим данным по предварительно определенному правилу F , т.е. $M = F(D)$. Соответственно, правило F , сформулированное на основании учета ряда дополнительных факторов, позволяющее определить свойства визуального образа данных, определим как функцию визуального представления. Определенная таким образом функция F является средством изменения результативности визуального анализа и должна определяться на основании необходимости увеличения общей результативности визуального исследования.

Выбор функции визуального представления зависит от свойств исследуемых данных D , допустимого уровня сложности T_{max} алгоритма решения (ограничение максимального времени решения) и характеристик исследователя $U = \{U_1 .. U_n\}$, существенных для

достижения исследователем необходимого ему уровня понимания изучаемых данных. Основным вопросом визуальной аналитики следует считать нахождение функции представления F , обеспечивающей визуальную модель M свойствами, достаточными для решения пользователем задачи исследования.

4. Семантическое взаимодействие. Увеличение размера входных данных является характерной чертой многих практических приложений визуальной аналитики. Усложнение моделей данных, происходящее как отклик на рост сложности задачи, имеет естественный предел, достижение которого делает бесполезными разработанные ранее средства визуальной аналитики. Перспективным направлением может стать организация совместного использования когнитивных и вычислительных ресурсов. Вариантом такого объединения может стать взаимное дополнение визуальной аналитики и методов машинного обучения [4].

Преимуществом систем анализа данных, ориентированных на реализацию семантического взаимодействия, является адаптация функции визуального представления данных к ментальным аналитическим конструкциям исследователя. Это позволяет объединить метафоры визуализации, управления и известные пользователю ассоциативные метафоры, создавая условия для снижения сложности визуальной модели. Развитие визуальной аналитики в этом направлении делает необходимым существование представления о закономерностях функционирования человеческого мышления. В данном случае, утверждается необходимость отхода от модели познания как модели «черного ящика», на вход которого поступают перцептивные данные, а на выходе формируются знания, интегрируемые в существующую систему понятий человека [2].

5. Принципы восприятия. Восприятие является процессом, обеспечивающим индивидуальную коммуникацию с внешними источниками информации. Кроме того, восприятие может быть определено как результат непосредственного воздействия предметов или явлений на органы чувств [10], выступающие в качестве анализаторов внешних источников информации. Среди существующих типов анализаторов, отвечающих за зрение, слух, осязание, обоняние, вкус и кинестезию, визуальный канал является наиболее информативным и нагруженным.

Образы внешних объектов, появившиеся в результате восприятия, являются субъективной формой, включающей, помимо данных, поступивших от анализаторов, аспекты индивидуального отношения к объекту: интерес, мотивацию, эмоциональную оценку [5]. Возможное влияние субъективных факторов на образ внешнего объекта приводит к изменчивости этого образа во времени или принципиально новому результату переоценки. Таким образом, анализ визуальной информации, целью которого является понимание ее смысла, зависит как от объективных характеристик визуального образа, так и от персональных качеств наблюдателя.

Визуальное восприятие, являясь одним из видов восприятия в целом, обладает рядом общих характеристик. К числу существенных для обоснования возможностей визуального анализа следует отнести обобщенность, предметность, целостность и константность восприятия. Обоснованное управление свойствами создаваемых визуальных образов, учитывающее эти характеристики, позволяет контролировать интерпретируемость визуализации. Таким образом, обоснованное использование субъективных аспектов восприятия создает условия для увеличения результативности визуального исследования.

6. Классификация задач анализа. В результате обобщения определений визуальной модели, следствий использования подходов к визуальному анализу, учитывающих семантическое взаимодействие и особенности визуального восприятия, формируется структура визуального анализа. Основой предлагаемой структуры является элементарный объект — структурная единица визуального анализа, а также соответствующая ему иконическая модель структурной единицы. Целью введения лаконичной формы представления структуры процесса визуализации является оперативное определение требований к визуальной модели исследуемых данных, выполнение которых позволит оценивать перспективность использования средств визуальной аналитики для решения конкретной задачи. Сокращение времени, затрачиваемого на этапе предварительного анализа задачи, увеличивает общую результативность средств визуальной аналитики.

Структурирование процессов визуального анализа позволяет установить соответствие между свойствами задач исследования и особенностями визуальных моделей, обладающих наилучшей результативностью при решении. Таким образом, становится возможным определение классификации задач анализа на основании их внешних признаков, к которым относятся свойства исследуемых данных, постановка цели исследования и некоторые дополнительные характеристики.

Появление классификации задач анализа позволит достичь следующих результатов:

- Быстрое принятие решения об особенностях средств визуальной аналитики, позволяющих достичь необходимых результатов.
- Накопление опыта применения средств визуального анализа и его систематизация, дополняющая классификацию.
- Прогностический анализ свойств визуальных моделей на основании сопоставления со свойствами существующих моделей и их применимости.
- Оценка результатов комбинирования структурных единиц, в том числе, различного вида, при создании комплексных решений.

7. Задачи принятия решения. В случае использования визуальных моделей в качестве средства, предназначенного для упрощения или ускорения процесса получения экспертного мнения, задача анализа имеет следующие особенности: необходимо принятие решения после изучения всей доступной информации, время получения решения имеет большое значение и должно быть минимальным, существует несколько возможных ответов и выбор одного из них не может быть сделан на основании формальной логики. Определяющим условием для задач этого типа становится активное использование знаний и опыта исследователя для получения новых сведений, оказывающих влияние на дальнейший ход решения. Оценка таких дополнительных сведений и построение гипотез на их основе становятся частью процесса решения и могут быть улучшены с помощью визуальных моделей решения [11].

Информационный вход визуальной модели принятия решений (IN) используется для приема сведений, необходимых для построения визуального образа. Одновременно с исследуемыми данными, на управляющий вход (IN2) поступает запрос, связанный с целью исследования, который определяет свойства визуального представления на основании цели исследования, свойств данных и субъективного опыта решений аналогичных задач. Основанием для такого изменения, по сравнению с другими типами задач, является активное

использование готовых ментальных моделей пользователя, совпадение с которыми результатов восприятия становится основанием для быстрого выбора и верификации гипотезы решения. Это означает, что воспринимаемые компоненты модели становятся объектами, инициирующими выбор гипотез ответа, и потому определение свойств визуального представления имеет принципиальное значение (Рис. 1 **Ошибка! Источник ссылки не найден.**).

Причиной значимости управляющих связей в задачах принятия решения, помимо согласования формы вопроса и модели, становится необходимость уточнения исследуемого образа в ходе решения задачи анализа. Основанием для этого становится появление дополнительных данных, являющейся результатом проверки промежуточных гипотез, возникающих в результате взаимодействия исследователя и визуальной модели. Следует отметить, что для задач принятия решения распространенной ситуацией является одновременное коллективное использование визуальной модели группой специалистов. В этом случае, определение функции визуального представления сталкивается с необходимостью уточнения конвенциональных параметров визуальной знаковой системы.



Рис. 1. Иконическая модель задачи принятия визуального решения

8. Время визуального исследования. Определим величину K_A как объем новой информации, получаемой исследователем при поиске и формулировании ответа на поставленный перед ним вопрос анализа. Информация является результатом взаимодействия пользователя с визуальной моделью данных, т.е. сообщением, полученным наблюдателем за ограниченный промежуток времени T_A .

$$K_A = \int_0^{T_A} P(I, K, t) dt = K_A t ,$$

где $P(I, K, t)$ — функция визуального восприятия, K — объем знаний пользователя.

Значение $K_A(T_A)$ может быть определено как когнитивная (информативная) ценность процесса визуального анализа и, соответственно, T_A — время решения, т.е. интервал времени, затрачиваемого наблюдателем для получения решения задачи анализа. Величина K^* , соответственно, является объемом информации, достаточным для достижения цели анализа, и может быть определена как сложность задачи – собственная характеристика задачи анализа. Таким образом, может быть определена обратная зависимость для времени решения T_A :

$$T_A = T_A(K^*, P, K_0, S),$$

связывающая время решения T_A со сложностью задачи K^* , начальным объемом знаний пользователя K_0 , процессом интерпретации $P(I, K)$, а также функцией визуального представления $S = S(t)$. Таким образом, количественной мерой, позволяющей оценивать возможности визуальной модели, действительно, способно выступать время анализа T_A , зависящее от задачи, для решения которой будет использована визуальная модель, а также от возможностей конкретного пользователя.

Количественная оценка объема информации в принятом сообщении, если оно является даже обычным текстом, представляет собой задачу, к решению которой можно подходить по-разному. Необходимо принимать во внимание тот факт, что последовательность измеряемых единиц информации, составляющих сообщение, не определяет его общую информационную ценность, т.к. для получателя сообщения эта величина имеет субъективный характер [3]. Следовательно, определение времени анализа в качестве меры результативности визуализации позволяет обойти проблему измерения информативной ценности визуализации при выборе функции визуального представления.

9. Экспериментальная оценка результативности визуального анализа. Полагая, на основании введенных определений, результативность процесса визуального исследования величиной, зависящей от времени, затраченного пользователем на достижение цели анализа, основной задачей для разработчиков средств визуального исследования данных следует считать сокращение времени анализа [9]. Таким образом, для оценки результативности средств визуальной аналитики необходимо иметь представление о зависимости времени исследования от любых факторов, оказывающих на него существенное влияние, на каждом этапе решения задачи анализа.

Для получения данных, позволяющих оценить достоинства конкретной функции визуального представления при решении соответствующей задачи анализа, разработана методика проведения тестовых решений. Предлагаемая методика состоит в проведении серий решений тестовых задач, характеризуемых рядом контролируемых ограничений, и измерении интервалов времени, соответствующих этапам взаимодействия исследователя и визуальной модели. Взаимодействие пользователя с моделью, в этом случае, подразумевает любые доступные пользователю операции, за исключением изменения функции визуального представления и набора визуализированных данных (рис. 2.).

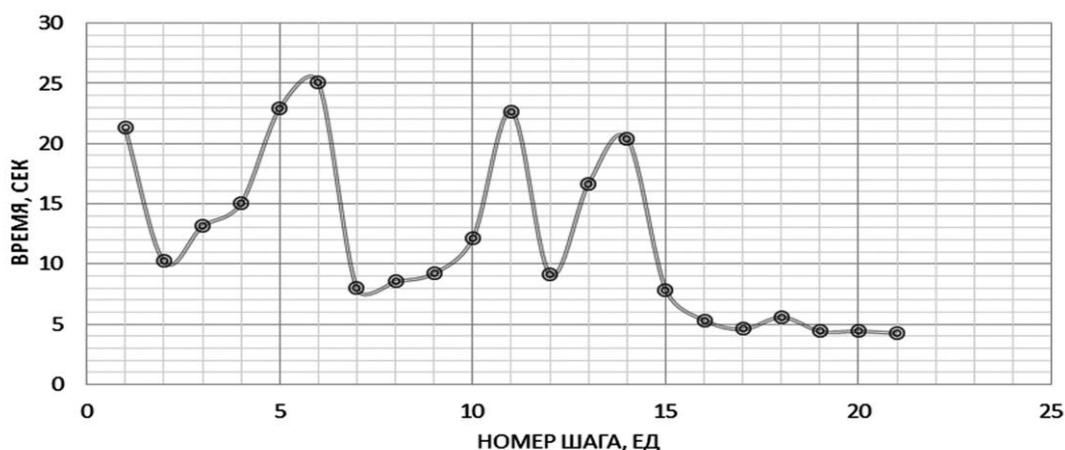


Рис. 2. Периодическое изменение скорости визуального анализа разных пользователей

Тестовое решение предполагает создание образа исследуемых данных с использованием predetermined функции визуального представления. Исследователю позволяет формулировать неограниченное число гипотез ответа на вопрос задачи, каждая гипотеза считается очередным шагом анализа. Формулирование верной гипотезы означает завершение решения задачи. Измерения интервалов времени, затрачиваемых исследователем на формирование каждой новой гипотезы, показали наличие стадий замедления и ускорения в ходе решения задачи анализа для большинства тестовых измерений. Циклическое повторение этих стадий может свидетельствовать о преобладании прямых действий над обратными и последующем изменении приоритета в противоположном направлении.

Этапы взаимодействия исследователя и визуальной модели, на протяжении которых время, затрачиваемое на формирование очередной гипотезы решения, постоянно сокращается, могут быть интерпретированы как интервалы «обучения». Основанием для этого является сокращение времени принятия нового решения при условии ошибочности предыдущей гипотезы. Таким образом, происходит накопление новых сведений, являющихся следствием верификации предлагаемых пользователем гипотез. Изменение уровня информированности K пользователя приводит к сокращению времени построения новой гипотезы, однако, после этапа накопления новых сведений, как показали результаты тестовых исследований, наступает этап переосмысления, характеризуемый существенным замедлением процесса построения новой гипотезы (рис. 3.).

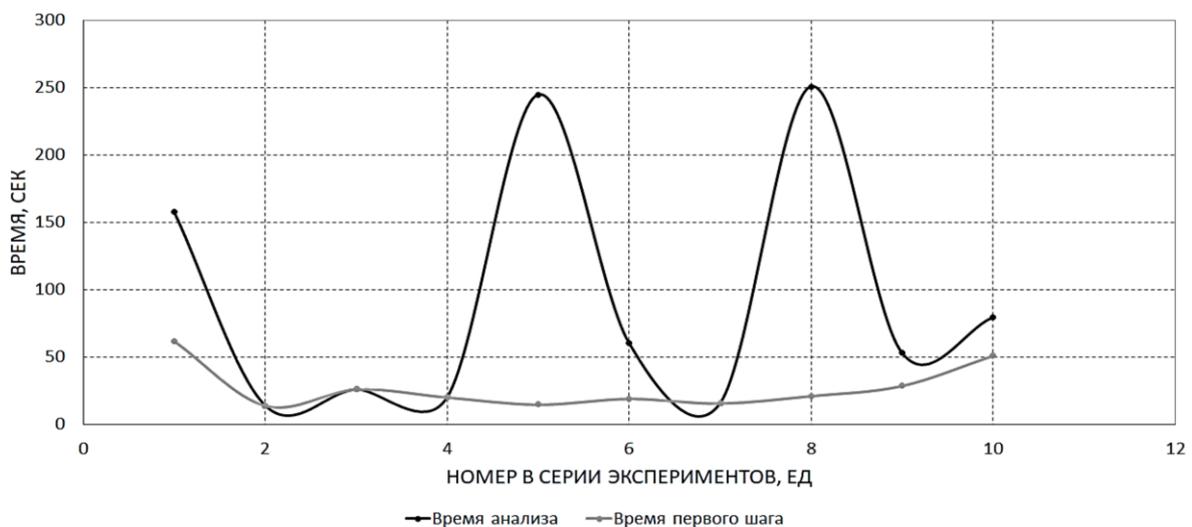


Рис. 3. Изменение времени принятия решения для различных функций представления

Особое значение в этом случае приобретает первый этап, характеризующийся увеличением скорости мышления пользователя, присутствующий в большинстве результатов тестовых измерений. В отличие от последующих аналогичных этапов, можно утверждать, что на протяжении первого этапа происходит ознакомление исследователя с образом данных и особенностями функции визуального представления. В пользу этого утверждения говорит сокращение этого этапа при повторном использовании исследователем той же функции визуального представления. Таким образом, полученные экспериментальные данные могут быть использованы для определения оценочных значений интервалов обучения и интерпретации информации для соответствующей функции визуального представления.

Заключение. В работе предложен способ определения результативности использования средств визуальной аналитики, позволяющий разрабатывать инструменты анализа экспериментальных данных для задач, характеризующихся наличием большого объема исследуемых данных, а также в случае изменений или накопления исследуемых сведений. На основании показанной связи между результативностью визуального исследования и временем, позволяющей делать качественные оценки применимости средств визуальной аналитики, проведены экспериментальные исследования, выявившие ряд особенностей взаимодействия пользователя и визуальной модели данных. Целенаправленное использование понимания особенностей визуального восприятия и структурированный подход к визуальному анализу делают перспективным использование средств визуальной аналитики при исследовании больших объемов эмпирических данных.

Работа выполнена при поддержке гранта Министерства образования и науки РФ № 2.1642.2017/4.6.

СПИСОК ЛИТЕРАТУРЫ

1. Шкляр А.В., Захарова А.А. Динамические знаковые структуры в информативных визуальных высказываниях // Труды 26-ой Международной конференции (GraphiCon2016). Россия, Нижний Новгород. 19–23 сентября 2016 г. Москва. ИПМ РАН. 2016. С. 131–136.
2. Chen C.C.C. An Information-Theoretic View of Visual Analytics // IEEE Comput. Graph. Appl. 2008. Т. 28. № February. Pp. 18–23.
3. Cui Q., Ward M., Rundensteiner E., Yang J. Measuring Data Abstraction Quality in Multiresolution Visualizations // IEEE Transactions on Visualization & Computer Graphics. 2007. vol. 12. no. 5. Pp. 709-716.
4. Endert A, North C., Chang R., Zhou M. Toward Usable Interactive Analytics: Coupling Cognition and Computation // ACM KDD 2014 Workshop on Interactive Data Exploration and Analytics (IDEA). 2014.
5. Mazza R. Introduction to information visualization. Springer. 2009. 139 p.
6. North C. Toward measuring visualization insight // IEEE Comput. Graph. Appl. 2006. Т. 26. № 3. Pp. 6–9.
7. Pirolli P., Card S. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis // Proc. Int. Conf. Intell. Anal. 2005. Т. 2005. Pp. 2–4.
8. Sacha D., Stoffel A., Stoffel F., Kwon B., Ellis G., Keim D. Knowledge Generation Model for Visual Analytics // Vis. Comput. Graph. IEEE Trans. 2014. Т. 20. № 12. Pp. 1604–1613.
9. Shklyar A., Zakharova A., Vekhter E., Pak A. Visual modeling in an analysis of multidimensional data // Journal of Physics: Conference Series. 2018. Т. 5. № 1. P. 125–128.
10. Wijk J.J. Van. The value of visualization // Visualization. 2005. Pp.79–86.
11. Zakharova A.A., Vekhter E. V., Shklyar A. V. Methods of solving problems of data analysis using analytical visual models // Scientific Visualization. 2017. Т. 9. № 4. Pp. 78–88.

EFFICIENCY OF VISUAL ANALYSIS IN DECISION-MAKING PROBLEMS

Aleksey V. Shklyar

Senior teacher, Department of automation and robotics

Tomsk Polytechnic University

84/3, Sovetskaya Str., 634034, Tomsk, Russia, e-mail: shklyarav@tpu.ru

Abstract. While studying results of experimental measurements related to most of areas of scientific researches, one of the problems with high relevance was creation of analysis tools that are applicable for studying large amount of initial data. Existing approaches to the analysis of experimental data have a number of disadvantages, which include significant resource capacity and high training requirements for specialists involved in the analysis. In this article some results of solving the problem of creating visual analysis tools for experimental data, aimed at increasing the effectiveness of empirical data research, have been showed.

Keywords: visual analytics, effectiveness of visual analysis, perception, complex visualization characteristic.

References

1. Shklyar A., Zakharova A.A. Dinamicheskie znakovye struktury v informativnyh vizual'nyh vyskazyvaniyah [Dynamic sign structures in informative visual statements.]. Protvino: Izdatel'stvo: Avtonomnaya nekommercheskaya organizaciya «Institut fiziko-tehnicheskoy informatiki» = Protvino: publishing House: Autonomous non-profit organization " Institute of physical and technical Informatics», 2016. Pp. 131–136. (in Russian).
2. Chen C.C.C. An Information-Theoretic View of Visual Analytics // IEEE Comput. Graph. Appl. 2008. T. 28. № February. Pp. 18–23.
3. Cui Q., Ward M., Rundensteiner E., Yang J. Measuring Data Abstraction Quality in Multiresolution Visualizations // IEEE Transactions on Visualization & Computer Graphics. 2007. vol. 12. no. 5. Pp. 709-716.
4. Endert A, North C., Chang R., Zhou M. Toward Usable Interactive Analytics: Coupling Cognition and Computation // ACM KDD 2014 Workshop on Interactive Data Exploration and Analytics (IDEA). 2014.
5. Mazza R. Introduction to information visualization. Springer. 2009. 139 p.
6. North C. Toward measuring visualization insight // IEEE Comput. Graph. Appl. 2006. T. 26. № 3. Pp. 6–9.
7. Pirolli P., Card S. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis // Proc. Int. Conf. Intell. Anal. 2005. T. 2005. Pp. 2–4.
8. Sacha D., Stoffel A., Stoffel F., Kwon B., Ellis G., Keim D. Knowledge Generation Model for Visual Analytics // Vis. Comput. Graph. IEEE Trans. 2014. T. 20. № 12. Pp. 1604-1613.
9. Shklyar A., Zakharova A., Vekhter E., Pak A. Visual modeling in an analysis of multidimensional data // Journal of Physics: Conference Series. 2018. T. 5. № 1. Pp. 125-128.
10. Wijk J.J. Van. The value of visualization // Visualization. 2005. P.79-86.
11. Zakharova A.A., Vekhter E. V., Shklyar A. V. Methods of solving problems of data analysis using analytical visual models // Scientific Visualization. 2017. T. 9. № 4. Pp. 78–88.