

УДК 004.9

Повышение эффективности анализа и визуализации данных на основе графовых моделей

Н. А. Бездетный, С. А. Зори

Донецкий национальный технический университет,
nekooolay@mail.ru, SPIN-код: 2472-1006

ik.ivt.rec@mail.ru, OrcID: 0000-0003-4018-234X, SPIN-код: 3565-6330

Аннотация

В статье рассматривается проблема повышения эффективности анализа и визуализации научных данных, обозначена актуальность исследования и разработки новых эффективных методов представления, анализа и визуализации данных в условиях возрастающих объемов информации. В качестве перспективного направления решения проблемы предлагается использование графовых моделей данных. Сформулирована цель исследования, определены основные задачи работы, направленной на развитие методов визуальной аналитики и повышение эффективности анализа данных различных прикладных областей.

Введение

Развитие информационных технологий привело к экспоненциальному росту объемов информации в современном обществе. Огромные массивы данных накапливаются во всех сферах человеческой деятельности - науке, экономике, производстве, социальных системах. Бурный рост информационных потоков поставил перед обществом ключевую задачу - разработать эффективные способы обработки и извлечения полезных знаний из огромных объемов накопленных данных [1].

Одним из наиболее перспективных направлений анализа данных является визуализация данных - графическое представление информации, позволяющее за счет особенностей человеческого восприятия выявлять скрытые закономерности и новые знания.

Применение средств компьютерной графики - области информатики, в которой рассматриваются методы, алгоритмы и технологии визуализации данных - позволяет по новому взглянуть на решение проблемы повышения эффективности анализа и визуализации научных данных для обеспечения реальных потребностей потенциальных пользователей, особенно с учетом роста этих потребностей потребителей и возможностей технических и алгоритмических средств [1].

С этой точки зрения можно выделить следующие актуальные области применения компьютерной графики:

- визуализация информации;
- моделирование процессов и явлений;
- проектирование технических объектов;

- организация пользовательского интерфейса для выше рассмотренных областей.

Таким образом, рассматриваемые в статье перспективные методы анализа и визуализации данных, определяют актуальность работы.

Использование аппарата графов для визуализации данных

Среди методов визуализации данных в последние годы все большее значение приобретают подходы, использующие аппарат теории графов [2]. Графы как математические объекты идеально подходят для моделирования связей и зависимостей в сложных системах. Их визуальное представление в виде диаграмм позволяет структурировать данные, выявлять топологию систем (рис. 1).

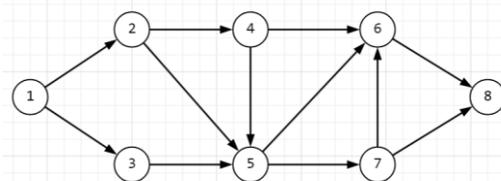


Рисунок 1 – Ориентированный граф

Особенно перспективно использование графовых диаграмм в задачах компьютерного моделирования и анализа процессов в сложных системах различной природы - технических, организационных, социальных и др. [3, 4, 5]. Однако методы визуализации на основе графов применительно к моделированию конкретных процессов исследованы не в полной мере.

Тематика работы описывает исследование методов визуализации с использованием графовых диаграмм, применительно к задачам

моделирования и анализа процессов в различных предметных областях. Работа направлена на расширение арсенала средств визуальной аналитики данных и повышение эффективности компьютерного моделирования.

Визуальная аналитика данных и компьютерное моделирование с графовыми диаграммами являются весьма актуальными направлениями исследований в современном высокотехнологичном мире [6]. Эффективная визуализация помогает справляться с экспоненциальным ростом объемов информации, выявлять скрытые закономерности и модели в сложных системах (рис. 2).



Рисунок 2 – Визуальная аналитика с применением графа

Разработка новых методов визуализации данных на основе графов, применительно к различным предметным областям позволит расширить аналитические возможности исследователей, инженеров, бизнес-аналитиков. Графовые диаграммы наглядно отражают структуру связей в системе, облегчают понимание процессов и причинно-следственных зависимостей. Комплексный подход к визуализации и моделированию на основе графов открывает новые горизонты для решения сложных научных и прикладных задач в IT, экономике, медицине, производстве и других областях.

Таким образом, исследования в рамках данной работы, связанные с усовершенствованием и разработкой новых

методов визуализации на основе графовых диаграмм, и применением современных технологий ее реализации, вносят ценный вклад в развитие перспективных инструментов интеллектуального анализа данных.

Анализ методов и технологий визуализации на основе графовых моделей

Существует множество подходов и инструментов для визуализации графов, которые различаются функциональностью, масштабируемостью, способами интерактивного взаимодействия.

В процессе анализа технологий и средств для визуализации данных на основе графовых моделей были рассмотрены такие решения, как Gephi, Linkurious, Graphistry и Cytoscape.js.

Gephi – мощный open-source инструмент для визуализации и анализа графовых структур. Позволяет строить графы, манипулировать их структурой, применять алгоритмы компоновки и фильтрации (рис. 3) [7].

Плюсы:

- open-source решение с открытым исходным кодом;
- большой выбор визуальных стилей и настроек;
- plugin архитектура для расширения функционала.

Минусы:

- ограниченные возможности для анализа очень больших графов;
- требует установки ПО на компьютер;
- неудобный интерфейс для новичков.

Linkurious – веб-платформа для интерактивной визуализации и анализа крупных графовых данных (рис.4) [8].

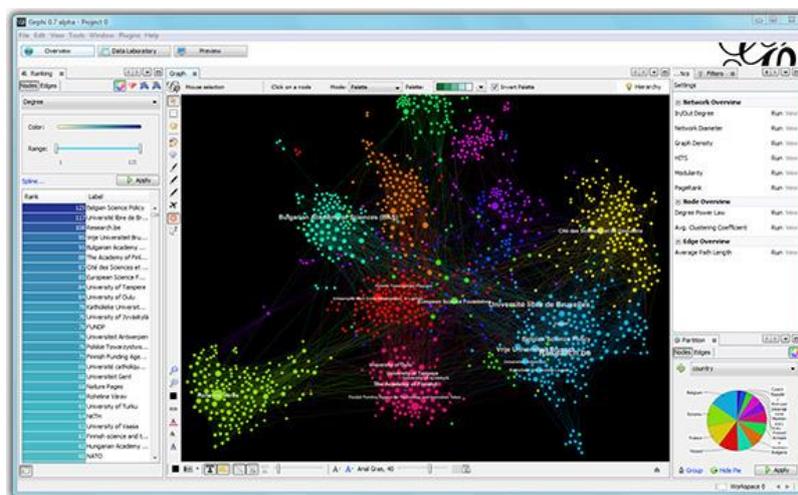


Рисунок 3 – Программное обеспечение Gephi

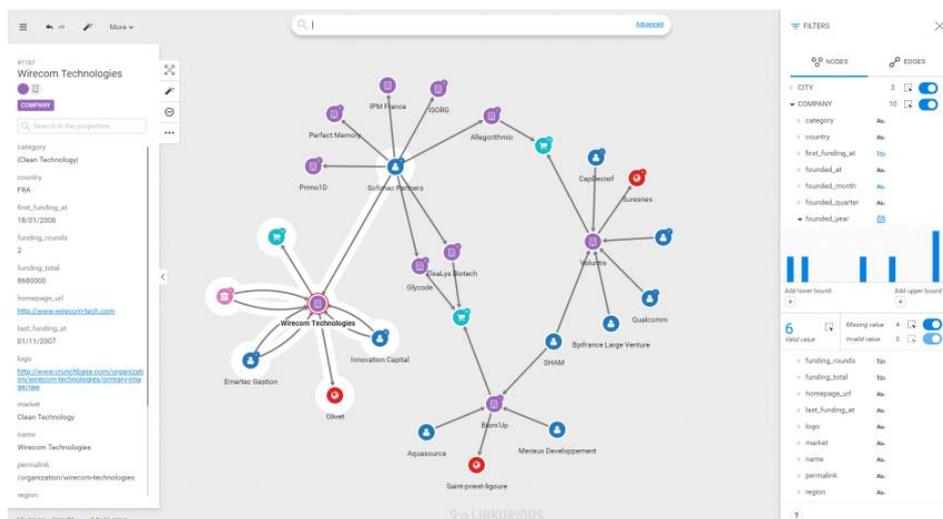


Рисунок 4 – Веб-платформа Linkurious

Плюсы:
– веб интерфейс, не требует установки ПО;
– хорошо масштабируется для анализа больших графов;
– расширенные аналитические инструменты (кластеризация, поиск пути и сообществ и т.д.)

Минусы:
– платная лицензия для использования;
– ограниченные возможности экспорта и сохранения результатов;
– сложности при интеграции с другими системами.

Graphistry – высокопроизводительный инструмент визуализации больших графов в браузере на основе GPU. Позволяет быстро строить изображение графа и манипулировать им (рис. 5) [9].

Плюсы:
– высокопроизводительная визуализация больших графов в браузере с помощью GPU
– плавная анимация перестроения графа
– API для интеграции со сторонними данными

Минусы:
– сложность настройки и развертывания, требует DevOps навыков
– ограниченные аналитические возможности в сравнении со специализированными платформами
– закрытый исходный код

Cytoscape.js - библиотека для построения интерактивных графов в веб-приложениях. Использует HTML5, canvas, SVG и WebGL (рис. 6) [10].

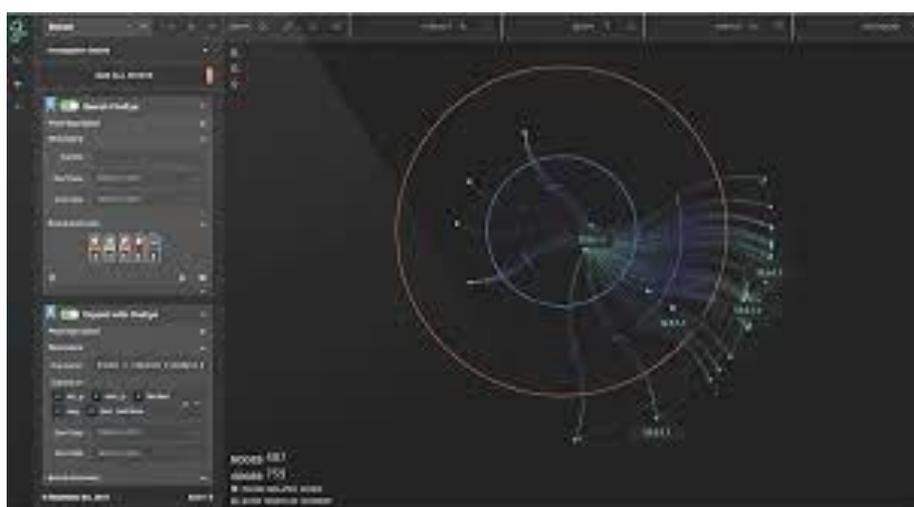


Рисунок 5 – Программное обеспечение Graphistry

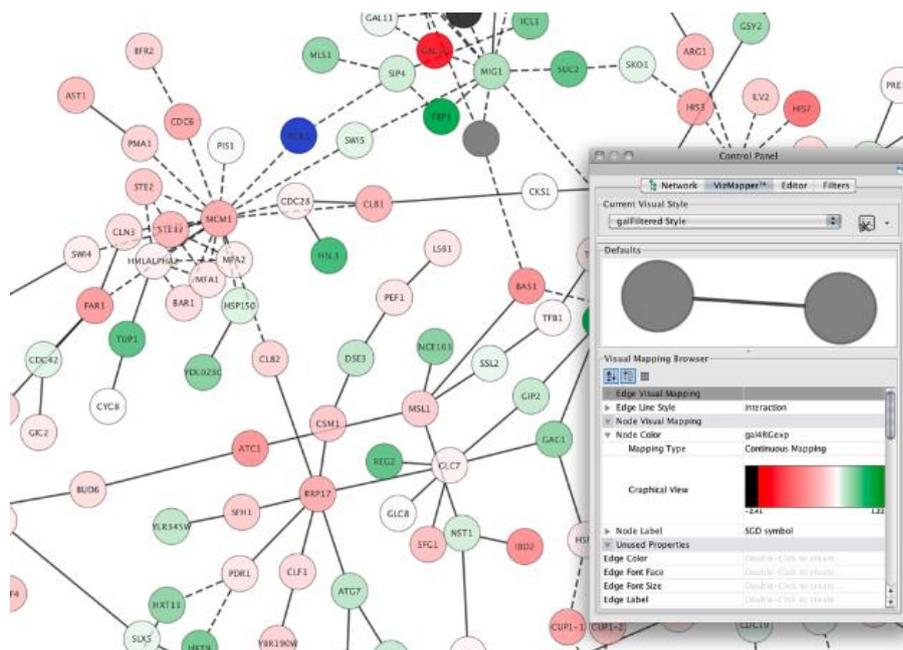


Рисунок 6 – Библиотека Cytoscape.js

Плюсы:
– open-source библиотека для разработчиков
– кроссбраузерная поддержка, использование HTML5 canvas/SVG/WebGL расширяемость и настраиваемость за счет модульной архитектуры, интеграция с внешними библиотеками (D3.js и др.)

Минусы:
– ограниченные встроенные аналитические возможности
– требует навыков JavaScript программирования
отсутствие готового desktop или веб-приложения

Проанализированные инструменты визуализации графовых данных, такие как Gephi, Linkurious, Graphistry и Cytoscape.js, обладают определенными возможностями, но также имеют свои ограничения.

Gephi, несмотря на гибкость визуализации, позволяет анализировать графы ограниченного размера. Linkurious специализируется на анализе крупных сетей, но имеет закрытый исходный код. Graphistry отличается производительностью, но сложен в настройке и интеграции. Cytoscape.js требует навыков программирования и расширения для полноценного анализа.

Таким образом, существует потребность в решениях, которые сочетали бы простоту, гибкость использования и визуальные возможности представления данных в виде графовых диаграмм. При этом проанализированные инструменты обладают различными возможностями и ограничениями.

Это указывает на необходимость комплексного решения, которое объединит их преимущества.

Предлагаемый в рамках данного исследования подход основан на создании гибких динамических моделей графовых структур и разработке интерактивных инструментов их визуализации и анализа. В основе подхода лежит представление данных в виде математических моделей с использованием методов теории графов. Это позволяет структурировать и формализовать графы, а также применять актуальные алгоритмы их анализа.

При создании механизма визуализации будут использованы интерактивные веб-интерфейсы с возможностями динамического отображения и манипулирования графами. Это обеспечит наглядность информации и удобство работы пользователя.

Данный комплексный подход, интегрирующий математическое моделирование и интерактивную визуализацию, позволит эффективно анализировать и представлять графо-ориентированные данные, что является актуальной научно-практической задачей.

Цель и задачи дальнейших исследований, новизна и значение работы

Объект исследования – средства анализа, обработки и визуализации информации в системах визуального компьютерного моделирования.

Предмет исследования – эффективные модели, методы и технологии обработки,

анализа и визуализации данных на основе графов.

Целью исследований является теоретическое обоснование и практическая реализация усовершенствованных эффективных методов и средств для обработки, анализа и визуализации данных и процессов компьютерного моделирования на основе графовых моделей.

Определены дальнейшие задачи исследований и разработок:

1. системный анализ существующих методов, технологий и средств для визуализации результатов компьютерного моделирования данных и процессов;

2. анализ способов их улучшения;

3. разработка эффективных средств для повышения эффективности систем визуального компьютерного моделирования, отличающейся простотой и наглядностью, удобством и функционалом интерфейса, а также возможностью интерактивного взаимодействия с моделью;

4. исследование возможности применения разработанных средств анализа, обработки и визуализации информации на основе графовых моделей для повышения эффективности компьютерного моделирования;

5. оценка характеристик и эффективности предложенных средств.

Методы исследований – решение сформулированных задач базируется на методах системного анализа, методах теории информации, цифровой обработки сигналов, методах математического и имитационного программного моделирования и способах реализации разработанных подходов и средств с использованием современного комплекса программных инструментов и технологий.

Предполагаемая научная новизна работы заключается в разработке новых эффективных методов визуального представления и анализа данных на основе динамических графовых моделей с возможностью интерактивного взаимодействия. Ожидается разработка усовершенствованного подхода к визуализации данных и процессов в виде диаграмм переходов, позволяющий наглядно отображать потоки данных и взаимосвязи элементов моделируемой системы.

Разработанные методы и модели, реализованные в виде программного инструментария, обеспечат повышение эффективности и скорости процессов анализа и оптимизации при визуальном компьютерном моделировании в различных предметных областях.

Практическое значение работы заключается в возможности создания на основе предложенных и разработанных средств

эффективных компьютерных программных систем анализа и визуального моделирования, позволяющих:

– повысить эффективность процессов анализа и оптимизации сложных систем в различных предметных областях за счет более наглядного графического представления информации и инструментов ее анализа;

– ускорить процессы принятия решений в науке и бизнесе благодаря возможности быстрого построения моделей и визуального исследования вариантов реализаций;

– снизить временные и финансовые затраты на разработку и отладку моделей сложных систем по сравнению с традиционными средствами;

– повысить качество обучения специалистов методам системного анализа и моделирования за счет наглядности и интерактивности графовых моделей;

– унифицировать подходы к визуальному компьютерному моделированию в различных предметных областях на основе единой графовой методологии, и созданию инструментария и компьютерных систем на этой основе.

Выводы

В работе рассмотрена проблема повышения эффективности анализа и визуализации научных данных, обозначена актуальность исследования и разработки новых эффективных методов представления, анализа и визуализации данных в условиях возрастающих объемов информации. В качестве перспективного направления решения проблемы предложено использование графовых моделей данных.

Выполнен анализ существующих методов, технологий и средств визуализации данных на основе графовых моделей, рассмотрены такие инструменты, как Gephi, Linkurious, Graphistry и Cytoscape.js, проанализированы их возможности и ограничения. На основе проведенного анализа были определены цель исследования, основные задачи исследований, сформулированы новизна и практическое значение работы, направленной на развитие методов визуальной аналитики и повышение эффективности анализа данных различных прикладных областей.

Литература

1. Компьютерная визуализация [Электронный ресурс]. – Режим доступа: http://aco.ifmo.ru/e1_books/computer_visualization/lectures/1.html - Загл. с экрана.

2. Теория графов. Термины и определения в картинках [Электронный ресурс]. – Режим доступа:

<https://habr.com/ru/companies/otus/articles/568026/>
- Загл. с экрана.

3. Информационные модели на графах [Электронный ресурс]. – Режим доступа: http://infoplaneta.ucoz.net/index/urok_23_ispolzovanie_grafov_prakticheskaja_rabota_10_sozdaem_diagrammy_grafiki_skhemy/0-141?ysclid=lnafles5rz498646207 - Загл. с экрана.

4. Макарова, А. Практическое применение теории графов (исследовательская работа). ГАПОУ НСО «Новосибирский педагогический колледж No2», 2016, Новосибирск, Россия.

5. Линник, Е. В. Графовая аналитика для решения ключевых проблем в банковской сфере / Е. В. Линник. — Текст : непосредственный // Молодой ученый. — 2018. — № 52 (238). — С. 128-134.

6. Графовый анализ — обзор и области применения [Электронный ресурс]. – Режим доступа:

<https://habr.com/ru/companies/glowbyte/articles/594221/> - Загл. с экрана

7. The Open Graph Viz Platform [Электронный ресурс]. – Режим доступа: <https://gephi.org> - Загл. с экрана.

8. Linkurious – Main Page [Электронный ресурс]. – Режим доступа: <https://linkurious.com> - Загл. с экрана.

9. Graphistry – 100X Investigations [Электронный ресурс]. – Режим доступа: <https://www.graphistry.com> - Загл. с экрана.

10. Cytoscape – Network Data Integration, Analysis, and Visualization in a Box [Электронный ресурс]. – Режим доступа: <https://cytoscape.org> - Загл. с экрана

Зори С. А., Бездетный Н. А. Повышение эффективности анализа и визуализации данных на основе графовых моделей. В статье рассматривается проблема повышения эффективности анализа и визуализации научных данных, обозначена актуальность исследования и разработки новых эффективных методов представления, анализа и визуализации данных в условиях возрастающих объемов информации. В качестве перспективного направления решения проблемы предлагается использование графовых моделей данных. Сформулирована цель исследования, определены основные задачи работы, направленной на развитие методов визуальной аналитики и повышение эффективности анализа данных различных прикладных областей.

Ключевые слова: визуальный анализ данных, визуализация, графовые модели, теория графов, диаграммы, интерактивное взаимодействие.

Zori S.A., Bezdetniy N.A. Improving the Efficiency of Data Analysis and Visualization Based on Graph Models. The article considers the problem of improving the efficiency of scientific data analysis and visualization, highlights the relevance of research and development of new effective methods for representing, analyzing and visualizing data in the face of increasing volumes of information. The use of graph data models is proposed as a promising solution to the problem. The aim of the study is formulated, the main objectives of the work are identified, which is aimed at developing visual analytics methods and improving the efficiency of analyzing data from various applied fields.

Keywords: visual data analysis, visualization, graph models, graph theory, diagrams, interactive interaction.

Статья поступила в редакцию 20.09.2023

Рекомендована к публикации профессором Мальчевой Р. В.