

УДК 37.031.4

Золотов Дмитрий Александрович,

Аспирант

*ФГАОУ ВО «Белгородский государственный национальный
исследовательский университет»*

Научный руководитель: Польщиков Константин Александрович

д.т.н., профессор

*ФГАОУ ВО «Белгородский государственный национальный
исследовательский университет»*

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ В ОБЛАСТИ ВИЗУАЛЬНОГО АНАЛИЗА ДАННЫХ

Аннотация: статья представляет собой обзор современных методов визуального анализа данных и их широкое применение в различных областях. Проводится анализ существующих графоаналитических моделей и выявляются их преимущества и недостатки. Работа охватывает аспекты развития визуального анализа данных, включая методы, модели и тенденции, способствующие эффективному визуальному представлению и анализу сложных данных.

Ключевые слова: визуальный анализ данных, методы визуального анализа, графоаналитические модели, визуализация данных, анализ сложных данных, информационная визуализация, инструменты визуализации данных.

Zolotov Dmitry Alexandrovich,

Graduate Student

Belgorod State National Research University

Scientific adviser: Polshchikov Konstantin Alexandrovich

Professor

Belgorod State National Research University

MODERN TRENDS IN THE FIELD OF VISUAL DATA

ANALYSIS

***Abstract:** The article provides an overview of contemporary methods of visual data analysis and their widespread application in various fields. An analysis of existing graph-analytical models is conducted, highlighting their advantages and disadvantages. The work covers aspects of the development of visual data analysis, including methods, tools, and trends contributing to effective visual representation and analysis of complex data.*

***Keywords:** visual data analysis, methods of visual analysis, graph-analytical models, data visualization, analysis of complex data, information visualization, data visualization tools.*

В современном информационном обществе, насыщенном объемами данных, визуальный анализ становится ключевым инструментом для понимания и извлечения ценной информации. С развитием новых технологий и появлением множества методов их визуализации, вопросы эффективного анализа и интерпретации сложных информационных структур остаются в центре внимания исследователей и практиков.

Графоаналитические модели, предоставляют средства для исследования связей и паттернов в данных, особенно в контексте сложных сетей и взаимодействий. Однако, несмотря на значительные достижения в этой области, существуют проблемы, такие как эффективность методов преобразования и обеспечение их адаптации к различным прикладным областям. Исходя из этого необходимо проанализировать известные методы и модели, определить текущие проблемы и тенденции их развития.

Информационная визуализация – это область, которая охватывает широкий спектр методов, направленных на визуализацию абстрактной и конкретной информации [1]. Визуальные элементы, такие как диаграммы, графики и карты, используются для представления данных в понятной форме. Этот подход активно применяется в бизнесе, науке, образовании и государственном управлении для облегчения принятия решений.

Графоаналитика, специализирующаяся на анализе графов и сетей, активно применяется в изучении социальных взаимодействий, биоинформатике и других областях, обеспечивая выявление паттернов и связей [1].

Следующий метод – это визуализация временных данных. Их представляют в виде временных графиков, тепловые карты, диаграммы Ганта и других инструментов, которые помогают отслеживать изменения в данных со временем. Этот метод активно используется в медицинской диагностике, финансах, метеорологии и многих других областях [2].

Важно отметить использование методов машинного обучения и искусственного интеллекта, которые значительно расширяют возможности анализа данных. Алгоритмы обучения с учителем и без учителя, а также нейронные сети, применяются для выявления закономерностей, кластеризации данных и предсказания трендов [1].

Далее необходимо рассмотреть существующие модели исследования графов, выявить их преимущества и недостатки, а также рассмотреть перспективы развития данной области.

В графоаналитики существуют многочисленные модели, разработанные для анализа и визуализации сложных сетевых структур. Одной из важнейших моделей является модель «PageRank» [3], предложенная Ларри Пейджем и Сергеем Брином в 1996 году. Используя алгоритм, основанный на вероятности перехода между веб-страницами, модель эффективно ранжирует страницы веб-сайтов, учитывая их взаимосвязи [1]. Эта модель нашла широкое применение в поисковых системах, но ее эффективность ограничена для анализа более сложных структур, таких как биологические или социальные сети [3].

Другой значимой моделью является «Centrality» – показатель, который измеряет степень важности узла в графе. Модель центральности применяется в областях социологии и анализ социальных сетей [3].

Однако, модель центральности имеет ограничения при работе с направленными графами и не всегда учитывает влияние среды или контекста на важность узла.

Также, стоит рассмотреть модель «Близость», которая оценивает, насколько узел близок к другим узлам в графе [1]. Суть заключается в том, чтобы измерить, насколько быстро узел может достичь других узлов в графе. Чем меньше среднее расстояние между данным узлом и остальными, тем выше его близость. Эта модель активно применяется в транспортных и социальных сетях, помогая выявлять узлы, которые находятся в центре связанных структур. Из недостатков стоит выделить чувствительность к выбору метрик расстояний между узлами, а также ограничения при работе с направленными графами.

Последняя рассматриваемая модель – это Hyperlink-Induced Topic Search (далее HITS). Она предложена в 1999 году Джоном Клейнбергом [2]. Она работает на основе идеи разделения узлов на два типа: хабы (hubs) и авторитеты (authorities). Хабы являются узлами, которые указывают на много других узлов, а авторитеты – теми, на которые много указывают. Алгоритм HITS позволяет выделить важные узлы, выступающие как хабы и авторитеты, что находит применение в ранжировании веб-страниц и анализе социальных сетей. Модель чувствительна к начальным значениям хабов и авторитетов и может быть неэффективной в графах с большим количеством узлов и связей [2].

Обсужденные модели отражают несколько общих тенденций в графоаналитике. Во-первых, стремление к учету разнообразных характеристик сетевых структур подчеркивает важность интеграции различных методов для более полного и точного анализа данных. Во-вторых, адаптивность моделей под конкретные сценарии задач становится ключевым направлением исследований. Также, в дальнейшем развитии будет акцент на создание универсальных инструментов, способных

эффективно оперировать с разнообразными структурами данных и учитывать их динамические характеристики.

Однако, ни одна из рассмотренных моделей не является универсальной, и их эффективность зависит от характеристик конкретных данных и постановки задачи. Интеграция различных моделей и их тщательная адаптация для конкретных сценариев становятся важными направлениями будущих исследований.

Использованные источники:

1. Крижановский, А.А. Математическое и программное обеспечение построения списков семантически близких слов на основе рейтинга вики-текстов [Текст] / Крижановский А.А – СПб.: Питер, 2008. – 188 с. (дата обращения 02.02.2024)

2. Ландэ, Д.В. Навигация в сложных сетях: модели и алгоритмы [Текст] / Ландэ Д.В., Снарский А.А. – М.: КД Либроком, 2009. – 264 с. (дата обращения 02.02.2024)

3. Меры центральности в Network Science [Электронный ресурс] – URL: habr.com/ru/articles/715386/ (дата обращения 01.02.2024)