

УДК 681.512

## Особенности средств информационного мониторинга многомерных объектов на основе технологий оперативного анализа данных

Н.А. Сенцов, А.П. Преображенский✉

*Воронежский институт высоких технологий, Воронеж, Россия*

*В данной работе рассматриваются возможности разработки средств, позволяющих осуществлять процессы информационного мониторинга многомерных объектов с привлечением технологий оперативного анализа данных. Проведен анализ интеллектуальных методов. Указаны характеристики многомерных хранилищ данных. Рассмотрены условия, при которых можно достичь повышения эффективности в ходе многомерного анализа.*

*Ключевые слова: информационный мониторинг, многомерные данные, большие данные, OLAP-кубы, оперативное решение, междисциплинарное сотрудничество, оптимизация процессов.*

## The characteristics of tools for information monitoring of multidimensional objects based on technologies of operational data analysis

N.A. Sentsov, A.P. Preobrazhenskiy✉

*Voronezh Institute of High Technologies, Voronezh, Russia*

*This paper discusses the possibilities of developing tools that allow to carry out the processes of information monitoring of multidimensional objects with the involvement of technologies for operational data analysis. An analysis of intellectual methods is carried out. Characteristics of multidimensional data stores are indicated. The conditions under which it is possible to achieve increased efficiency during multivariate analysis are considered.*

*Keywords: information monitoring, multidimensional data, big data, OLAP cubes, operational solution, interdisciplinary cooperation, process optimization.*

Анализ современных информационных систем показывает, что в настоящее время большое значение приобретает разработка перспективных инструментов мониторинга и глубокого мультифакторного анализа данных, применимых при изучении чрезвычайно сложных многопараметрических систем, где большое число взаимосвязанных характеристик и показателей формируют высокоорганизованные структуры, полно и точно описываемые лишь в высокоразмерном пространстве данных [1, 2]. Исследование таких объектов требует применения интегрированных информационных платформ нового поколения, способных не только накапливать и обрабатывать колоссальные объемы поступающей информации, но и проводить ее глубокий оперативный анализ с целью выделения скрытых закономерностей и смыслов.

Разработка данных высокотехнологичных систем должна опираться на последние достижения в области обработки больших данных, машинного обучения, искусственного интеллекта и когнитивных технологий. Лишь синергетическое взаимодействие и интеграция этих перспективных научных направлений может обеспечить создание инновационных платформ нового поколения, позволяющих осуществлять комплексный мониторинг и глубокий анализ сложных многомерных объектов в реальном времени. Это, несомненно, станет мощным толчком для прогресса

фундаментальных наук и разработки революционных интеллектуальных технологий будущего [3, 4].

Основной задачей таких интеллектуальных платформ является обеспечение пользователей своевременной и глубоко структурированной информацией, отражающей актуальное состояние и динамику изменений сложнейших многопараметрических объектов исследования в режиме реального времени. Это позволяет оперативно реагировать на любые отклонения и тенденции в поведении данных систем, своевременно корректировать технологические режимы их функционирования и принимать обоснованные решения на основе глубокого анализа имеющейся информации [5, 6].

Основными методами интеллектуального анализа данных являются, в первую очередь, методы, основанные на переборе. Обычный перебор всех вариантов занимает  $O(2^N)$  операций (где  $N$  – общее количество объектов), а это значит, что с ростом числа объектов перебора вычислительная сложность растёт экспоненциально, что при значительном количестве объектов делает решение такой задачи практически невыполнимым.

Для того чтобы уменьшить количество переборов и снизить вычислительную сложность таких алгоритмов, методы интеллектуального анализа данных используют различные эвристические подходы. Одним из преимуществ методов интеллектуального анализа данных является их лёгкость, как для восприятия, так и для программной реализации.

Недостатком методик Data Mining считается отсутствие в них формализованного математического аппарата, а также строго оформленная теория, на которой происходит построение методик. Вследствие указанного недостатка, данные методики развиваются с трудом. Укажем главные типы знаний, добываемых методами интеллектуального анализа:

- ассоциативные правила;
- деревья решений;
- кластеры;
- математические функции.

Методики поиска данных знаний применяют для самых разных научных областей. Это теория вероятностей, математический метод анализа, статистический метод анализа, теория множеств, нейронные сети, программирование генетического и эволюционного типа.

Для достижения наивысшей эффективности мониторинга и удобства восприятия полученных результатов такие платформы оснащаются широким спектром инструментов визуализации и представления результатов, обеспечивающих наиболее ясное и целостное понимание происходящих процессов [7, 8].

Создание подобных прикладных инструментариев требует постоянного междисциплинарного взаимодействия масштабных научных коллективов, объединяющих ведущих специалистов в области математического моделирования, статистического и машинного анализа данных, информационных технологий и искусственного интеллекта. Помимо этого, критически важно тесное сотрудничество с ведущими экспертами соответствующих предметных областей из разных областей науки и техники.

Лишь такой широкий междисциплинарный синергетический подход позволяет не только адекватно формализовать задачи и функциональные требования к системе, но и обеспечить последующую адаптацию ее возможностей под конкретные запросы прикладных областей.

Одним из наиболее ярких примеров реализации такого подхода являются платформы для мониторинга и прогнозирования глобальных экологических процессов планетарного масштаба. Они позволяют систематизировать и интерпретировать большие массивы данных от спутников и датчиков различных типов, выполнять моделирование климатических и иных природных систем, а также обеспечить раннее определение потенциально опасных экологических ситуаций и разработку мер по их предотвращению или смягчению последствий [9, 10].

В современном мире, где объемы цифровой информации растут в геометрической прогрессии, роль инновационных интеллектуальных систем сбора, хранения и анализа больших данных становится критически важным. Среди основных элементов таких систем выделяются многомерные базы данных, способные обеспечить эффективную работу с потоками информации петабайтных масштабов [11].

Одной из ключевых особенностей многомерных баз данных является их уникальная архитектура – распределённая структура для хранения данных в виде компактных фрагментов, адресуемых с помощью мощных алгоритмов, что обеспечивает максимальную производительность процессов поиска и извлечения информации. Также данные технологии позволяют интегрировать и обрабатывать сведения, собранные из разнородных источников, используя сложные методы статистического, машинного и нейронного анализа для выявления ранее неизвестных закономерностей [12, 13].

Ключевую роль в таких системах играют OLAP-технологии – многомерные хранилища данных, формируемые на базе OLAP-кубов. Представляя информацию в многомерном формате по принципу сложных математических объектов, OLAP-кубы обеспечивают эффективное и компактное её хранение, а также высокопроизводительную обработку сложных многокритериальных запросов. Совместное использование всех этих технологий позволяет создавать инновационные платформы для работы с большими данными в науке, бизнесе и различных сферах деятельности человека [2].

Структура OLAP-куба способствует легкому пониманию отношений между данными, представляя информацию в удобной для анализа статистической форме. Это позволяет пользователям выполнять разнообразные операции анализа, углубляясь в детали и выявляя скрытые взаимосвязи.

Хотя кубы могут иметь практически неограниченный размер, они не всегда физически трехмерны, что может влиять на производительность компьютера при обработке больших объемов данных. Тем не менее OLAP-куб служит логической моделью для представления многомерных данных, где каждый элемент данных или «гиперкуб» характеризуется через набор индикаторов и измерений [13, 14].

Индикаторы в гиперкубе (например, продажи, расходы) представлены набором значений, количественно описывающих процессы или явления. Величины, связанные, например, с производственным процессом, такие как время работы подразделений, местоположение объектов производства, параметры производства, определяют контекст для индикаторов, организуя их в структурированную и иерархическую систему. Эта структура помогает создавать оси гиперкуба, облегчая анализ многомерных данных.

В конечном итоге, OLAP-кубы предлагают уникальную возможность для глубокого анализа данных, предоставляя многоуровневый взгляд на информацию и обеспечивая основу для управленческих решений [15].

Важным аспектом технологии OLAP является её «прозрачность». Это означает, что пользователь должен легко воспринимать многомерный куб, управление им

должно быть интуитивно понятным, а используемая в анализе терминология – знакомой.

Анализ данных в OLAP проходит через несколько ключевых этапов: консолидацию, агрегирование (группировку данных), вырезание и вращение. Консолидация объединяет данные из различных источников, а агрегирование позволяет сгруппировать эти данные для упрощенного представления и анализа. Вырезание относится к процессу выбора конкретного сегмента данных из куба, что эффективно уменьшает его размер и упрощает анализ. Вращение изменяет ориентацию данных в кубе, предоставляя разные перспективы для анализа [3].

Создание «срезы» куба позволяет фокусироваться на определенных данных, сокращая объем анализируемой информации и оставляя только релевантные измерения. Это погружение во «внутренний куб» облегчает детальный анализ.

Такой гибкий подход, заключающийся в возможности настройки представления OLAP-куба под специфику любых конкретных аналитических задач путем творческой манипуляции его структурой и свойствами, предоставляет аналитику значительную свободу действий в принятии обоснованных решений, основанных на извлечении наиболее ценной и релевантной информации из структурированных данных без ограничений жесткой формализации.

Взамен шаблонных стандартизированных форм отчетности такой подход позволяет проводить глубокое творческое исследование данных для достижения максимально эффективных результатов анализа. Это значительно повышает комфорт и продуктивность работы аналитика.

OLAP-куб представляет собой многомерный массив структурированных данных. Эти данные являются разнородными, они собраны из различных реляционных источников. На основе OLAP-куба есть возможности для последующего интерактивного визуального анализа таких данных и исследования любых произвольных подмножеств с целью извлечения наиболее существенных скрытых закономерностей, моделей и новых знаний [4].

Измерения (атрибуты), формируя многомерное аналитическое пространство OLAP-куба, обеспечивают свободу оценки и исследования обширного спектра количественных и качественных характеристик гигантских объемов разной природы данных, извлеченных из разнородных источников для целей последующего всестороннего анализа.

Иерархически структурированная система вложенных друг в друга кубов, охватывающих различные масштабы обобщенности информации, позволяет с легкостью получать доступ к содержимому на множестве уровней детализации через многоуровневую навигацию.

Иерархия измерений предоставляет гибкие алгоритмы агрегирования исходных данных на основе различных критериев классификации для последующего удобного анализа информации на различных временных, пространственных и иных интервалах обобщения.

Набор количественных метрик и показателей отражает в оперативной памяти множество разноплановых аспектов изучаемых данных с целью выявления закономерностей на основе сравнения и сопоставления данных.

Система символических классификаторов формализует размерности и объемы структурированной в OLAP-кубе информации.

В качестве примера можно привести визуальное представление OLAP-куба продаж с осями «Месяц», «Товар» и «Точка продаж», где каждая элементарная ячейка несет в себе конкретный количественный показатель, обеспечивающий удобный детальный многоаспектный анализ [5].

OLAP-куб представляет собой мощнейший инструментальный аппарат, позволяющий в высшей степени оптимизировать процессы агрегирования и классификации колоссальных объёмов строго структурированных данных в хранилищах. Он обеспечивает возможность проведения разнообразнейших аналитических трансформаций для максимального повышения эффективности извлечения наиболее ценной и полезной информации посредством:

1. Сегментирования, способствующего извлечению специфических подмножеств данных, которые нацелены под решение конкретных исследовательских задач.

2. Переориентации, позволяющей трансформировать представление данных для оптимального восприятия посредством адаптации их структуры под цели анализа.

3. Детализации, обеспечивающей доступ к наиболее точной и подробной информации по запросу пользователя.

4. Консолидации, осуществляющей объединение разрозненных данных для формирования всеобъемлющего обзора.

Качество аналитических выводов, получаемых на основе OLAP-куба, зависит от качества исходных данных и чёткости построения мультиаспектной аналитической модели, включающей комплекс взаимосвязанных OLAP-субмоделей, нацеленных на разрешение определённых классов исследовательских задач с применением соответствующих алгоритмов обработки [11].

Для достижения наивысшей эффективности OLAP-аналитики критически важным является соблюдение ряда фундаментальных условий:

1. Скорость обработки должна с чрезвычайно высокой степенью точности оставаться постоянно оптимальной вне зависимости от объёмов обрабатываемой информации, полностью исключая риск каких-либо задержек при извлечении любых имеющихся данных.

2. Гибкость подразумевает потенциальную возможность проведения всех вообразимых классов анализа, включая самые тонкие статистические, математические и иные виды расчётов, обеспечивая тем самым максимально полное изучение предметной области.

3. Обеспечение многопользовательского доступа позволяет бригадам аналитиков одновременно эффективно взаимодействовать с единой базой данных для оперативного принятия сбалансированных решений.

4. Постоянная доступность означает возможность моментального извлечения любой запрошенной информации в любой момент времени.

5. Масштабируемость измерений предполагает способность работать с практически неограниченными объёмами данных посредством анализа без ограничений в разветвлённости аспектов исследования.

Это позволяет довести производительность, гибкость и функциональность OLAP-систем до абсолютного оптимума для решения самых сложнейших аналитических задач.

Таким образом, OLAP-структура с её многомерными базами данных становится ключевым элементом для оперативного анализа данных, удовлетворяя высоким требованиям современного аналитического процесса.

Заключение. Разработка средств информационного мониторинга многомерных объектов на основе технологий оперативного анализа данных открывает новые возможности для улучшения качества управления и принятия решений в различных областях, от экологии до промышленности и финансов. Это требует комплексного подхода к разработке и внедрению таких систем, а также постоянного совершенствования технологий обработки и анализа данных. В работе продемонстрированы возможности и подходы, связанные с обработкой многомерной информации и условий для их реализации.



## СПИСОК ИСТОЧНИКОВ

1. Паневин Р.Ю. Структурные и функциональные требования к программному комплексу представления знаний / Р.Ю. Паневин, Ю.П. Преображенский // Вестник Воронежского института высоких технологий. – 2008. – № 3. – С. 61-64.
2. Ермолова В.В. Оптимизация взаимодействий компонентов человеко-машинной системы цифровизации / В.В. Ермолова, Я.Е. Львович, Ю.П. Преображенский // Моделирование, оптимизация и информационные технологии. – 2023. – Т. 11. – № 2. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1374> (дата обращения: 15.03.2024).
3. Кострова В.Н. Оптимизация распределения ресурсов в рамках комплекса общеобразовательных учреждений / В.Н. Кострова, Я.Е. Львович, О.Н. Мосолов // Вестник Воронежского государственного технического университета. – 2007. – Т. 3. – № 8. – С. 174-176.
4. Ермолова В.В. Оптимизация человеко-машинной среды управления организацией с применением киберфизической системы / В.В. Ермолова, Я.Е. Львович, Ю.П. Преображенский // Моделирование, оптимизация и информационные технологии. – 2023. – Т. 11. – № 3. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1360> (дата обращения: 15.03.2024).
5. Гусев П.Ю. Структуризация многофункциональной цифровизированной системы и управление ею на основе оптимизационных моделей деагрегации ресурсов и объемов деятельности / П.Ю. Гусев, Я.Е. Львович // Моделирование, оптимизация и информационные технологии. – 2023. – Т. 11. – № 4. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1441> (дата обращения: 15.03.2024).
6. Борзова А.С. Особенности построения системы принятия решений при многовариантной оптимизации структуры цифрового управления логистическим процессом в организационной системе на основе имитационного моделирования / А.С. Борзова, В.В. Муха // Моделирование, оптимизация и информационные технологии. – 2021. – Т. 9. – № 3. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1048> (дата обращения: 15.03.2024).
7. Муха В.В. Оптимизация цифровой нити логистических цепочек в практике управления организационными системами / В.В. Муха // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 1. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1132> (дата обращения: 15.03.2024).
8. Львович К.И. Оптимизация адаптационного процесса при организации командной деятельности персонала в условиях цифровизации управления многокомпонентной организационной системой / К.И. Львович, А.П. Преображенский // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 4. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1244> (дата обращения: 17.03.2024).
9. Львович А.И. Алгоритмизация процесса визуально-экспертного моделирования при оптимизации управления развитием организационных систем с использованием мониторинговой информации / А.И. Львович, А.П. Преображенский // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 2. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1154> (дата обращения: 17.03.2024).
10. Шульга Т.Э. О подходе к обеспечению транзитивности свойств в веб-онтологиях / Т.Э. Шульга, А.А. Сытник // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 4. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1265> (дата обращения: 17.03.2024).

11. Львович Я.Е. Характеристики хранилища данных в корпоративных системах / Я.Е. Львович, А.П. Преображенский // Вестник Воронежского института высоких технологий. – 2020. – Т. 14. – № 1 (32). – С. 31-33.

12. Преображенский Ю.П. Анализ некоторых характеристик спутниковых систем связи / Ю.П. Преображенский, Т.В. Аветисян, Е. Ружицкий // Вестник Воронежского института высоких технологий. – 2022. – Т. 16. – № 1 (40). – С. 82-85.

13. Preobrazhenskiy Yu.P. The Problems of Process Control in Computer Systems / Yu.P. Preobrazhenskiy, Yu.L. Chuprinskaya, E. Ruzhicky // Вестник Воронежского института высоких технологий. – 2022. – Т. 16. – № 1 (40). – С. 92-94.

14. Preobrazhenskiy Yu.P. Some Characteristics of Computer Networks / Yu.P. Preobrazhenskiy, K.M.V. Azer, D. Dzhumageldiev // Вестник Воронежского института высоких технологий. – 2022. – Т. 16. – № 1 (40). – С. 86-88.

15. Мельникова Т.В. Моделирование обработки больших массивов данных в распределенных информационно-телекоммуникационных системах / Т.В. Мельникова, М.В. Питолин, Ю.П. Преображенский // Моделирование, оптимизация и информационные технологии. – 2022. – Т. 10. – № 1. – URL: <https://moitvvt.ru/ru/journal/pdf?id=1117> (дата обращения: 17.03.2024).

### ИНФОРМАЦИЯ ОБ АВТОРАХ

**Сенцов Николай Андреевич**, аспирант, Воронежский институт высоких технологий, Воронеж, Россия.

*e-mail:* [sentsov673\\_rrt@yandex.ru](mailto:sentsov673_rrt@yandex.ru)

**Преображенский Андрей Петрович**, доктор технических наук, профессор, Воронежский институт высоких технологий, Воронеж, Россия.

*e-mail:* [app@vvt.ru](mailto:app@vvt.ru)