УДК 004.89

А.В. КОСЬКИН, А.А. МИТИН, А.В. АРТЕМОВ

A.V. KOSKIN, A.A. MITIN, A.V. ARTEMOV

**КОНЦЕПЦИЯ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ С ВЫБОРОМ МЕТОДОВ И СРЕДСТВ АНАЛИЗА ДАННЫХ ДЛЯ ОБРАБОТКИ ИНФОРМАЦИИ**

**THE CONCEPT OF FORMATION AN INTELLIGENT SYSTEM WITH A CHOICE OF METHODS AND TOOLS FOR ANALYZING DATA FOR INFORMATION PROCESSING**

*В статье описывается концепция интеллектуальной информационно-аналитической системы для поддержки принятия решений. Представлена структура такой системы, описаны составляющие её компоненты и их функциональные возможности. Сформулированы требования к интеллектуальной информационно-аналитической системе.*

*Ключевые слова: информационно-аналитическая система, интеллектуальная система, анализ данных, обработка информации.*

*The article describes the concept of an intellectual information and analytical system for supporting decision-making. The structure of such a system is presented, its constituent components and their functional capabilities are described. The requirements for an intelligent information and analytical system are formulated.*

*Keywords:* *information-analytical system, intellectual system, data analysis, data processing.*

В связи с увеличением объема данных, развитием и совершенствованием технологий их хранения и обработки, развитием технологии «больших данных» резко выросли информационные потоки, которые необходимо эффективно обрабатывать, анализировать, интерпретировать с целью получения качественной и достоверной информации. В силу этих причин в области информационных технологий особую актуальность приобретают вопросы построения информационно-аналитических систем. Информационно-аналитические системы (ИАС) как класс информационных систем ориентированы на решение задач принятия решений на основе разнородных данных, хранящихся в различных источниках. ИАС – это современный высокоэффективный инструмент принятия тактических, стратегических и оперативных решений на основе предоставления всей необходимой совокупности данных пользователям, ответственным за принятие решений.

Проведение анализа требует доступа к большому объёму данных и обычно целью анализа является выявление определенных тенденций: сначала осуществляется анализ исторических данных, а затем выявленная тенденция экстраполируется на будущий период времени.

Основными требованиями к ИАС являются:

1. Наличие доступа к большим объемам данных, хранящихся в различных источниках.
2. Взаимодействие с внешними системами операционной обработки данных и источниками данных.
3. Наличие средств проверки непротиворечивости и полноты данных.
4. Наличие средств просмотра, визуализации и интерпретации данных, генерации отчетов.

Одним из стратегических направлений развития информационно-аналитических систем является их интеллектуализация, заключающаяся в построении формализованных процедур обработки данных и выбора метода для их анализа, интерпретации и представления информации в виде некоторых показателей и их дальнейшее использование непосредственно для принятия решения в целевой области, т.е. выбора одного из альтернативных вариантов действий пользователем.

Таким образом, интеллектуальная информационно-аналитическая система (ИИАС) – это информационно-аналитическая система, которая основана на концепции выбора процедур обработки данных и методов, средств и технологий анализа данных в зависимости от качественного содержания и структурированности данных в процессе принятия решений.

Интеллектуальный анализ данных чаще реализуется автономными программными системами в связи со сложностью решаемых задач. В то же время OLAP-системы частично выполняют самые распространенные и легко реализуемые функции интеллектуального анализа. Интеллектуальные информационно-аналитические системы являются своего рода надстройкой для традиционных систем обработки данных (СОД). Они используют предоставляемые СОД данные, а также дополнительную информацию и данные для анализа, проводимого с использованием собственных приложений, результаты которого фактически могут являться знаниями, т.е. структурированной информацией, содержащей оценки взаимосвязи между параметрами описания объекта управления.

В обобщенном виде структура интеллектуальной информационно-аналитической системы представлена на рисунке 1.

1

Репозиторий методов анализа данных

Конфигурация

информационно-аналитической системы

Интеллектуальная

информационно-аналитическая система

Источники

данных

Базовая модель информационно-аналитической системы

Средства

OLAP

СУБД

Рисунок 1 –Структура интеллектуальной информационно-аналитической системы

Базовая модель информационно-аналитической системы включает:

1. Средства импорта и загрузки данных из операционных баз и других источников информации, взаимодействующие с различными операционными системами и СУБД.
2. Средства преобразования данных, осуществляющие проверку на правильность, преобразование структур, агрегирование.
3. Набор программных инструментальных средств, которые выполняют операционные функции оперативного анализа.
4. Средства графического и визуального построения отчетов, рассчитанные на конечного пользователя.
5. Средства удаленного доступа, коллективного доступа, обеспечения работы в распределенном режиме.
6. Средства моделирования объектов и процессов.
7. Средства администрирования.

Репозиторий методов анализа данных должен включать набор моделей, методов и технологий, которые могут быть использованы для получения некоторых закономерностей для данных, хранящихся в источниках данных. При этом модели могут иметь вид математических моделей или представлять собой экспертные системы, базирующиеся на логической модели предметной области, реализованной в форме базы знаний и механизма вывода. В репозитории могут быть представлены следующие методы анализа данных: ассоциация, классификация, кластеризация, прогнозирование, метод последовательных моделей, нейросетевые методы и технологии, методы деревьев решений, методы на основе прецедентов и ряд других статистических методов [1].

Источники данных – это структурированные и неструктурированные наборы данных, которые включают описания или характеристики некоторых объектов или процессов. В общем случае данные, необходимые для поддержки принятия решений, можно разделить на следующие виды:

1. первичные данные, количественно характеризующие состояние объекта (процесса) управления;
2. результаты обработки первичных данных, которая выполняется по алгоритмам, принятым в конкретной системе исходя из её функционального назначения, представляющие агрегированные (усредненные) данные;
3. обобщенные показатели функционирования объекта за определенный период времени, характеризующие эффективность его целевого функционирования.

В общем виде задачу, которую необходимо решать в интеллектуальной информационно-аналитической системе, можно сформулировать как преобразование данных типа 1), 2) и 3) в сводные показатели свойств объекта управления, которые не поддаются непосредственному измерению. Эти показатели затем могут интерпретироваться как некоторое качество объекта управления в целом или определенных его свойств.

Типовыми задачами средств OLAP являются:

1. Обеспечение полноты и достоверности хранимых данных – «очистка данных», включая проверку на непротиворечивость.
2. Обеспечение доступа к сложным многомерным данным в любом заданном разрезе (формирование сложных запросов).
3. Обеспечение отображения сложных многомерных данных в удобном для восприятия виде.

Конфигурация информационно-аналитической системы формируется посредством выбора методов и технологий анализа из репозитория. После этого происходит обработка, анализ данных и интерпретация результатов в соответствии с процедурами выбранного метода. Результаты визуализируются и представляются пользователю.

К основным задачам интеллектуального анализа данных относятся:

* классификация (распознавание) данных и ситуаций;
* кластеризация данных – разбиение массивов данных на группы (кластеры) по близости структур значений признаков;
* построение обобщенных показателей эффективности и качества;
* прогнозирование.

Применение существующих методов обработки многомерных данных выявило ряд недостатков их практического использования. Одним из важных среди них является практическая невозможность учета неопределенности в исходных данных при их обработке, оценке эффективности системы, прогнозировании развития объектов без принятия допущений, проверка которых просто невозможна или сопряжена с большими затратами.

В рамках существующих подходов снятие неопределенности достигается путем распространения аксиоматики теории вероятностей на формализацию процессов жизнедеятельности объектов. Субъективизм в использовании вероятностной меры приводит к расхождениям, получаемым на моделях результатов и экспериментальных данных. Анализ методов учета неопределенности при построении обобщенных характеристик объектов управления показал, что использование для этой цели фундаментального принципа максимума энтропии позволяет получить ряд важных преимуществ. При этом не возникает формальных и вычислительных сложностей в решении задач идентификации состояний объектов, описываемых многомерными данными, по двум причинам:

1. Энтропия, являясь мерой неопределенности, обладает свойством, что её максимальное значение одновременно представляет собой меру близости, обладающую основными свойствами метрики.
2. Полученные формальные представления энтропии позволяют разработать простые и эффективные вычислительные алгоритмы оценки обобщенных характеристик объектов.

Таким образом, для выполнения функций поддержки принятия решений ИИАС должна обеспечить:

* предварительную аналитическую обработку, направленную на получение сводных показателей рассматриваемого процесса;
* интеллектуальную аналитическую обработку, заключающуюся в выборе метода и технологии анализа для данных в зависимости от их качественного содержания и особенностей и построении интегральных показателей рассматриваемого процесса за любой временной период;
* представление данных и результатов в любом заданном разрезе в графическом виде.

**СПИСОК ЛИТЕРАТУРЫ**

1. Митин А.А. Методы и средства интеллектуального анализа данных. – Информационные системы и технологии № 1(105) 2018. – Орел: ОГУ имени И.С. Тургенева, 2018. – с. 34-38.
2. Белов В.С. Информационно-аналитические системы. Основы проектирования и применения. Учебное пособие. – М: МЭСИ, 2004.
3. Филиппов, В.А. Интеллектуальный анализ данных: методы и средства [Текст] / М.: Эдиториал УРСС. – 2001. – 52 с.
4. Осипов, Г.С. Методы искусственного интеллекта [Текст] / Физматлит, 2011. – 296 с.
5. Волков В.Н., Стычук А.А., Митин А.А. Разработка информационной модели представления данных о регламенте электронной услуги. – Информационные системы и технологии № 4(84) 2014. – Орел: [Орловский государственный университет им. И.С. Тургенева](https://elibrary.ru/publisher_titles.asp?publishid=7199), 2014. – c. 21-30.
6. Ельцова Н.С., Еременко В.Т., Коськин А.В., Мишин Д.С., Мишин Я.Д. Интерпретация моделей процессов получения и обработки информации в порталах органов исполнительной власти на основе операторных схем. – Информационные системы и технологии № 4(96) 2016. – Орел: ОГУ имени И.С. Тургенева, 2016 – c. 42-50.
7. Коськин А.В., Ужаринский А.Ю. Механизмы доступа к данным на основе единой интегрирующей схемы данных. – Информационные системы и технологии № 1(87) 2015. – Орел: ОГУ имени И.С. Тургенева, 2015. – с. 38-48.
8. Коськин А.В., Ужаринский А.Ю. Методика формирования интегрирующей модели данных на основе имеющихся разнородных источников данных. – Информационные системы и технологии № 2(82) 2014. – Орел: ОГУ имени И.С. Тургенева, 2014 – c. 19-27.

**Коськин Александр Васильевич**

ФГБОУ ВО «Орловский государственный университет имени И.С. Тургенева»

Д.т.н., профессор кафедры информационных систем

Тел.: +79065692020

E-mail: koskin@ostu.ru

**Митин Александр Александрович**

ФГБОУ ВО «Орловский государственный университет имени И.С. Тургенева»

К.т.н., доцент кафедры информационных систем

Тел.: +79065692020

E-mail: mcc77@yandex.ru

**Артемов Андрей Владимирович**

ФГБОУ ВО «Орловский государственный университет имени И.С. Тургенева»

К.т.н., доцент кафедры программной инженерии

Тел.: +79065692020

E-mail: avladar1984@yandex.ru