

## ПРОБЛЕМЫ МОДЕЛИРОВАНИЯ ПРЕДМЕТНОЙ ОБЛАСТИ ПРИ ПРОЕКТИРОВАНИИ ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИХ СИСТЕМ

АРИФЖАНОВ А.Ш., НАБИЕВ О.М.

*E-mail:* [aarifganov@mail.ru](mailto:aarifganov@mail.ru).

Вступление человечества в XXI век требует перехода к новой стратегии развития общества на основе знаний и высокоэффективных технологий. Широкое применение передовых информационно-коммуникационных технологий и информационно-аналитического обеспечения принятия решений в государственных структурах, экономике и образовании становится одним из приоритетных направлений развития общества. Создание информационно-аналитических систем (ИАС), реализующих аналитические технологии поддержки принятия решений, стало возможным благодаря возникновению новой корпоративной технологии накопления и оперативной обработки сверхбольших объемов ретроспективной информации BI- Business Intelligence с применением концепции Хранилища Данных (ХД) [1- 4].

Объекты управления, для которых реализуются крупномасштабные проекты создания современных ИАС, характеризуются, как правило, рядом особенностей, среди которых [1]:

- структурная сложность (многоуровневая иерархическая структура организации) и территориальная распределенность;
- функциональная сложность (многоуровневая иерархия и большое количество функций, выполняемых организацией; сложные взаимосвязи между ними);
- информационная сложность (большое количество источников и потребителей информации (министерства и ведомства, местные органы власти, организации-партнеры), разнообразные формы и форматы представления информации, сложная информационная модель объекта - большое количество информационных сущностей и сложные взаимосвязи между ними), сложная технология прохождения документов;
- сложная динамика поведения, обусловленная высокой изменчивостью внешней среды (изменения в законодательных и нормативных актах, нестабильность экономики и политики) и внутренней среды (структурные реорганизации, текучесть кадров).

Указанные особенности объектов внедрения ИАС обуславливают наряду, с чрезвычайно большим объемом данных в ХД, и их существенную разнородность. В таких условиях обеспечение достаточности и полноты анализируемой информации, т.е. целостности информационной базы ИАС для принятия адекватных реальному состоянию объекта управления решений, является основной проблемой в создании эффективной ИАС [5]. Потому что нагромождение ХД ИАС информацией, не относящейся к предметной области и задачам стратегического управления объектом контроля, резко снижает её эффективность. Информационная база ИАС по своим параметрам (функциональным возможностям) должна полностью соответствовать представлениям (концептуальной интерпретации) процесса управления должностных лиц (лиц, принимающих решения – ЛПР), для которых создается данная ИАС. Только при наличии достаточного соответствия между базовыми параметрами объекта управления и данными в информационной базе ИАС об их состояниях (значениях), можно говорить о функциональной пригодности ИАС в информационном аспекте, а именно, о полноте информации (или более высокого уровня – знаний как интеллектуально преобразованной информации). В этом случае можно говорить о соответствии требований процесса принятия решения (со стороны ЛПР как органа принятия управляющих решений) и характеристик информационного процесса (достаточности, достоверности и оперативности обработки данных) в управляющей части с учетом частоты поступления сведений о состоянии объекта управления.

Однако, несмотря на все более широкое распространение ИАС, почти половина проектов по их реализации оказывается неуспешной (согласно данным исследовательской организации Gartner Group [6]). Основной причиной этого является недостаточная разработка математической теории описания, представления и моделирования предметных областей ПрО в информационных системах (ИС). Проблема состоит как в недостаточной выразительности имеющихся средств представления семантики ПрО в ИС, так и неформализованности процесса моделирования ПрО, что делает



Этап семантической локализации ПрО чаще всего осуществляется на интуитивном уровне, однако он является сложной задачей и зависит от целей построения ИС над ПрО. Как отмечено в [7], пересечение свойств предметных областей заданного семантического пространства может иметь место и в том случае, если это целесообразно с точки зрения решаемых над ПрО задач. В этом случае принято говорить про распределенные по определенному признаку ПрО [7]. В качестве примера можно рассматривать распределение ПрО:

- в пространстве;
- по классам решаемых задач;
- по типам ПрО (физические, абстрактные, социальные) и т.д.

Таким образом, после этапа локализации критерий (1) описывает распределенные ПрО (рис.1).



*Рис. 1. Локализация в семантическом пространстве предметных областей*

Главной характеристикой модели ПрО является способность давать новую семантическую информацию об объекте-оригинале. Таким образом, говоря о моделировании ПрО, мы подразумеваем, что модель ПрО в ИС является средством для получения новой информации о свойствах реальной ПрО.

Это возможно благодаря составляющему основу метода моделирования принципу подобия. Однако в случае построения ИС над ПрО подобие является косвенным или условным. Между описанием и моделью ПрО нет прямого соответствия, в зависимости от поставленной цели для построения модели ПрО могут быть использованы различные описания. То же можно сказать про различные способы представления логической модели ПрО на физическом уровне в разных моделях ИС.

В общем случае методы построения ИС с заданными свойствами для конкретной ПрО зависят от:

- структуры и свойств ПрО (например, по характеру преобладающих закономерностей можно выделить статистические или динамические ПрО);
- целей построения ИС и класса решаемых над ПрО задач;
- доступных технических средств.

Вне зависимости от этих факторов можно выделить [7] такие два инвариантных этапа построения ИС над ПрО, как описание и моделирование, результаты которых имеют определенное представление в ИС (рис. 2).

*Рис. 2. Этапы и методы построения ИС над ПрО*

Описание является совокупностью утверждений о свойствах и поведении объектов ПрО, которым можно приписать оценку истинности и представить в виде предикатов после введения соответствующих переменных. Данная оценка истинности является мерой адекватности описания и модели ПрО действительности. Однако заметим, что истинная валидность распределения вероятностей проявления свойств модели ПрО может быть проверена только на практике. Чаще всего



модель ПрО проверяется на адекватность описанию ПрО, являющемуся первоисточником информации о свойствах ПрО.

На основе описания ПрО строится модель ПрО. Про модель предметной области можно говорить, когда в представлении ПрО в ИС пройден этап формализации связей между объектами ПрО.

Существующие на сегодняшний день методологии и технологии разработки ИАС недостаточно формализуют моделирование предметной области, начиная с его первого этапа – анализа и описания ПрО. Представленные в форме графических диаграмм или текстовых спецификаций модели ПрО не дают возможности для формальной проверки их свойств, в первую очередь, на функциональную полноту и

логическую целостность [1,5].

Однако свойство функциональной полноты отнюдь не означает, что в модели ПрО должны быть отражены все функциональные и информационные элементы, а также их связи, имеющие место в обследуемой предметной области [1,5]. Модель ПрО необходима не сама по себе, а для решения определенного класса задач, то есть должна включать лишь необходимые и достаточные для этого свойства.

Таким образом, функциональная полнота модели ПрО подразумевает не максимально полное отражение свойств объектов ПрО, а фиксацию лишь тех свойств, которые являются необходимыми и достаточными для решения задач (автоматизации управления, информационного поиска, принятия решений, проектирования архитектуры, дедукции знаний и др.). Таким образом, критерий функциональной полноты модели ПрО зависит от класса решаемых над ПрО задач и в свою очередь требует определения критерия глубины детализации предметной области.

Таким образом, можно выделить следующие свойства модели ПрО:

- зависимость от целей разработки ИС и класса решаемых над ПрО задач;
- формальность;
- адекватность действительности (описанию ПрО, если рассматривать его как первоисточник информации для построения модели);
- способность давать новую информацию;
- функциональная полнота;
- логическая целостность.

Постановка целей разработки ИАС и определение класса решаемых над ПрО задач являются первичными по отношению к описанию и построению модели ПрО в ИАС. Именно класс решаемых задач определяет методы построения ИС с заданными свойствами, а также конкретные свойства, структуру и функции ПрО, которые необходимо отразить в описании и модели.

Таким образом, первым этапом для описания и построения ИС над ПрО должен стать этап построения модели задачи. Именно формальная формулировка задач, решаемых над ПрО, позволяет выделить свойства, моделирование которых является необходимым и достаточным для их решения.

Исходя из этого, актуальным является рассмотрение средств формального представления моделей задач, которые необходимо решить над ПрО, поиск их соответствия функциональной и информационной моделям ПрО. Модель ПрО в этом случае можно определить как средство для организации и управления действиями пользователя по решению класса задач над ПрО. Класс решаемых над ПрО задач зависит также от возможностей ИС, например, положенной в ее основу модели данных. Рассматривая ИАС как систему, основанную на знаниях, основным фактором, определяющим класс решаемых над ПрО задач, будет являться положенная в основу ИАС база знаний. Основной проблемой, возникающей при преобразовании семантической модели ПрО в модель ИС, является потеря свойств семантики ПрО.

## Литература

1. Арифжанов А.Ш. Проблемы формирования целостной информационной базы информационно-аналитической системы // Материалы Международной научно-практической конференции «Актуальные проблемы математики, информатики, механики и теории управления». - Алматы, 2009, часть 1, С. 254-260.
2. Codd E. F., Codd S. B., Salley C. T. Providing OLAP (On-Line Analytical Processing) to User-Analysts: An IT Mandate. - E. F. Codd & Associates, 1993.
3. Щавелёв С.Д. Оперативная аналитическая обработка данных: концепции и технологии. [http://olap.ru/basic/olap\\_and\\_ida.asp](http://olap.ru/basic/olap_and_ida.asp).
4. Львов В. Создание систем поддержки принятия решений на основе хранилищ данных // СУБД, №3. 1997. С.30-40.
5. Арифжанов А.Ш. Целостность информационной базы в информационно-аналитической системе. Материалы Республиканского научно-практического семинара «Моделирование и управление в реальном секторе экономики Республики Узбекистан», Ташкент, 11-13 октября 2007 года, С. 13-18.
6. URL: <http://www.gartner.com>
7. Востров Г.Н., Межуев В.И.. Проблемы построения информационных систем над предметными областями //Журнал «Искусственный интеллект», № 4, 2008, С. 736-746
8. Коннолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация, сопровождение. Теория и практика. – М.: Вильямс, 2000. С. 1120.
- 9 Цаленко М.Ш. Моделирование семантики в базах данных. – М.: Наука, 1989. С. 288.



