

На правах рукописи

ЗОТОВ Алексей Алексеевич

**Разработка методики построения хранилищ данных
для систем финансового учета**

Специальность 05.13.11 –
Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Автореферат

диссертации на соискание ученой степени кандидата технических наук

Санкт – Петербург 2004

Работа выполнена в Государственном образовательном учреждении высшего профессионального образования «Санкт-Петербургский государственный политехнический университет».

Научный руководитель – доктор технических наук, профессор
Кракау Татьяна Константиновна

Официальные оппоненты – доктор технических наук, профессор
Мелехин Виктор Федорович
- кандидат технических наук, профессор
Котляров Всеволод Павлович

Ведущая организация - НПО “Импульс”

Защита состоится «11» ноября 2004 г. в 16 часов на заседании диссертационного совета Д 212.229.18 в Государственном образовательном учреждении высшего профессионального образования “Санкт-Петербургский государственный политехнический университет” по адресу: 195251, Санкт-Петербург, Политехническая ул., д.29, 9 уч. корп., ауд. 325.

С диссертацией можно ознакомиться в Фундаментальной библиотеке университета.

Автореферат разослан «___» октября 2004 г.

Ученый секретарь
диссертационного совета

Шашихин В.Н.

Общая характеристика работы

Актуальность

На сегодняшний день в области информационных технологий особую актуальность приобретают вопросы построения информационно – аналитических систем. Особенно это отражается на такой области, как банковская деятельность.

Ситуация на рынке аналитической банковской деятельности в России такова, что более половины банков не имеют соответствующих информационно-аналитических систем (ИАС). Вопросы анализа информации в таких банках решаются полуавтоматизированным способом или же вручную. Это приводит к понижению конкурентоспособности конкретного банка и нестабильности банковской системы страны в целом. Существующие же аналитические системы решают весьма ограниченный круг конкретных задач анализа, не унифицированы и привязаны к технологическим процессам и методикам определенного банка.

Построение ИАС по сути своей очень трудоемкий и дорогостоящий процесс, особенно для России. Постоянное изменение законодательства, нестабильная финансовая и политическая ситуация страны обуславливают необходимость периодического внесения соответствующих, порой существенных изменений в архитектуру и бизнес процессы аналитических систем. Зачастую такие изменения требуют принципиальных доработок системы, а в ряде случаев вызывают необходимость ее перепроектирования. Для того чтобы избежать таких проблем и обеспечить возможность проектирования достаточно унифицированных ИАС, легко адаптируемых к изменению внешних условий, необходима теоретическая и методологическая база и соответствующие инструментальные средства ее реализации.

В данной работе на основе анализа существующих концепций создания систем данного класса и тенденций развития технологии и информационно – аналитической деятельности в целом, разработаны теоретическая база и методика построения ИАС. Разработанная методика обеспечивает возможность построения ИАС, адаптируемой к быстроизменяющимся требованиям и пригодной для обеспечения деятельности большой группы предприятий, в сравнительно короткие сроки. Предложенная теоретическая база позволяет автоматизировать процесс построения хранилищ данных (ХД) для аналитических систем на основании формальной модели.

Вышеизложенное свидетельствует об актуальности данной работы.

Цель и задачи работы

Целью работы является создание методики построения и формального аппарата описания универсального ХД для повышения производительности труда разработчиков и повышения эффективности ИАС за счет:

- формализации процесса построения ИАС и ХД с учетом отраслевой специфики, обеспечивающей возможность их автоматизированного проектирования;
- разработки практической методики построения ИАС и ХД, позволяющей обеспечивать в сравнительно короткие сроки их проектирование и адаптацию к внешним условиям в соответствии с быстроменяющимися требованиями на основе универсального ХД, пригодного для удовлетворения информационно – аналитических потребностей большой группы предприятий.

Методы исследования

В диссертации используются методы теории множеств, реляционной алгебры, реляционного исчисления, проектирования и нормализации баз данных (БД) а так же методы системного структурного анализа.

Достоверность полученных результатов подтверждается использованием утверждений, доказанных в реляционной алгебре, корректным доказательством и непротиворечивостью собственных утверждений и достаточной аргументацией приведенных в работе выводов и рекомендаций, а также практическим внедрением результатов исследований и их апробацией.

Научная новизна работы

Научную новизну работы представляют следующие результаты:

1. Для формализации требований к ИАС предложен механизм формирования иерархии функциональных требований. На его основе сформированы методы выделения результирующих, промежуточных и базовых требований, позволяющие повысить скорость проектирования на этапе анализа предметной области.

2. Предложен механизм построения информационной структуры ИАС высокого уровня, базирующийся на основе сформированных базовых требований, позволяющий повысить показатели внедряемости.

3. Предложен новый подход к проектированию логической структуры хранилища данных, основывающийся на использовании готовых структур внешних источников, позволяющий увеличить скорость разработки ХД.

4. Разработана формальная модель хранилища данных, на базе которой проводится расширение реляционной алгебры и предлагаются правила преобразования структур данных, обеспечивающие возможность автоматизированного проектирования хранилища.

5. На базе построенной формальной модели ХД и доказанных утверждений введены правила “денормализации”, с помощью которых проектируется ХД.

6. Впервые использованы правила “объединения” и “всеобщего объединения” применительно к уже существующим структурам данных.

7. На основании введенных правил и функциональных требований сформулирован набор критериев для оценки необходимости и эффективности применения соответствующих правил преобразования.

Практическая значимость работы

Полученные в работе результаты были использованы при построении ИАС для решения широкого круга аналитических задач для группы коммерческих банков. На основании формальной модели было построено универсальное ХД, позволяющее решать большинство аналитических задач, присущих банковской деятельности.

Результаты, полученные при реализации предложенных методов построения ИАС, использованы в разработке программной системы формирования отчетности и анализа банковских данных – SBR&A. Так же результаты применения разработанной методики были использованы при построении ХД для торгового предприятия.

Хочется выразить глубокую благодарность за чрезвычайно ценные замечания и советы Дробинцеву Д.Ф., послужившие более четкому и полному изложению материала.

Апробация работы

Основные результаты диссертационной работы обсуждались в рамках научно-практических конференций СПбГПУ, а также на семинарах «29-я, 31-я, 32-я неделя науки» СПбГПУ и были доложены на II Международной научно-практической конференции «Информационные технологии в моделировании и управлении» и на IV Всероссийской научно-методической конференции "Фундаментальные исследования в технических университетах".

Внедрение

Разработанная методика, инструментальные средства поддержки и готовые программные комплексы внедрены в ООО «Деловые консультации, СПб», ЗАО

«Телекомбанк», АКБ «Связь-банк», ЗАО «РБД» и в рамках учебного курса «Базы данных» кафедры «Информационные и Управляющие Системы» СПбГПУ.

Публикации

По материалам диссертации опубликовано 10 работ.

Структура и объем работы

Диссертация содержит 180 страниц основного текста, 31 рисунок, 30 таблиц и состоит из введения, четырех глав, заключения, списка литературы и двух приложений.

Содержание работы

Во введении обосновывается актуальность работы, формулируются цель, задачи, объект и предмет диссертационного исследования, показана научная новизна и практическая ценность полученных результатов. Приводится структура диссертации.

Первая глава посвящена изучению различных аспектов аналитической деятельности предприятий, на примере коммерческих банков.

Очевидно, что наличие в банке ИАС, которая решает разнообразные аналитические задачи, на порядок повышает эффективность работы. Но, тем не менее, более половины банковских институтов Российской Федерации не имеют полноценных ИАС. Для того, чтобы понять, почему складывается такая неблагоприятная ситуация, в работе был проведен сравнительный анализ ИАС банковской деятельности.

Среди достоинств рассматриваемых ИАС можно выделить высокую функциональность и широкие возможности анализа данных, а так же высокую скорость реакции на запросы пользователей. Недостатками является то, что все эти системы привязаны к транзакционным системам, технологическим процессам и методикам расчета и анализа конкретного предприятия. Изменение требований применительно к таким системам влечет за собой значительные затраты на доработки и модернизацию.

Автор диссертационной работы считает, что основная причина недостатков состоит в том, что в процессе построения ИАС разработчики не имеют возможности опираться на формальные модели структур данных и единую методику их построения. Именно из этих суждений следует постановка задачи и цели работы.

Наличие формальных моделей позволит более эффективно управлять параметрами ИАС в процессе ее построения, сократить затраты на разработку и, как следствие, получить дополнительные возможности повышения качества системы. Наличие единой методики построения ИАС даст возможность проектировать структуры данных, не привязываясь к

бизнес-процессам конкретного предприятия. Это позволит уменьшить затраты на создание ИАС и сократить стоимость ее сопровождения.

В рамках первой главы проведен сравнительный анализ концепций построения ИАС и выбран ряд концепций для дальнейшего использования в работе. Среди выбранных концепций следует отметить концепцию хранилищ данных и витрин данных (рис. 1).

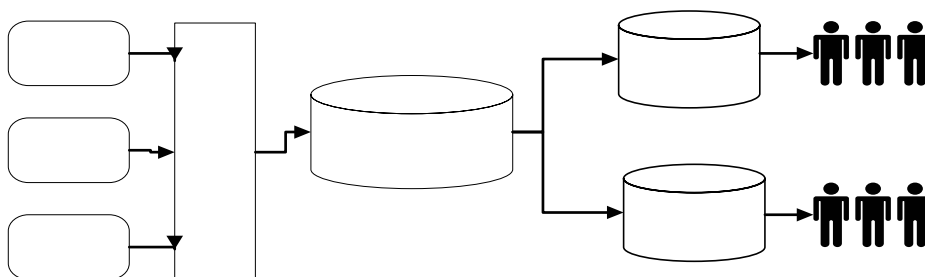


Рисунок 1. Концепция ХД

Используя данную концепцию при реализации ИАС разработчикам важно определить стратегию ее построения и средства разработки. В работе предлагается спиральная (итерационная) стратегия и различные критерии выбора программных средств реализации ИАС.

Во второй главе автор предлагает к рассмотрению собственную методику построения ХД и ИАС, полученную на основе анализа:

- концепций организации ИАС и ХД;
- банковской деятельности и требований к анализу и отчетности;
- принципов и примеров построения подобного рода систем;

В рамках сравнительного анализа различных методик построения ХД были выделены классическая методика Б.Инмона и методика построения корпоративных ХД Э. Спирли. Анализ показал, что все рассматриваемые методики имеют следующие недостатки:

- привязка проекта ХД к конкретному предприятию;
- отсутствие формальных моделей построения ХД;
- построение ХД только для крупных предприятий.

Эти недостатки перекликаются с недостатками конкретных ИАС банковской деятельности, которые рассмотрены в главе 1, поэтому можно говорить о том, что рассматриваемые ИАС так или иначе построены с использованием данных методик. Диссертационная работа как раз и направлена на устранение вышеуказанных недостатков существующих методик.

Сформулированы основные цели, которые должна преследовать предложенная автором методика:

1. ИАС, построенные на основе данной методики не должны быть привязаны к конкретному предприятию. Гибкость настройки и учет специфики деятельности должен позволять применять построенную ИАС к решению аналогичных задач в различных предприятиях (мобильность).
2. При изменении требований, законодательства и других факторов необходимо минимизировать существенные изменения структуры хранилища (сопровождаемость).
3. Построенная ИАС должна обладать достаточно высокими характеристиками, которые отражают объемы, времена ввода, обработки, хранения и выдачи информации (эффективность).
4. Система должна обладать минимальной избыточностью и высоким качеством данных за счет достаточно подробной формализации проектирования и функциональных процессов, протекающих в ИАС.

На этапе формирования методики за основу был взят спиральный жизненный цикл, на каждой итерации которого используется схема Захмана. В рамках данной модели различные участки проекта, такие как данные, функции, сеть и взаимодействие с внешней средой, рассматриваются в разрезе отдельных уровней жизненного цикла ИАС: контекстного, исследовательского, концептуального, логического, физического и пользовательского.

Данная схема позволяет разделить сложные методологии и правила на более простые и понятные. Кроме того, на ее базе можно реализовать и управление проектом. Описание предлагаемой методики в рамках схемы Захмана приведено в табл. 1.

Таблица 1 Схема Захмана, как отображение одной итерации построения ХД

	Информационная модель. Данные	Функциональная модель. Функции	Сетевая модель. Взаимодействие.
Контекстный уровень	Модель данных высокого уровня	Функциональные требования	Описание внешней среды
Концептуальный уровень	Архитектура	Функции архитектурных блоков	Модель метаданных
Логический уровень	База данных	Логика	Метаданные
Физический уровень	Программное обеспечение и реализация	Аппаратное обеспечение и реализация	Вспомогательное обеспечение
Пользовательский уровень	Пользовательский интерфейс	Технология и функции системы	Настройки

Предложенная методика разбита на 6 этапов и описывает процесс от определения необходимости построения ХД и ИАС до разработки и тестирования готовой системы (табл. 2). Выделенные этапы являются наиболее значимыми при построении ИАС, и именно на этих этапах автором диссертационной работы был *впервые* предложен аппарат проектирования, использование которого позволило устранить недостатки существующих методик.

Таблица 2. Методика построения ХД и ИАС.

В рамках предложенного аппарата следует выделить процесс построения **иерархии функциональных требований (ИФТ)** на *этапе 2*. Разработчики, сталкиваясь с новой предметной областью для автоматизации, должны провести детальное исследование и формирование требований к будущей системе. Аналитическая деятельность сложна и многообразна, поэтому процесс ее обследования должен носить четкий формальный и систематизированный характер. Для обследования предметной области и построения функциональной модели предлагается использовать матрицу ИФТ (рис. 1).

На вершине ИФТ расположены результирующие показатели, на следующем уровне сущности, на основе которых рассчитывается результирующий показатель. Одной из “контрольных точек” ИФТ являются базовые показатели, которые включают сущности, используемые при расчете результирующего показателя.

Результирующие и промежуточные показатели выявляются путем анализа предметной области. Базовые показатели из ИФТ выявляются по максимальному соответствию с понятиями транзакционных систем.

№ Название этапа

1 Определение

2 необходимости

3 построения ХД

4 Выбор и

5 обследование

6 Выбор

Гло
Вос

Поскольку базовые показатели напрямую связаны с понятиями внешних источников, они будут определять **информационную структуру** ХД. Промежуточные показатели будут определять **технология расчета и подготовки данных** для анализа. Результирующие показатели будут определять **подготовленные данные**, над которыми будут проходить процессы анализа.

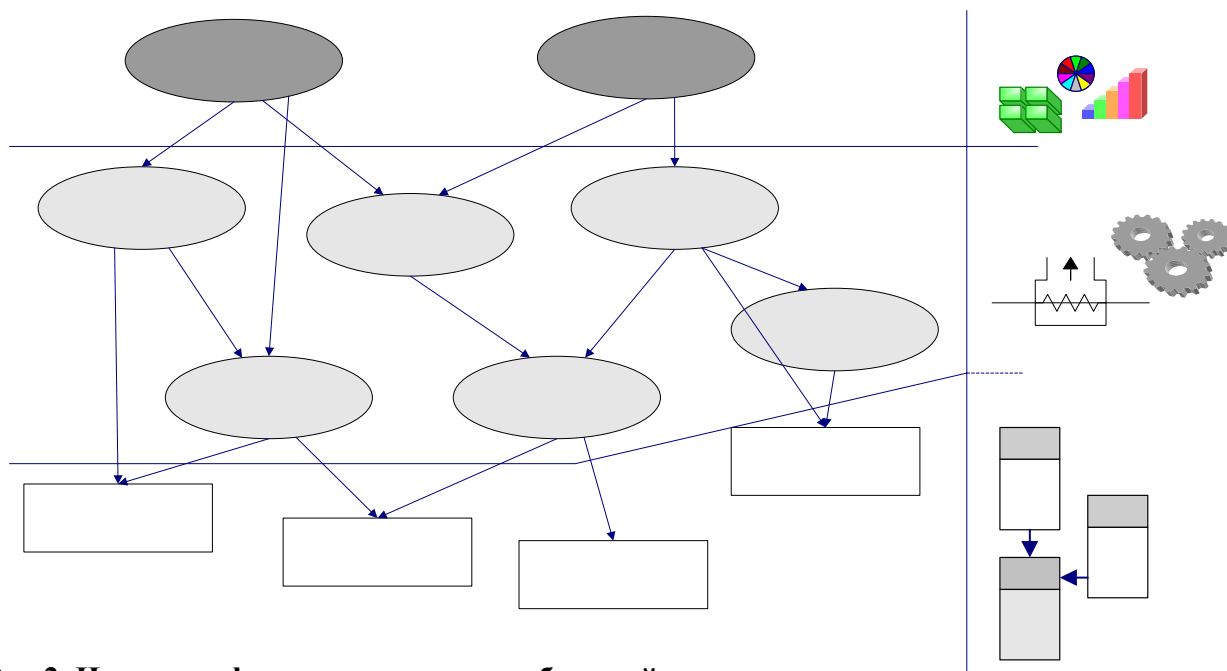


Рис 2. Иерархия функциональных требований.

Следуя основным принципам, на основе которых разрабатывается данная методика можно отметить, что понятия внешних источников являются **базовыми** понятиями банковской деятельности, целевая функция которых практически не изменяется со временем, поэтому информационная структура ХД, построенная на основе данных понятий, будет также неизменной, что приведет к устойчивости ХД к изменению требований.

На *2 этапе* разработчики определяют модели (функциональную, информационную и сетевую) высокого уровня, а на *4 этапе* они преобразуются в модели более низкого уровня - логические. Корректно построенная логическая архитектура будет определять важнейшие параметры ИАС, такие как время реакции, объемы данных, качество данных, время загрузки данных. Для решения задач данного этапа автором предложена формальная модель и правила проектирования ХД.

При проектировании ХД перед разработчиком стоит задача создания базы данных, отвечающей поставленным требованиям и содержащей данные из множества внешних источников предприятия. При условии, что все исходные данные можно представить с использованием реляционной модели, а ХД вторично по отношению в ВИ, можно использовать **новый** подход при проектировании ХД, предложенный автором

диссертационной работы. Это подход заключается в описании ХД на основании зафиксированных структур данных ВИ и правил формального преобразования.

Данные из ВИ состоят из объектов. Объект ВИ является отношением (или в терминах современной СУБД таблицей), определенной на множестве ВИ.

Отношение может быть представлено следующей тройкой (рис. 3а): $O = \{R, \Delta, \Lambda\}$, где:

R - описание структуры отношения или атрибутов, входящих в отношение;

Δ - множество потенциальных ключей отношения, которое определяет набор атрибутов, на которые ссылаются другие отношения;

Λ - множество внешних ключей отношения, которое определяет ссылки данного отношения на другие.



Рис. 3. Отношение (а) и отношение в функционально-зависимой форме (b)

Любое отношение можно представить в форме функциональной зависимости (рис 3б). $O = \{D \rightarrow F, \Delta, \Lambda\}$ (1)

, где $D \rightarrow F$ есть выражение, описывающее структуру отношения. D называют детерминантом отношения, а F - функционально зависимой частью.

Определение: D функционально определяет F , тогда и только тогда, когда каждое значение D связано в точности с одним значением F .

Функциональные зависимости широко применяются при проектировании базы данных (нормализация путем синтеза и через декомпозицию). Но в данном случае, выражение отношений в форме функциональных зависимостей, применяется к уже спроектированным данным (отношениям).

Необходимо формализовать процессы преобразования данных внешних источников в данные хранилища. Для того чтобы формально описать процессы преобразования, отношения во ВИ должны быть приведены к виду функциональной зависимости, который определен в (1), причем атрибуты отношения, входящего в функциональную зависимость должны быть приведены к упрощенной форме, то есть имя домена должно совпадать с именем атрибута.

Задача преобразования сводится к процессам объединения описанных отношений в отношения хранилища данных. Причем это объединение должно соответствовать

требованиям, которые предъявляются к ХД. Кроме того, все отношения должны быть приведены к единому формату по доменам и по системам отсчета.

Для описания процессов преобразования данных были использованы правила, выведенные на основании правил Армстронга и Дарвена.

Правило №1 Простое объединение.

Пусть имеются два отношения:

$$O1 = \{(D1 \rightarrow F1), \Delta1, \Lambda1\} \text{ и } O2 = \{(D2 \rightarrow F2), \Delta2, \Lambda2\}$$

Если детерминанты этих отношений совпадают $D1 = D2$, то данные отношения можно объединить, в результате чего получится отношение объединения (рис 4).

$$O12 = \{(D1 \rightarrow F1F2), \Delta1\Delta2, \Lambda1\Lambda2\} \quad (2).$$

У отношения объединения детерминанты пересекаются, и при пересечении остается один из детерминантов (поскольку они равны), функционально зависимые части объединяются, при этом дублирующие атрибуты так же объединяются, потенциальные ключи и внешние ключи объединяются.

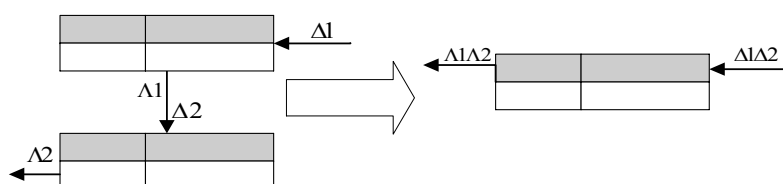


Рис 4. Правило простого объединения

Замечание. Операция объединения дублирующих атрибутов представляется следующим образом: пусть A, B два множества атрибутов, тогда $AB = A \cup B - A \cap B$.

Правило №2 Всеобщее объединение

Пусть имеются два отношения:

$$O1 = \{(D1 \rightarrow F1), \Delta1, \Lambda1\} \text{ и } O2 = \{(D2 \rightarrow F2), \Delta2, \Lambda2\}$$

Тогда существует отношение, которое является результатом всеобщего объединения двух отношений.

$$O12 = \{(D1(D2 - F1) \rightarrow F1F2), \Delta1\Delta2, \Lambda1\Lambda2\} \quad (3)$$

У отношения всеобщего объединения детерминанты объединяются, причем из второго детерминанта вычитаются атрибуты первой функционально зависимой части, функционально зависимые части объединяются, так же объединяются потенциальные и внешние ключи.

Это правило говорит о том, что можно объединить два любых отношения.

Правило №3 Денормализация вертикальных структур

Вертикальная структура используется для хранения неких параметров и их свойств, с учетом того, что свойств может быть неограниченное количество. В основном такие структуры состоят из 3 отношений (рис. 5a):

- Отношение описания параметра **P**;
- Отношение описания свойств всех параметров **S**;
- Отношение значений свойств для параметров **V**.

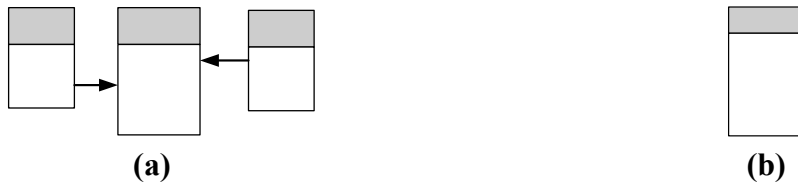


Рис. 5 Вертикальная структура данных (a) и горизонтальная структура данных (b)

При денормализации вертикальных структур особое внимание следует обращать на отношение V - именно оно содержит всю необходимую информацию и имеет т.н. вертикальную структуру.

$$V = \{ (p_id , s_id \rightarrow val), (p_id , s_id), ((p_id), (s_id)) \}$$

Детерминант этого отношения есть идентификатор параметра и идентификатор свойства. Функционально зависимая часть содержит значение конкретного свойства для конкретного параметра. Необходимо в первую очередь денормализовать данное отношение, т.е преобразовать его в горизонтальный вид, а затем соединить полученное отношение с отношением параметров **P**.

Сначала необходимо выявить, какие свойства мы будем преобразовывать в горизонтальный вид, определим эти свойства как

$$Gs = \{gs1, \dots, gsN\} \quad (4).$$

Набор получившихся свойств имеет тип данных такой же, как тип данных атрибута val отношения **V**. При необходимости можно применить к полученному списку атрибутов операцию приведения к единому формату $gsi' = F(gsi)$, $Gs' = \{gs'1, \dots, gs'N\}$.

Типы данных для приведения можно формально определить либо по отношению **S**, либо исходя из смыслового содержания свойства. Далее будем считать, что данный набор атрибутов уже приведен к единому формату по типам данных.

После выявления этого списка преобразуем отношение V в отношение Gv , которое будет иметь т.н. горизонтальную структуру

$$Gv = \{ (p_id \rightarrow gs1, \dots, gsN), (p), (p) \}$$

Детерминантом этого отношения является идентификатор параметра и м.б. другие атрибуты, функционально зависимая часть представлена списком атрибутов, который мы определили выше. Внешний ключи ссылаются на отношение параметров, потенциальный ключ есть детерминант. Далее полученное отношение G_V соединяется с отношением P по правилу всеобщего объединения. В результате соединения этих двух отношений получим искомое денормализованное отношение (рис. 5b)

$$GVP = \{(p_id \rightarrow gs1, \dots, gsN, \dots), \Lambda, \Delta\}; \quad (5)$$

Правила объединения и денормализации данных, описанные выше, необходимо применять осознанно, принимая во внимание требования к ИАС, предельные объемы данных хранилища и адекватные времена ответа на запросы, формируемые к данным. Если при проектировании ХД эти аспекты не принимаются во внимание, то можно получить ХД не удовлетворяющее поставленным задачам.

В работе сформулированы некоторые критерии объединения отношений. Отношения могут объединяться если:

- они соответствуют условиям правил объединения и ИФТ;
- мощность пересечения предполагаемого объединения по запросам достаточно высока;
- объем предполагаемого объединения менее предельно допустимого объема, определенного для данного объединения;
- совокупное время реакции по запросам, выполняемым над данным объединением, должно быть минимальным.

Учитывая данные критерии можно увидеть, как меняются характеристики отношений и запросов при применении правил объединения и денормализации.

Применение формальной модели при построении ХД дает разработчикам следующие преимущества:

- возможность предварительной оценки объемов ХД;
- возможность управления параметрами ХД;
- дополнительные возможности оптимизации (уменьшение избыточности, управление индексацией, увеличение времен реакции);
- возможность автоматизации процесса проектирования логической структуры ХД;
- увеличение скорости проектирования.

В третьей главе показаны реальные примеры использования разработанной методики и формальной модели для построения ИАС банковской деятельности. Предложенная методика была внедрена в процесс организации – разработчика программных систем, в результате чего был разработан ряд ИАС. Среди них хочется отметить несколько систем,

автоматизирующих аналитическую деятельность банков. Это три системы, построенные в рамках одного жизненного цикла на различных его итерациях – ИАС “Бюджет клиента”, ИАС “Регистры налогового учета” и ИАС “Ликвидность”. Так же была построена ИАС для обеспечения аналитической деятельности торгового предприятия, что показывает возможность применения предложенной методики не только для построения ИАС банковской деятельности, но и для других отраслей.

В рамках применения методики в главе 3 подробно рассматривается процесс создания ИАС “Бюджет клиента”. Основная задача этой системы формирование бюджетов клиентов в виде отчетных форм, их анализ и, как следствие, управление процентными ставками по кредитам и депозитам.

На *этапе 2* была построена ИФТ для выбранной предметной области, что позволило максимально быстро и эффективно выявить данные для анализа (результатирующие показатели), промежуточные показатели, определившие технологию расчета, и базовые показатели, определившие модель данных высокого уровня.

На *этапе 4*, при формировании информационной логической структуры ХД применялись правила и критерии, определенные в формальной модели, в результате чего была построена структура ХД и оптимизированы основные параметры ИАС.

Говоря о количественной оценке применения формальной модели, следует отметить, что при использовании правил, сформулированных на ее базе, удалось максимально сократить количество отношений ХД, что позволило уменьшить избыточность и увеличить качество данных, добиться оптимизации времени реакции ХД, времени загрузки данных в хранилище и объемов данных. Так, например, для объекта “счета”, нам удалось 16 отношений ВИ преобразовать в одно, для объекта “клиенты” из 5 отношений, представляющих собой вертикальную структуру, сформировано одно отношение горизонтальной структуры.

В результате внедрения и промышленной эксплуатации этой ИАС в банках были получены следующие результаты:

- С внедрением "Бюджета клиента" время расчета стало в 26 раз меньше, чем время расчета бюджета без использования ИАС.
- Время анализа и "подгонки" стало в 31 раз меньше.
- В процессе расчета бюджета стало участвовать менее половины сотрудников, которые были задействованы в этом процессе ранее.
- Внедрение ИАС принесло банку дополнительную выгоду в части, касающейся более оперативного принятия стратегически важных решений, а именно регулирования процентных ставок по кредитам и депозитам.

Так же следует заметить, что при разработке других ИАС в рамках этого жизненного цикла, логическая структура ХД практически не изменялась. То есть можно говорить о том, что применение формальной модели и механизма ИФТ позволяет строить ХД универсальной структуры, в рамках выбранной предметной области.

Проекты по созданию ХД, выполненные на базе предложенной методики показали ее работоспособность.

В главе 4 для получения оценки эффективности был проведен анализ критериев качества ИАС и возможностей их оптимизации в рамках разработанного аппарата формального описания. Среди них выделены эффективность и трудоемкость.

В рамках оценки эффективности были использованы следующие критерии, значения которых определяются экспериментальным путем:

- среднее время реакции для решения типичных задач;
- среднее время закачки данных в хранилище при постоянном объеме ХД;
- объемы ХД в различные периоды функционирования ИАС.

А так же были использованы такие критерии, как степень избыточности, денормализации и структурная сложность, значения которых определяются метрически.

В рамках оценки трудоемкости были использованы следующие критерии:

- Коэффициент переиспользования $K_{n1 \rightarrow 2} = t_{\text{сд}} \div (t_{\text{хд}} + t_{\text{зд}} + t_{\text{сд}})$;
- Стоимость сопровождения $T_c = N * A * t_{\text{хд}}$;
- Стоимость внедрения ИАС $T_e = N * A * (t_{\text{зд}} + t_{\text{сд}})$;

При сравнении ИАС, построенных с использованием предлагаемой методики (группа1) с ИАС сторонних производителей (группа2) по сформированным критериям были получены следующие результаты:

- Среднее время реакции, время загрузки данных, объемы данных ИАС группы 1 превышают данные параметры ИАС группы 2, что говорит о более высокой эффективности разработанных систем.
- Эти результаты обусловлены сравнительно низкой степенью избыточности, структурной сложности и высокой степенью денормализации, за счет применения формальной модели.
- Трудоемкость разработки среди ИАС группы 1 на первой итерации сравнительно выше, на последующих гораздо ниже соответствующих коэффициентов ИАС группы

2. Это говорит о том, что предложенную методику целесообразно применять для разработки целого ряда ИАС в рамках одного жизненного цикла.

В заключении дается итоговая оценка проделанной работы, и формулируются основные результаты работы.

Основные результаты работы

В работе проведен сравнительный анализ существующих ИАС и методик их построения, указаны их основные достоинства и недостатки. Рассмотрены различные концепции построения и реализации ХД и ИАС, проведен сравнительный анализ подходов к построению ХД, в ходе которого выявлена необходимость применения единых методик и формальных подходов к построению ИАС.

Разработана методика построения ХД, в рамках которой предлагается использование схемы Захмана на каждой итерации жизненного цикла.

Предложен механизм построения ИФТ, который дал разработчикам возможность формального описания предметной области в привязке к функциональности ИАС.

Предложен новый подход к проектированию ХД, который заключается в использовании структур данных внешних источников.

Разработана формальная модель ХД, в рамках которой проведено расширение реляционной алгебры, в результате которого выведены правила и критерии проектирования ХД.

Разработанная методика и формальная модель были успешно внедрены в процесс организации – разработчика, в результате чего были построены ряд ИАС для торговых и кредитных организаций. Это показывает работоспособность методики.

Проведена оценка предлагаемой формальной модели, результаты которой показали преимущества ее применения для построения ХД, среди которых следует выделить уменьшение времени реакции и уменьшение планируемых объемов данных.

Проведено сравнение построенных ИАС с ИАС сторонних производителей, результаты которой показали высокие показатели качества.

На основе анализа результатов внедрения методики делается вывод о целесообразности ее применения для построения ИАС для решения многих аналитических задач, применительно к выбранной предметной области.

Содержание диссертации отражено в следующих публикациях:

1. Зотов А.А., Абдрахманов Р.Л., Басин В.В., Дробинцев Д.Ф. Разработка системы управления банковским счетом через Интернет // XXIX неделя науки СПбГТУ. Ч.IV: Материалы межвузовской научной конференции. СПб.: СПбГТУ, 2000. С. 15 - 16.
2. Зотов А.А., Басин В.В., Дробина Ю.А. Использование архивирования в комплексных системах автоматизации. // Информационные технологии в моделировании и управлении: Труды II Международной научно-практической конференции, 20-22 июня 2000 года. СПб: СПбГТУ, 2000. С. 11 - 12.
3. Зотов А.А., Капустин В.В. Некоторые аспекты разработки информационно-аналитических систем банковской деятельности // Материалы IV Всероссийской научно-методической конференции "Фундаментальные исследования в технических университетах". СПб.: СПбГПУ, 2000. С. 18 - 20.
4. Зотов А.А., Кракау Т.К., Капустин В.В. Архитектура хранилища данных многофилиального банка. // XXXI неделя науки СПбГПУ. Ч.IV: Материалы межвузовской научной конференции. СПб.: СПбГПУ, 2003. С. 10 - 11.
5. Зотов А.А., Бутенко И.В. Архитектура построения автоматизированной банковской системы для многофилиального банка. // XXXI неделя науки СПбГПУ. Ч.IV: Материалы межвузовской научной конференции. СПб.: СПбГПУ, 2003. С.8 - 9.
6. Зотов А.А., Дробинцев Д.Ф., Бутенко И.В. Разработка подсистемы формирования данных для Интернет-банка. // Конкурс – конференция студенческих работ в области современных технологий программирования в СПбГПУ. СПб.: СПбГПУ, 2003. С. 9-10.
7. Зотов А.А., Пархоменко С.А., Иванов А.В. Разработка набора СОМ-объектов для MS DTS, реализующего загрузку информации из текстовых файлов в таблицы метаданных. // XXXII неделя науки СПбГПУ. Ч.V: Материалы межвузовской научно-технической конференции. СПб.: СПбГПУ, 2004. С. 29 - 30.
8. Зотов А.А., Пархоменко С.А., Воинов Н.В. Реализация подсистемы взаимодействия метаданных библиотеки разработки СКАУТ v4.5 и объектов Microsoft Analysis Services. // XXXII неделя науки СПбГПУ. Ч.V: Материалы межвузовской научно-технической конференции. СПб.: СПбГПУ, 2004. С. 37 - 38.
9. Зотов А.А., Пархоменко С.А., Ульянов С.Г. Разработка шаблонов реализации СОМ-объектов MS DTS. // XXXII неделя науки СПбГПУ. Ч.V: Материалы межвузовской научно-технической конференции. СПб.: СПбГПУ, 2004. С. 36 - 37.
10. Зотов А.А. Принципы проектирования аналитической системы в условиях быстро меняющихся требований // XXXII неделя науки СПбГПУ. Ч.V: Материалы межвузовской научно-технической конференции. СПб.: СПбГПУ, 2004. С. 38 - 40.