

На правах рукописи

АНТОНОВ Вячеслав Викторович

**ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА
НА ОСНОВЕ ЗНАНИЙ РЕГИСТРАЦИОННОГО УЧЕТА НАСЕЛЕНИЯ**

**Специальность: 05.13.10 «Управление в социальных
и экономических системах»**

**АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук**

Уфа 2007

Работа выполнена на кафедре автоматизированных систем управления
Уфимского государственного авиационного технического
университета

Научный руководитель

д-р техн. наук, проф.

КУЛИКОВ Геннадий Григорьевич

Официальные оппоненты

д-р техн. наук, проф.

КАБАЛЬНОВ Юрий Сергеевич

канд. техн. наук, доц.

ДУЛЕНКО Вячеслав Алексеевич

Ведущее предприятие

**Башкирский государственный
университет**

Защита диссертации состоится « 18 » мая 2007 г. в 10⁰⁰ часов
на заседании диссертационного совета Д-212.288.03
в Уфимском государственном авиационном техническом университете
по адресу: 450000, г. Уфа, ул. К. Маркса, 12

С диссертацией можно ознакомиться в библиотеке университета

Автореферат разослан « 13 » апреля 2007 г.

Ученый секретарь
диссертационного совета
д-р техн. наук, проф.



В.В. Миронов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Традиционно в большинстве организаций непроизводственной и социальной сферы формировались независимые центры актуализации данных. Это обуславливалось в первую очередь необходимостью автоматизации независимых друг от друга проблемно-ориентированных технологических процессов, составляющих основные направления деятельности организации. Ввиду важности информации для каждого отдельно взятого заинтересованного подразделения, создавались локальная база данных и свой центр ее актуализации. Исторически сложившиеся технологии обработки данных, значительные инвестиции в аппаратное и программное обеспечение не позволяют кардинально решить проблему их дифференциации путем механического переноса данных в центральную базу данных.

Актуальной темой становятся технологии и программные продукты, которые могут обеспечить эффективную интеграцию баз данных для представления информации в виде, удобном для анализа и использования при принятии управленческих решений в интересах всей организации, обеспечения не только унифицированного доступа к продолжающим функционировать центрам актуализации данных, но и обеспечивающим создание инфраструктуры для доступа к данным, опирающихся на единые стандарты и единые принципы сетевого взаимодействия.

Накопленные к настоящему времени колоссальные объемы информации, в совокупности с непрерывно увеличивающимися темпами ее роста определяют актуальность и значимость исследований в данной области.

Целью диссертационной работы является концептуальная разработка распределенной учетной системы формирования хранилища данных для аналитического анализа по правилам нечеткой логики для информационной поддержки принятия управленческих решений для органов государственной власти и местного самоуправления.

Задачи исследования. Для достижения цели работы поставлены и решены следующие задачи:

1. Разработать структуру распределенной учетно-аналитической системы для формирования хранилища данных в режиме мониторинга.
2. Разработать структуру хранилища данных и правила семантического анализа для формирования множества локальных проблемно-ориентированных баз знаний на основе нечеткой логики с применением OLAP технологий.
3. Провести анализ и разработать требования к программной реализации модели учетно-аналитической системы на основе нечеткой логики и OLAP технологий.
4. Оценить эффективность предложенной системы для информационной поддержки принятия управленческих решений для органов государственной власти и местного самоуправления на примере системы регистрационного учета населения.

Методы исследования. Обоснованность результатов диссертационной работы основывается на корректном применении методов системного анализа, теории организационного управления, теории процессных методов управления, теории автоматизированного проектирования информационных систем, теории организации баз данных и баз знаний, технологии высокоуровневого программирования.

На защиту выносятся:

1. Метод формирования распределенной учетной системы на основе структурной формализации и дальнейшей интеграции бизнес-процессов и процессов управления по семантическим правилам для формирования хранилища данных.

2. Модель хранилища данных для автоматизированной учетной системы на основе процессного анализа и классификации атрибутов предметной области по правилам нечеткой логики.

3. Технология программной реализации распределенной учетно-аналитической системы отвечающей заданным требованиям надежности, защиты и производительности.

4. Оценки эффективности предложенных решений на примере регистрационного учета населения.

Научная новизна работы:

1. Предложен новый метод формирования структуры информационной системы в виде совокупности взаимодействующих вертикальных и горизонтальных семантически определенных и формализованных объектов, связанных друг с другом иерархическими отношениями классов атрибутов определяющих их бизнес-процессы, что открывает возможность целенаправленного развития системы.

2. Новизна предложенной концептуальной модели многоуровневой архитектуры хранилища данных, основанной на логическом и далее платформенном разделении функций учета и отчетности, заключается в применении нечетких правил перекрытия взаимодействующих объектов учета для формирования четких измерений OLAP куба.

3. Для обеспечения функциональной надежности хранилища данных, в отличие от традиционной программной реализации, предложено дополнительно использовать промежуточный склад данных, который является копией учетной базы данных.

4. Новизна предложенного метода оценки эффективности системы регистрационного учета населения, заключается в том, что появляется возможность количественной оценки снижения потерь ВВП за счет снижения времени получения справок населением.

Практическая ценность работы определяется возможностью использования результатов работы при проектировании и внедрении новых территори-

ально распределенных многоуровневых автоматизированных учетных систем, а также возможностью использования отдельных разработанных методов при разработке и исследовании широкого круга информационных задач.

Внедрение результатов. Основные результаты диссертационной работы внедрены в МВД РБ при решении задач автоматизации централизованных учетов, в учебный процесс УГАТУ при постановке учебных курсов «CASE технологии, теория систем и системный анализ», в учебный процесс УЮИ МВД РФ при преподавании ряда дисциплин кафедры управления в ОВД.

Апробация работы. Положения диссертации и результаты исследований докладывались на следующих всероссийских и международных конференциях:

- IX международная научная конференция «Информатизация правоохранительных систем», 2000 г., г. Москва.

- Республиканская конференция «Актуальные проблемы борьбы с преступностью в Республике Башкортостан», 2000 г., г. Уфа.

- НПК «Актуальные проблемы совершенствования правоохранительной деятельности в современных условиях», 2001 г., г. Челябинск.

- Всероссийское семинар-совещание «О задачах информационных центров МВД, ГУВД, УВД субъектов РФ по реализации концепции развития информационно-вычислительной системы МВД России на 2002-2006 годы», 2003 г., г. Казань.

- Международные научно-практические конференции «Computer Science & Information Technology, CSIT 2005» г. Уфа, УГАТУ и «CSIT 2006», г. Карлсруэ (Германия).

- Управление экономикой: методы, модели, технологии. Шестая Всероссийская научная конференция с международным участием, 2006 г., г. Уфа.

Публикации. Список публикаций по теме диссертации содержит 11 работ, в том числе 1 в рецензируемом журнале из списка ВАК.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, основных результатов и выводов, списка литературы, содержит 146 листов машинописного текста и включает 32 рисунка, 10 таблиц, 118 наименований использованных литературных источников.

СОДЕРЖАНИЕ РАБОТЫ

Во введении дается общая характеристика работы: цель исследований, актуальность решаемых задач, определяется научная новизна и практическая значимость защищаемых результатов.

Первая глава посвящена анализу существующих подходов и методов к построению учетных систем и систем аналитического анализа. Проводится анализ основных проблем при создании автоматизированных учетных систем в органах государственного управления. Находящаяся в распоряжении государства информация используется для определения, защиты или ограничения прав и

свобод граждан, и, тем самым, достоверность оперативность и безопасность сбора этой информации оказывают влияние на качество функционирования государственных институтов. С ведением того или иного вида учета связано выполнение большинства государственных функций, в том числе оказание всех государственных и социально-экономических услуг. Официальность учета в обязательном порядке обеспечивается нормативными актами. При отсутствии средств автоматизации для выполнения функций управления и принятия решений в социально-экономических системах неизбежно присутствие функциональных вертикалей, которые обеспечивают управление «сверху вниз» и в целом контроль, за каждым шагом управляющего воздействия. С точки зрения степени формализуемости государственные процессы и регламенты могут быть классифицированы на следующие 4 категории, приведенные на рис.1. В первую очередь подлежат автоматизации административные регламенты, рутинные и хорошо структурированные. Этот тип государственных процессов и регламентов очень похож на бизнес-процессы, характеризуется высокой степенью повторяемости, например выдача паспорта.



Рис. 1 – Категоризация административных регламентов и процессов с точки зрения степени формализуемости

Любой административный регламент и процесс может быть разбит (декомпозирован) на отдельные подпроцессы, операции и микрооперации.

Анализ типового, обобщенного административного процесса позволяет выявить общие для многих регламентов и процессов характеристики и на этой основе идентифицировать общие, базовые компоненты и сервисы. Интегрирование OLAP и учетных систем может осуществляться по-разному, возможно создание дублированных баз данных, витрин и хранилищ данных. Использование Хранилищ данных позволяет решить ряд проблем, возникающих при интеграции распределенных баз данных учетной системы. В архитектуре, одновременно использующей реляционные и многомерные системы, данные хранятся на OLAP-сервере или OLAP-структуры используются в качестве кэша для реляционных данных. Можно использовать комбинацию двух этих подходов, мини-

мизируя объем данных, перемещаемых из реляционной среды в многомерную и обратно. Обе модели дают возможности анализа.

Анализ существующих решений для комплексного внедрения автоматизированных территориально-распределенных учетных систем совместно с продуктами оперативного анализа (OLAP) для использования в органах государственного управления, не выявил готовых и апробированных решений в данной области, пригодных к внедрению без существенных доработок, что во многом связано с отсутствием установленных стандартов на ряд технологических процессов информационных технологий при решении задач управления и принятия решений в социальных и экономических системах. Решение указанных задач востребовано и актуально не только в области построения учетных информационно-поисковых систем, но и вообще в системах обработки и анализа информации. При реализации проектов по построению хранилищ данных возникает ряд общих задач, независимых от предметной области: проектирование структуры иерархических измерений; проектирование структуры медленно меняющихся измерений; проектирование и актуализация агрегатных значений. Все это обуславливает актуальность и значимость исследований в области процессного анализа предметной области и разработки формальных методов определяющих структуры хранилищ данных.

В ходе анализа предметной области **во второй главе** предложена структура информационной системы в виде совокупности взаимодействующих вертикальных и горизонтальных семантически определенных и формализованных объектов, связанных друг с другом иерархическими отношениями классов атрибутов определяющих их бизнес-процессы, предложен «эволюционный» подход внедрения, связанный с структурированием деятельности по сбору информации в форме совокупности организационно – функциональных моделей для бизнес-процессов, что позволяет применить методологию функционально-стоимостного анализа ABC (Activity Based Costing) для более адекватной оценки экономических затрат.

На основании результатов, приведенных анализов, сформулированы основные принципы концепции эволюции существующих систем учета предназначенных для повышения их эффективности:

Широкое применение электронных средств, требует повышения степени формализации организационной структуры. Основные информационные объекты, используемые в бизнес-процессах структурного подразделения сведены в матрицу операций над ресурсами и приведены в табл.1. Для информационного объекта определяется структура его жизненного цикла, назначаются роли и их исполнители в соответствии с организационной структурой и штатным расписанием. Для инициализации и управления исполнением бизнес-процессов они дополняются необходимыми функциями управления (нормирование, планирование, регулирование, учет) и соответственно ролями и исполнителями этих

функций. Таким образом, за каждым бизнес-процессом с информационным объектом закрепляется процесс управления:

$$\begin{aligned} MOP &= \{ИО\} \cap \{ПИП\} & МФУ &= \{ИО\} \cap \{ФУ\} & МУР &= \{Роли\} \cap \{ФУ\} \\ БП &= \{ИО, ПИП, ШС\} & ПУ &= \{БП, ФУ, ШС\}, \end{aligned} \quad (1)$$

где ИО – множество информационных объектов,
 ПИП – множество первичных информационных процессов,
 ФУ – множество функции управления,
 МОР – матрица операций над ресурсами,
 МФУ – матрица функций управления информационными объектами,
 МУР – матрица ролей для функции управления,
 БП – бизнес-процесс,
 ПУ – процесс управления,
 ШС – множество штатных единиц (штатная структура).

Таблица 1. Фрагмент матрицы операций над ресурсами

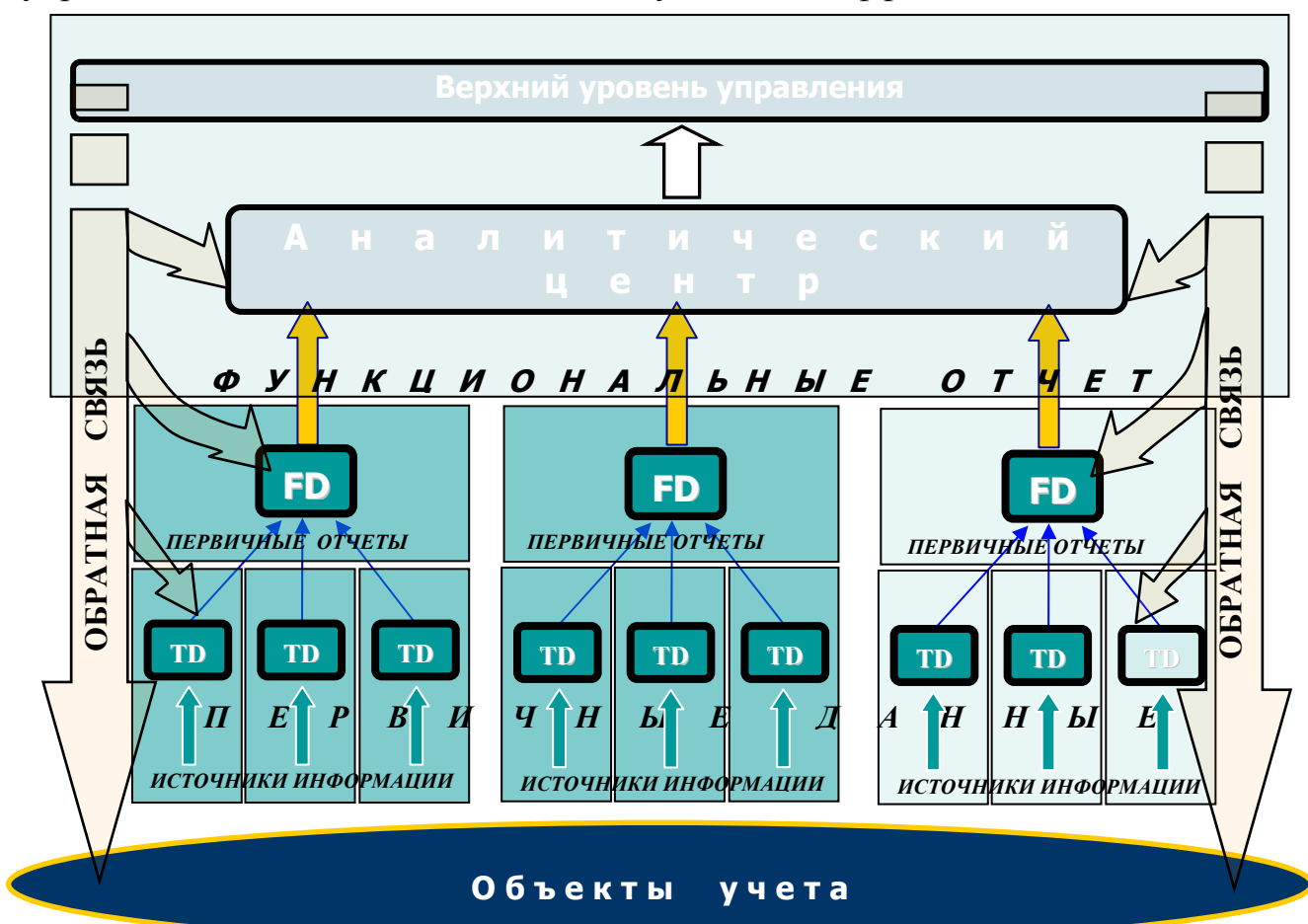
Первичные информационные процессы		сбор	преобразование	кодирование	...	использование
Обслуживаемая территория						
Вид учета 1						
	Объект учета 1	V		V		V
	Объект учета 2		V	V		V
Вид учета 2						
	Объект учета 1	V	V	V		
	Объект учета 2		V	V		V
	Объект учета 3					V
Вид учета 3						
	Объект учета 1	V	V	V		V
	Объект учета 2		V	V		V
	Объект учета 3					V
....						

2. Повышение степени структурной и параметрической адекватности модели объекта учета, путем регулярной организации анализа функциональной составляющей системы учета. Моделью знаний объекта учета является хранилище данных, актуализируемое во времени и пространстве горизонтальными и вертикальными бизнес-процессами. Таким образом, может быть реализован

адекватный процессный мониторинг за счет изменения, как самой структуры, так и хранилища информации.

3. Интеграцию репозитория моделей и хранилища данных на основе правил нечеткой логики для аналитического анализа с применением OLAP технологий.

Получена структура с условным разделением на «объекты» (рис.2) для каждого, из которых, возможен дифференцированный подход и установлены значения учетных регистров. Управление любой сложной социально-экономической системой невозможно без обратной связи, которая заключается в отслеживании и анализе данных, отражающих состояние этой системы и ситуацию вокруг нее. Постоянная доступность актуальной информации дает возможность оценить текущее положение дел, а обзор изменения конкретных характеристик во времени позволяет обнаружить тенденции развития. Такое управление основано на знании и потому наиболее эффективно.



TD – территориальное подразделение **FD** – функциональное подразде-

Рис.2 – Структура организации на основе формального разнесения по семантическим правилам на объекты

Определив бизнес-процессы выделенных объектов в качестве информационных объектов рассматриваемой модели, и рассматривая каждый независимый

бизнес-процесс как отдельную часть информационной системы, можем получить формализованное описание информационной схемы каждого объекта и множества всех информационных систем объектов. Проведя классификацию атрибутов: идентификационные (ключевые) атрибуты (обозначим A^1), функциональные атрибуты (обозначим A^2), неформализованные атрибуты (задающиеся вербальным описанием, имеющие качественные характеристики) (обозначим A^3), рассматриваемый бизнес процесс можем определить как множество упорядоченных пар:

$$x = \left\{ \langle a_1^1, d_1^1 \rangle, \dots, \langle a_k^1, d_k^1 \rangle, \langle a_1^2, d_1^2 \rangle, \dots, \langle a_l^2, d_l^2 \rangle, \langle a_1^3, d_1^3 \rangle, \dots, \langle a_m^3, d_m^3 \rangle : \right. \\ \left. k, l, m \in I, |x| = k + l + m \right\}. \quad (2)$$

Классификация бизнес-процессов по отношению к друг к другу при интеграции приведена на рис.3.

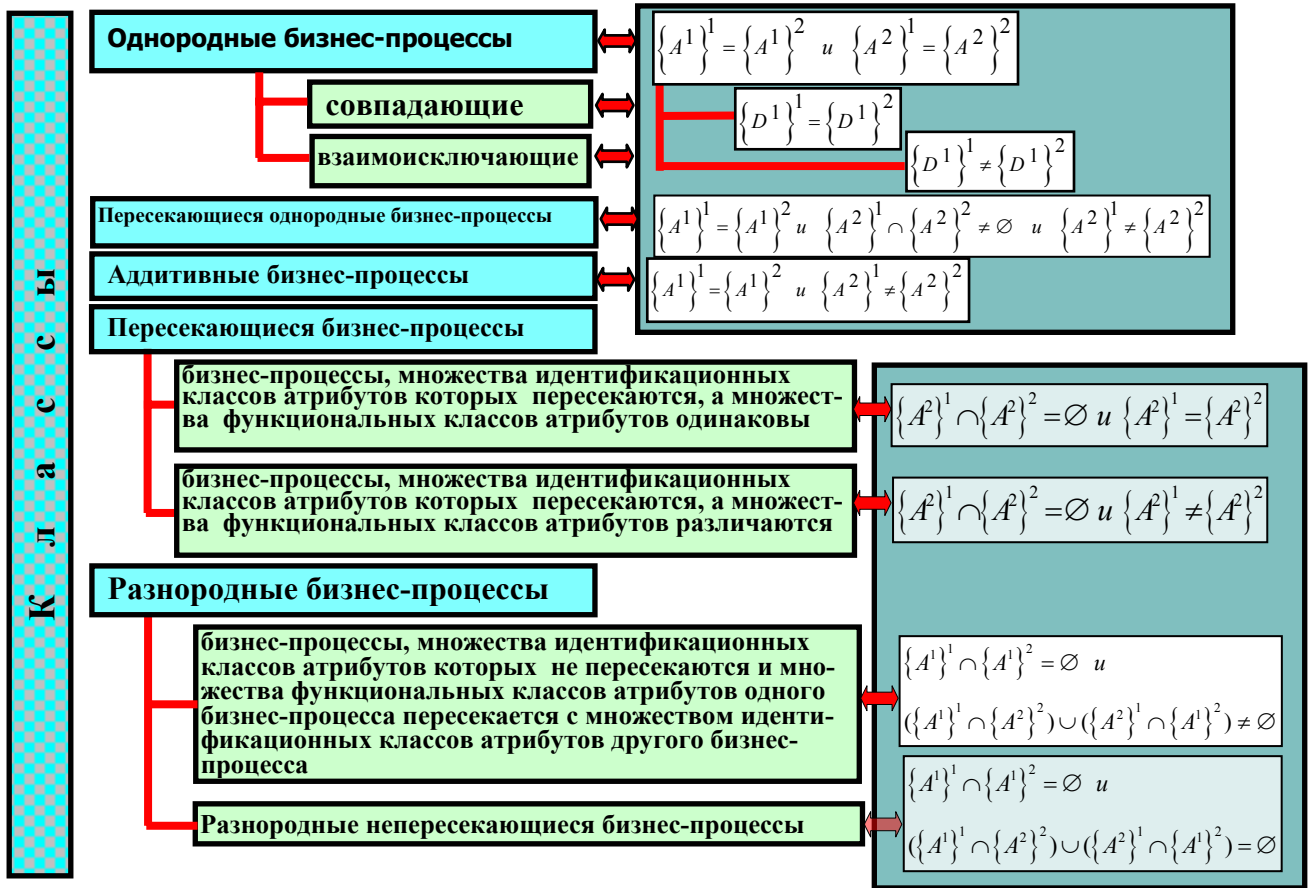


Рис.3 – Классификация по взаимодействию

В зависимости от принадлежности к тому или иному классу взаимодействия к подлежащим интеграции бизнес-процессам будет применяться соответствующее правило (рис.3) и для получаемого интегрированного бизнес-процесса обуславливаться изменением числа атрибутов множества идентификационных и функциональных классов:

- Для однородных бизнес-процессов и пересекающихся однородных бизнес-процессов интеграция производится без изменения числа атрибутов множества идентификационных и функциональных классов.

- Для аддитивных бизнес-процессов интеграция производится без изменения числа атрибутов множества идентификационных классов. Множество функциональных классов будет определяться формулой

$$\{A_1^2\}^1 = \{A^2\}^1 \cup \{A^2\}^2, \quad (3)$$

где A_1 – атрибут интегрированного бизнес-процесса.

Число атрибутов множества функциональных классов будет увеличено на количество несовпадающих атрибутов.

- Для пересекающихся бизнес-процессов интеграция производится по совпадающим атрибутам идентификационных классов, при этом множество идентификационных классов будет определяться формулой

$$\{A_1^1\}^1 = \{A^1\}^1 \cup \{A^1\}^2. \quad (4)$$

Число атрибутов множества идентификационных классов будет увеличено на количество несовпадающих атрибутов. Интеграция производится без изменения числа атрибутов множества функциональных классов.

- Для пересекающихся бизнес-процессов с элементами аддитивности интеграция производится по совпадающим атрибутам идентификационных классов, при этом множество идентификационных классов будет определяться формулой

$$\{A_1^1\}^1 = \{A^1\}^1 \cup \{A^1\}^2. \quad (5)$$

Число атрибутов множества идентификационных классов будет увеличено на количество несовпадающих атрибутов. Множество функциональных классов будет определяться формулой

$$\{A_1^2\}^1 = \{A^2\}^1 \cup \{A^2\}^2. \quad (6)$$

Число атрибутов множества функциональных классов будет увеличено на количество несовпадающих атрибутов.

- Для разнородных пересекающихся бизнес-процессов интеграция производится по совпадающим атрибутам идентификационных классов одного бизнес-процесса со значениями функциональных атрибутов другого бизнес-процесса. При этом множество идентификационных классов будет определяться формулой

$$\{A_1^1\}^1 = \{A^1\}^1 \cup (\{A^1\}^1 \cap \{A^2\}^2). \quad (7)$$

Число атрибутов множества идентификационных классов будет увеличено на количество совпадающих атрибутов идентификационных классов одного бизнес-процесса со значениями функциональных атрибутов другого бизнес-процесса. Интеграция производится без изменения числа атрибутов множества функциональных классов. Схематично данный процесс представлен на рис.4.

• Для разнородных непересекающихся бизнес-процессов интеграция производится в виде совместного хранения полностью независимых учетов.

Таким образом, отношения между взаимодействующими бизнес-процессами оцениваются двумя типами свойств: те, которые можно непосредственно измерить, и те, которые являются качественными и требуют попарного сравнения объектов, обладающих оцениваемым свойством, чтобы определить их место по отношению к рассматриваемому понятию.

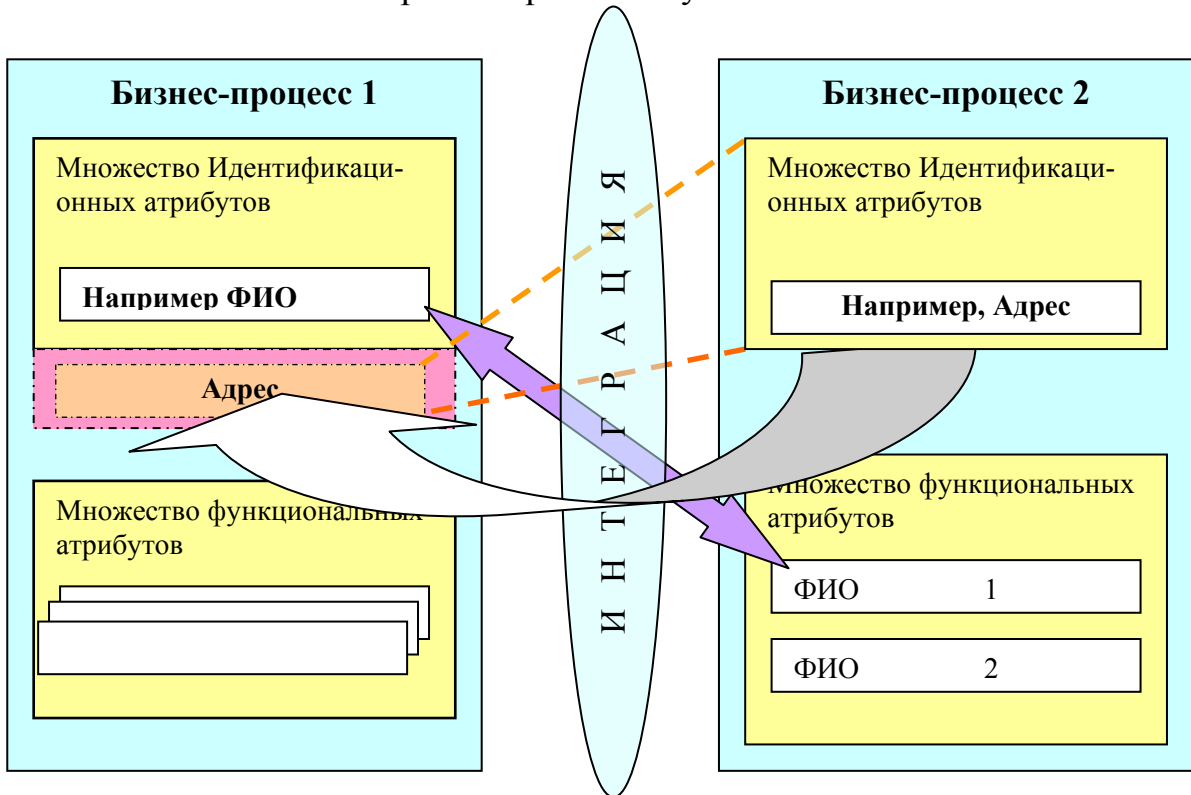


Рис. 4 – Интеграция разнородных, пересекающихся бизнес-процессов

Обработка и представление информации, применительно к этапам рассматриваемого процесса для количественных показателей атрибутов, может производиться применением прямого метода для одного эксперта построения функции принадлежности нечеткого множества, например предложенным Ч. Осгудом методом семантических дифференциалов.

Вопрос об интеграции на базе неформализованных атрибутов и непротиворечивости при этом информационной системы верхнего уровня сводится к рассмотрению зависимостей между типами входящих нее элементов. Каждому типу взаимодействующих объектов поставим в соответствие множество семантических зависимостей с другими типами объектов $C_i = \{c_{ij} : j \in I\}$.

Каждый тип взаимодействующих объектов T_i , элементы которого удовлетворяют свойству C_i , определяется как множество

$$T_i = \left\{ \langle A^1, D^1 \rangle, \langle A^2, D^2 \rangle, \langle A^3, \mu_{c_{ij}}(A^3) \rangle : j \in I \right\}, \quad (8)$$

где A – множество атрибутов, D – множество значений атрибутов.

Определенные таким образом правила перекрытия (пересечения) взаимодействующих объектов учета и полученные на пересечении множеств новые множества определяют, в конечном счете, возможность использования OLAP технологии на всем информационном пространстве интегрированного хранилища данных. Учитывая, что сами семантические ограничения (правила) могут быть заданы набором признаков в виде попарных сравнений, не для всех из которых существует однозначность ответа, для их анализа наиболее целесообразно использовать аппарат нечеткой логики.

Проведенный анализ позволяет свести задачу построения информационной аналитической системы к формальному алгоритму.

В третьей главе представлен анализ программных средств реализации многомерных хранилищ данных, рассмотрены основные модели данных и их характеристики. Отмечено, что эволюционно основными классическими моделями данных были иерархическая, сетевая и реляционная модели данных. В настоящее время развиваются и так называемые постреляционные подходы. Сделан обзор моделей данных, в ходе которого рассмотрены области применения и возможности иерархических и сетевых, реляционных и объектно-ориентированных моделей данных.

Приведена классификация систем управления базами данных (СУБД) по степени универсальности и модели данных. Рассмотрен алгоритм формирования OLAP куба на основе нечетких правил. Отмечено, что хранилища данных практически всегда создаются в качестве настройки над множеством существующих операционных систем, производящих сопровождение реляционных баз данных. Поэтому, наиболее целесообразно, оставаясь на реляционной платформе, воспользоваться преимуществами иерархического и сетевого подхода в процессах семантического моделирования структуры хранилища данных. OLAP-куб создается из соединения таблиц с применением схемы звезды. Модель данных состоит из двух типов таблиц: одной таблицы фактов, которая является центром звезды, и нескольких таблиц измерений, по числу измерений в модели данных - лучей звезды. В центре находится таблица, которая содержит ключевые факты, по которым делаются запросы. Множественные таблицы с измерениями присоединяются к таблице «фактов» и показывают возможные варианты анализа агрегированных данных. Иерархии в измерениях необходимы для возможности агрегации и детализации значений показателей согласно иерархической структуре. Базируясь на предложенной в исследовании классификации атрибутов и методе формирования структуры информационной системы в виде совокупности взаимодействующих вертикальных и горизонтальных семантически определенных и формализованных объектов, связанных друг с другом иерархическими отношениями классов атрибутов определяющих их бизнес-процессы, предложено использовать в качестве измерений при построении ку-

ба множеств, полученных согласно правил перекрытия взаимодействующих объектов учета.

Таблица фактов при этом содержит целочисленные колонки, дающие числовую характеристику каждого, определенного таким образом измерения, и несколько целочисленных колонок-ключей для доступа к таблицам измерений, которые их расшифровывают.

Для каждого измерения составляем список уникальных значений из элементов, хранящихся в столбцах и производим предварительное агрегирование фактов для записей, имеющих одинаковые значения размерностей. Используя промежуточные таблицы (так называемые кросс - таблицы) можем связать элементы разных таблиц между собой, для чего каждой записи в таблицах измерений поставим в соответствие список, элементами которого будут номера фактов, при формировании которых использовались эти измерения. Для фактов соответственно каждой записи поставим в соответствие значения координат, по которым она расположена в гиперкубе. Измерения имеют иерархическую структуру, состоящую из одного или нескольких уровней, на основании которой осуществляются операции свертки или детализации.

При введенных обозначениях:

$L_i : i=1, \dots, k$ – количество уровней иерархической структуры i -го измерения, k – количество измерений.

$N_i^j : j=1, \dots, L_i, i=1, \dots, k$ – количество агрегатов для j -го уровня i -го измерения.

Степень агрегации куба может быть представлена отношением реального количества агрегированных значений показателей измерений v^G к максимально возможному V^G

$$V_{ag} = \frac{v^G}{V^G} = \frac{\prod_{i=1}^k (\sum_{j=0}^{L_i} N_i^j) - \prod_{i=1}^k N_i^0}{\prod_{i=1}^k (\sum_{j=0}^{L_i} N_i^j + 1) - \prod_{i=1}^k N_i^0}. \quad (9)$$

Выбор метода агрегации данных зависит от конкретной решаемой задачи, подсчет агрегатов выполняется с использованием «карт агрегации», включающих стандартные методы агрегации. Учитывая предложенный способ построения измерений, подразумевающий наличие первичных данных на нижнем уровне иерархии, наиболее целесообразно использование при агрегации метода суммирования детализированных данных.

Реализация выбранного механизма группировки атрибутов полученных измерений при рассмотрении сформированного куба в более крупные измерения способствует существенному снижению его мерности и облегчению восприятия информации.

В связи с наличием нескольких центров актуализации, жестко регламентированных сроках и видах регламентной отчетности (по исходящей снизу вверх), наиболее целесообразно обеспечить одновременное хранение в рамках

интегрированной информационной системы верхнего уровня детализированных данных, метаданных и агрегированных в виде аналитических таблиц данных. Учитывая, наличие несовпадающих и независимых регламентов отчетности, необходимо использование многоуровневой архитектуры хранилища данных. Решение данной проблемы может быть получено путем введения слоя навигационных метаданных между физическим хранилищем и программными приложениями его обработки, который к тому же обеспечивал бы оперативную трансляцию реляционных данных. Предложено использование промежуточного склада данных, который является копией базы данных учетной системы. Данные промежуточного склада используются для создания OLAP-кубов и витрин данных по разным направлениям отчетности, с различным регламентом поступления и подготовки информации. Таким образом, построение хранилища данных выполняется в трехуровневой архитектуре:

На первом уровне расположены источники данных — внутренние регистрирующие, учетные и справочные системы.

На втором уровне содержится:

- центральное хранилище, куда стекается информация с первого уровня;
- промежуточный склад данных, который не содержит исторических данных и является источником аналитической информации для создания OLAP-кубов и витрин данных по разным направлениям регламентной отчетности.
- На третьем уровне расположен набор предметно-ориентированных витрин данных, источником информации для которых является центральное хранилище данных.

Сделан анализ путей повышения эффективности запросов к разреженному гиперкубу с заключением необходимости использования алгоритмов, позволяющих фиксирование получения пустой выборки на этапе формирования запроса.

В четвертой главе приведена оценка эффективности предложенной системы на примере регистрационного учета населения. Показано, что изложенная концепция создания территориально распределенной интегрированной системы учета населения не затрагивает и не ограничивает свободы развития каждой из взаимодействующих систем, которые продолжают работать штатно, основываясь на существующей правовой и нормативно методической базе. Благодаря этому, каждая из систем может развиваться и совершенствоваться независимо. Однозначная идентификация осуществляется на основании полного набора основных и дополнительных идентификационных данных. Список реквизитов, с помощью которых производится идентификация, формируется на основании специализированных алгоритмов процедуры идентификации, основанных на идентификации по степени схожести основных идентификационных данных. Алгоритмы идентификации личности гарантируют вероятность корректной идентификации, пренебрежительно мало отличающейся от стопроцентной, при

условии корректности исходных данных, что не всегда выполняется. Наибольшее число ошибок и неточностей встречается в реквизитах отражающих место рождения (включая синтаксические ошибки). Проблемы таких несоответствий решаются с помощью альтернативных классификаторов.

С точки зрения информационного обеспечения широкого круга потенциальных абонентов, интеграция центральной и территориальных автоматизированной системы учета населения в единое информационное пространство, создает новое качество, которое заключается в появлении новых возможностей оперативного решения разнообразных статистических и аналитических задач, формирования прогнозных оценок по социально-демографическому развитию любой из обслуживаемых территорий. Повышается оперативность получения комплекса сводных персональных данных на любом уровне использования автоматизированной системы. При этом возрастает роль, как всей интегрированной сети, так и каждой из входящих в нее систем.

Экономический эффект от создания автоматизированной системы учета населения складывается из двух составляющих: экономии средств населения и экономии бюджетных средств.

В целом, в макроэкономическом, смысле обе составляющие являются тесно взаимосвязанными и обеспечивают общий эффект, который может измеряться в масштабе общего объема валового внутреннего продукта. Количественный эффект в данном случае определяется экономией на затратах, связанных с обеспечением межведомственного взаимодействия в целях организации ведения учета населения в условиях отсутствия целостной системы; дублированием действий различных органов власти при сборе, хранении и использовании персональной информации; необходимостью получения населением разнообразных справок в различных органах власти и др.

Одним из важнейших экономических эффектов, является снижение издержек, связанных с необходимостью получения гражданами различных справок. В работе приведена оценка, сделанная на основе имеющихся статистических данных.

Проведенные исследования позволили обосновать и разработать принципы системного построения автоматизированной системы регистрационного учета населения, определить требования к ее основным элементам, характеру и топологии взаимодействия этих элементов.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

1. На основании результатов, проведенных анализов предложен метод формирования структуры информационной системы в виде совокупности взаимодействующих вертикальных и горизонтальных семантически определенных и формализованных объектов, связанных друг с другом иерархическими отношениями классов атрибутов, определяющих их бизнес-процессы, отличающийся от традиционных методов тем, что на основе предложенной классификации ат-

рибутов, позволяет повысить формализованность правил взаимодействия объектов учета, что открывает возможность целенаправленного развития системы.

2. На основе проведенного анализа тенденций развития информационных систем, предложена концептуальная модель многоуровневой архитектуры хранилища данных, основанная на логическом и далее платформенном разделении функций учета и отчетности, новизна которой заключается в применении методов теории нечетких множеств к нечетким правилам перекрытия взаимодействующих объектов учета для формирования четких измерений OLAP кубов интегрированного хранилища данных.

3. На основе анализа методов взаимодействия структурных компонентов предложено использование трехуровневого хранилища данных, которое в состоянии обеспечить достаточную, функциональную надежность хранилища данных и хорошую возможность расширения за счет быстрого добавления витрин данных. Для создания OLAP-кубов и витрин данных по разным направлениям отчетности, с различным регламентом поступления и подготовки информации, упрощения работы потребителей аналитической информации, в архитектуру трехуровневого хранилища данных введен промежуточный склад данных, создав этим новую архитектуру трехуровневого хранилища данных.

4. Предложенная методика оценки эффективности учетной системы действующей в органах государственного управления, позволяет количественно оценить экономию отдельных затрат государства. Новизна предложенного метода оценки эффективности учетной системы, заключается в том, что появляется возможность количественной оценки снижения потерь ВВП на основе статистических данных о снижении временных затрат населения связанных с получением различных документов из нее.

ПУБЛИКАЦИИ, ОТРАЖАЮЩИЕ РЕЗУЛЬТАТЫ РАБОТЫ

В рецензируемом журнале из списка ВАК

1. Метод формирования структуры хранилища данных для автоматизированной учетной системы на основе процессного анализа предметной области / Г.Г. Куликов, В.В. Антонов // Вестник УГАТУ : науч. журн. Уфимск. гос. авиац. техн. ун-та. 2006. Т. 8, № 1(17). С. 60–67.

В других изданиях

2. Проблемы использования современных информационных технологий в управлении / Н.Д. Андреев, В.В. Антонов // Труды Саратовского юридического института МВД России. СЮИ МВД РФ. 1997. С. 83–86.

3. Современные информационные технологии в деятельности МВД РБ / Н.Д. Андреев, В.В. Антонов // Вестник УЮИ. Уфа. 1999, №1. С. 57–60.

4. Совершенствование информационного обеспечения МВД РБ на основе инструментального средства ФАКТ. Достижения и перспективы / В.В. Анто-

нов, Н.Д. Андреев // Информатизация правоохранительных систем : матер. IX междунар. науч. конф. Москва: АН МВД РФ, 2000. С. 127–134.

5. Инструментальное средство ФАКТ, проблемы и перспективы / В.В. Антонов, Н.Д. Андреев // Совершенствование информационного обеспечения МВД РБ : матер. респ. конф. «Актуальные проблемы борьбы с преступностью в Республике Башкортостан». Уфа, 2000. С. 18–20.

6. Информационно-вычислительная сеть МВД РБ / Н.Д. Андреев, В.В. Антонов, В.И. Михайлов // Опыт работы и перспективы развития : матер. межвузовской НПК «Актуальные проблемы совершенствования правоохранительной деятельности в современных условиях». Челябинск: Челябинский юридический институт МВД России, 2001. С. 14–17.

7. Практика взаимодействия ИЦ МВД РБ с подразделениями ГУИН МЮ в вопросах предоставления информации / В.В. Антонов // Матер. всерос. семинар - совещания «О задачах информ. центров МВД, ГУВД, УВД субъектов РФ по реализации концепции развития информационно-вычислительной системы МВД России на 2002 - 2006 годы». Казань, 2003. С. 52–57.

8. Применение современных информационных технологий. Работа с информационно-справочными базами данных с использованием единой информационно-телекоммуникационной системы / В.В. Антонов // Сборник Информационно-аналитических материалов №31 МВД РБ : матер. учеб.-метод. сборов начальников ОВД. Уфа, 2005. С. 25–52.

9. Один из методов повышения эксплуатационных характеристик применяемых автоматизированных информационных систем в МВД Республики Башкортостан / В.В. Антонов // Computer Science & Information Technology, CSIT'2005 : матер. VII междунар. конф. по компьютерным наукам и информационным технологиям. Уфа: УГАТУ, 2005. Т. 3. С. 148–151. (Статья на англ. яз.).

10. Комбинированный подход решения проблемы перехода от оперативных баз данных специального назначения к интегрированным информационным системам корпоративного уровня / В.В. Антонов, Г.Г. Куликов // Computer Science & Information Technology, CSIT'2006 : матер. VIII междунар. конф. по компьютерным наукам и информационным технологиям. Карлсруэ (Германия): УГАТУ, 2006. Т. 2. С. 7–14. (Статья на англ. яз.).

11. Определение структуры многоуровневой территориально-распределенной учетной системы на основе формальной интеграции бизнес-процессов объектов по семантическим правилам / Г.Г. Куликов, В.В. Антонов // Управление экономикой: методы, модели, технологии : матер. шестой всерос. науч. конф. с междунар. участием. Уфа, 2006. Т. 1. С. 228–233.

Соискатель



В.В. Антонов

АНТОНОВ Вячеслав Викторович

ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА НА
ОСНОВЕ ЗНАНИЙ РЕГИСТРАЦИОННОГО УЧЕТА НАСЕЛЕНИЯ

Специальность: 05.13.10 «Управление в социальных
и экономических системах»

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Подписано к печати 11.04.2007 г. Формат 60×84 1/16.

Печать плоская. Бумага офсетная. Гарнитура Times New Roman Сур.

Усл. печ.л.1,0. Усл. кр.-отт. 1,0. Уч.-изд.л. 0,9.

Тираж 100 экз. Заказ № 173.

ГОУ ВПО Уфимский государственный авиационный технический университет

Центр оперативной полиграфии
450000, Уфа-центр, ул. К.Маркса, 12