

УДК 004.65

Г. Г. КУЛИКОВ, В. В. АНТОНОВ

МЕТОД ФОРМИРОВАНИЯ СТРУКТУРЫ ХРАНИЛИЩА ДАННЫХ ДЛЯ АВТОМАТИЗИРОВАННОЙ УЧЕТНОЙ СИСТЕМЫ НА ОСНОВЕ ПРОЦЕССНОГО АНАЛИЗА ПРЕДМЕТНОЙ ОБЛАСТИ

Рассмотрены проблемы интеграции информации в распределенных системах для поддержки принятия управленческих решений и пути решения данных проблем. В целях снятия большей части организационных конфликтов предложена «эволюционная» модель внедрения системы, обеспечивающая поэтапную «вертикальную» автоматизацию бизнес-процессов при использовании процессного подхода. *Вертикальная автоматизация; процессный анализ; бизнес-процесс; непротиворечивость системы*

ВВЕДЕНИЕ

Исторически в большинстве организаций непроизводственной сферы, в виду ряда причин, как правило, административного характера, были сформированы независимые центры актуализации данных. Это обуславливалось в первую очередь необходимостью автоматизации независимых друг от друга технологических операций и процессов, составляющих основные направления деятельности организации. Ввиду важности информации для каждого отдельно взятого заинтересованного подразделения, ими создавались собственная база данных и свой собственный центр ее актуализации. Информация из данной системы предоставлялась только по официальным запросам. Исторически сложившиеся технологии обработки данных, значительные инвестиции в аппаратное и программное обеспечение каждого такого центра не позволяют кардинально решить проблему их разобщенности путем механического переноса данных в центральную базу данных. Актуальной темой становятся технологии и программные продукты, которые могут обеспечить безболезненную интеграцию баз 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, опирающуюся на единые стандарты и единые принципы сетевого взаимодействия.

1. ЗАДАЧИ ИССЛЕДОВАНИЯ

Помимо задачи получения интегрированных данных существует задача обеспечения целостности всей совокупности данных и предоставления данных для аналитического анализа. Существуют различные подходы к решению указанных задач. В работах многих авторов – Ф. Бруссар, Б. Голд-Бернштейн, В. Ю. Добрынин, А. Г. Дубинский, С. Крэггс, J. Kleinberg, Дж. Мэдден, Т. Мэттьюс, А. Г. Назаров, А. А. Сухобоков, Дж. Тейлор предлагаются различные способы формирования интегрированных ин-

формационных систем. Необходимо, чтобы часть из независимых систем были объединены в так называемую «федерацию» взаимодействующих баз данных, при этом подразумевается сопровождение данного процесса тем или иным вариантом реинжиниринга. Учитывая, что идеология управления при «ручном» учете заметно отличается от идеологии управления при автоматизированном учете, необходимо уже при формировании системы управления определить, какие функции будут реализованы ручными методами, какие полностью автоматизированными, а какие смешанными и иметь функциональные модели процедур. В любом случае, процесс внедрения характеризуется в первую очередь появлением новой системы управления, и во вторую очередь ее реализацией в виде автоматизированной системы.

2. ПОСТРОЕНИЕ МОДЕЛИ СИСТЕМЫ

Для решения данной проблемы рассмотрим возможность использования так называемой «эволюционной» модели внедрения системы, что не должно привести к большому увеличению стоимости технического решения, а скорее наоборот позволит более продуманно определить необходимую конфигурацию системы, отказаться от излишней функциональности, за счет чего уменьшить данные расходы. Помимо этого «эволюционная» модель внедрения системы позволяет постепенно внедрять систему по частям, снимая большую часть организационных конфликтов. Рассмотрим вопрос предпочтения того или иного подхода к решению задачи управления. Очевидно, что деятельность любой устойчиво работающей организации основана на бизнес-процессах, которые обладают свойством устойчивости. Система управления, построенная на базе этих процессов, обеспечивает устойчивость функционирования предприятия. При рассмотрении организации, основные виды деятельности которой регламентированы действующим законодательством, бумажный документооборот обеспечивает документальное подтверждение произведенных операций, моделирование автоматизированных бизнес-процессов должно быть основано на базе действующих бизнес-процессов, т.е. с сохранением контролируемого

ручного контура. При отсутствии средств автоматизации для выполнения функций управления неизбежно присутствие функциональных вертикалей, которые обеспечивают управление «сверху вниз» и контроль за исполнением, а информация, требуемая для управления, создается, обрабатывается и передается через подразделения, обеспечивающие эту информацию на функциональной основе. Таким образом, для обеспечения функционирования организации в целом обеспечивается контроль за каждым шагом управляющего воздействия.

Применение процессного подхода в данном случае приводит к ситуации, когда сотрудник будет производить оценку и контролировать самого себя, что не способствует устойчивости процессов в течение времени и их необходимую согласованность.

При наличии автоматизации информационного управления использование процессного управления предполагает существенное изменение ряда действующих бизнес-процессов. Это в первую очередь обеспечение автоматизированных функций контроля и согласованности, которые должны быть выведены из «ручного контура». С другой стороны, обеспечение прямого обращения исполнителей к централизованному автоматизированному учету позволяет снять ограничительный характер на получение информации, не связанной с функциональной основой данного подразделения. Процессный подход в чистом виде встречается очень редко и в основном представляет способ управления, когда определенный сквозной процесс связывает несколько функциональных подразделений и должностное лицо, несущее ответственность за ход и результаты, которое не является руководителем этих подразделений. Процессный подход требует применения специальных средств описания и классификации процессов, составляющих деятельность организации, например методологии функционального моделирования IDEF0, которая является стандартом для функционального моделирования в ряде стран и поддерживается компьютерными программами. Поэтому возникает сложная задача определения оптимального соотношения между функциональными возможностями внедряемой системы, затратами предприятия и масштабом всей совокупности проблем, связанных с информационной поддержкой процессного управления, экспорта данных из баз различных форматов.

Вообще говоря, построение осмысленной, архитектурно простой и эффективной программной инфраструктуры в типичной современной организации, где скопилось множество компьютеров различных моделей и производителей, и эксплуатируются, как правило, СУБД от нескольких поставщиков, является ключевой задачей при создании информационной системы.

Присутствует большое число предприятий непромышленной сферы (преимущественно государственные структуры), в деятельности которых наиболее характерно выражены следующие аспекты:

1. Достаточно четкая регламентация выполняемых действий на каждом уровне иерархии.

2. Наличие территориально распределенных подчиненных иерархий и сильная связь с вышестоящей иерархией.

3. Постоянный аудит со стороны вышестоящих иерархий.

4. Потребность глубоких аналитических исследований и анализа поступающей информации на всех уровнях иерархии.

5. Строго регламентированные информационные потоки.

6. Наличие регламентированной специфики информационных взаимодействий на каждом уровне иерархии.

Типовая организационно-функциональная модель организации, описанной выше, приведена на рис.1.

В качестве центров актуализации выступают территориальные и функциональные подразделения.

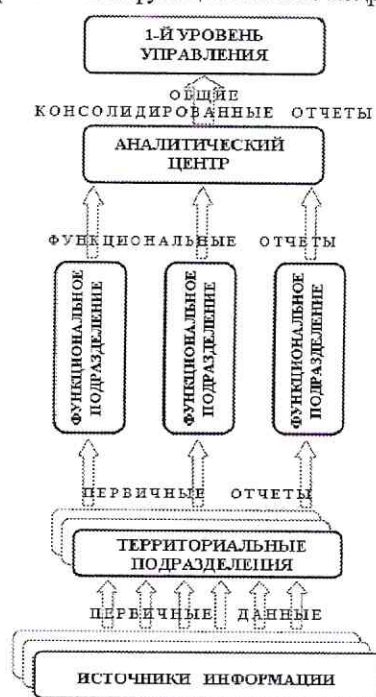


Рис.1

При традиционном подходе к автоматизации бизнес-процессов одним из самых существенных недостатков является то, что внедряемая система автоматизации позволяет получать результаты от работы системы в первую очередь исполнителям нижнего уровня. Только после завершения автоматизации всех бизнес-операций на нижнем уровне возможно получение отчетов всех видов на более высоких уровнях. Применение процессного подхода требует введение автоматизации «сверху вниз». При этом результаты от работы системы в первую очередь появляются на верхнем уровне управления через довольно непродолжительный промежуток времени после начала работ. Дальнейшая постройка системы происходит в соответствии с потребностями высшего управляющего звена организации и может развиваться по мере развития потребностей. На начальном этапе сбор информации осуществляется на основе различных сводных отчетностей подразделений предприятия, которые вводятся в систему на

соответствующем уровне иерархии. Вводимые низовыми подразделениями данные не могут быть проверены на достоверность, так как отделены от первичных данных, на основании которых они получены. Достоверность вводимых данных увеличивается по мере автоматизации нижних иерархий. Процесс же создания и активного использования полноценного банка данных, основанного на первичных документах и метаданных, откладывается до автоматизации нижнего уровня.

Рассмотрим возможность совмещения поэтапной автоматизации с процессным подходом. Т.е. первоначально проведем обследование, целями которого будет выявление независимых функциональных процессов организации. Далее производится автоматизация выбранного процесса на всех уровнях иерархии. Информация становится доступной руководителю для анализа сразу же после ее ввода исполнителем. При этом полнота информации, доступной на всех уровнях иерархии, растет по мере увеличения числа автоматизированных бизнес-процессов. Очередность автоматизации бизнес-процессов определяется потребностями управляющего звена верхнего уровня.

Приведенная на рис.1 структура содержит несколько вертикальных семантических блоков, которые обладают свойством независимости.

Учитывая предполагаемую территориальную разобщенность и самостоятельность подразделений в выполнении возложенных на них функций, определенную действующими регламентами, разнесем на три уровня приведенную структуру и по горизонтали. В результате получим структуру с условным разделением на «блоки» – для каждого из которых возможен дифференцированный подход и установлены значения учетных регистров:

1. Зафиксирован перечень организационных звеньев и их иерархия.
2. Определен перечень функций, которые выполняются в организации, и их взаимосвязи по иерархии - функциональные модели.
3. Определено, каким образом функции, реализуемые в компании, закрепляются за организационными звеньями или «кто за что» отвечает.

Перечень всех функций дает положение об этом организационном звене, которое традиционно называют Положением о подразделении. Внутри каждого «блока» при введении автоматизации бизнес-процессы легко преобразуются в сквозные, причем без существенных материальных затрат, а все уровни соединяются между собой в единую функциональную систему. При этом будут разрешены следующие вопросы:

- на всех уровнях организована единая информационная среда источников и потребителей информации, что позволит кардинально изменить назначение бумажного документа и рассматривать его не как носитель информации, а как метаданные, сформированные на основе соответствующего информационного объекта базы данных;
- ряд бумажных документов преобразуется в носитель юридического статуса, и станет представлять

собой набор метаданных из базы, распечатанный на бланке или в виде бланка. При этом электронный документ может быть воспроизведен из централизованного архива в любое время по установленной связи с объектом базы данных, на основании которого он был получен;

- на всех уровнях информационной системы пользователи работают с соответствующим информационным объектом напрямую в рамках предоставленных прав;
- в связи с изменением статуса бумажного документа и организацией реальной совместной работы с информацией на нижнем уровне, возможен переход от последовательного способа обработки информации к параллельному (распараллеливание бизнес-процесса);

«Источники информации» в рассматриваемой структуре являются низовым звеном, на уровне которого производится фиксирование информации на первичных документах (как в электронном виде, так и в бумажном варианте). Под «1-м УРОВНЕМ УПРАВЛЕНИЯ» понимается вышестоящая организация, для которой рассматриваемая структура является в свою очередь низовым звеном или «Источником информации». Таким образом, рассмотренная схема является упрощенным описанием одного иерархического звена крупной территориально-распределенной иерархической организации.

Типовая организационно-функциональная модель рассматриваемого уровня иерархии с разнесением на условные семантические блоки приведена на рис.2.

Вариант поэтапной автоматизации независимых бизнес-процессов приведен на рис. 3.

Построение информационной системы указанным образом позволит минимизировать затраты на ряд формальных мероприятий, сократить время от начала работ до получения практических результатов и обеспечить практически одновременный эффект от системы для всех групп пользователей от низшего звена до руководителей организации.

В настоящее время одним из действенных механизмов интеграции процессов информатизации органов государственной власти является Федеральная целевая программа «Электронная Россия», согласно которой доступ к информации государственных организаций регламентируется целым рядом взаимосвязанных концепций, рассматриваются вопросы юридической значимости учетной информации, в том числе в связи с предоставлением ее иным лицам [17, 18]. Одним из основных вопросов, решаемых в рамках данной программы, является «Концепции информационного регулирования».

В связи с переходом к сбору и хранению информации в электронной форме в настоящей концепции вводится специальный термин «раскрываемая информация» – информация, опубликованная в сети общего пользования и доступная по адресу, установленному законом.

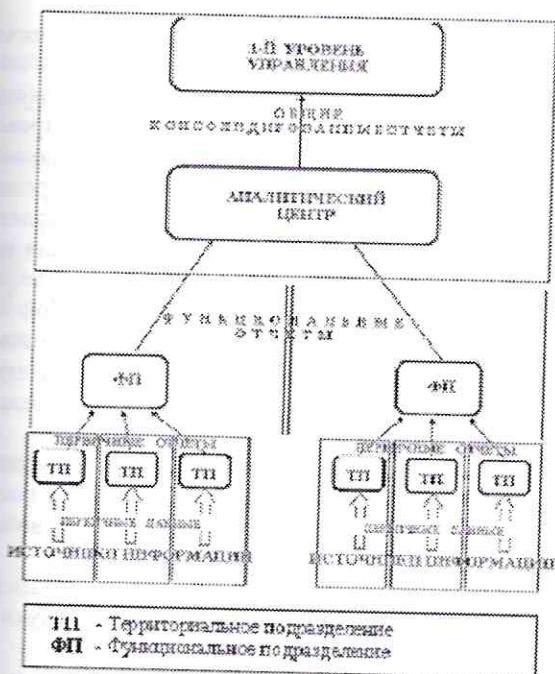


Рис. 2

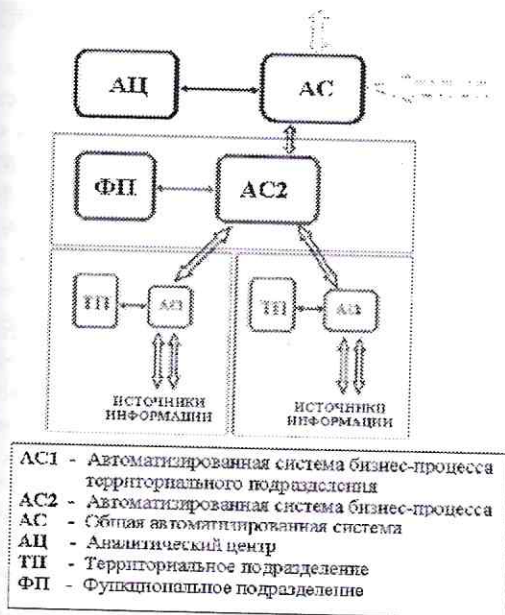


Рис.3. Автоматизация одного бизнес-процесса

Для предоставления такой информации применяется термин «публичное раскрытие» или просто «раскрытие» исключительно в электронном виде, с возможностью использования раскрытых данных гражданами в качестве юридически значимой информации без дополнительных обращений к государственным служащим. Права доступа к различным типам информации могут различаться. Для взаимодействующих сторон, при необходимости организации информационного обмена, описание структур всех обменных данных должно быть открытым и доступным, а сам обмен данными между системами должен быть построен на основе информационных объектов, снабженных отчуждаемыми метаописаниями, используемыми для поиска и обработки.

Мы рассматриваем специальный тип информационных систем – учетную информационную систему, то есть систему, которая предназначена для хранения, обработки и предоставления информации.

Единое информационное пространство будем рассматривать как совокупность взаимодействующих и взаимосвязанных учетных систем выделенных блоков.

В качестве информационных объектов рассматриваемой модели определим независимые бизнес-процессы выделенных блоков, которые характеризуются значениями своих атрибутов. Поэтому каждый рассматриваемый бизнес-процесс может быть определен как множество упорядоченных пар

$$x = \{ \langle a_i, d_i \rangle : i \in I \}, \text{ где } I - \text{ множество индексов, } a - \text{ имя атрибута, } d - \text{ значение атрибута.}$$

Учитывая многоуровневую модель нашей ИС и цель – построение интегрированной системы на верхнем уровне - сделаем следующие допущения: на нижнем и среднем уровне нашей системы откажемся от интеграции бизнес-процессов, каждый независимый бизнес-процесс будем рассматривать как отдельную (или обособленную) часть информационной системы.

Тогда под информационной системой выбранного «блока» будем понимать «информационную схему блока», описывающую характеристики входящих в нее бизнес-процессов и множество самих бизнес-процессов.

Введем обозначения:

$\varphi: A \rightarrow D$ – отображение, ставящее в соответствие каждому атрибуту множество его возможных значений;

$T = \{t_i : i \in I\}$ – множество типов независимых бизнес-процессов (типов объектов);

$\delta: T \rightarrow 2^A$ – отображение, задающее для каждого типа бизнес процесса (типа объекта) множество его атрибутов;

$U = \{x_i : i \in I\}$ – множество информационных объектов (независимых бизнес-процессов);

$\gamma: U \rightarrow T$ – отображение, ставящее в соответствие объекту его тип.

«Информационная схема[6,7] блока» S будет определяться кортежем

$$S = \langle A, D, T, \varphi, \delta \rangle \quad (1)$$

Информационная система блока, построенная по схеме S , будет определяться кортежем

$$U^S = \langle S, U, \gamma \rangle \quad (2)$$

Определим множество всех информационных систем блоков

$$U^{\bar{S}} = \{ U_i^{S_j} : i, j \in I \}, \quad (3)$$

$$\text{где } \bar{S} = \{S_j : j \in I\}$$

Рассмотрим несколько частных случаев.

При автоматизации одного (первого) бизнес-процесса получим множество информационных систем нижнего уровня (АС1 рис. 3)

$$U^{\bar{S}} = \{U_i^{S_j} : i, j \in I\}, \text{ информационная система}$$

среднего уровня (АС2 рис. 3) будет автоматизировать только один бизнес-процесс и определяться объединением информационных систем нижнего уровня

$$U^{\bar{S}} \subseteq \bigcup_{i \in I} U^{S_i} \quad (4)$$

Информационная система верхнего уровня (АС рис. 3) будет эквивалентна (АС2 рис. 3)

$$U_{AC} \leftrightarrow U^{\bar{S}} \quad (5)$$

Введем обозначения :

$$\bar{A} = \{A_j : j \in I\}, \quad \bar{D} = \{D_j : j \in I\},$$

$$\bar{T} = \{T_j : j \in I\}, \quad \bar{U} = \{U_j : j \in I\}$$

Каждая из информационных систем блоков будет непротиворечивой по определению, так как не взаимодействует с внешними объектами учета (бизнес-процессами)[6].

Очевидно, что полученное множество информационных систем также непротиворечиво ввиду того, что:

$$\exists \bar{\varphi}, \quad \bar{\varphi} : \bar{A} \rightarrow \bar{D}, \quad \bar{\varphi} = \{\varphi_i : i \in I\}$$

$$\exists \bar{\delta}, \quad \bar{\delta} : \bar{T} \rightarrow 2^{\bar{A}}, \quad \bar{\delta} = \{\delta_i : i \in I\}$$

$$\exists \bar{\gamma}, \quad \bar{\gamma} : \bar{U} \rightarrow \bar{T}, \quad \bar{\gamma} = \{\gamma_i : i \in I\}$$

Информационная система U' , определенная как $U' = \langle \bar{S}, \bar{U}, \bar{\gamma} \rangle$, будет интегрирована на нашем множестве информационных систем блоков. Полученная информационная система полностью обеспечивает соответствие совокупности семантических ограничений и зависимостей, которые определены в информационных системах блоков.

Рассмотрим, что изменится при автоматизации второго (третьего и т.д.) бизнес-процесса. Для информационных систем нижнего уровня (АС1 рис. 3) изменится только их количество за счет создания новых территориальных АС1, обладающих теми же свойствами. Появятся дополнительные системы среднего уровня (АС2). Но, согласно нашему допущению, об отсутствии интеграции бизнес-процессов

на среднем уровне они также будут непротиворечивы.

Рассмотрим информационную систему верхнего уровня (АС рис. 3). Ввиду того, что рассмотрение вопроса о непротиворечивости информационной системы верхнего уровня сводится к рассмотрению зависимостей между типами входящих в нее элементов, необходимо понижение уровня абстракции модели, применение функциональной декомпозиции, конечный продукт которой представляет собой иерархию функций. Необходимо выделение основных классов объектов единого информационного пространства, для которых должна быть определена структура, свойства и т. д.

Оставаясь на верхнем уровне абстракции, можем выделить два класса объектов учета (бизнес-процессов), характеризующихся основными свойствами:

1. Свободные объекты учета, т.е. бизнес-процессы, объекты учета которых не взаимодействуют с объектами учета других бизнес-процессов.
2. Взаимодействующие объекты, т.е. бизнес-процессы, объекты учета которых взаимодействуют с объектами учета других бизнес-процессов.

В первом случае в связи с отсутствием пересечения множеств объектов учета информационных систем блоков результат тождественен результату автоматизации одного бизнес-процесса. Каждому типу взаимодействующих объектов поставим в соответствие множество семантических зависимостей с другими типами объектов $C = \{c_i : i \in I\}$.

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.

В нашем случае, непротиворечивость рассматриваемой системы будет определяться отсутствием правил, имеющих сходные посылки и различные или взаимоисключающие следствия.

Введем обозначения:

Пусть E – универсальное множество, x – элемент E , а C – множество установленных правил (свойств). Подмножество A универсального множества E , элементы которого удовлетворяют свойству C , определяется как множество упорядоченных пар

$A = \{ \langle c_i, \mu_A(c_i) \rangle : i \in I \}$, где $\mu_A \in M$ – характеристическая функция принадлежности [10,11].

Если значения функций принадлежности нам известны, например, $\mu_A(c_i) = w_i$, то попарные сравнения можно представить матрицей отношений

$$A = \{a_{ij}\}, \quad \text{где} \quad a_{ij} = \frac{w_i}{w_j} \quad i, j \in I.$$

В общем случае задача сводится к поиску вектора w , удовлетворяющего уравнению вида $A \cdot w = \lambda_{\max} w$, где λ_{\max} – наибольшее собственное значение матрицы A . Поскольку матрица A положительна по построению, решение данной задачи существует и является положительным.

Степень непротиворечивости i -го и k -го правил можно задавать величиной

$$C_{ik} = \left| \max_x (\mu_{A_i}(x) \wedge \mu_{A_k}(x)) - \max_y (\mu_{B_i}(y) \wedge \mu_{B_k}(y)) \right| \quad (6)$$

Суммируя по k и i , получаем оценку непротиворечивости установленных правил в системе:

$$C = \sum_i \sum_k C_{ik}, \quad \text{где} \quad i \neq j, \quad i, j \in I, \quad (7)$$

Каждое правило может иметь свой вес в зависимости от того, насколько это правило является важным. В приведенной формуле (7) весовые коэффициенты приняты равными 1. Однако условие применения семантического ограничения (правила) может быть связано с дополнительными условиями, зависящими от содержательной части не взаимодействующих объектов учета бизнес-процесса, т.е. тематической близости установленному правилу. Необходимо количественно оценить степень тематической принадлежности обрабатываемого документа к установленному правилу. В этом случае представим обрабатываемый объект учета бизнес-процесса и сопоставляемое ему правило в виде тематических классов заданных множествами ключевых элементов [14].

Введем следующие обозначения:

S_C^{hl} – множество ключевых элементов определенных правил для H – бизнес-процесса, для определения критерия применимости к объекту учета бизнес-процесса L :

$$S_C^{hl} = \{ \langle k_{ci}^{hl}, i \rangle : i \in I \}, \quad \text{где коэффициенты}$$

k_{ci}^{hl} – весовые коэффициенты, определяющие значимость данного элемента в установленном правиле.

S_f^{hl} – множество ключевых элементов обрабатываемого документа определенных правил для H – бизнес-процесса, для определения критерия применимости к семантическому ограничению взаимодействия с бизнес-процессом L :

$$S_f^{hl} = \{ \langle k_{fi}^{hl}, i \rangle : i \in I \}, \quad \text{где коэффициенты}$$

k_{fi}^{hl} – весовые коэффициенты, определяющие значимость данного элемента в связи с контекстом обрабатываемого документа.

Для сопоставимости показателей нормируем по 1 единице суммы этих весовых коэффициентов для каждого множества. Тематическую близость по каждому из информационных элементов будем вычислять как функцию отклонения весовых коэффициентов, при этом коэффициент общей тематической близости будет соответствовать формуле

$$\omega^{hl} = \sum_i \frac{\min(k_{ci}^{hl}, k_{fi}^{hl})}{\max(k_{ci}^{hl}, k_{fi}^{hl})} \times k_{ci}^{hl} \quad (8)$$

Очевидно, что значение ω^{hl} лежит в диапазоне от 0 до 1. Множитель k_{ci}^{hl} введен для учета степени значимости данной разницы по отношению к установленному правилу.

Сопоставляя данный коэффициент с соответствующим семантическим ограничением, можем принять однозначное решение о применимости правила к данному объекту учета бизнес-процесса.

Традиционно создатели хранилищ данных наибольшее внимание уделяют функциям, направленным на реализацию бизнес-процессов предприятий, документооборота, учета, планирования и других сфер. Поэтому в первую очередь оптимизируется ввод новых учетных данных в систему. При этом на второй план уходит возможность использования накопленных данных, развитие встроенной системы отчетности и, что наиболее важно, преобразование массивов данных в информацию для принятия управленческих решений. Все хранилища данных имеют подсистему формирования отчетов, в составе которой поставляется базовый минимум отчетов по предметным разделам. Номенклатура, уровень достаточности и глубина штатных отчетов определяются разработчиками без учета потребностей конкретного конечного пользователя продукта. Эти отчеты выпускаются из базы данных, на которой работает учетная система. Скорость формирования таких отчетов находится в обратной зависимости от их информационного состава, количества и сложности фильтров и группировок. Создать дополнительные отчеты можно с помощью инструментария для разработки отчетов, который поставляется в составе хранилища данных, что и выполняется силами программистов (или подготовленных пользователей) и занимает значительное время, существенно зависящее от сложности работы. При этом любое изменение структуры отчета требует серьезных затрат временных и человеческих ресурсов.

Перечислим основные факторы, которые препятствуют свободному получению информации из учетной системы:

- ограниченный набор базовой отчетности;
- необходимость привлечения программистов для настройки дополнительных отчетов;
- низкая скорость формирования сложных отчетов из общей с системой ввода базы данных;

Учитывая наличие нескольких центров актуализации (практически в каждом выделенном нами блоке), 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. Часть детализированных данных можно будет получать из оперативных баз данных «среднего и нижнего» уровней по информации, полученной из хранилища данных, – так называемый «ссылочный массив». Учитывая, что на разных уровнях (слоях) и в каждом выделенном нами блоке может действовать свой собственный регламент отчетности, необходимо использование многоуровневой архитектуры хранилища данных.

Выходом из этой ситуации может быть платформенное разделение учета и отчетности, при котором учетная система оптимизирована для отражения бизнес-процессов, а отчетная является надстройкой над данными учетной системы и предназначена для быстрого создания и выпуска разнообразных отчетов. Тогда все системы сбора информации работают в штатном режиме, т. е. вводят данные и выпускают текущие отчеты для проверки корректности выполненных транзакций. Параллельно планировщик задач с заданным периодом (обычно совпадающим с периодичностью регламентной отчетности – день, месяц, квартал и т.д.) запускает перекачку данных в промежуточный склад данных (слой данных) для разделения системы учета и системы отчетности. Перегрузка происходит в фоновом режиме без отключения пользователей. Созданный промежуточный склад данных является копией базы данных хранилища. Данные промежуточного склада используются для создания OLAP-кубов витрин данных по разным направлениям отчетности, с различным регламентом поступления и подготовки информации.

Важный момент связан с агрегированными значениями: это их создание по имеющимся детализированным данным при подготовке OLAP-таблиц, и поступающих из оперативных баз данных в виде различных форм отчетности. Создание и настройка OLAP-отчетов осуществляется службами Аналитического центра. Готовые кубы содержат в себе как метаданные, представляющие собой настройки внешнего вида отчетов, доступных пользователю измерений и фильтров и т. д., так и просто данные – выборки из промежуточного склада данных. Любая базовая отчетность хранилища данных стандартно представляется в виде OLAP-таблицы, поэтому управление формой и содержанием отчета не требует специальных знаний и осуществляется самим пользователем (например, группировки и детализации данных, фильтрации по измерениям и периодам), т.е.

конечный пользователь имеет доступ к данным: ему предоставляются стандартные средства гибких запросов и отчетов.

Концептуальная модель «хранилища данных» приведена на рис. 4.

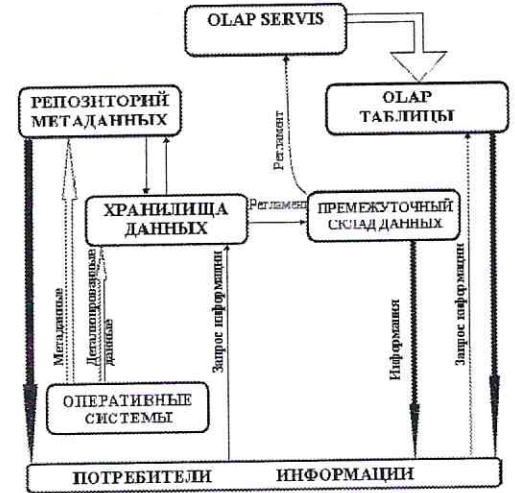


Рис.4

Таким образом, описываемый вариант хранилища данных создается в виде самостоятельной системы для «верхнего слоя» и в качестве надстройки над множеством существующих систем обработки данных «среднего и нижнего слоев», где могут функционировать и проблемно-ориентированные (ведомственные) реляционные базы данных и унаследованные системы. Для каждого свободного бизнес-процесса может создаваться отдельная OLAP-таблица, логическое связывание которых и построение OLAP-таблиц следующего уровня возможно уже только на уровне агрегированных аналитических показателей, приведенных к сопоставимой форме. Для бизнес-процессов, соединенных правилами взаимодействия, может создаваться общая OLAP-таблица, работающая на всем объединенном множестве объектов учета взаимодействующих бизнес-процессов. При этом поддержка принятия управленческих решений на основе накопленной информации может осуществляться в трех основных областях:

1. Область детализированных данных (включая данные, хранящиеся в базах данных среднего и низшего уровней).

2. Область агрегированных показателей. Комплексный взгляд на собранную в хранилище данных информацию, ее обобщение и агрегация, гиперкубическое представление и многомерный анализ являются задачами систем оперативной аналитической обработки данных (OLAP).

3. Область закономерностей, где обработка производится методами интеллектуального анализа данных, при этом осуществляется поиск функциональных и логических закономерностей, прогнозирования рассматриваемых процессов.

ЗАКЛЮЧЕНИЕ

1. В статье предложен формально-математический метод формирования структуры хранилища данных для автоматизированной учетной системы на основе процессного анализа предметной области.

2. За счет создания виртуальной модели предприятия в виде независимых вертикальных и горизонтальных семантических блоков представляющих собой бизнес-процессы, реализована возможность совмещения поэтапной автоматизации с процессным подходом.

3. Предложена концептуальная модель «хранилища данных», основанная на платформенном разделении учета и отчетности, при котором учетная система оптимизирована для отражения бизнес-процессов, а отчетная является надстройкой над данными учетной системы и предназначена для быстрого создания и выпуска разнообразных отчетов.

4. Разработанная модель, метод и алгоритмы поэтапной автоматизации позволяют преобразовать действующие бизнес-процессы в сквозные для выбранного семантического блока, а все блоки соединить между собой в единую, функциональную систему. Весь анализ произведен на верхнем уровне абстракции, где в качестве информационных объектов рассматриваемой модели определены независимые бизнес-процессы семантических блоков, что определяет возможность использования отдельных методов и алгоритмов при разработке и исследовании широкого круга информационных задач.

СПИСОК ЛИТЕРАТУРЫ

1. **Абаев, Л. Ч.** Выбор вариантов в нечеткой среде: бинарные отношения и нечеткая декомпозиция. (Русскоязычная версия статьи: Abaev L. Ch. Choice of variants in a fuzzy environment: binary relations and fuzzy decomposition // *Soft Models in Decision-Making: Special Issue of International Journal of General Systems* / Ed. by I. Batyrshin and D. Pospelov. 2001. Vol. 30. P. 53–70.
2. **Антонов, В.В.** Применение современных информационных технологий в ОВД / В. В. Антонов // *Использование возможностей ЕИТКС ОВД : сб. информац.-аналит. материалов МВД РБ. Уфа, 2005. №31. С. 25–52.*
3. **Анфилатов, В.С.** Системный анализ в управлении / В. С. Анфилатов, А. А. Емельянов, А. А. Кукушкин. М. : Финансы и статистика. 2006. 367 с.
4. **Борисов, А.Н.** Принятие решений на основе нечетких моделей. Примеры использования / А. Н. Борисов, О. А. Крумберг, И. П. Федоров. Рига : Зинатне, 1990. 184 с.
5. **Боровиков, В.** STATISTICA. Искусство анализа данных на компьютере: для профессионалов / В. Боровиков. СПб. : Питер, 2003. 688 с.
6. **Воробьева, М. С.** Математическое моделирование и технологии интеграции данных в УИС : дис. ... канд. техн. наук / М. С. Воробьева. Тюмень : ТюмГУ, 2006.
7. **Воробьева, М. С.** Построение модели интеграции данных в информационно-управляющих системах / М. С. Воробьева // *Модернизация образования в условиях глобализации: круглый стол «образование через науку и инновации», 14–15 сентября. 2005. Тюмень : ТюмГУ, 2005. С. 26–28.*
8. **Дубинский, А. Г.** Некоторые вопросы применения векторной модели представления документов в информационном поиске / А. Г. Дубинский // *Управляющие системы и машины. 2001. № 4. С. 77–83.*
9. **Когаловский, М. Р.** Энциклопедия технологий баз данных / М. Р. Когаловский. М. : Финансы и статистика, 2002. 800 с.
10. **Куликов, Г.Г.** Интеллектуальные информационные системы / Г. Г. Куликов, Т. В. Брейкин, В. Ю. Арьков, Уфа : УГАТУ, 1999. 129 с.
11. **Пивкин, В.Я.** Нечеткие множества в системах управления / В. Я. Пивкин, Е. П. Бакулин, Д. И. Кореньков // http://idisys.iae.nsk.su/fuzzy_book/content.html.
12. **Паклин, Н.В.** Нечеткая логика – математические основы / Н. В. Паклин // <http://basegroup.ru/fuzzylogic/math.htm>
13. **Тухватов, М.Б.** Лекции по общей математике (Множества и их отображения. Дискретная математика) / М. Б. Тухватов. Уфа : БГАУ, 2002. 393 с.
14. **Чугреев, В.Л.** Модель структурного представления текстовой информации и метод ее тематического анализа на основе частотно-контекстной классификации : дис. ... канд. техн. наук / В. Л. Чугреев. С.-Петербург. гос. электротехн. ун-т им. В. И. Ульянова (Ленина) «ЛЭТИ», 2003.
15. **Antonov, V.** One of methods of increase of operational characteristics of the used automated information systems in the Ministry of Internal Affairs of Republic Bashkortostan / V. Antonov // *The 7th Int. Workshop on Computer Science and Information Technologies (CSIT'2005). Ufa. Russia. 2005. Vol. 2. P. 125–127.*
16. **Cordon, O.** A General study on genetic fuzzy systems / O. Gordon, F. Herrera // *Genetic Algorithms in engineering and computer science, 1995. P. 33–57.*
17. **Правительство РФ.** Положения федеральной целевой программы «Электронная Россия (2002 – 2010 г.г.)», утвержденные Постановлением Правительства Российской Федерации от 28.01.2002 г. №65.
18. **Правительство РФ.** Концепция использования информационных технологий в деятельности федеральных органов исполнительной власти до 2010 года, одобренная распоряжением Правительства Российской Федерации от 27.09.2004. №1244.