

Рис. 3. Структура радиальной сети

Для обучения НС используется гибридный алгоритм [4]. Процесс обучения делится на два этапа: 1) подбор линейных параметров сети (весов выходного слоя); 2) адаптация нелинейных параметров радиальных функций (и характеристики ширины этих функций). В качестве обучающей выборки использовались данные наблюдений за пользователем на протяжении 10 дней, в течение всего рабочего дня.

Эффективность распознавания ИПП отображена процентом случаев правильного распознавания 5 пользователей в 1000-кратном эксперименте. Распознаваемость составила 90-93%. Увеличение количества зарегистрированных пользователей на одном рабочем месте (более 5) уменьшает вероятность распознавания в случае отсутствия пополнения обучающей выборки, т.е. без внесения дополнительной информации в компьютерные образы пользователей.

СПИСОК ЛИТЕРАТУРЫ

1. Власов А.Н. Способ представления координатной составляющей информационного почерка пользователя // Мягкие вычисления: Материалы междунар. науч. конф. Т. 1. СПб., 2003. С. 116-119.
2. Гузик В.Ф., Галуев Г.А., Десятерик М.Н. Биометрическая нейросетевая система идентификации пользователя по особенностям клавиатурного почерка // Нейрокомпьютеры. Разработка, применение. 2001. № 7-8. С. 104-118.
3. Диденко С.М., Шапцев В.А. Исследование динамики работы пользователя с манипулятором «мышь» // Математическое и информационное моделирование. Тюмень: Изд-во ТюмГУ, 2004, С. 295-304.
4. Оссовский С. Нейронные сети для обработки информации / Пер. с польского И. Д. Рудинского. М.: Финансы и статистика, 2002, 250 с.
5. Obaidat M.S., Sadoun B. Keystroke dynamics based authentication / Monmouth University W. Long Branch. NJ, 2000
6. Shivani Hashia, Authentication by Mouse Movements / CS 297 Report. 2004. May.
7. Секерин А.Б. Метод оценки устойчивости нейронно-сетевых моделей. 2005 (<http://zhurnal.apelarn.ru/articles/2005/031.pdf>).

*Денис Богданович КЕПЕЩУК —
аспирант кафедры информационных систем*

УДК 004.652.4

РАЗРАБОТКА МОДЕЛИ ОБМЕНА ДАННЫМИ В ГЕТЕРОГЕННЫХ РАСПРЕДЕЛЕННЫХ БАЗАХ ДАННЫХ

АННОТАЦИЯ. В статье приводятся основные результаты разработки модели обмена данными в распределенных базах данных с гетерогенными составляющими.

In the article are given the main results of the research of the model of data exchanging in distributed databases with heterogeneous constituents.

Введение

Современные компании, корпорации, учреждения имеют сложную структуру, включающую производственные объединения, научные центры, менеджмент верхнего уровня и т. д. Бизнес становится все более мобильным. Инфокоммуникационная среда территорий и стран интенсивно развивается как под влиянием государства (ФНТП «Электронная Россия», в частности), так и эволюции мирового сообщества к информационному обществу. Этот процесс вынужден проходить в условиях наличия баз данных разнообразного предметного содержания и в разных инструментальных оболочках. Формируются корпоративные банки данных. При этом корпоративный банк данных (КБД) должен быть мобильным и доступным в том месте, где это может понадобиться бизнесу, менеджменту, структурам учреждения. Главное свойство КБД — распределенность, т. к. сам бизнес распределен, как правило, как географически, так и функционально.

На пути формирования распределенных корпоративных банков и хранилищ данных возникает проблема интеграции (стыковки) и совместного использования разных баз данных (БД), появившихся в разное время и на разных уровнях технологической культуры.

Аналогичная ситуация наблюдается и в нефтегазовом комплексе [1]. Предприятиями нефтегазовой промышленности, а также сторонними компаниями разработано либо находится в стадии разработки или модернизации огромное количество систем по сбору, накоплению и анализу разнообразных данных. Многие системы работают с информацией сходного содержания, и эффективность применения подобных систем значительно повышается при их взаимной интеграции.

1. Постановка задачи

Традиционно задача синхронизации распределенных данных решается с помощью репликации данных [2]. Такие решения разрабатываются для заранее определенных структур БД. После изменения структуры одной из баз, участвующей в репликации, обычно требуется модификация сценариев репликации. Если же отдельные БД распределенной системы развиваются отдельными разработчиками, необходимо время для изучения ими структуры БД-поставщика и затраты труда по модификации существующих сценариев репликации. Это препятствует динамичному развитию распределенной информационной системы. Поэтому актуальна проблема поиска такого решения, которое, при минимуме затрат, позволило бы подсистемам сложной организационной структуры получать данные от БД поставщиков не зависимо от структуры данных последних.

2. Построение математической модели

Определим распределенную базу данных с гетерогенными составляющими как универсальную группу U , состоящую из следующих элементов:

$$U = \langle X, R_2 \rangle, \quad (1)$$

где X — множество подсистем-составляющих (совокупностей схем отношений), R_2 — множество связей между ними.

Определим множество подсистем как множество схем совокупностей отношений X , состоящее из конечного числа элементов.

$$X = \{X_i\}, i = \overline{1, N}, \quad (2)$$

где N — общее количество подсистем гетерогенной распределенной БД. Тогда схему совокупностей отношений можно представить в виде группы множеств:

$$X_i = \langle T, R, D, S_{proc}, S_{view} \rangle, \quad (3)$$

где T — конечное множество отношений, R — конечное множество связей, D — конечное множество доменов (типов информационных элементов), используемых в данной схеме совокупностей отношений, S_{proc} — конечное множество хранимых процедур в совокупности отношений, S_{view} — конечное множество отображений, заданных на отношениях из совокупности отношений. Конечное множество отношений:

$$T = \{T_j\}, j = \overline{1, M_i}, \tag{4}$$

где M_i — количество отношений в совокупности отношений X_i . Отсюда $|X_i| \equiv M_i$. Определим отношение как группу

$$T_j = \langle F, P \rangle, \tag{5}$$

в которой: F — конечное множество атрибутов отношения, P — первичный ключ (набор атрибутов, полностью и однозначно идентифицирующий группу элементов в отношении).

Пусть конечное множество доменов D , введенное в формуле (3), определяется как

$$D = \{d_q\}, q = \overline{1, Q_i}, \tag{6}$$

где Q_i — общее количество доменов в схеме отношений X_i . Тогда

$$F = \{d_{a_l}\}, a_l = \{a_1, a_2, \dots, a_{fc}\}, 0 \leq fc \leq Q_i, 1 \leq a_l \leq Q_i, F \subset D. \tag{7}$$

Аналогично определяется первичный ключ

$$P = \{f_{a_m}\}, a_m = \{a_1, a_2, \dots, a_{pc}\}, 0 \leq pc \leq fc_i, 1 \leq a_m \leq fc_i, P \subset F. \tag{8}$$

3. Специализированный сервер-коммутатор

В гетерогенной информационной среде целесообразно использовать специальный сервер-коммутатор, хранящий информацию о структурах всех баз данных-поставщиков, участвующих в обмене данными и преобразующий информацию в вид, необходимый подсистеме-потребителю. При такой схеме работы подсистемы-потребители не взаимодействуют непосредственно с подсистемами-поставщиками, а получают всю необходимую им информацию в нужном виде от сервера-коммутатора. Рис. 1 иллюстрирует схему потоков данных в описанном виде.



Рис. 1. Схема потоков данных при использовании сервера-коммутатора

Сервер-коммутатор решает следующие задачи [3]:

1. Хранение прав доступа отдельных подсистем к данным и контроль доступа.
2. Хранение прав доступа разработчиков подсистем для администрирования структур баз данных и контроль доступа.

3. Хранение и изменение метаданных о структурах баз данных-поставщиков и метаданных преобразования структуры.

4. Хранение и изменение информации об используемых драйверах и способах подключения к базам данных-поставщиков.

5. Обработка запросов от подсистем-потребителей и возврат результатов запросов (по протоколу SOAP).

6. Автоматическое присвоение глобальных уникальных идентификаторов (GUID) записям таблиц баз данных-поставщиков, выполнение запросов с использованием присвоенных глобальных уникальных идентификаторов.

Архитектурно сервер-коммутатор подразделяется на следующие составляющие (рис. 2).

6. *Анализатор запросов* разбивает блок запросов на отдельные запросы, производит синтаксический разбор запросов и запускает преобразователь запросов.

7. *Преобразователь запросов* обращается к хранилищу метаданных структуры, строит запрос к базе данных поставщику, посылает запрос в пул соединений, преобразует результаты запроса к виду, запрошенному потребителем, возвращает результаты потребителю.

8. *Хранилище метаданных структуры* — база данных, хранящая метаданные о структурах баз данных-поставщиках.

9. *Пул соединений* управляет соединениями со всеми зарегистрированными базами данных-поставщиками; с помощью соответствующего компонента доступа к базам данных, выполняет построенный запрос, возвращает данные преобразователю запросов.

Для абстрагирования от структур баз данных-поставщиков используется подход выделения ключевых сущностей, при котором, в отличие от реляционной структуры данных поставщика, присутствует только небольшое число основных (главных) сущностей (например, студенты, приказы) с большим набором атрибутов, полученных из справочников. При этом атрибут присоединяется к сущности независимо от того, через сколько отношений «таблица — внешний ключ — таблица» необходимо пройти, чтобы достигнуть данный атрибут.

4. Реализация на примере ИНГРИС ТюмГУ

Основные результаты данной работы реализуются в информационной системе ИНГРИС, разрабатываемой в Тюменском государственном университете.

Между отдельными компонентами распределенной среды ИНГРИС существует взаимодействие. Подсистемы, накапливающие данные одной сферы, используются другими подсистемами как источники информации, исключая дополнительную работу по ручному вводу данных.

На рис. 3 показаны потоки данных между отдельными подсистемами системы ИНГРИС.

Для решения проблемы сложности независимой модификации структур баз данных отдельных подсистем информационной системы использовали специальный сервер-коммутатор, описанный ранее. Рис. 4 иллюстрирует схему потоков данных в информационной системе ИНГРИС ТюмГУ, полученную при использовании сервера-коммутатора.

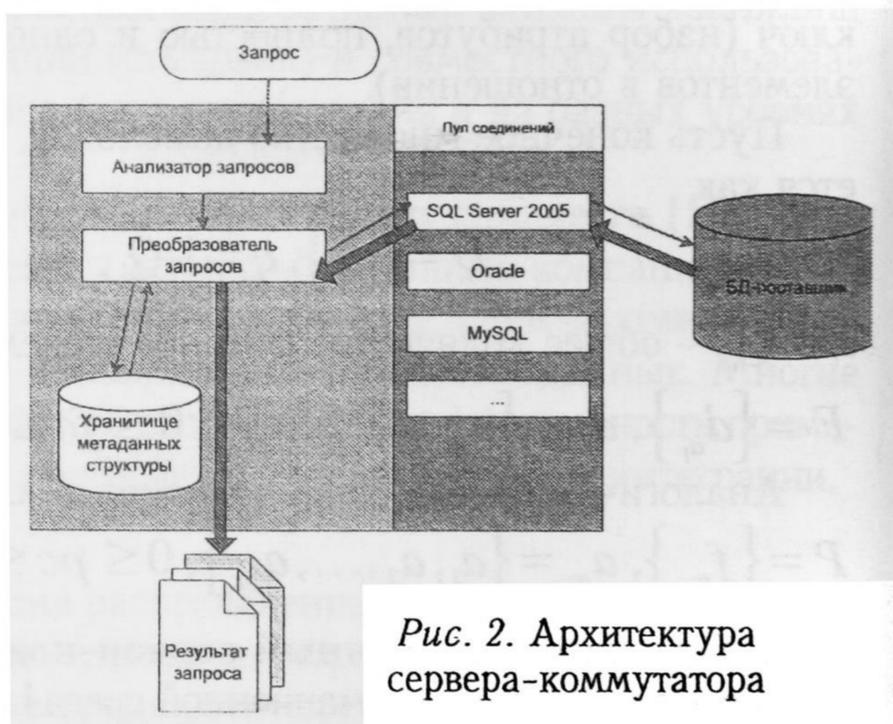


Рис. 2. Архитектура сервера-коммутатора

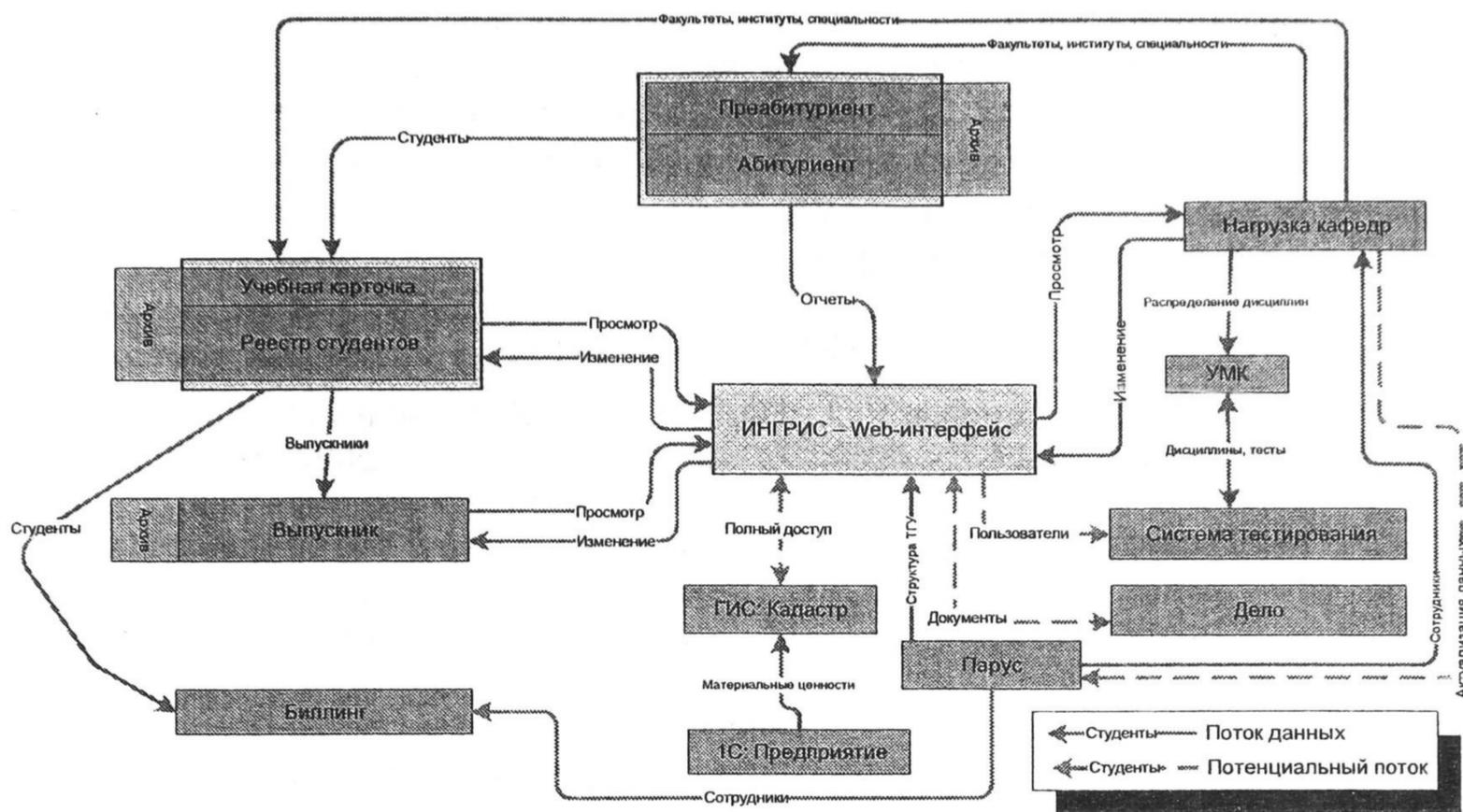


Рис. 3. Потоки данных в информационной системе ИНГРИС



Рис. 4. Схема потоков данных при использовании специального сервера-коммутатора ИНГРИС

Для обмена данными используется открытый протокол SOAP, основанный на универсальном языке описания данных XML. Так как язык является универсальным, он требует наложения специфических правил с целью использования в конкретных задачах. Для запросов к серверу-коммутатору был разработан специализированный язык запросов MetaQuery, основанный на XML. Для возвращаемых данных используется синтаксис, принятый в среде NET (XSD-схема для описания структуры возвращаемых данных и XML-файл для самих данных).

5. Основные результаты

1. Анализ существующих методологий и технологий, применяемых в настоящее время для обмена данными между подсистемами информационной системы с распределенными гетерогенными составляющими, показал, что обмен данными в РБД с гомогенной или гетерогенной структурой проработан достаточно полно лишь для случая неизменности структур узлов. Средство же, позволяющее решить проблему независимой модификации узлов гетерогенных РБД, отсутствует.

2. Построена математическая модель обмена данными между составляющими гетерогенной распределенной базы данных в виде группы множества совокупностей схем отношений и множества связей между ними, позволяющая реализовать систему обмена данными, не чувствительную к модификации отдельных хранилищ данных.

3. Разработан комплекс программ в виде специализированного сервера-коммутатора для обмена данными в гетерогенных распределенных базах данных, позволивший производить модификацию структур составляющих гетерогенных распределенных хранилищ данных без модификации кода подсистем потребителей данных. Комплекс программ обеспечил работу гетерогенной среды БД ТюмГУ в системе ИНГРИС [3].

СПИСОК ЛИТЕРАТУРЫ

1. Коровин С. Я., Артамонов Р. А., Назаров В. Ю. Информационная нефтепромышленная система нефтегазовой компании // Нефтяное хозяйство. 2002. № 8. С. 113-118.
2. Кепещук Д. Б., Шапцев В. А. Проблемы гетерогенных распределенных баз данных // Математическое и информационное моделирование. 2005. № 7. С. 161-166.
3. Ганопольский Р. М., Кепещук Д. Б. Представление знаний в гетерогенных распределенных базах данных на примере ИНГРИС ТюмГУ // Вестник кибернетики. 2006. № 5.

*Евгений Сергеевич ОХОТНИКОВ —
ассистент кафедры программного обеспечения*

УДК 519.711.3,519.872

МАТЕМАТИЧЕСКИЕ МОДЕЛИ СИСТЕМ ОПЕРАТИВНОГО МОНИТОРИНГА ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ НЕФТЕГАЗОДОБЫВАЮЩИХ ПРЕДПРИЯТИЙ

АННОТАЦИЯ. Рассмотрены моделирование и классификация систем мониторинга с позиций теории массового обслуживания. Предложены новые модель поллинга при отсутствии очередей и обобщенная модель.

The subject of the article is classification and modelling monitoring systems with queueing theory methods. New polling model without queueing and common model are offered.

Системы мониторинга технологических процессов нефтедобывающей отрасли (СМТПН) — территориально распределенные информационные системы контроля, диагностики и управления, основной целью применения которых является повышение эффективности и безопасности нефтегазодобывающего производства благодаря следующим факторам [1]: непрерывному мониторингу распределенных технологических объектов, управлению процессами добычи, транспортировки и учета готовой продукции, замене физически и морально устаревших средств автоматизации и систем управления, повышению безопасности производства за счет средств диагностики и улучшения экологической обстановки в нефтедобывающем регионе, снижению трудоемкости управления технологическими процессами нефтедобычи.

Внедрение таких систем, как отмечается в [2], «на предприятиях нефтедобычи, нефтепереработки и нефтехимии приобретает особое значение, так как позволяет обеспечить эффективную работу предприятий в заданных режимах, повысить качество выпускаемых продуктов, обеспечить безаварийность и экологическую безопасность производств, повысить производительность труда». Однако сами СМТПН являются сложными многокомпонентными системами, что делает целесообразным применение методов математического моделирования при их проектировании и эксплуатации.

В данной статье рассматривается задача моделирования процессов возникновения и передачи информации в СМТПН, для решения которой используем методы теории массового обслуживания (ТМО). Эта задача особенно актуальна в случае территориально распределенных систем с низкоскоростными УКВ-