


На правах рукописи



Гребенщиков Николай Николаевич

**МЕТОДИКА ОЦЕНКИ ЭФФЕКТИВНОСТИ СПОСОБОВ
РЕЛЯЦИОННОГО МОДЕЛИРОВАНИЯ СИСТЕМ УПРАВЛЕНИЯ
ИЕРАРХИЧЕСКИМИ ДАННЫМИ**

АВТОРЕФЕРАТ

диссертации на соискание учёной степени кандидата
технических наук по специальности 05.13.11 –
«Математическое и программное обеспечение
вычислительных машин, комплексов и
компьютерных сетей»

Красноярск – 2008

Работа выполнена в Институте информатики и телематики Хакасского государственного университета им. Н.Ф. Катанова на кафедре программного обеспечения вычислительной техники и автоматизированных систем

Научный руководитель:	кандидат технических наук, доцент Швец Сергей Викторович
Официальные оппоненты:	доктор технических наук, профессор Легалов Александр Иванович кандидат технических наук, профессор Сорокин Владимир Афанасьевич
Ведущая организация:	Сибирский государственный технологический университет

Защита состоится «4» декабря 2008 г. в 14-00 часов на заседании диссертационного совета ДМ 212.099.05 при Сибирском федеральном университете по адресу: 660074, г. Красноярск, ул. Киренского, 26, ауд. УЛК 115.

С диссертацией можно ознакомиться в библиотеке Сибирского федерального университета по адресу: ул. Киренского, 26, ауд. Г 2-74.

Автореферат разослан «3» ноября 2008 г.

Ученый секретарь
диссертационного совета,
кандидат технических наук



Вейсов Е.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы

В ходе создания программного обеспечения разработчикам приходится принимать множество ответственных решений. Наиболее важные из них связаны с архитектурой разрабатываемой программной системы.

Способ реализации управления данными является важным архитектурным решением. Благодаря простоте структуры данных и естественности операций, реляционные СУБД заняли доминирующее положение на рынке и используются при создании большого круга программных систем. Однако при разработке программного обеспечения часто приходится сталкиваться с данными, которые имеют иерархическую структуру. Такие данные обычно являются дополнительными иерархическими онтологиями, которые имеют тесную связь с данными, расположенными в реляционной СУБД. Вследствие такого положения прикладываются значительные усилия по изучению вопросов реляционного моделирования иерархий. В настоящий момент можно с уверенностью констатировать, что существует целый спектр теоретических разработок, технологических решений и конкретных реализаций в области построения иерархических справочников. Этот факт позволяет утверждать, что данное направление находится в активной стадии своего развития. Большой вклад в решение проблемы реляционного моделирования иерархических систем внесли Виноградов С.А., Курчидис В. А., Чиркунов В.А., Назанский А.С., Гладков М.В., Сажин А.С., Голованов М., J. Celko, T. Haughey, M. Hillyer, L. Jons-son, S. Deleurme, D. Forbes, N. Gassiep, V. Tropashko, I. Ben-Gan. Однако, окончательное формирование данного направления еще не произошло. Можно выделить ряд теоретических и технологических проблем, требующих разрешения, и еще большее количество конкретных, прикладных задач, которые ждут реализации.

Для реляционного моделирования иерархий существует несколько способов. Каждый из способов основан на использовании различных свойств иерархии. Поэтому эффективность разных способов при решении одной и той же задачи может быть различной.

Целью диссертационного исследования является разработка методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными в контексте определенных для системы требований.

В соответствие с целью работы были поставлены и решены следующие научные задачи:

- изучение спектра современных теорий, информационных технологий и конкретных реализаций с целью выбора математических решений, наиболее пригодных для использования в задачах реляционного моделирования систем управления иерархическими данными;
- разработка способа математического моделирования требований к системам управления иерархическими данными;
- определение критериев эффективности способов реляционного моделирования систем управления иерархическими данными;
- разработка методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными;
- апробация методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными в процессе разработки реальных программных приложений.

Методы исследования. Для решения поставленных в диссертационной работе задач были применены методы системного анализа, методы объектно-ориентированного анализа и проектирования, методология UML, методы реляционного моделирования, теория чисел, статистические и численные методы, элементы имитационного моделирования.

Объектом исследования является моделирование систем управления иерархическими данными.

Предмет исследования: оценка эффективности способов реляционного моделирования систем управления иерархическими данными.

Научную новизну работы составляют:

- модель иерархического отношения, отличительной особенностью которой является использование непрерывных функций плотности распределения элементов по уровням иерархии и распределения дочерних элементов среди элементов одного уровня иерархии;
- модель функциональности системы управления иерархическими данными, где каждая функциональная единица описывается множеством иерархических операций с заданным для них законом распределения аргументов, что позволяет приближенно описывать на ранних стадиях разработки взаимодействия пользователя с системой.
- алгоритм генерации дерева и алгоритм генерации сценария нагрузочного тестирования по разработанным математическим моделям, основной чертой которых является применение подхода «разделяй и властвуй» для распределения дискретной величины (количество элементов) в соответствии с непрерывными функциями плотности ее распределения.

Практическая значимость

Разработана и внедрена автоматизированная система «Бизнес-справочник SIB-INFO.RU». Данная автоматизированная система является центральным звеном организации регионального бизнес-справочника по Республике Хакасия. На нее возложены задачи по хранению справочной информации, учету клиентов бизнес-справочника и обеспечению процесса расчетов, публикации справочной информации в сети Интернет, формирование информационного макета печатного издания бизнес-справочника.

Применение разработанной методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными при разработке автоматизированной системы «Бизнес-справочник «Товары и услуги Хакасии и юга Красноярского края» позволило эффективным способом реализовать иерархический справочник видов деятельности.

Разработана и внедрена автоматизированная система «Реестр субъектов малого предпринимательства Республики Хакасия». Данная автоматизированная система построена по принципам объектно-ориентированного анализа, проектирования и программирования. Для организации объектно-реляционного отображения, которое бы позволило осуществлять группировку данных, была использована древовидная структура, помещенная в реляционную базу данных. Хранение иерархии в реляционной базе данных позволяет использовать как преимущества иерархии, так и плюсы реляционных баз данных в обращении с линейными структурами.

Применение разработанной методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными при разра-

ботке автоматизированной системы «Реестр субъектов малого предпринимательства» позволило эффективным образом реализовать реляционно-иерархическое хранилище объектной модели.

На защиту выносятся:

- методика оценки эффективности способов реляционного моделирования систем управления иерархическими данными на основе нагрузочного тестирования, включающая построение иерархической модели, преобразование иерархической модели в реляционную, тестирование и анализ результатов;
- модель нагрузки для систем управления иерархическими данными, описывающая внутреннее состояние системы, взаимодействие с ней пользователей и критерии ее эффективности;
- программное обеспечение системы тестирования реляционных реализаций систем управления иерархическими данными, алгоритмы преобразования моделей в наборы данных для нагрузочного тестирования;
- способ применения разработанной методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными для адаптации архитектуры таких систем к условиям эксплуатации.

Апробация работы

Основные положения и результаты работы докладывались на следующих научно-технических конференциях: Конференция-конкурс «Технологии Microsoft в теории и практике программирования» (23 февраля 2006 года, г.Новосибирск), Российская научно-техническая конференция «Информатика и проблемы телекоммуникаций» (27 апреля 2006 года, г. Новосибирск), Конференция-конкурс «Технологии Microsoft в теории и практике программирования» (25 февраля 2007 года, г.Новосибирск).

Теоретические результаты, отдельные положения, а также результаты конкретных прикладных исследований и разработок обсуждались на научных семинарах Института Информатики и Телематики Хакасского Государственного Университета им. Н.Ф. Катанова (2003-2008 гг.)

Результаты исследований обсуждались на многочисленных рабочих совещаниях и семинарах с профильными специалистами и научными сотрудниками институтов города Абакана (Республика Хакасия) и города Красноярска (Красноярский край).

Публикации

Основное содержание диссертационной работы отражено в 10 печатных работах. В том числе одна работа опубликована в журнале, рекомендуемом ВАК.

Личный вклад автора

Автором проведен анализ подходов к моделированию иерархических систем, разработан способ моделирования нагрузки для систем управления иерархическими данными, разработана методика оценки эффективности способов реляционного моделирования иерархических систем, разработан способ перехода от математической модели нагрузки системы управления иерархическими данными к реляционной, проведена оценка эффективности реляционных реализаций систем управления иерархическими данными в ходе работы над несколькими программными приложениями.

Автор участвовал в создании автоматизированных систем основанных на способе оценки эффективности реляционных реализаций систем управления иерархическими данными, разработанном в данной диссертационной работе.

В ходе разработки автоматизированной системы «Бизнес-справочник SIB-INFO.RU» автор участвовал в постановке задачи для Интернет версии справочника, создании методики сбора и заполнения справочной информации, разработке системы управления прайс-листами компаний, разработал общую технологическую архитектуру системы, схемы баз данных, программное обеспечение для трансформации существующей структуры базы данных, технологию и программное средство выгрузки данных бизнес-справочника для создания печатных версий справочника, внутренние технические задания на элементы программного обеспечения, участвовал в оформлении документации на систему.

В ходе разработки автоматизированной системы «Реестр субъектов малого предпринимательства» автор участвовал в постановке задачи, анализе предметной области, разработал архитектуру системы, схемы баз данных, реализовал основные программные модули системы.

Структура работы

Диссертационная работа выполнена на 108 страницах машинописного текста, включающего в себя введение, четыре раздела, заключение, 30 рисунков, 10 таблиц и библиографический список из 129 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В диссертационной работе проводится анализ состояния вопроса моделирования иерархических структур. Разрабатывается методика оценки эффективности способов реляционного моделирования систем управления иерархическими данными на основе нагрузочного тестирования. Описывается модель нагрузки для систем управления иерархическими данными. Приводятся критерии эффективности. Описывается апробация разработанной методики в реальных проектах по разработке программных систем.

Во введении обоснована актуальность работы, сформулированы цели и задачи исследования, перечислены основные результаты работы.

Первый раздел диссертации посвящен изучению состояния вопроса разработки реляционных моделей иерархических систем и реляционной реализации иерархических операций.

В разделе выполняется анализ таких понятий как иерархия и иерархическая операция. Под иерархической операцией понимается операция, которая производится над иерархией. Описывается реляционная реализация иерархических операций, для которых в качестве аргументов выступают: исходное реляционное отношение, описывающее иерархию, и некоторое множество значений. В результате выполнения операции может произойти изменение исходного реляционного отношения, а в качестве возвращаемого результата получается реляционное отношение.

В основе реляционного моделирования системы управления иерархическими данными лежит построение такого реляционного отношения, по атрибутам которого можно восстановить установленные в иерархии связи. Также подразумевается существование внешнего отношения, в котором содержится информация, описывающая элементы иерархии. Для каждого элемента данного отношения вводятся первичный ключ, который используется для связи с элементами иерархического отношения.

Среди методов реляционного моделирования древовидных иерархий выделяют две группы: иерархические/рекурсивные расширения SQL и методы кодирования элементов дерева. Также можно выделить метод, который не относится к описанным выше двум группам. Данный метод подразумевают хранение всех пар элементов, между которыми определено транзитивное отношение. Этот метод будем называть методом транзитивного замыкания.

В заключение раздела, основываясь на различиях существующих способов реляционного моделирования систем управления иерархических данных и важности выбора такого способа для архитектуры разрабатываемой системы, обосновывается необходимость разработки методики оценки эффективности способов для конкретной системы. Под эффективностью способа реляционного моделирования системы управления иерархическими данными понимается эффективность функционирования построенной с помощью данного способа конкретной системы.

Основными критериями функционирования системы управления иерархическими данными являются: производительность системы и необходимая для работы системы память.

Во втором разделе рассматривается нагрузочное тестирование как метод оценки эффективности способов реляционного моделирования систем управления данными, и предлагается способ моделирования нагрузки для их тестирования, выделяются такие аспекты моделирования как моделирование данных и моделирование функциональности системы.

Модель данных определяет, какая информация будет находиться в системе управления иерархическими данными. В модели используются функции плотности распределения элементов иерархии. Модель данных представлена следующей формулой.

$M_D = \{n, f_{level}(x), C, MEM_{max}\}$, где n – общее количество элементов в иерархии, $f_{level}(x)$ – функция плотности распределения элементов по уровням и $x \in [0,1]$, $C = \{f_{children,1}(x), f_{children,2}(x), \dots, f_{children,m}(x)\}$ – множество, а $f_{children,j}(x)$ – функция плотности распределения количества сыновей у элементов j -го уровня и $x \in [0,1]$, MEM_{max} – верхнее ограничение по использованию памяти в пересчете на один элемент иерархии.

Для идентификации элемента в дереве применяются следующие координаты: номер уровня, на котором находится элемент, и порядковый номер элемента среди элементов данного уровня. На рис. 1. представлен пример дерева с расставленными координатами элементов.

Модель функциональности системы управления иерархическими данными определяет задачи, которые необходимо выполнять данной системе.

Такие задачи выражаются в терминах элементарных иерархических операций. Модель функциональности представлена на следующей формуле.

$M_F = \{F_1, F_2, \dots, F_i, \dots, F_{n-1}, F_n\}$, где F_i – единица функциональности системы управления иерархическими данными, а n – общее количество таких функций.

$F_i = \{k_i, T_{max,i}, H_i\}$, где k_i – весовой коэффициент важности данной функции для системы, $T_{max,i}$ – верхнее ограничение по времени выполнения рассматриваемой функции, H_i – множество описаний вызовов элементарных иерархических операций, $H_i = \{(h_1, g_1(x, y), o_1), \dots, (h_m, g_m(x, y), o_m)\}$, где m – количество описы-

ваемых в множестве вызовов, o_j – вид элементарной иерархической операции, h_j – весовой коэффициент в контексте задачи F_j для j -го вызова операции, $g_j(x, y)$ – двумерная функция плотности распределения выборки для j -го вызова операции.

В качестве критериев эффективности способов реляционного моделирования систем управления иерархических данных приводятся критерии эффективности функционирования таких систем. Основными критериями эффективности функционирования системы управления иерархическими данными являются: производительность системы – обратная величина времени выполнения функций системы; необходимая для работы системы память – размер базы данных.

Исходя из модели функциональности, описанной выше, критерий производительности рассчитывается следующим образом. Пусть Q^{F_i} – производительность системы при выполнении единицы функциональности F_i , а k_i является весовым коэффициентом важности данной единицы функциональности для системы. Тогда общая производительность системы рассчитывается по следующей формуле.

$$Q^F = \sum_{i=1}^n (k_i \cdot Q^{F_i}) \quad (1)$$

Производительность функциональной единицы системы рассчитывается на основе данных измерений. Пусть T_j^i – время выполнения элементарной иерархической операции o_j^i заданное количество раз, с учетом, распределенных по закону $g_j^i(x, y)$, входных данных. Тогда производительность при выполнении функциональной единицы системы рассчитывается по следующей формуле.

$$Q^{F_i} = \sum_{j=1}^n \frac{h_j^i}{T_j^i} \quad (2)$$

Критерием использования памяти системой управления иерархическими данными является обратная величина необходимой памяти для хранения одного элемента иерархии.

При применении разработанной модели в процессе нагрузочного тестирования возникают следующие задачи:

- задача построения реального дерева на основании заданной модели;
- задача построения списка тестовых испытаний для построенного по разработанной модели дерева.

Алгоритм построения дерева по заданным функциям плотности распределения элементов дерева по уровням и по количеству детей среди элементов уровня.

Для построения дерева, у которого количество элементов было бы равно n и распределено согласно f_{level} , а количество детей у элементов на каждом уровне согласно $f_{children}$, где f_{level} и $f_{children}$ определены на промежутке $[0, 1]$ применяется следующий алгоритм:

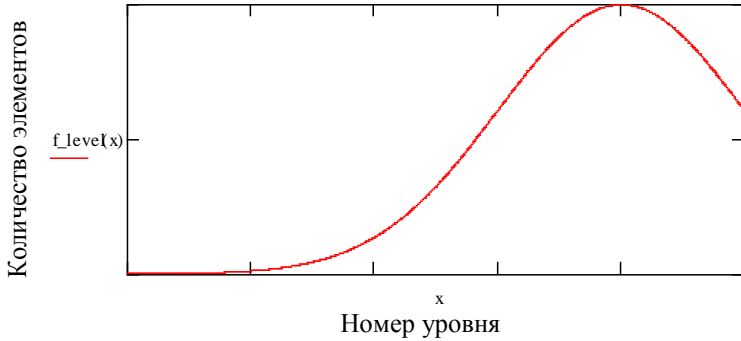
Шаг 1. Определение количества уровней в дереве.

Шаг 2. Определение количества элементов на каждом уровне дерева.

Шаг 3. В цикле для элементов каждого уровня определяется количество детей.

$$M_D = \{20, f_{level}(x) = \frac{1}{0.2 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-0.8)^2}{2 \cdot 0.2^2}}, \{f_{children}(x) = 1\}, \infty\}$$

Распределение элементов по уровням



Распределение детей среди элементов одного уровня

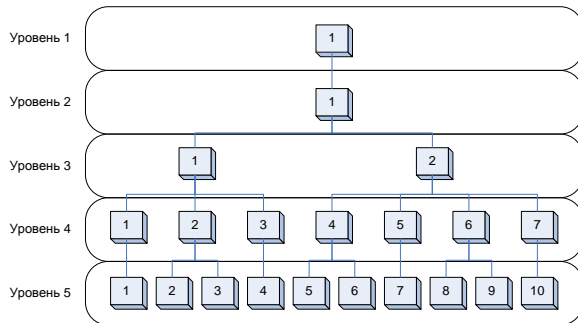
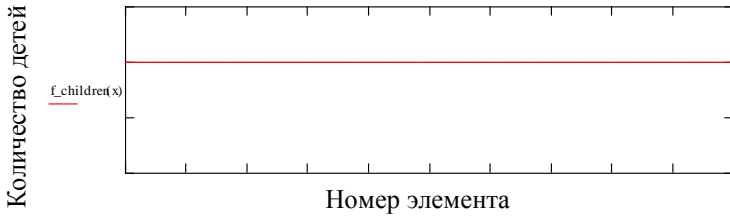


Рис. 1. Пример модели данных системы управления иерархическими данными

Шаг 1. Определение количества уровней в дереве.

На основе первоначальных данных можно сказать, что $f_{level}(x)$ определяет распределение по a уровням $\int_0^a f_{level}(x)dx$ элементов. Соответственно необходимо

найти такой коэффициент k , что $\int_0^{k \cdot a} k \cdot f_{level}\left(\frac{x}{k}\right)dx = n$. Пусть $k \cdot f_{level}\left(\frac{x}{k}\right) = g(x, k)$.

Для поиска k используем алгоритм, основанный на бинарном поиске.

Шаг 2. Определение количества элементов на каждом уровне

Данная задача может быть перефразирована следующим образом. Дано X шаров, N пронумерованных от 1 до N корзин и некоторый закон распределения заданный функцией плотности. Необходимо распределить шары по корзинам по данному закону распределения.

Решение задачи:

Для решения данной задачи используется прием «разделяй и властвуй». Суть его заключается в следующем – для известного интервала номеров корзин и количества шаров, которое должно попасть в данный интервал необходимо узнать, сколько шаров будет попадать в левый и правый интервалы, если данный интервал разделить пополам по середине.

Пусть $f(x)$ – функция плотности распределения определена на промежутке $[a, b]$. Тогда алгоритм решения рассматриваемой задачи можно записать следующей рекурсивной функцией (см. рис. 3.), принцип которой рассмотрен на рис. 4. Пусть a, b – целые, неотрицательные числа, $m = b - a, m > 0$. Тогда, если $m=1$, то на a -ом уровне будет находиться n элементов, иначе распределяем m . Определяем, что в промежутке

$\left[a, \left\lfloor \frac{a+b}{2} \right\rfloor \right]$ находится $\left\lfloor \frac{S_1 \cdot m}{S_1 + S_2} \right\rfloor$ шаров, а в промежутке

$\left[\left\lfloor \frac{a+b}{2} \right\rfloor, b \right]$ находится $\left\lfloor \frac{S_2 \cdot m}{S_1 + S_2} \right\rfloor$. Остаток распределяется в зависимости от величины дробной части. В нашем случае, для распределения элементов по уровням, описанная выше функция должна быть вызвана следующим образом *РаспределитьШары* ($1, a \cdot k, n$).

Шаг 3. Определение количества детей у каждого элемента

Распределение детей среди элементов уровня также сводится к решению задачи размещения шаров по корзинам с заданной функцией распределения.

Количество элементов на текущем уровне будет равно количеству корзин, а количество элементов на следующем уровне количеству шаров.

Алгоритм построения псевдослучайной последовательности тестирования иерархической операции, построенной по заданному закону распределения.

Задача построения псевдослучайной последовательности тестирования иерархической операции, заключается в следующем. Дано:

- множество элементов дерева заданных двумя координатами (номером уровня и номером внутри уровня), номер максимального уровня max_level , максимальный номер элемента внутри уровня max_on_level ;

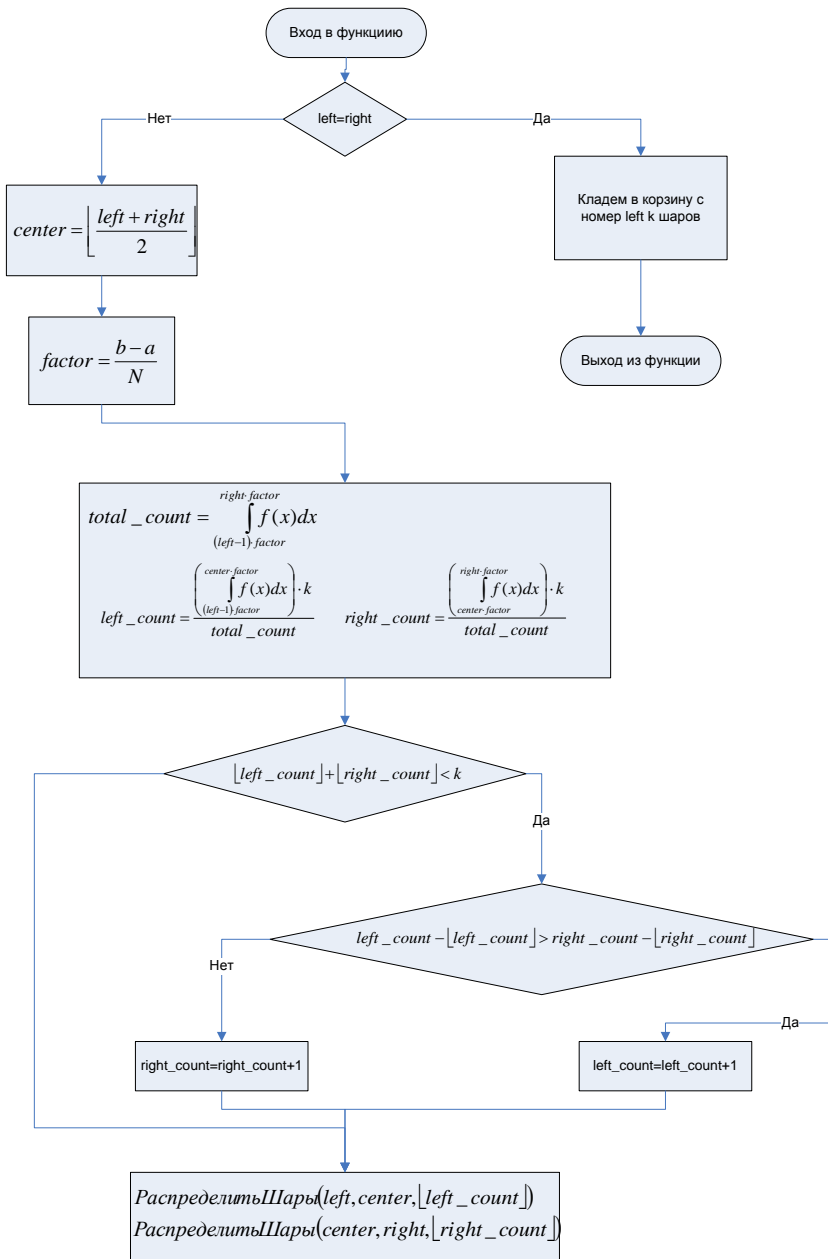


Рис. 3. Функция *РаспределитьШары(left, right, k)*

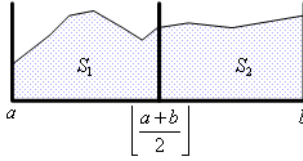


Рис. 4. Иллюстрация работы функции *РаспределитьШары* (a, b, n).

- двумерная функция плотности распределения выборки $f(x, y)$, заданная на прямоугольной области $[0, a][0, b]$;

- необходимое количество тестов n .

Необходимо найти псевдослучайную последовательность элементов заданного дерева, закон распределения, которых совпадал бы с законом заданным функцией плотности $f(x, y)$.

Для решения данной задачи также используется прием «разделяй и властвуй». Она похожа на задачу распределения шаров по корзинам с заданным распределением, за исключением того, что теперь закон распределения и расположение шаров двумерно. Перед началом распределения элементов необходимо нормировать $f(x, y)$ для прямоугольной области $[1, \max_level][1, \max_on_level]$. Пусть $g(x, y)$ будет нормированной функцией плотности распределения, и находится следующим образом:

$$factor = \max\left(\frac{\max_level}{a}, \frac{\max_on_level}{b}\right) \quad g(x, y) = factor \cdot f\left(\frac{x}{factor}, \frac{y}{factor}\right)$$

Алгоритм распределения тестов может быть описан рекурсивной функцией *РаспределитьТесты*($f(x, y), tree, n, a, b, c, d$) (см. рис. 5). В данном вызове $f(x, y)$ – функция плотности распределения выборки; $tree$ – двумерный массив, определяющий дерево; n – количество тестов; $[a, b][c, d]$ – область дерева, на которой необходимо распределить n тестов по закону распределения, описываемому функцией плотности $f(x, y)$. После выполнения описанной выше функции мы получим список элементов дерева. Теперь необходимо равномерно перемешать данный список. Этого можно добиться, перемещая элементы в другой список, выбирая элементы из существующего списка случайным образом с равномерным распределением.

В третьем разделе рассматривается методика оценки эффективности способов реляционного моделирования систем управления иерархическими данными на основе нагрузочного тестирования. Описывается система тестирования, исходными данными для которой являются: модель нагрузки и способы реляционного моделирования систем управления иерархическими данными. Тестирующая система дает ответ на вопрос, какой способ реляционного моделирования в наибольшей степени подходит для реализации тестируемой модели.

В начале раздела описывается процесс нагрузочного тестирования реляционных реализаций систем управления иерархическими данными, который можно разбить на следующие этапы:

- создание программных модулей реляционных моделей для использования в системе тестирования;
- описание модели нагрузки;
- создание плана тестирования;
- тестирование;

- получение оценок критериев эффективности в разрезе функциональных единиц;
- получение интегральной оценки критериев эффективности.

Затем анализируется процесс выполнения реляционных операций в рамках системы «клиент-сервер». В основе процесса выполнения операций в системе «клиент-сервер», где в качестве сервера выступает реляционная СУБД, является SQL-запрос, обращенный от клиента серверу, и ответ на него, особенно важно рассмотреть процесс обработки такого запроса. Обычно процесс выполнения SQL-запроса обобщают, выделяя следующие четыре фазы: декомпозиция, оптимизация, генерация кода, выполнение. В нашем случае система управления иерархическими данными содержит в себе РСУБД (далее, СУБД) как самостоятельную подсистему. Значит, в обобщенный процесс обработки SQL-запроса следует добавить еще три фазы: подготовка запроса, посылка запроса, получение ответа. Соответственно, стоимость выполнения SQL-запроса можно представить в следующем виде:

$$T_{sql} = T_{prepare} + T_{send} + T_{optimization} + T_{perform} + T_{receive} \quad (4)$$

где T_{sql} – общее время выполнения SQL-запроса, $T_{prepare}$ – время на подготовку SQL-запроса внешней по отношению к СУБД системой, а также сюда следует включить время на генерацию различных производных от хранящихся в базе данных параметров необходимых для обработки запроса, T_{send} – время затрачиваемое на посылку запроса (возможно инициализацию соединения с СУБД), $T_{optimization}$ – время на оптимизацию запроса самой СУБД, $T_{perform}$ – время обработки запроса (выполнения операций реляционной алгебры), $T_{receive}$ – время затрачиваемое, на получение результатов запроса.

Следует обратить внимание на то, что некоторые иерархические операции требуют выполнения нескольких запросов. Это не может не влиять на производительность данных операций. Соответственно, общее время выполнения операции можно представить в следующем виде:

$$T_{operation} = \sum_{i=1}^n (T_{sql})_i \quad (5)$$

где $(T_{sql})_i$ – время выполнения i -го SQL-запроса; $T_{operation}$ – общее время выполнения операции; n – количество запросов, необходимое для выполнения операции.

Исходя из оценки $T_{operation}$ и описанных формул разработчики могут найти причину недостаточной производительности системы.

Далее описывается программная система для оценки описанных ранее критериев эффективности. Ключевым требованием к архитектуре системы тестирования является абстрагирование от метода реляционного моделирования иерархий, то есть тестирующая система должна выполнять эксперимент одинаково, вне зависимости от того какой метод тестируется.

Также важно обеспечить возможность задания различных планов тестирования; абстрагирование от используемой системы управления базами данных; возможность использования различных исходных данных, также и возможность генерирования исходных данных; совместимость формата выходных данных тестирующей системы с математическими пакетами программ. Например, MathCad.

Исходя из поставленных задач, систему тестирования можно разделить на два основных компонента (см. рис. 6.): преобразователь математической модели в реляционный вид и подсистема тестирования реляционной модели. В задачи

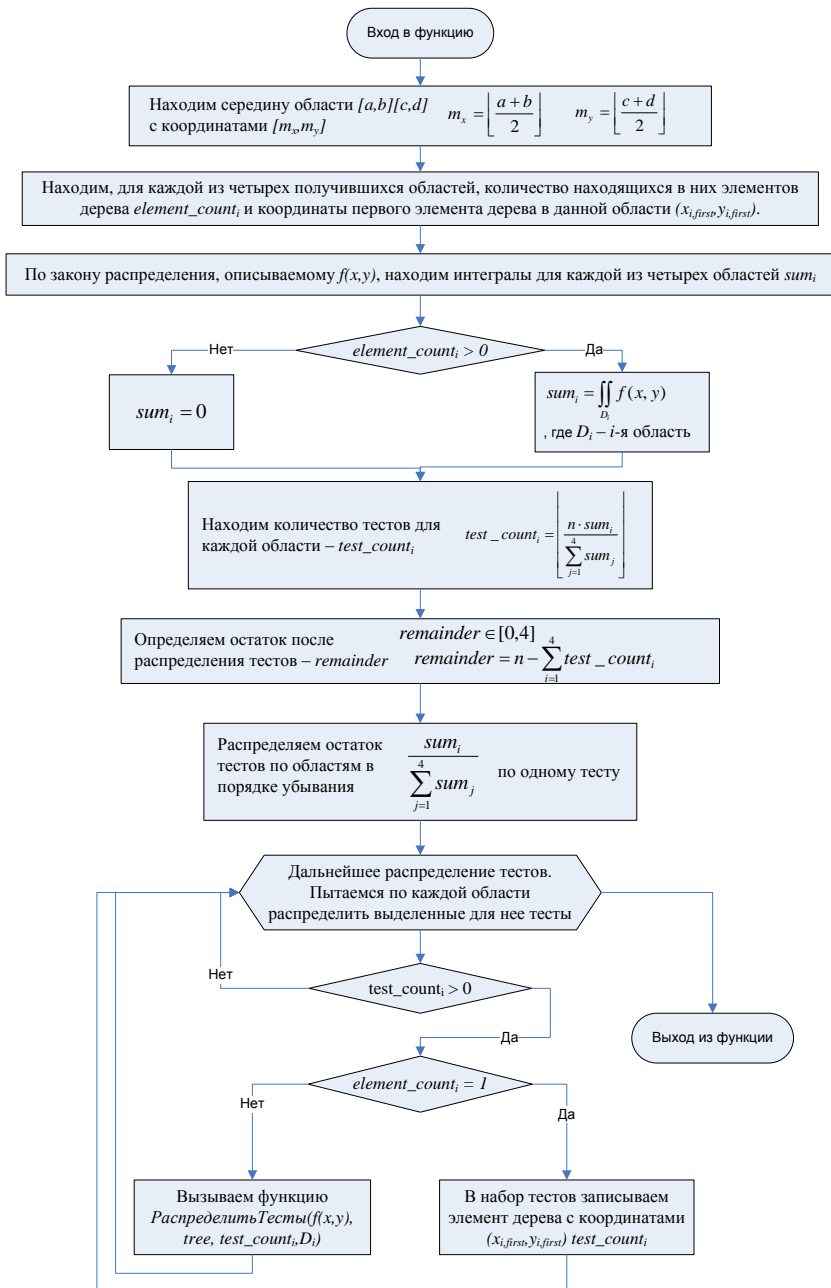


Рис. 5. Функция *РаспределитьТесты*($f(x,y)$, $tree$, n , a , b , c , d)

преобразователя математической модели в реляционный вид входит: построение на основе математической модели промежуточной реляционной модели, по которой уже будут строиться модели в фазе тестирования, и составление на основе математической модели и плана тестирования последовательности вызовов иерархических операций с аргументами из промежуточной реляционной модели.

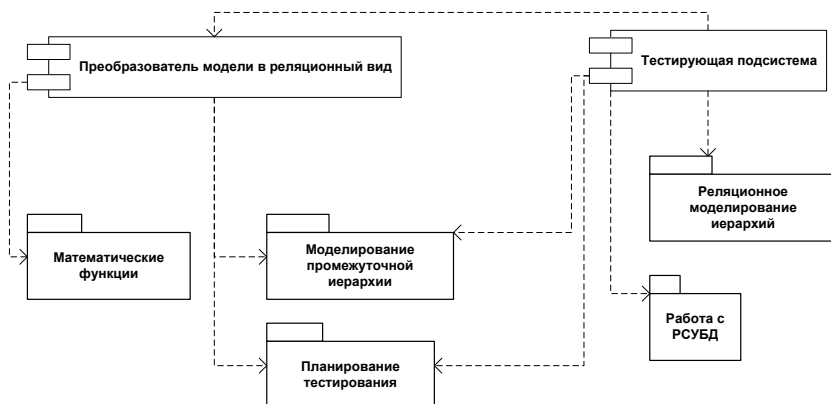


Рис. 6. UML диаграмма компонентов системы тестирования

Задачей подсистемы тестирования реляционной модели является выполнение плана тестирования на сгенерированном по модели дереве, замер времени выполнения и вычисление критериев производительности.

В заключение раздела рассматривается возможность использования разработанной методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными в качестве средства адаптации архитектуры таких систем к условиям среды эксплуатации (см. рис. 7). В ходе мониторинга системы управления иерархическими данными производится замер изменяемых характеристик данных и использования функциональности. По данной информации производится оценка эффективности способов реляционного моделирования, и дается рекомендация выбора того или иного способа. Что приводит к замене соответствующего модуля в архитектуре.

В четвертом разделе описывается место методики оценки эффективности в процессе разработки приложений, описываются решаемые с помощью нее задачи. Также в разделе приводятся сведения о внедрении результатов диссертационной работы. Разработанная методика оценки эффективности способов реляционного моделирования систем управления иерархическими данными была использована при принятии архитектурных решений в ходе разработки нескольких программных систем.

В ходе разработки автоматизированной системы «Бизнес-справочник SIB-INFO.RU» разработанная методика оценки эффективности способов реляционного моделирования систем управления иерархическими данными была использована для выбора способа реализации иерархического справочника видов деятельности.



Рис. 7. Адаптация архитектуры системы управления иерархическими данными основанной на РСУБД к среде эксплуатации

В ходе разработки автоматизированной системы «Реестр субъектов малого предпринимательства Республики Хакасия» разработанная в данной работе методика оценки способов реляционного моделирования систем управления иерархическими данными была использована для оценки эффективности способов реализации реляционно-иерархического хранилища объектной модели.

Ниже описывается использование разработанной методики для разработки автоматизированной системы «Реестр субъектов малого предпринимательства Республики Хакасия». Для определения данных хранилища объектной модели была использована следующая модель:

$$M_D = \{160000, f_{level}(x) = 120000 \cdot x + 20000, \{f_{children}(x) = \frac{1}{0.5 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-0.5)^2}{2 \cdot 0.5^2}}\}, \infty\}$$

$$M_F = \{F_{detail}, F_{general}, F_{direct_detail}\}$$

$$F_{general} = \left\{ 1, \infty, \left\{ 1, f(x, y) = \frac{1}{0.0001 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-1)^2}{2 \cdot 0.0001^2}} + 1, o_{ancestors} \right\} \right\}, \text{ где } o_{an-}$$

cestors – операция поиска множества предков.

$$F_{detail} = \left\{ 1, \infty, \left\{ 1, f(x, y) = \frac{1}{0.0001 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-1)^2}{2 \cdot 0.0001^2}} + 1, o_{descendants} \right\} \right\}, \text{ где } o_{descen-}$$

dants – операция поиска множества потомков.

$$F_{direct_detail} = \left\{ 1, \infty, \left\{ 1, f(x, y) = \frac{1}{0.0001 \cdot \sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-0.5)^2}{2 \cdot 0.0001^2}} + 1, o_{children} \right\} \right\}, \text{ где } o_{children} -$$

операция поиска множества сыновей.

На основании описанной выше модели было произведено тестирование реляционных реализаций реляционно-иерархического хранилища объектной модели автоматизированной системы «Реестр субъектов малого предпринимательства». Результаты оценки производительности и использования памяти представлены в таблицах 1 и 2.

Таблица 1.
Результаты оценки производительности автоматизированной системы «Реестр субъектов малого предпринимательства»

Методы	Оценка эффективности для функций			Интегральная оценка эффективности
	Поиск родителей	Поиск сыновей	Поиск потомков	
Список смежности	0.171	0.064	0.147	0.617
Материализация пути	0.217	0.135	0.122	0.826
Вложенные множества	15.724	178.475	0.112	388.51
Транзитивное замыкание	0.075	0.066	0.06	0.341
Вложенные интервалы	41.046	30.562	30.504	173.716
Объединенный метод списка смежности и материализации пути	0.071	0.064	0.13	0.4

Таблица 2.
Результаты оценки использования памяти автоматизированной системы «Реестр субъектов малого предпринимательства»

Методы	Размер, Кб
Список смежности	20144
Материализация пути	45320
Вложенные множества	56086
Транзитивное замыкание	36634
Вложенные интервалы	53904
Объединенный метод списка смежности и материализации пути	35232

По данным тестирования производительности наилучшие результаты показали: метод транзитивного замыкания и объединения методов списка смежности и материализации пути. По данным тестирования использования памяти наилучший результат показал метод списка смежности. Еще два метода показали несколько худший результат. На основании проведенного тестирования было принято решение о реляционной реализации реляционно-иерархического хранилища объектной модели автоматизированной системы «Реестр субъектов малого предпринимательства» по объединенному методу списка смежности и материализации пути.

В заключении перечислены основные научные и практические результаты работы.

В приложениях описываются методы реляционного моделирования иерархий.

ОСНОВНЫЕ НАУЧНЫЕ И ПРАКТИЧЕСКИЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИОННОЙ РАБОТЫ

1. Разработана модель иерархического отношения, отличительной особенностью которой является использование непрерывных функций плотности распределения элементов по уровням иерархии и распределения дочерних элементов среди элементов одного уровня иерархии. Применение непрерывных функций обеспечивает простоту масштабирования заданной с помощью них структуры иерархии. Данная модель легла в основу описания внутреннего состояния системы управления иерархическими данными при нагрузочном тестировании.

2. Разработана модель функциональности системы управления иерархическими данными, где каждая функциональная единица описывается множеством иерархических операций с заданным для них законом распределения аргументов. Данная модель позволяет уже на ранних стадиях разработки приблизительно описывать взаимодействие пользователя с системой и сценарии нагрузочного тестирования.

3. Разработаны алгоритм генерации дерева и алгоритм генерации сценария нагрузочного тестирования по разработанным математическим моделям, основной чертой которых является применение подхода «разделяй и властвуй» для распределения дискретной величины (количество элементов) в соответствии с непрерывными функциями плотности ее распределения. Применение данного подхода позволяет пропорционально распределить точное количество элементов.

4. Разработана методика оценки эффективности способов реляционного моделирования систем управления иерархическими данными на основе нагрузочного тестирования, включающая построение иерархической модели, преобразование иерархической модели в реляционную, тестирование, анализ результатов. Разработано математическое и программное обеспечение системы тестирования, алгоритмы преобразования моделей в наборы данных для нагрузочного тестирования. Данный комплекс программных средств и методических указаний обеспечивает возможность оценки эффективности способов реляционного моделирования систем управления иерархическими данными уже на ранних стадиях их разработки.

5. Разработан способ применения методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными для адаптации архитектуры таких систем к условиям эксплуатации. Применение рассматриваемой методики позволяет оценить эффективность различных методов реляционного моделирования иерархий для сложившейся вовремя эксплуатации ситуации и выбрать наиболее эффективный.

6. Произведена апробация разработанной методики оценки эффективности способов реляционного моделирования систем управления иерархическими данными в процессе разработки реальных программных приложений. Применение рассматриваемой методики оценки позволило эффективным способом реализовать иерархический справочник видов деятельности для автоматизированной системы «Бизнес-справочник «Товары и услуги Хакасии и юга Красноярского края» и ре-

ляционно-иерархического хранилище объектной модели для автоматизированной системы «Реестр субъектов малого предпринимательства».

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Гребенщиков Н.Н., Применение современных Интернет-технологий при разработке сайта ХГУ. Вестник Хакасского государственного университета им. Н.Ф. Катанова. Выпуск 4. Серия 1: Информатика – Абакан: Издательство Хакасского государственного университета им. Н.Ф. Катанова, 2001. – С. 19-22.

2. Гребенщиков Н.Н., Пример подхода к процессу разработки интерактивного Интернет-сайта. Вестник Хакасского государственного университета им. Н.Ф. Катанова. Выпуск 4. Серия 1: Информатика – Абакан: Издательство Хакасского государственного университета им. Н.Ф. Катанова, 2003. – С. 23-25.

3. Гребенщиков Н.Н., Способ сравнения эффективности методов реляционного моделирования иерархических структур. Вестник Хакасского государственного университета им. Н.Ф. Катанова. Информатика – Абакан: Издательство Хакасского государственного университета им. Н.Ф. Катанова, 2006. – С. 28-31.

4. Гребенщиков Н.Н., Сравнение эффективности методов реляционного моделирования древовидного классификатора бизнес-справочника. Вестник Хакасского государственного университета им. Н.Ф. Катанова. Информатика – Абакан: Издательство Хакасского государственного университета им. Н.Ф. Катанова, 2006. – С. 31-33.

5. Гребенщиков Н.Н., Масштабируемое реляционно-иерархическое хранилище объектной модели. // Тезисы докладов конференции конкурса «Технологии Microsoft в теории и практике программирования» – Новосибирск: НГУ, 2006. – С. 81-83.

6. Гребенщиков Н.Н., Метод реляционного моделирования иерархий. // Информатика и проблемы телекоммуникаций. – Новосибирск: СибГУТИ, 2006. – С. 27-31.

7. Нагрузочное тестирование реляционных реализаций систем управления иерархическими данными [Текст]: свидетельство об отраслевой регистрации разработки №7440 / Н.Н. Гребенщиков. – № 50200700071; заявл. 18.12.2006; опубл. 26.12.2006; Инновации в науке и образовании № 12(23). – С. 28.

8. Гребенщиков Н.Н., Моделирование систем управления иерархическими данными для нагрузочного тестирования на начальных этапах разработки программных систем [Электронный ресурс] // МИТС-НАУКА: международный научный вестник: сетевое научное издание. – 2007. – №1. Зарегистрировано ФГУП НТУ «Информрегистр» 13.02.2007 №0420700032/0007.

9. Гребенщиков Н.Н., Метод генерации древовидных данных // Тезисы докладов конференции конкурса «Технологии Microsoft в теории и практике программирования» – Новосибирск: НГУ, 2007. – С. 104-106.

10. Гребенщиков Н.Н., Представление древовидной зависимости в реляционной базе данных // Программные продукты и системы. – 2008. – №1 (81). – Тверь: НИИ ЦПС. – С. 41-45.

Соискатель

A handwritten signature in black ink, appearing to read "А. С. Сидоров" (A. S. Sidorov), written in a cursive style.

Подписано в печать _____.2008
Усл. Печ. Л. 1. Тираж ____ экз.