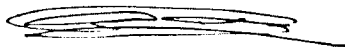
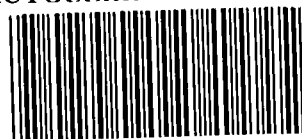


На правах рукописи



Горшков Виктор Сергеевич

**МОДЕЛИ И МЕТОДИКИ ПОСТРОЕНИЯ
РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ ХРАНИЛИЩ
АВТОМАТИЗИРОВАННЫХ СИСТЕМ МЧС РОССИИ**



4849361

05.25.05 – информационные системы и процессы

9 ИЮН 2011

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

Санкт-Петербург – 2011

Работа выполнена в Санкт-Петербургском университете Государственной противопожарной службы МЧС России

Научный руководитель: доктор технических наук, профессор
Иванов Александр Юрьевич

Официальные оппоненты: доктор технических наук, профессор
Малыгин Игорь Геннадьевич;
кандидат технических наук
Василевич Евгений Владимирович

Ведущая организация: Северо-Западный региональный центр
МЧС России

Защита состоится «28» апреля 2011 года в 12 часов на заседании совета по защите докторских и кандидатских диссертаций Д 205.003.02 при Санкт-Петербургском университете Государственной противопожарной службы МЧС России (196105, Санкт-Петербург, Московский проспект, д.149).

С диссертацией можно ознакомиться в библиотеке Санкт-Петербургского университета Государственной противопожарной службы МЧС России.

Автореферат разослан «23» марта 2011 г.

Ученый секретарь
диссертационного совета Д 205.003.02
доктор технических наук, профессор



А.Ю. Иванов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Предупреждение чрезвычайных ситуаций (ЧС), их устранение и ликвидация последствий являются важными составляющими обеспечения безопасности населения, экономической инфраструктуры и страны в целом. Для борьбы с поражающими факторами ЧС создаются группировки сил и средств МЧС России. Координацию действий этих сформированных осуществляют органы управления различных уровней иерархии, к основным из которых относятся Центры управления в кризисных ситуациях (ЦУКС).

Результативность применения сил и средств МЧС в существенной мере определяется способностью органов управления к выработке решений, адекватных развивающейся в зоне ЧС обстановке. Требуемый уровень оперативности и обоснованности решений в современных условиях не может быть достигнут без автоматизации управленческой деятельности. Поэтому важнейшим элементом системы управления в ЧС являются автоматизированные системы (АС).

Сложность и нетривиальность процессов выработки решений по действиям в ЧС определяют характер функциональных задач управления, решаемых должностными лицами ЦУКС в автоматизированном режиме: большинство наиболее ответственных задач относится к классу информационно-аналитических, целью решения которых является поддержка принятия решений по управлению силами и средствами МЧС. Специфический характер названных задач требует соответствующего информационного обеспечения. Классические базы данных (БД) не могут быть использованы в этом качестве, поскольку по своей сущности они являются отражением временного среза предметной области и не предполагают учета динамики изменения состояния объектов учета.

Устранить этот недостаток способны информационные хранилища (ИХ), строящиеся как многомерные структурированные совокупности данных. Эти хранилища ориентированы на решение задач, связанных с анализом и прогнозом различных процессов, и служат информационной основой систем поддержки принятия решений.

Существенный вклад в становление и развитие концепции информационного хранилища внесли видные ученые Э. Кодд (*Codd E.*), К. Дейт (*Date C.*), У. Инмон (*Inmon W.*). Необходимо отметить отечественных специалистов в этой области, таких как: Артамопов В.С., Куватов В.И., Кузнецов С.Д., Львов В.В., Поляков А.С., Сахаров С.С., Сасенко И.Б., Таранцев А.А., Щербаков О.В. и другие.

Сформировавшаяся концепция ИХ, к тому же имеющая многочисленные приложения, в существующем виде не может быть рекомендована в качестве теоретической основы для построения информационного обеспечения решения задач анализа и прогноза в АС МЧС России. Здесь слабым звеном является централизация хранилища как источника данных для распределенных систем обработки, к которым относятся названные автоматизированные системы. Такой характер размещения данных не позволяет ожидать высоких значений показателей оперативности и надежности решения информационно-аналитических задач.

В сложившейся ситуации возникает *проблема* совершенствования информационного обеспечения автоматизированных систем МЧС России путем организации распределенных информационных хранилищ (РИХ). Анализ публикаций свидетельствует, что в настоящее время отсутствуют научные подходы к решению названной проблемы. Этим определяется *актуальность* темы диссертационного исследования.

Цель работы состоит в повышении эффективности функционирования АС МЧС России за счет построения и внедрения распределенных информационных хранилищ, обеспечивающих своевременный и непрерывный доступ должностных лиц к актуальной и непротиворечивой аналитической информации, необходимой для принятия решений по управлению силами и средствами в операциях по предупреждению и ликвидации ЧС.

Объект исследования – информационные хранилища как основа построения систем, связанных с автоматизацией процессов анализа и прогнозирования развития ЧС и применения сил и средств для их предупреждения и ликвидации.

Предмет исследования – модели и методы организации распределенных информационных хранилищ.

Научная задача заключается в разработке моделей и методик структурной организации и оценки распределенных информационных хранилищ автоматизированных систем МЧС России.

Частные научные задачи исследования:

1. Формирование основ технологии распределенных информационных хранилищ.

2. Разработка структурных решений на построение распределенных информационных хранилищ.

3. Разработка последовательности проектирования распределенных информационных хранилищ.

4. Оценка эффективности построения распределенных информационных хранилищ.

5. Выработка рекомендаций по применению распределенных информационных хранилищ в автоматизированных системах МЧС России.

Методы исследования. Для решения научной задачи использовались методы общей теории систем, теории баз данных и теории реляционных баз данных в частности, теории эффективности, теории массового обслуживания, теории распределений, а также методы математического моделирования.

Результаты исследования. Основными результатами диссертационной работы, выносимыми на защиту, являются:

1. Концептуальная модель распределенного информационного хранилища.

2. Модель представления данных в распределенных информационных хранилищах.

3. Методика проектирования распределенных информационных хранилищ.

4. Методика оценки эффективности распределенных информационных хранилищ.

Научная новизна результатов диссертационного исследования обусловлена формированием общего облика РИХ как логически интегрирован-

ного и физически распределенного информационного образования, обеспечивающего независимость различных способов представления сведений о предметной области при решении задач аналитической обработки данных; адаптацией многомерно-реляционного представления данных к сетевой топологии, что позволяет децентрализовать информационно-аналитическое обслуживание пользователей; учетом специфики архитектурного построения распределенных информационных хранилищ, определяющей структуру их проектирования и содержание этапов этого процесса, а также согласованием структурного построения РИХ с процессом автоматизированного решения функциональных задач, что обуславливает объективность оценки эффективности распределенных информационных хранилищ.

Достоверность научных результатов обеспечивается использованием апробированного методологического аппарата сформировавшихся теорий в области математических методов и информационных технологий, а также корректностью основных положений и обобщений проведенного исследования.

Практическая значимость полученных результатов определяется их важностью для организации автоматизированного управления в условиях ЧС. Предпосылкой применения разработанных моделей и методик является стремление к повышению оперативности и обоснованности решений за счет реализации информационно-аналитических методов.

Проверка предложенного в работе методологического инструментария, связанного с организацией распределенных информационных хранилищ, показала его результативность и возможность использования в научно-исследовательских и опытно-конструкторских работах по автоматизации управления силами и средствами МЧС. Определенная общность результатов позволяет сделать вывод о том, что модели и методики имеют научно-прикладное значение для других предметных областей, принятие решений в которых проходит в условиях дефицита времени и требует глубокого анализа и аргументированного прогноза развития обстановки.

Публикации по теме диссертации. Результаты диссертационного ис-

следования опубликованы в шести работах, в том числе в одном издании по перечню ВАК. Список публикаций приведен в конце автореферата.

Реализация. По результатам работы получены акты реализации от следующих организаций: Департамент пожарно-спасательных сил, специальной пожарной охраны и сил гражданской обороны МЧС России, Санкт-Петербургский университет ГПС МЧС России.

Апробация результатов исследования. Основные положения исследования докладывались и обсуждались в период с 2009 г. по 2010 г. на заседаниях научно-тематического семинара Санкт-Петербургского университета ГПС МЧС России, VI Международной научно-практической конференции «Технические средства противодействия террористическим и криминальным взрывам» (Санкт-Петербург, 2010), международном семинаре «Предупреждение пожаров и организация надзорной деятельности» (Санкт-Петербург, 2010).

Структура и объем работы. Диссертация состоит из введения, трех глав, заключения, списка литературы и приложений. Общий объем диссертации составляет 115 страниц основного текста, в том числе 26 рисунков, 5 таблиц и список литературы из 71 наименования.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснованы выбор темы диссертации и ее актуальность, приведены основные атрибуты исследования, отражены научные результаты, выносимые на защиту, и их характеристика, а также указаны сведения об апробации и реализации результатов диссертационного исследования.

В первой главе «Анализ проблемы построения распределенных информационных хранилищ автоматизированных систем МЧС России» проведено исследование наиболее важных и масштабных автоматизированных систем МЧС, изучено состояние их информационного обеспечения, дана характеристика предметной области, связанной с ЧС, и решаемых в автоматизированном режиме функциональных задач по управлению силами и средствами, а также выявлена и сформулирована проблема диссертационной ра-

боты и обозначено генеральное направление ее решения.

Постоянные угрозы возникновения ЧС природного и техногенного характера определили необходимость создания в Российской Федерации Единой государственной системы предупреждения и ликвидации чрезвычайных ситуаций (РСЧС), объединяющей органы управления, силы и средства федеральных органов исполнительной власти, органов исполнительной власти субъектов РФ, органов местного самоуправления и организаций, в полномочия которых входит решение вопросов в области защиты населения и территорий от ЧС. Управление системой осуществляет МЧС России. Стержневым элементом системы управления РСЧС выступает иерархическая совокупность ЦУКС и оперативно-дежурных служб органов управления МЧС.

В целях своевременного и полного обеспечения информацией должностных лиц в РСЧС создана и функционирует автоматизирующая информационно-управляющая система (АИУС). В АИУС автоматизирован сбор и обработка информации о ЧС и проводимых мероприятиях, их учёт и накопление. Функциональные комплексы и задачи этой системы призваны обеспечить решение информационно-аналитических, прогностических, функциональных и специальных технологических задач управления.

АИУС РСЧС тесно взаимодействует с Национальным центром управления в кризисных ситуациях (НЦУКС), который представляет собой территориально-распределенный информационно-управляющий комплекс с периферийными элементами, позволяющими управлять силами, средствами и ресурсами РСЧС и гражданской обороны в условиях кризисов и ЧС. НЦУКС включает в себя соответствующие органы повседневного управления системы МЧС России (ЦУКС МЧС России, ЦУКС региональных центров, ЦУКС главных управлений МЧС России по субъектам РФ), их пункты управления, сети связи и передачи данных, а также автоматизированную систему. Основными структурными элементами АС НЦУКС являются стационарные объектовые комплексы, которые строятся на базе современных аппаратно-программных средств с использованием архитектуры локальных вычисли-

тельных сетей и состоит из комплекса средств автоматизации, комплекса средств связи и телекоммуникаций и центра оповещения.

Анализ видов обеспечения АИУС РСЧС и АС НЦУКС показал, что эти системы характеризуются иерархичностью и территориально-распределенным структурным построением. К числу их функций относится полное и своевременное обеспечение информацией о ЧС должностных лиц органов управления РСЧС и МЧС, в частности. Поддержка принятия решений осуществляется путем решения функциональных задач управления. Одно из приоритетных направлений этого процесса – прогнозирование развития ЧС и выработка планов применения сил и средств при их ликвидации.

Автоматизированное решение названных задач требует информационного обеспечения, которое адекватно отражает предметную область, связанную с возникновением, протеканием и ликвидацией последствий ЧС. Созданные к настоящему времени в рамках АС МЧС информационные системы носят преимущественно узконаправленный характер. Построенные по принципу «снизу-вверх», путем непрерывного наращивания и увязки старых и новых технологий, существующие информационные системы не формируют единую информационную среду. Отчасти это обусловлено следованием традиционным подходам к построению информационных систем, основу которых составляют базы данных (БД). Построенные по традиционной технологии БД не приспособлены для решения аналитических задач, которые предполагается использовать в процессе управления силами и средствами в ЧС.

Предметная область, связанная с ЧС, является сложной, многогранной и неоднородной. Ее основные кластеры применительно к информационному обеспечению АИУС РСЧС и АС НЦУКС представлены в таблице 1. Каждый кластер представляет собой совокупность групп показателей, каждая из которых включает в себя некоторое число информационных единиц. Указанные группы связаны между собой ссылками. Многие группы и показатели имеют привязку ко времени. Эти показатели служат базисом решения задач управления силами и средствами в операциях по ликвидации ЧС. Отличи-

тельной чертой названных задач является их информационно-аналитический характер, требующий реализации новых подходов к обработке данных, накапливаемых в автоматизированных системах МЧС.

Таблица 1 – Кластеры предметной области, связанной с ЧС

АИУС РСЧС	АС НЦУКС
Территориальные образования (включая проживающее население)	Характеристики регионов и субъектов РФ
Объекты экономики (в первую очередь потенциально опасные объекты)	Критически важные для национальной безопасности объекты и другие потенциально-опасные объекты
Силы и средства предупреждения и ликвидации ЧС, материально-технические, продовольственные, медицинские и другие ресурсы	Силы, средства и ресурсы
Окружающая среда (атмосфера, акватории, земная поверхность, недра)	Материалы дистанционного зондирования земли
Статистика ЧС	Прошедшие ЧС и документы, которые были отработаны по данным ЧС
Мероприятия по предупреждению и ликвидации ЧС	Типовые сценарии, планы ликвидации типовых ЧС и смягчения последствий циклических рисков
Органы и системы управления	

Эффект от внедрения информационных систем в практику автоматизированного управления определяется возможностью решения должностными лицами органов управления задач, связанных с аналитической обработкой данных. Этот класс задач базируется на концепции оперативной аналитической обработки (ОАОД) и интеллектуального анализа данных (ИАД). Под ОАОД понимают технологии комплексного многомерного анализа данных для поддержки принятия решений, а ИАД основан на статистических методах и служит для выявления заранее неизвестных закономерностей. Основу ОАОД и ИАД составляет представление информации в виде гиперкубов (рисунок 1).

БД, построенные на основе гиперкубов, относят к категории аналитических (АБД). Совместно с классическими реляционными БД, называемыми операционными, АБД образуют информационное хранилище (ИХ). Во многих случаях под термином «информационное хранилище» именно АБД.

Модификациями АБД являются темпоральные и многоверсионные базы данных (ТБД и МВБД соответственно). При этом ТБД – это базы данных, по-

строенные с дополнительным измерением «Время», а МВБД рассматриваются как модификация ТБД, в которых снято ограничение упорядоченности временных рядов и измерение «Варианты» сопоставлено измерению «Время». При решении функциональных задач по управлению силами и средствами МЧС востребованными являются все названные варианты АБД. При этом предполагается комплексное использование указанных БД и традиционных операционных баз данных.

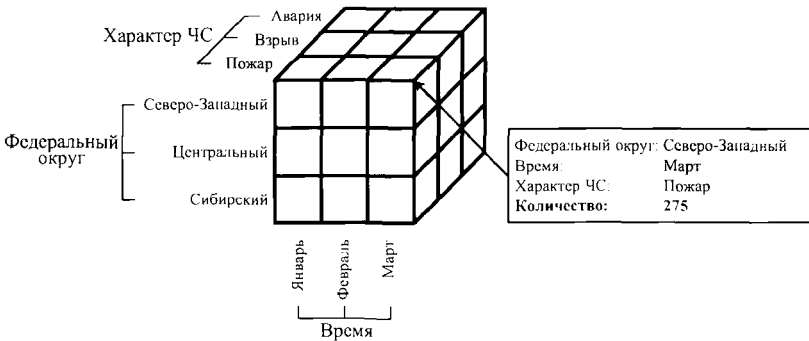


Рисунок 1 – Представление информации о предметной области в виде гиперкуба

Используемые в настоящее время ИХ не способны обеспечить требования по оперативности и надежности выполнения пользовательских запросов в распределенных системах в силу своего централизованного построения. В то же время ориентация на аналитическую обработку накапливаемых данных определяет необходимость их использования в практике автоматизированного управления силами и средствами в ЧС. Выход из сложившейся проблемной ситуации видится на пути построения распределенных информационных хранилищ (РИХ) АС МЧС России. Имеющийся научный и практический опыт построения и использования распределенных баз данных создает определенные предпосылки к формированию аналогичного класса информационных хранилищ.

Вторая глава «Построение распределенных информационных хранилищ» посвящена изложению концептуальных положений по организации распределенных информационных хранилищ, в частности принципов их построения и архитектурного облика, разработке многоуровневой модели пред-

ставления данных в распределенных информационных хранилищах, как развитию многомерно-реляционной модели, а также формированию методики проектирования распределенных информационных хранилищ.

Информационным хранилищем АС МЧС называется специальным образом организованная совокупность БД, предназначенная для обеспечения должностных лиц ОУ МЧС информацией в интересах ОАОД и поддержки принятия решений по применению сил и средств в ЧС. Тогда *распределенное информационное хранилище* – это логически интегрированное хранилище, физически размещаемое на нескольких территориально распределенных компьютерах сети.

Целью построения РИХ АС МЧС является обеспечение оперативного доступа должностных лиц к многомерной агрегированной информации, характеризующей состояние сил и средств МЧС, а также окружающей среды, в настоящее время и в ретроспективе. *Задачами*, решаемыми должностными лицами ОУ МЧС с помощью РИХ, являются анализ текущей и предыдущих ситуаций, а также прогнозирование развития ЧС и действий в них сил и средств на ближайшую и удаленную перспективы.

Сложность РИХ и важность решаемых с их помощью задач определяют необходимость обоснования *принципов* построения таких хранилищ. Принципы устанавливают соответствие между объективными требованиями к организации РИХ и субъективной деятельностью разработчиков во избежание принятия ошибочных решений на построение и управление хранилищами. В более узком аспекте принципы можно рассматривать как общие требования, которым должны удовлетворять хранилища. Все множество принципов можно разделить на два подмножества (таблица 2).

Архитектура РИХ определяется совокупностью решений по его структурному построению (составу и взаимной связи входящих в него компонентов) на различных уровнях рассмотрения. Это построение можно рассматривать на нескольких уровнях представления данных, таких как: концептуальный, логический, фрагментный, распределенный и физический.

На *концептуальном уровне* РИХ рассматривается с точки зрения конечного пользователя вне зависимости от программных и аппаратных инструментальных средств. Результатом представления информации на этом уровне является концептуальная модель данных. Эта модель является многомерной, имеющей несколько информационных измерений. По каждому измерению может быть организована одна или несколько иерархий элементов, разделенных по уровням. Уровни иерархии необходимы для агрегирования детализированных данных.

Таблица 2 – Принципы построения распределенных информационных хранилищ

Организационные принципы	Технические принципы	
1. Единство методических, организационных и технологических основ разработки и применения РИХ по назначению	1. Многомерное концептуальное представление данных	7. Динамическая обработка распределенных данных
	2. Прозрачность данных	8. Поддержка многопользовательского режима
2. Соответствие уровня сервиса, полноты, актуальности и оперативности информации, содержащейся в РИХ, требованиям эффективности реализации функций управления	3. Прозрачность распределения	9. Поддержка кроссмерных операций
	4. Устойчивая производительность	10. Интуитивное манипулирование данными
3. Оценка эффективности от создания и внедрения РИХ	5. Клиент-серверная архитектура	11. Гибкий механизм генерации отчетов
4. Координация работ по созданию и применению РИХ со стороны должностных лиц, ответственных за применение АС МЧС	6. Равноправие измерений	12. Неограниченное количество измерений и уровней агрегации

Логический уровень представления данных в РИХ определяет его структуру с учетом системы управления информационным хранилищем (СУИХ), в качестве которой может выступать традиционная или специализированная СУБД, ориентированная на организацию хранилищ. Результатом представления данных является глобальная логическая структура (схема) РИХ. Возможны различные подходы к построению названной структуры.

Первый подход основан на использовании многомерной модели представления данных, являющейся расширением реляционной модели. В многомерной модели таблицы преобразуются в гиперкубы, вдоль ребер которых

лежат информационные измерения.

Второй подход связан с использованием реляционной модели данных и состоит в эмуляции многомерной модели с помощью реляционной СУБД.

Третий подход заключается в гибридном сочетании первых двух и предполагает возможность одновременной поддержки многомерной и реляционной моделей.

Фрагментный уровень представления данных используется для определения непересекающихся подмножеств данных, называемых логическими фрагментами. При использовании многомерной модели расчленение целостной схемы данных может проводиться по информационным измерениям, а при ориентации на реляционную модель – путем горизонтальной и/или вертикальной фрагментации.

На *уровне распределения* данных определяется географическое расположение экземпляров каждого логического фрагмента. Уровень допускает существование нескольких физических фрагментов, соответствующих одному логическому фрагменту.

Физический уровень соответствует описанию той части хранилища, которая существует в конкретном узле. На этом уровне и реляционные таблицы, и гиперкубы данных преобразуются, как правило, к одной и той же форме представления – табличной.

Модель данных, свойственная ИХ, основана на многомерно-реляционном представлении данных (МРПД). Цель МРПД – это реализация многомерных АБД в среде СУБД, поддерживающих традиционную реляционную модель, а также интеграция АБД и операционных баз данных под управлением единой реляционной СУБД.

Под многомерными данными понимают такие атрибуты (атрибуты-факторы), значения которых зависят от значений, принимаемых сразу несколькими другими атрибутами (атрибутами-измерениями). Значения атрибутов-факторов интересуют пользователей БД в разрезе нескольких информационных измерений.

Пусть в БД находятся атрибуты $\{A_i^d\}$, обладающие домспами $Dom(A_i^d)$, и атрибуты $\{A_j^f\}$, обладающие доменами $Dom(A_j^f)$, а на доменах атрибутов выполняются следующие множества отображений:

$$MD_j : Dom(A_{i_1}^d) \times Dom(A_{i_2}^d) \times \dots \times Dom(A_{i_j}^d) \rightarrow Dom(A_j^f) \quad (1)$$

Тогда атрибуты $\{A_j^f\}$ – это атрибуты-факторы, а атрибуты $\{A_i^d\}$ – их атрибуты-измерения. Зависимости, соответствующие (1) называют *многомерными зависимостями*.

Пусть имеется множество атрибутов-измерений $\{A_i^d\}$, $i=1, \dots, n$, и множество атрибутов-факторов $\{A_j^f\}$, $j=1, \dots, m$. При этом значения обоих типов атрибутов удовлетворяют следующему множеству отображений

$$MD_j : Dom(A_1^d) \times Dom(A_2^d) \times \dots \times Dom(A_n^d) \rightarrow Dom(A_j^f), \quad j=1, \dots, m. \quad (2)$$

В этом случае считается, что между атрибутами $\{A_i^d\}$ и $\{A_j^f\}$ существует *многомерное отношение* M , имеющее схему

$$Sch(M) = \langle A_1^d, A_2^d, \dots, A_n^d, A_1^f, A_2^f, \dots, A_m^f \rangle. \quad (3)$$

Многомерное отношение M может быть представлено в виде реляционного отношения R со схемой $\mathbf{R} = \langle A_1^d, A_2^d, \dots, A_n^d, A_1^f, A_2^f, \dots, A_m^f \rangle$, причем ключом этого отношения является $\mathbf{K} = \langle A_1^d, A_2^d, \dots, A_n^d \rangle$.

Реляционный объект, отображающий набор многомерных зависимостей, образующих схему одного многомерного отношения, называют *многомерно-реляционным набором* (МРН) данных. Для многомерного отношения $M^{(n)}$ степени n многомерно-реляционным набором $S^{(n)}$ степени n является совокупность реляционных отношений $S^{(n)} = \langle \{R_i^f\}, \{\{R_{1j}^d\}\}, \dots, \{\{R_{nj}^d\}\} \rangle$, которая отображает многомерное отношение $M^{(n)}$. В состав МРН $S^{(n)}$ входят два типа реляционных отношений, называемых отношениями (таблицами) факторов (R_j^f) и отношениями (таблицами) измерений (R_{ij}^d) .

Сущность операций над МРН может быть описана следующим образом.

Проекция предназначена для сокращения числа учитываемых в МРН атрибутов и выдачи таблиц только с теми атрибутами, которые интересуют пользователя. *Селекция* предназначена для выборки из таблиц МРН кортежей, значения атрибутов которых удовлетворяют логическому условию. *Соединение* МРН позволяет формировать новый набор на основе двух исходных, имеющих одно или несколько общих информационных измерений. С помощью операции *среза* МРН возможно отображение значений атрибутов-факторов по фиксированным измерениям и значениям атрибутов-измерений. Операции *подъема* и *спуска* предназначены для отображения атрибутов-факторов на уровнях иерархии, соседних по отношению к заданному.

Одной из проблем, решаемых при построении РИХ, является представление многомерных данных в хранилище. За основу такого представления принята многоуровневая модель, определяющая архитектуру РИХ. В соответствии с таким подходом *логический уровень* представления данных в РИХ предполагает использование многомерно-реляционной модели. Ее базовым элементом является МРН.

Фрагментный уровень представления данных используется для формирования непересекающиеся подмножеств РИХ, называемых логическими фрагментами. Возможны следующие виды фрагментации.

Вертикальная фрагментация первого порядка возможна применением к МРН операции вертикального среза:

$$S2 = VSlice(S1|k_{slice}), \quad k_{slice} = \{k_1, k_2, \dots, k_r\}, \quad (4)$$

где $S1$ – исходный МРН, $S2$ – результирующий МРН, k_{slice} – вектор индексов r информационных измерений, по которым проводится операция.

Для осуществления *вертикальной фрагментации второго порядка* необходимо задействовать операцию проекции МРН:

$$S2 = Project(S1|A_\pi), \quad (5)$$

где A_π – список атрибутов исходного МРН как условие проекции.

В отличие от вертикального среза, который понижает степень исходного МРН, проецирование не обязательно приводит к такому же результату.

Горизонтальная фрагментация первого порядка реализуется операцией горизонтального среза, которая не меняет степень исходного набора:

$$S2 = HSlice(S | P_{\sigma}(A^d)), \quad (6)$$

где $P_{\sigma}(A^d)$ – условие поиска на атрибутах измерений.

При выполнении *горизонтальной фрагментации второго порядка* требуется использовать операцию селекции:

$$S2 = Select(S | P_{\sigma}(A_{\sigma})), \quad (7)$$

где $P_{\sigma}(A_{\sigma})$ – условие селекции.

Отличием селекции от горизонтального среза является то, что при селектировании степень результирующего набора может понижаться в зависимости от условия, заданного в операции.

На *уровне распределения* копии логических фрагментов (физические фрагменты), размещаются в узлах сети. Известны три варианта размещения фрагментов данных: разделение, дублирование, комбинирование.

В любом случае представление данных на уровне распределения устанавливает отношение между узлами сети и логическими фрагментами распределенного информационного хранилища. Это описание может задаваться в виде матрицы соответствия $X = \|x_{ij}\|$, где x_{ij} – булева переменная, такая что

$$x_{ij} = \begin{cases} 1, & \text{если } j\text{-й фрагмент подлежит размещению в } i\text{-м узле сети;} \\ 0, & \text{если } j\text{-й фрагмент не подлежит размещению в } i\text{-м узле сети,} \end{cases}$$

$j = 1, \dots, N$; $i = 1, \dots, M$; N – число логических фрагментов, M – число узлов.

На *физическом уровне* данные представляются в виде совокупности реляционных таблиц, хранящихся в каждом узле сети. При этом многомерные наборы, полученные при фрагментации исходного МРН, преобразуются к табличной форме представления. Это позволяет считать представление данных на физическом уровне соответствующим классическому реляционному подходу.

Процесс проектирования РИХ целесообразно рассмотреть как совокуп-

ность этапов, определяющих общую схему построения хранилища. Основные этапы этой последовательности представлены на рисунке 2.



Рисунок 2 – Общая схема проектирования РИХ

Этап *концептуального моделирования* предназначен для изучения, как предметной области, так и задач ОАОД и ИАД, а также для формирования концептуальной модели хранилища. Этап *логического проектирования* связан с созданием логической структуры информационного хранилища. Этап *фрагментирования* логической структуры нацелен на получение набора логических фрагментов. Этап *распределения* копий логических фрагментов необходим для формирования содержательного наполнения узлов сети необходимыми данными. Этап *физической реализации* имеет целью преобразование фрагментированной структуры РИХ в структуру данных локального уровня.

Каждый этап характеризуется своим комплексом работ, исходными данными для их проведения и получаемым результатом. Процесс проектирования РИХ является многоэтапным и итеративным. При реализации этапов этого процесса могут использоваться формальные и эвристические методы как специфические для информационных хранилищ, так и общие для классических операционных баз данных, в том числе распределенных.

Третья глава «Применение распределенных информационных хранилищ» содержит решения по оценке эффективности распределенных информационных хранилищ на основе математических моделей, отражающих процесс решения функциональных задач по управлению силами и средствами МЧС, а также охватывает материал, связанный с выработкой рекомендации по практическому использованию этого класса информационного обеспечения автоматизированных систем.

Оценку эффективности структурной организации РИХ рекомендуется проводить на основе внешнего подхода, сущность которого применительно к настоящему исследованию состоит в следующем.

В качестве критерия эффективности используется $\max P(t_p \leq t_{дон})$, где t_p – время решения задачи (комплекса задач), $t_{дон}$ – допустимое время решения.

Исполнение запросов к РИХ формально представляется как обработка заявок в стохастической сети массового обслуживания (СМО), узлы которой интерпретируют элементы АС, в частности серверы баз данных, на которых хранятся фрагменты РИХ. Структура СМО зависит от варианта топологии РИХ. Модель процесса реализации запросов построена как имитационная. С ее помощью получают значения числовых характеристик случайной величины – времени обработки запросов, в частности – первые четыре начальных момента.

Процесс решения функциональных задач рассматривается как последовательность операций по обращению к РИХ и интерпретируется оргграфом, узлы которого отображают операции по обработке запросов к РИХ, а дуги – связи между операциями (рисунок 3). С позиций теории массового обслуживания граф может быть представлен как СМО, в которой последовательность решения задачи описывается движением заявки от узла 0 к узлу k .

Для определения функции распределения случайной величины (времени пребывания заявок в сети) достаточно знать первые четыре центральных момента этой величины. Дальнейший переход к искомой функции осуществляется с помощью семейств кривых. Преобразование начальных моментов времени пребывания заявки на каждом из узлов сети в центральные моменты времени прохождения за-

явкой сети выполняется в соответствии с методами теории распределений.

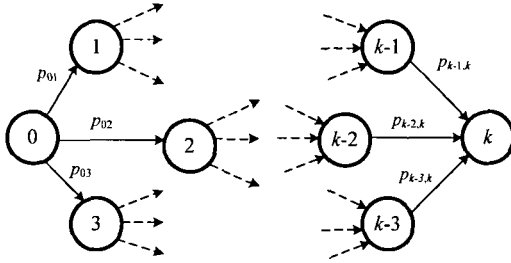


Рисунок 3 – Представление процесса решения задачи в виде графа

На базе рассмотренных моделей разработана методика оценки эффективности РИХ, в процедурном плане представляющая собой последовательность этапов, таких как: определение исходных данных, настройка имитационной модели на конкретный вариант размещения фрагментов РИХ, проведение эксперимента на этой модели, формирование графового представления процесса решения функциональных задач, определение временных параметров этого процесса, а также расчет значений показателей эффективности решения комплекса задач.

Приведенный в главе пример построения макета РИХ демонстрирует возможность реализации методики проектирования. В ходе разработки сформирована концептуальная модель предметной области, построена и фрагментирована применительно к иерархии НЦУКС логическая структура РИХ, а также проведено распределение фрагментов по узлам сети. В результате определен структурный состав локальных БД реляционного типа, предполагаемых к использованию в НЦУКС.

Проведенные модельные испытания позволили выявить преимущества распределенных информационных хранилищ перед централизованными путем сравнения характера изменения зависимостей параметров хранилищ от внешних факторов, влияющих на результаты их применения (рисунки 4–6). Это позволило выработать рекомендации по применению РИХ:

1) распределенные информационные хранилища должны по возможности охватывать все центры (пункты) управления силами и средствами МЧС, задействованные в операциях по устранению ЧС и ликвидации их последствий;

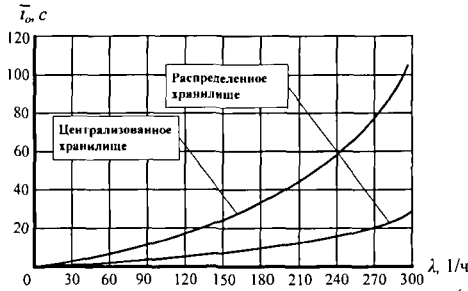


Рисунок 4 – Зависимость среднего времени обработки (\bar{t}_o) запросов от интенсивности их поступления (λ)

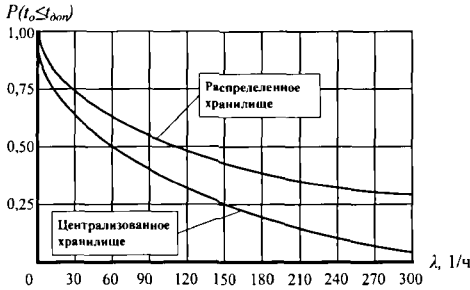


Рисунок 5 – Зависимость вероятности своевременной обработки запросов $P(t_o \leq t_{дон})$ от интенсивности их поступления (λ)

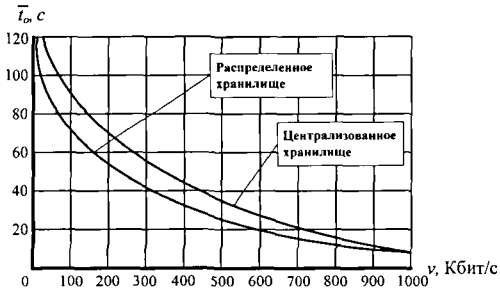


Рисунок 6 – Зависимость среднего времени обработки (\bar{t}_o) запросов от скорости каналов передачи данных (v)

2) наибольший эффект от применения РИХ наблюдается при повышенной нагрузке на АС в период выработки в автоматизированном режиме решений по управлению силами и средствами в ЧС;

3) РИХ целесообразно использовать в АС, базирующихся на низкоскоростные каналы передачи данных.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

В диссертационной работе представлены решения, связанные с разработкой общего подхода, моделей и методик построения и применения распределенных информационных хранилищ автоматизированных систем МЧС России с целью повышения эффективности функционирования этих систем.

Предложена концептуальная модель распределенного информационного хранилища, которая определяет принципиальные и архитектурные основы его организации.

Разработана многоуровневая модель представления многомерных данных в распределенных информационных хранилищах, являющаяся основой выделения и структуризации информационных объектов хранилища с учетом их сетевого размещения.

Сформирована методика проектирования распределенных информационных хранилищ, позволяющая реализовывать комплекс работ по переходу от концептуального описания предметной области к формированию информационного наполнения узлов сети, на которую базируется хранилище.

Разработана методика оценки эффективности распределенных информационных хранилищ, ориентированная на получение значений показателей своевременности решения функциональных задач по управлению силами и средствами МЧС с учетом характера структурного построения информационного базиса.

Представленный в работе инструментарий рекомендуется использовать при разработке информационного обеспечения процессов автоматизированного решения задач аналитической обработки данных для поддержки принятия решений в органах управления МЧС России.

Основные работы, опубликованные по теме диссертации

Издания по Перечню ВАК РФ

1. Горшков В.С. Концепция распределенных информационных хранилищ / В.С. Горшков, А.Ю. Иванов // Проблемы управления рисками в техносфере. – 2010. – № 1 [13]. (0,85/0,6 п.л.).

Ведомственные издания

2. Горшков В.С. Модель представления данных в распределенных информационных хранилищах / В.С. Горшков // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России. Научный электронный журнал. <http://vestnik.igps.ru>. – 2010. № 1 (0,8 п.л.).

3. Горшков В.С. Методика оценки эффективности распределенных информационных хранилищ / В.С. Горшков, А.Ю. Иванов, С.Л. Исаков // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России. Научный электронный журнал. <http://vestnik.igps.ru>. – 2010. № 2. (0,8/0,4 п.л.).

4. Горшков В.С. Методика проектирования распределенных информационных хранилищ / В.С. Горшков // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России. Научный электронный журнал. <http://vestnik.igps.ru>. – 2010. № 4. (0,8 п.л.).

5. Горшков В.С. К вопросу информационного обеспечения систем контроля доступа на объекты / В.С. Горшков, А.Ю. Иванов, И.И. Попивчак // Технические средства противодействия террористическим и криминальным взрывам: тезисы докладов VI Междунар. науч.-прак. конф. / Санкт-Петербургский университет ГПС МЧС России. – СПб., 2010. (0,2/0,1 п.л.)

6. Горшков В.С. Информационный базис систем анализа и прогнозирования пожарной обстановки / В.С. Горшков // Предупреждение пожаров и организация надзорной деятельности: материалы Междунар. семина. / Санкт-Петербургский университет ГПС МЧС России. – СПб., 2010. (0,2 п.л.).