

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

На правах рукописи

Степанов Сергей Юрьевич

Разработка геоинформационной системы на основе использования разнородной
пространственно-распределенной информации в интересах управления территориями

Специальность 25.00.35 - Геоинформатика

Диссертация
на соискание учёной степени
кандидата технических наук

Научный руководитель:
доктор технических наук,
профессор Е.П. Истомин

Научный консультант:
кандидат географических наук,
А.А. Фокичева

Санкт-Петербург – 2017

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
1. Использование разнородных данных в геоинформационных системах при решении задач управления территориями	10
1.1. Теоретические аспекты интеграции разнородных данных	10
1.2. Обработка и размещение разнородных геопространственных данных исполнительных органов государственной власти в системе РГИС	15
1.3. Анализ прикладных систем и технологий обработки разнородных данных	27
1.4. Обзор программных инструментов для доступа к обработке баз данных	42
1.5. Разработка модели обработки разнородных данных для решения прикладных задач	51
2. Разработка методики управления разнородными пространственно-распределенными данными в геоинформационной системе на основе метеорологической информации	66
2.1. Алгоритм управления разнородными метеорологическими данными для подготовки варианта оптимального управленческого решения	67
2.2. Использование методики обработки разнородной метеорологической информации на примере оптимизации отпуска тепла ТЭЦ	78
2.2.1 Разработка функции потерь для теплоэнергетики с использованием модели управления разнородными данными	80
2.2.2. Разработка погодо-хозяйственного регламента с использованием разнородных метеорологических данных	90
2.2.3. Оценка экономической полезности прогнозов температуры воздуха для теплоэнергетики	98
3. Разработка геоинформационной системы поддержки принятия управленческих решений ...	102
3.1. Методика проектирования геоинформационной системы поддержки принятия решений на основе модели обработки разнородных данных	103
3.2. Верификация модели обработки и управления разнородными данными на примере задачи оптимизации регулирования отпуска тепла на ТЭЦ в геоинформационной системе.....	121
3.2.1 Формирование базы разнородных метеорологических данных по территории Санкт-Петербурга	123
3.2.2 Разработка геоинформационной системы с использованием разнородных пространственно-распределенных метеорологических данных	131
ЗАКЛЮЧЕНИЕ	140
Список сокращений и условных обозначений.	142
СПИСОК ЛИТЕРАТУРЫ.....	143
ПРИЛОЖЕНИЕ	153

ВВЕДЕНИЕ

Актуальность исследования. Одной из тенденций последнего десятилетия в исследованиях стал комплексный подход к обработке данных: совместное использование данных различных инструментов, в том числе различных источников данных. Поскольку каждый из них может обладать как преимуществами, так и ограничениями в применении, которые имеют различную природу, объединение данных представляет собой бесспорный выигрыш в сравнении с раздельной обработкой, за счет расширения объема получаемой информации благодаря как взаимодополняемости, так и взаимозаменяемости.

Объединение данных различного происхождения для последующих совместного анализа содержат разнородные данные, представляемые различными форматами, структурами и реализуемыми на разнотипных платформах. Главной задачей технологии слияния разнородных данных является решение содержательных задач: принятие решений, классификация, определение состояния объектов, оценка ситуации и т.д. Среди собственных задач, связанных с объединением данных, наиболее значимыми являются:

- создания и развития метамodelей распределенных источников данных;
- создания и развития моделей объединения решений;
- построения и упорядочивания архитектуры данных.

Управление данными на современном предприятии характеризуется наличием большого количества разнородных источников данных, не связанных едиными механизмами управления, в том числе и слабоструктурированных или неструктурированных данных. Раздельно существуют системы оперативного управления данными, системы управления документами, а также системы аналитической обработки. Различные задачи требуют использования различных моделей представления данных.

В современных условиях все отрасли экономики испытывают возрастающую потребность в метеорологической информации. Особую актуальность приобретает разработка путей экономически выгодного применения гидрометеорологической информации в технологических процессах отраслей экономики.

В современных условиях быстро развивается техника, совершенствуется структура всех отраслей производства, и вместе с этим растут запросы к службе погоды в объективной прогностической информации. Особое значение приобретают прогнозы для таких отраслей, как энергетика, сельское хозяйство, строительство, всех видов транспорта. Региональные разработки экономико-метеорологических моделей позволяют оптимально реализовать информацию о погоде в хозяйственной практике.

Разработка и внедрение экономико-метеорологических моделей, позволяющих включить метеорологическую информацию в процесс принятия управленческих решений, позволяют эффективно реализовывать информацию о погоде в различных областях хозяйственной деятельности и значительно снизить издержки, обусловленные неблагоприятным влиянием погодных условий. Важнейшей проблемой, решению которой в значительной степени посвящена данная работа, является разработка методических подходов к поддержке принятия решений с использованием разнородной метеорологической информации. В этом состоит актуальность данной работы, что диктуется необходимостью применения современных программных решений к использованию гидрометеорологической информации в различных сферах производственной деятельности.

Объем доступной для анализа информации в настоящее время нарастает на порядки быстрее чем вычислительные мощности. Традиционные математические методы и модели в такой ситуации становятся неприменимы. Необходимо создание новой модели, адаптированной под новые соотношения между данными и вычислительными ресурсами.

Из чего следует научная задача, решению которой посвящена диссертационная работа: «Разработка информационных технологий интеграции разнородных геоинформационных данных для обеспечения принятия решения по управлению территорией» [12, 25].

Степень разработанности проблемы. В процессе исследования были изучены труды отечественных и зарубежных ученых по рассматриваемой проблеме, таких как Акальцов В.П., Алексеев В.В., Берлянт А.М, Бескид П.П., Бурковский В.Л., Дорофеев А.Н., Духин С.В., Истомина Е.П., Кошкарев А.В., Татарникова Т.М., Фуфаев Э.В., Хандожко Л.А. и других.

Проанализированы результаты научных исследований, нормативно-правовые документы, связанные с управлением территорий, развитием регионов и проблемами использования больших массивов данных в задачах оптимизации и принятия управленческих решений.

Проблема интеграции пространственных данных в различных форматах в России изучена недостаточно. Не в полной мере разработаны технологии и методы обработки геоинформации на основе распределенных, разнородных баз данных. Таким образом, необходимость совершенствования механизмов работы с распределенными геоданными различных форматов определяет объект и предмет, цель и задачи диссертации.

Объектом исследования являются геоинформационные системы обработки пространственно-распределенной разнородной информации.

Предмет исследования: технология использования пространственно-распределенной разнородной информации в геоинформационных системах.

Цель диссертационной работы: автоматизация процесса обработки разнородной пространственно-распределенной информации для обеспечения принятия решения по управлению территорией с использованием гидрометеорологической информации.

Реализация поставленной цели исследования предопределяет постановку следующих **задач**:

1. Модифицировать модель обработки и управления разнородной геопространственной информацией;
2. Усовершенствовать методику проектирования геоинформационной системы с использованием модели обработки разнородной геопространственной информацией;
3. Разработать методику управления разнородными данными для автоматизации подготовки варианта управленческого решения;
4. Разработать геоинформационную систему для поддержки принятия управленческого решения, на основе использования разнородных геоданных.

Теоретические и методические основы исследования. Теоретической основой исследования являются труды отечественных и зарубежных ученых в области анализа механизмов доступа к разнородным данным, проблематика использования пространственных геоданных в различных геосистемах, аналитические обзоры, нормативно-правовые документы, связанные с методикой построения информационных систем. Методической основой исследования является обобщение и анализ существующего опыта создания распределенных информационных и геоинформационных систем с разнородными данными, системный анализ, концептуальное и математическое гео моделирование.

Обоснованность и достоверность результатов исследования, выводов и рекомендаций обеспечивается:

- использованием для достижения цели работы нормативных документов, программ, документов федеральных и региональных органов власти, касающихся методики построения информационных систем;
- внутренней непротиворечивостью результатов исследования и их соответствием теоретическим положениям фундаментальных исследований в области технологий сбора, регистрации, хранения, передачи и обработки геоинформации с использованием вычислительной техники, телекоммуникационных систем распространения пространственно-временной геоинформации, технологий хранения и использования геоинформации на основе распределенных баз данных и знаний;

- применением принципов системного анализа и концептуального моделирования, аналитических исследований, математического моделирования и других современных научных методов;
- апробацией результатов исследования на научно-практических конференциях и отражением основных результатов диссертации в открытой печати.

При решении поставленных в работе задач получены следующие **результаты, выносимые на защиту**:

1. Модель обработки и управления разнородной геопространственной информацией;
2. Методика проектирования геоинформационной системы с использованием модели обработки разнородной геопространственной информации, модифицированная с учетом требований к системе поддержки принятия решений;
3. Методика управления разнородными данными для автоматизации подготовки варианта управленческого решения;
4. Геоинформационная система управления пространственными разнородными данными для поддержки принятия управленческого решения, применительно к задачам управления территориями.

Научная новизна.

Автором проведен широкий анализ механизмов, моделей и систем обработки распределенной пространственной информации с использованием информационных технологий и определено, что ни один представленный инструмент не может полностью удовлетворить требованиям обработки разнородных данных при создании геоинформационных систем.

Построена модель обработки и управления разнородной геопространственной информацией, основанная на технологиях доступа к данным, представлена схема потоков информации и описан механизм эффективного управления данными. Приведен алгоритм модуля обработки данных, который позволяет получить доступ к любым источникам информации, необходимым для принятия управленческого решения.

Усовершенствована методика проектирования геоинформационной системы с использованием модели обработки разнородной геопространственной информации, с учетом требований к системе поддержки принятия решений. Автором были предложены требования к создаваемой системе управления данными.

Впервые применена методика управления разнородными пространственными данными в геоинформационной системе на примере задачи регулирования отпуска тепла в зависимости от температуры окружающей среды. Методика применена в новом качестве, её результат может

стать новым видом информационного продукта, который учитывает неопределенность реализации текста прогноза и позволяет оптимизировать управление ресурсами.

Представлена геоинформационная система поддержки принятия решений с использованием разнородных геопространственных метеорологических данных для оптимизации процесса управления теплоисточниками в интересах управления территориями. Методика предложенная для верификации ГИС впервые применяется в исследуемых системах.

Практическая и научная значимость работы заключается в том, что решена научно-техническая задача, имеющая существенное значение для моделирования геоинформационных систем и системного анализа разнородной информации, а так же для принятия решений по управлению территориями; исследованы технологии и способы обработки разнородной информации; определена возможность совершенствования технологий хранения и использования геоинформации на основе распределенных разнородных баз данных по средствам применения новых методик и моделей построения ГИС, что соответствует п.п. 3, 7, 9 области исследования паспорта специальности 25.00.35 «Геоинформатика». Предложенные автором разработки могут быть использованы при создании специализированных систем, использующих разнородные пространственные данные. В работе выделены основные этапы проектирования геоинформационных систем с учетом обработки разнородных данных. Основные выводы и положения диссертационного исследования были использованы в научно-исследовательских работах и учебном процессе.

Практическая ценность работы заключается в повышении эффективности управления разнородными данными за счет улучшения методов интеграции данных в геоинформационных системах.

Практическое значение результатов, разработанные в работе модели и созданные на их основе программные средства, дадут возможность усовершенствовать процесс обработки пространственных разнородных данных.

Личный вклад автора. Автор принимал непосредственное участие в подготовке и обработке исходных материалов, анализе и обобщении полученных результатов.

Апробация работы.

1. Международная конференция «Региональная информатика (РИ-2012)», секция «информационные технологии в гидрометрии» (24-26 октября 2012г.), доклад по теме: «Создание и размещение геопространственной информации исполнительных органов государственной власти в региональной геоинформационной системе Санкт-Петербурга (РГИС)».

2. Международная научно-практическая конференция «Инфогео 2013», секция «геоинформационные системы» (26-28 ноября 2013г.), публикация по теме: «Анализ статистических данных космических измерений с приполярной территории для разработки стохастической модели оценки георисков».
3. Международная научно-практическая конференция «Инфогео 2014», секция «геоинформационные системы» (3-6 октября 2014г.), доклад с публикацией по теме: «Обзор технологических методов доступа к базам данных распределенных ГИС».
4. Международная научно-практическая конференция “Инфогео 2015”, секция “геоинформационные системы” (24-26 ноября 2015 г.), публикации по темам: «Разработка модели доступа и технологий обработки гетерогенных баз данных для использования в прикладных ГИС», «Технический анализ механизмов доступа к гетерогенным данным для использования в прикладных ГИС», «Реализация модели подготовки гетерогенных данных в автоматизированной системе».
5. Свидетельство о регистрации базы данных «SeaIce version 1.0» №19868, дата регистрации 10.01.2014 года.
6. Свидетельство о регистрации базы данных «Ice_Concentration» №2014621110, дата государственной регистрации в реестре баз данных 07.08.2014 года.
7. Свидетельство о регистрации базы данных «База данных метеорологических параметров» №2016620986, дата государственной регистрации в реестре баз данных 20.07.2016 года.
8. НИР №1223 «Разработка и развитие методов, моделей и систем геоинформационного управления пространственно-распределенными объектами», 2014-2016 год.
9. НИР №74.20.56 «Разработка методических основ геоинформационного управления рисками развития рекреационных приморских территорий», 2013 год.

Структура и объем диссертации. Диссертация состоит из введения, трех глав, заключения, списка использованной литературы, списка сокращений и условных обозначений, приложений. Объем работы составляет 152 листа, в том числе 45 рисунков, 19 таблиц и 9 приложений. Библиография диссертационной работы составляет 114 наименований.

Во **введении** раскрывается актуальность решения научной задачи, сформулированы объект, предмет, цели и задачи исследований, показана теоретическая и практическая значимость работы, приведено краткое содержание работы по разделам, и перечислены основные научные результаты, выносимые на защиту.

В **первой главе** «Использование разнородных данных в геоинформационных системах при решении задач управления территориями» приводятся результаты анализа современных механизмов доступа к данным, заключающихся в определении нехватки простого метода

работы с разноформатными данными. Рассмотрены основные проблемы при обработке пространственных данных. Проведен анализ разрабатываемых прикладных геоинформационных систем, результатом которого стали обоснования к модернизации модели обработки разнородных данных. На основе изучения существующих систем предложены инструментальные средства обработки данных, учитывающие разноформатность и распределенность данных, способные улучшить работу по использованию пространственной информации.

Во *второй главе* «Разработка методики управления разнородными пространственно-распределенными данными в геоинформационной системе на основе метеорологической информации» разрабатывается методика управления пространственной метеорологической информацией для автоматизации поддержки принятия управленческого решения. В основу методики закладывается обработка разнородных массивов пространственных метеорологических данных для подготовки вариантов принятия управленческих решений по управлению территорией. Разработанная методика применяется на примере задачи регулирования отпуска тепла ТЭЦ в зависимости от температуры окружающей среды.

В *третьей главе* «Разработка геоинформационной системы поддержки принятия управленческих решений» разрабатывается методика проектирования геоинформационной системы поддержки принятия управленческих решений на основе стандартов проектирования информационных систем. Разрабатывается геоинформационная система на основе оптимизации использования метеорологических данных и модели обработки разнородных данных. Приводятся результаты информационной системы поддержки принятия решений, подтверждающие возможность реализации модели обработки разнородных пространственных данных.

В *заключении* сформулированы выводы и предложения, вытекающие из результатов исследования, определены направления дальнейших исследований.

1. Использование разнородных данных в геоинформационных системах при решении задач управления территориями

В современном мире с вхождением в век информатизации человечеству стало доступно чрезмерно огромное количество информации. Большие массивы данных используются ежедневно в различных целях, будь то наука, исследование, любительское познание, поиск информации, развлечение, работа, хобби и многое другое. Обработка информации предоставляет человечеству возможность быстрого доступа к многочисленным данным. В свою очередь мало кто задумывается, где на самом деле хранятся предоставляемые данные. Самое главное для потребителя, что не зависимо от места своего расположения, информацию возможно получить.

Решение задач развития и управления территориями требует привлечения большого количества информации, различающейся по степени упорядоченности, форме закрепления, содержанию и т.п. Доступная для анализа информация накапливается и хранится в больших объемах. Для обработки информации могут использоваться различные виды хранилищ данных: сменные носители, flash карты, жесткие диски, серверы данных, облачные хранилища и web-сервисы. Интерфейс хранилищ различен и отличается друг от друга, как и формат записи данных, их структура, тип данных и связи.

Для выбора оптимальных стратегий управления пространственно-распределенными объектами социально-экономических систем необходим комплексный анализ всего объема накопленной информации об объектах управления, включая спутниковые данные, данные полевых наблюдений, картографическую и другую необходимую информацию. Сбор информации для анализа возможен из различных источников, зачастую не связанных едиными механизмами управления и обусловлен отдельным применением внутренних и внешних сервисов и систем, а зачастую и вовсе отсутствием таковых.

1.1. Теоретические аспекты интеграции разнородных данных

Практика решения задач развития территорий показывает, что они имеют комплексный характер и требуют привлечения большого количества разнообразной информации, получаемой из разных источников, относящихся к разным видам деятельности. Более того, для оперативной аналитической обработки требуется привлечение внешних источников данных, которые тем более могут обладать разными форматами, а так же иметь разные представления и могут быть

несогласованными. Источники информации содержат данные, представляемые различными форматами, структурами, реализуемыми на разнотипных платформах. Комплексное использование такой информации связано с большими проблемами, обусловленными, прежде всего, разнородностью данных.

Основными факторами разнородности данных и их источников являются:

1. Различные типы данных (логический, целый, вещественный, объектный и др.);
2. Различная природа данных (числовые массивы, тексты, видеоданные);
3. Различные типы (модели) баз данных (БД) - реляционные, иерархические, объектно-ориентированные, сетевые, многомерные и т.п.;
4. Различные форматы представления данных;
5. Различия в степени распределенности систем хранения данных;
6. Различной природы данных (например, тексты, изображения, сигналы);
7. Различная степень достоверности и точности данных, измеряемых в различных масштабах и единицах измерения;
8. Различная степень и форма структурированности данных и т.п.

Выделяют следующие основные виды разнородности данных: синтаксическая, структурная и семантическая.

1. Синтаксическая разнородность данных, обусловлена различными типами данных, форматами их хранения, допустимыми значениями, единицами измерения и т. п.;
2. Структурная разнородность данных, обусловлена различными моделями данных, использованными при моделировании объектов реального мира (векторная, растровая, реляционная, объектная и т. п.);
3. Семантическая разнородность данных, проявляется в виде несоответствий в способах отображения одних и тех же объектов в компьютерной среде. Этот вид разнородности включает в себя:
 - 3.1. Концептуальную (когнитивную) разнородность, обусловленную тем, что разные люди при моделировании определенного явления создают собственное представление (абстракцию);
 - 3.2. Формализационную разнородность, которая является результатом использования различных языков и терминов при формировании концепций;
 - 3.3. Контекстную разнородность, которая относится к изменению смысла концептуальных построений под влиянием различных условий и обстоятельств.

Безусловно, такое разнообразие видов разнородности данных является основным проблемным фактором, в процессе интеграции информации [17, 28, 41]. В последние годы проблеме интеграции и совместного использования разнородных данных уделяется повышенное внимание. В контексте разработки геоинформационной системы разнородными данными в первую очередь считаются источники, требующие различных технических способов доступа к данным. Интеграция разнородных данных в единую информационную среду обеспечивает возможность их комплексного анализа и позволяет получить качественно новые знания об объекте исследования. Интеграция представляет пользователю единообразный взгляд на разнородные источники данных, что предполагает общую модель и общее отношение к семантике, с тем чтобы обеспечить возможность для доступа к данным, а в случае необходимости разрешать конфликтные ситуации [22, 101].

Рассмотрим интерпретированную схему классификаций технологий интеграции данных, разработанную Клаусом Дитрихом:

1. Common Data Storage (Общие системы хранения). Осуществляется за счет слияния данных из разных систем хранения данных в одну общую. Сегодня мы бы объединили эти два уровня в один и назвали бы его виртуализацией систем хранения.
2. Uniform Data Access (Унифицированный доступ к данным). На этом уровне осуществляется логическая интеграция данных, различные приложения получают единообразное видение физически распределенных данных. Такая виртуализация данных имеет свои несомненные достоинства, но гомогенизация данных в процессе работы с ними требует значительных ресурсов.
3. Integration by Middleware (Интеграция средствами ПО промежуточного слоя). ПО этого слоя играет посредническую роль, его составляющие способны к выполнению отдельных предписанных им функций, в полном объеме интеграционная задача решается во взаимодействии с приложениями.
4. Integration by Applications (Интеграция средствами приложений). Обеспечивает доступ к разным источникам данных и возвращает пользователю обобщенные результаты. Сложность интеграции на этом уровне объясняется большим разнообразием интерфейсов и форматов данных.
5. Common User Interface (Общий пользовательский интерфейс). Дает возможность единообразного доступа к данным, например, с помощью браузера, но при этом данные остаются неинтегрированными и неоднородными.
6. Manual Integration (Интеграция вручную). Пользователь сам объединяет данные, применяя различные типы интерфейсов и языки запросов.

Приведенная схема Дитриха интересна тем, что позволяет связать вместе интеграцию данных с интеграцией информации – по мере продвижения снизу вверх простые атомарные данные обретают семантику. Таким образом, автор утверждает, что в настоящее время ни о какой реальной интеграции информации речи пока быть не может: все, что существует сегодня, сводится лишь к консолидации данных на том или ином уровне абстракции.

В системе каждый из источников использует специфические термины и таблицы. В общем случае данные от разных источников обычно между собой не согласованы, а сами источники могут работать под управлением различных СУБД. Вследствие этого при обращении к данным нескольких источников через единую программную компоненту возникают проблемы интеграции информации, связанные с наличием или отсутствием каких-либо признаков в источниках, несоответствиями в типах представленных данных и т.д [102].

Кент В. [112] пишет, что вне зависимости от выбранных технологий и метода интеграции данных, остаются вопросы, связанные с их смысловой интерпретацией и различиями в представлении одних и тех же вещей. Приходится разрешать несоответствие схем данных и несоответствие самих данных. В несоответствии схем данных он выделяет структурные и семантические проблемы:

1. *Различие в типах данных.*

Некоторый домен в одном источнике может представляться числом, в другом - строкой фиксированной длины, в третьем - строкой переменной длины.

2. *Различие в единицах измерения.*

В одной БД указана величина в сантиметрах, в другой - в дюймах. В этом случае существует отображение 1:1.

3. *Различие в множестве допустимых значений.*

Один и тот же признак может определяться разными наборами констант. Например, выполнение задания одним источником может оцениваться по четырех балльной шкале (неудовлетворительно, удовлетворительно, хорошо, отлично), другим - по трех балльной (-, ±, +), третьим - по сто балльной. Отображение не является 1:1, оно может быть многозначным, может не иметь обратного, может зависеть от сторонних данных (например, 30 по математике соответствовать «удовлетворительно», а по русскому языку - «неудовлетворительно»).

4. *Различие «домен-отношение».*

Домен в одной БД (например строковое значение) соответствует таблице в другой БД (записи из таблицы-справочника).

5. *Различие «домен - группа доменов».*

В одном источнике адрес записывается одной строкой, в другом - отдельные поля для улицы, дома, строения, квартиры.

6. *Различие «данные-схема».*

Данные одной БД соответствуют схеме (метаданным) другой. В одной БД «инженер» - значение атрибута «должность» отношения «работник», в другой «инженеры» - отношение, содержащее некоторых работников, в то время как «бухгалтеры» содержит других.

7. *Отсутствующие значения.*

В каком-то из источников может отсутствовать информация, имеющаяся в большинстве других.

В данный перечень проблем несоответствия схем данных можно добавить еще ряд пунктов:

1. *Различия в наименованиях.*

Различия в обозначениях данных в разных источниках, несущих одну и ту же нагрузку на концептуальном уровне, т.е. когда одни и те же сущности (или поля) называются разными именами.

2. *Различия в наполнении.*

Одно и то же понятие, описываемое в каждом из источников, может толковаться своеобразно. Сюда же можно отнести различия в смысловой наполненности полей МЕМО (комментариев).

3. *Искажение сущностей.*

Например, механические опечатки, ошибки восприятия на слух сложно произносимых имен/названий, отсутствие единых стандартов транскрипции с иностранных языков.

4. *Потеря актуальности данных.*

Например, изменение актуальности было произведено только в одном источнике данных: была произведена смена фамилии, места работы, паспорта или др. информации.

5. *Различия в представлении значений.*

Например, одна и та же организация может быть записана в разных источниках разному: «ООО АГРО Русь», «АГРОРУСЬ», «АГРО-РУСЬ»

6. *Различия формата данных.*

Например, одна и та же запись в таблице может быть записана по разному: «ул. Маркса, дом 4» или «улица Маркса, д.4»; «8(812) 111-11-11» или «8-812-111-11-11»

Исследуя проблемы интеграции и доступа к разнородным данным, можно сделать следующий вывод, использование разнородных данных для комплексного анализа потребует большого количества специфичного программного обеспечения, а так же знаний языков запросов к данным и их семантику [101, 104]. Исследования и разработки программного обеспечения в данной области востребованы и являются актуальными уже на протяжении нескольких лет.

Таким образом, анализ в области исследований интеграции разнородных данных позволяет прийти к выводу, что большое количество вопросов семантической интеграции все еще остаются открытыми.

В ходе проведения исследований по обработке пространственных данных, была изучена географическая система Санкт-Петербурга. Региональная геоинформационная система Санкт-Петербурга (РГИС) представляет собой единое хранилище пространственных данных об объектах недвижимости Санкт-Петербурга, включая достоверные сведения о земельных участках, зданиях (сооружениях), градостроительных зонах, объектах культурного наследия и охранных зонах. В результате проводимых исследований, при интеграции и геокодировании данных, были выявлены многочисленные ошибки, связанные с разнородностью данных. Описание проводимых работ представлено в следующей главе.

1.2. Обработка и размещение разнородных геопространственных данных исполнительных органов государственной власти в системе РГИС

Региональная геоинформационная система (РГИС) Санкт-Петербурга - первая в Российской Федерации геоинформационная система, которая в сети общего доступа интернет публично предоставляет объединенные пространственные данные, содержащиеся в государственных информационных ресурсах, таких как городской кадастр объектов недвижимости, система обеспечения градостроительной деятельности, водный реестр, охрана памятников и других. РГИС изначально создаваясь как информационный портал исполнительных органов государственной власти города при поддержке Правительства Санкт-Петербурга, является визитной карточкой города в реализации приоритетной задачи создания электронного правительства в регионах РФ.

Пользователями региональной геоинформационной системы Санкт-Петербурга могут стать: физические и юридические лица, федеральные органы власти, органы власти Санкт-Петербурга, органы юстиции, службы спасения, архитекторы, кадастровые инженеры,

девелоперы, проектировщики, инвесторы. Система реализована в виде информационного сайта, размещенного по интернет адресу - <http://rgis.spb.ru>. Система протестирована в интернет-браузерах Internet Explorer 7, Opera 10, Mozilla Firefox 3.66, Google Chrome 5.03, Safari 4.02. Дизайн сайта выражен в строгом стиле, с символикой Правительства Санкт-Петербурга, преобладают сине-голубые тона. Элементы интерфейса выполнены в едином стиле, хорошо подчеркивая основную цветовую схему дизайна сайта. Пользователь кроме стандартных возможностей подобных проектов (изменение масштаба, перемещение по карте, измерение расстояния) может получить информацию по всем информационным слоям, сформировать отчет по объектам, попадающим в произвольную область, напечатать схему, подать заявление в электронной форме и многое другое. Поиск по кадастровому номеру может осуществляться по зданию, земельному участку, краткосрочной аренде, помещению, кадастровому кварталу. Доступен поиск здания по адресу. Для быстрого перехода к интересующей области карты в следующих сеансах работы можно воспользоваться сервисом формирования ссылки на эту область (геолинк). Обширный инструментарий позволяет пользователю осуществлять множество операций в окне карты. Это и измерение линейных расстояний, получение характеристик площадных объектов произвольной формы и в виде окружности (площадь и периметр), выбор объектов в произвольной области. По выбранным объектам можно просмотреть информацию, сформировать отчет в электронном виде в формате PDF. Получение информации осуществляется дифференцировано – по умолчанию выдается информация о подключенных в настоящее время информационных слоях, но выбирая соответствующую опцию в информационном окне можно мгновенно получить информацию по всем слоям в данном месте. Использование обзорной карты помогает быстро перейти к интересующей области. Любой посетитель сайта может воспользоваться интегрированными сервисами других сайтов органов государственной власти Санкт-Петербурга, не покидая РГИС (Комитет по управлению городским имуществом, Фонд имущества, Управление инвестиций, Государственное учреждение инвентаризации и оценки имущества, Комитет по тарифам, Комитет экономического развития, промышленной политики и торговли).

Используя указанные сервисы, можно перейти к подробному описанию объекта недвижимости, выставленному на продажу или перейти в реестр объектов, свободных от арендных отношений, воспользоваться сервисами оценки объектов недвижимости, выполнить расчет коэффициента месторасположения в конкретной точке, посмотреть информацию о нежилых помещениях в зданиях, узнать информацию о котельных города и многое другое. Кроме того, посетитель портала может сформировать в электронном виде документы, необходимые для подачи инвестиционного заявления в соответствующий орган

государственной власти, получить информацию в виде распечатки информации по рассматриваемой территории, а также сформировать и сохранить на своем компьютере электронный отчет с информацией по интересующей территории. Особую значимость имеет социальная информация, размещенная в разделе «Адресные программы». Например, представлен слой «Ремонт жилых домов», с помощью которого любой горожанин может узнать, вошел ли его дом в программу капитального ремонта и, что именно предполагается ремонтировать. Реализованная на портале система авторизованного доступа предоставляет авторизованным пользователям воспользоваться рядом дополнительных функций и сервисов. Авторизованные пользователи представляют как физические так и юридические лица. К юридическим лицам относятся топогеодезические и проектные организации, исполнительные органы государственной власти. Для этих организаций предусмотрена возможность выгружать картографический материал в границах проектирования или границах выполнения кадастровых работ по государственным контрактам в векторном виде, а также вести архив выгружаемой информации. Всем авторизованным пользователям доступно создание пользовательских меток и сохранение набора используемых слоев.

Органы государственной власти (на текущий момент в лице Комитета экономического развития, промышленной политики и торговли) получили возможность ведения собственных геоинформационных слоев в РГИС. На настоящий момент зарегистрировано более 200 авторизованных организаций и 300 авторизованных пользователей. Для обеспечения информацией зарубежных пользователей портал ведется параллельно на 2-х языках – русском и английском.

Система оснащена богатым справочным материалом. Руководство пользователя доступно для скачивания и обновляется на регулярной основе. Внутри системы существуют постоянно обновляемые разделы «Что нового?» и «Советы дня». Оперативное взаимодействие с активными пользователями осуществляется с помощью социальных сетей – twitter, facebook, ВКонтакте. Обратная связь с клиентами поддерживается путем отправки (получения) писем по e-mail и телефонной поддержкой пользователей.

Задачами создания РГИС являются:

- создание единой базы данных Санкт-Петербурга, как источника актуальных и достоверных сведений обо всех пространственных данных, созданных и разработанных за счет средств бюджета Санкт-Петербурга;
- создание условий, обеспечивающих свободный доступ органов государственной власти, органов местного самоуправления, организаций и граждан к обобщенным пространственным данным Санкт-Петербурга и их эффективное использование;

- создание эффективного средства информационной поддержки процессов принятия решения органами государственной власти при исполнении ими государственных функций и услуг;
- снижение бюджетных расходов на создание пространственных данных за счет исключения дублирования работ;
- упрощение административных процедур, связанных с получением сведений через Интернет.

Основными задачами создания системы было объединение информационных ресурсов различных органов власти для предоставления информации всем гражданам в свободном доступе. Создание РГИС знаменует собой значительный шаг в рамках реализации программы «электронного правительства» и сокращения времени получения информации, а использование сети интернет позволяет воспользоваться всеми преимуществами публичного распространения информации. Для многих организаций и физических лиц сайт РГИС стал каждодневным рабочим инструментом, в рабочие дни независимый счетчик Yandex фиксирует около 1500 визитов на сайт в сутки, среднее время пребывания пользователя на сайте составляет более 13 минут.

Целями создания РГИС являются:

- объединение информационных ресурсов органов государственной власти, органов местного самоуправления в Санкт-Петербурге и организаций;
- комплексное распространение и представление гражданам и организациям информации о пространственных объектах из единого источника информации;
- повышение инвестиционной привлекательности Санкт-Петербурга за счет представления актуальной и достоверной комплексной информации о пространственных объектах.

Оператором РГИС назначен Комитет по земельным ресурсам и землеустройству Санкт-Петербурга. Комитет по земельным ресурсам и землеустройству Санкт-Петербурга (далее - Комитет) является исполнительным органом государственной власти Санкт-Петербурга, образованным в соответствии с законодательством Санкт-Петербурга для проведения государственной политики и осуществления государственного управления в сфере земельных отношений и ведения кадастра недвижимости в Санкт-Петербурге, а также для координации деятельности в этой сфере территориальных и отраслевых исполнительных органов государственной власти Санкт-Петербурга. Комитет подчинен Правительству Санкт-Петербурга.

Постановлением Правительства Санкт-Петербурга от 30.11.2009 № 1387 создана региональная информационная система. В промышленную эксплуатацию РГИС введена приказом Комитета по земельным ресурсам и землеустройству Санкт-Петербурга от 08.10.2010 № 295. Система РГИС зарегистрирована в Едином реестре информационных ресурсов и информационных систем Санкт-Петербурга (регистрационный номер 000478/10). Объединение информационных ресурсов осуществляется на основе Соглашений об информационном обмене и сотрудничестве между Комитетом и организациями”.

Из приказа № 1387 о создании системы: “В соответствии с Федеральным законом "Об информации, информационных технологиях и о защите информации", Законом Санкт-Петербурга от 01.07.2009 № 371-70 "О государственных информационных системах Санкт-Петербурга" и Законом Санкт-Петербурга от 30.04.2008 № 278-47 "О разграничении полномочий органов государственной власти Санкт-Петербурга в сфере регулирования земельных отношений на территории Санкт-Петербурга" постановлением Правительства Санкт-Петербурга от 30.11.2009 № 1387 создана региональная информационная система, содержащая сведения об объектах недвижимости и объектах землеустройства "Геоинформационная система Санкт-Петербурга" (РГИС).

Данная система получила высокую оценку Губернатора Санкт-Петербурга В.И. Матвиенко на заседании Правительства Санкт-Петербурга 02.11.2010. Было отмечено, что круглосуточное представление информации в открытом и бесплатном доступе создает благоприятный инвестиционный климат, помогает решать вопрос получения актуальной и достоверной информации без излишних временных затрат. В.И. Матвиенко также настоятельно рекомендовала всем руководителям отраслевых комитетов и районных администраций научиться пользоваться «Геоинформационной системой Санкт-Петербурга».

В РГИС включаются:

1. Сведения об административно-территориальном устройстве Санкт-Петербурга, а также избирательных округах, границах территорий судебных участков мировых судей Санкт-Петербурга;
2. Классификаторы и кодификаторы адресов Санкт-Петербурга;
3. Объектно-адресная система Санкт-Петербурга;
4. Сведения о водных объектах на территории Санкт-Петербурга;
5. Сведения об объектах землеустройства, земельных участках, в том числе обременениях земельного участка, ограничениях его использования;

6. Сведения о зонах с особыми условиями использования, в том числе охранных зонах (водоохранных зонах водных объектов, охранных зонах инженерных коммуникаций, охранных зонах памятников природы и т.д.);
7. Материалы градостроительного планирования и градостроительного зонирования, в том числе Генерального плана Санкт-Петербурга, правил землепользования и застройки, проектов планировки и проектов межевания;
8. Сведения об объектах недвижимости, в том числе зданиях, сооружениях, линейных объектах, помещениях, расположенных на территории Санкт-Петербурга;
9. Сведения о территориях объектов культурного наследия (памятниках истории и культуры) народов Российской Федерации, в том числе вновь выявленных объектах культурного наследия;
10. Тематические карты;
11. Сведения о зеленых насаждениях общего пользования;
12. Материалы лесоустройства;
13. Картографические материалы;
14. Иные сведения, определяемые Комитетом.

В настоящее время пользователю доступна информация более чем по 120 информационным слоям, что не имеет аналогов в России по качеству и количеству предоставляемой информации. Информация, представленная на РГИС, сгруппирована по тематикам:

- Административные границы
- Сведения кадастра недвижимости
- Сведения информационной системы обеспечения градостроительной деятельности
- Адресные программы
- Объекты культурного наследия и др.

Тематическая группировка позволяет пользователю быстро находить слои и при необходимости включать (выключать) их, изменять порядок их отображения. Для быстрой настройки существуют predefined рабочие наборы – “стартовый” для инвесторов, проектировщиков, кадастровых инженеров. Набор включает для показа определенные геоинформационные слои.

Уникальность данной системы заключается в оперативности и достоверности представленной в ней информации. Достоверность обеспечивается предоставлением юридически значимой информации от исполнительных органов государственной власти Санкт-Петербурга и внутренним контролем со стороны Комитета. Оперативность обеспечивается

обновлением кадастровой информации периодичностью 1 раз в час и загрузкой на информационный сайт слоев, получаемых в соответствии с Соглашениями об информационном обмене, в течение не более 1 суток с момента их получения.

Картографические данные, отображаемые на сайте, разработаны с помощью географической информационной системы картографирования MapInfo. Это развитая система настольной картографии, профессиональное средство для создания, редактирования и анализа картографической и пространственной информации. В дополнение к традиционным для СУБД функциям, MapInfo позволяет собирать, хранить, отображать, редактировать и обрабатывать картографические данные, хранящиеся в базе данных. С учетом пространственных отношений объектов, решать сложные задачи географического анализа, такие как создание районов, связь с удаленными базами данных, включение графических объектов в другие приложения, создание тематических карт и многое другое.

Комитет по земельным ресурсам и землеустройству Санкт-Петербурга поставил задачу: создать средствами MapInfo тематический слой - объекты физкультуры и спорта, для дальнейшего размещения информации на интернет-сервисе - региональная информационная система "Геоинформационная система Санкт-Петербурга". Данные, для выполнения поставленной задачи, были получены от Комитета по физической культуре и спорту.

Комитет по физической культуре и спорту является исполнительным органом государственной власти Санкт-Петербурга - структурным подразделением Правительства Санкт-Петербурга, созданным для проведения государственной политики и осуществления государственного управления в сфере физической культуры и спорта Санкт-Петербурга, а также для координации деятельности в этой сфере территориальных и иных отраслевых исполнительных органов государственной власти Санкт-Петербурга.

Информационные данные для геокодирования были предоставлены в разнородных форматах (.doc), (.xls), (.rtf). В таблице 1 приведены документы с указанием типа данных и размера.

Таблица 1 – Разнородные данные исполнительных органов государственной власти

Район (имя документа).	Тип.	Размер. (кило байт)
Адмиралтейский на 25.01.2011	Microsoft Word 97-2003	570
Василеостровский – 2011	Microsoft Word 97-2003	666
Выборгский – 2011	Microsoft Word 97-2003	875
Калининский – 2011	Microsoft Word 97-2003	1042

Продолжение таблицы 1 – Разнородные данные исполнительных органов государственной
власти

Кировский район 2011	Microsoft Word 97-2003	993
Колпинский район – 2011	Microsoft Word 97-2003	679
Красное село – 2011	Microsoft Excel 97-2003	167
Красногвардейский - 2011 с удаленными	Microsoft Word 97-2003	883
Кронштадтский	Microsoft Word 97-2003	313
Курортный	Microsoft Word 97-2003	587
Московский-январь.2011	Microsoft Word 97-2003	785
Невский район – 2011	Microsoft Word 97-2003	695
Петроградский район	Microsoft Excel 97-2003	124
Петродворцовый	Microsoft Word 97-2003	407
Приморский – 2011	Microsoft Word 97-2003	1000
Пушкинский-2011	Microsoft Word 97-2003	69
Фрунзенский район – 2011	Формат RTF	2980
Центральный район 29.01.2011	Microsoft Word 97-2003	1408

В соответствии с «Соглашением об информационном обмене сведениями региональной информационной системы, содержащей сведения об объектах недвижимости и объектах землеустройства «Геоинформационная система Санкт-Петербурга»» в приложение 5 к Соглашению об информационном обмене и сотрудничестве: «Регламент обмена геопространственной информацией, классификаторами и базами данных» – указаны требования к данным, согласно которым, картографические данные должны быть представлены в формате таблиц ГИС MapInfo (tab) или в формате обменного файла ГИС MapInfo (mid/mif). В случае использования в информационных ресурсах Взаимодействующей стороны иного формата данных Взаимодействующая сторона производит конвертацию формата самостоятельно.

Исходя из вышеописанных требований к картографическим данным, в задачи входит:

1. Конвертация форматов.
2. Создание тематического слоя - объекты физкультуры и спорта, далее «объекты_кфкс», средствами MapInfo.

Конвертация в нужный формат привела к многочисленным ошибкам из-за не соответствия адресов в исходной и целевой таблицах, что является одной из самых важных проблем для геоинформационной системы при геокодировании. Вследствие ошибки при

конвертации форматов было принято решение перейти к выполнению второй задачи: создание тематического слоя объекты_кфкс средствами MapInfo.

Из-за не возможности выполнить “автоматическое” или “ручное” геокодирование поставленная задача усложнилась. По этой причине предстояло создать тематический слой объекты_кфкс и нанести исходные данные вручную.

В рамках проделанной работы был создан тематический слой объекты_кфкс, ниже представлены объекты физкультуры и спорта северной и южной части Санкт-Петербурга при масштабе 1:50 000 (рисунок 1-2).

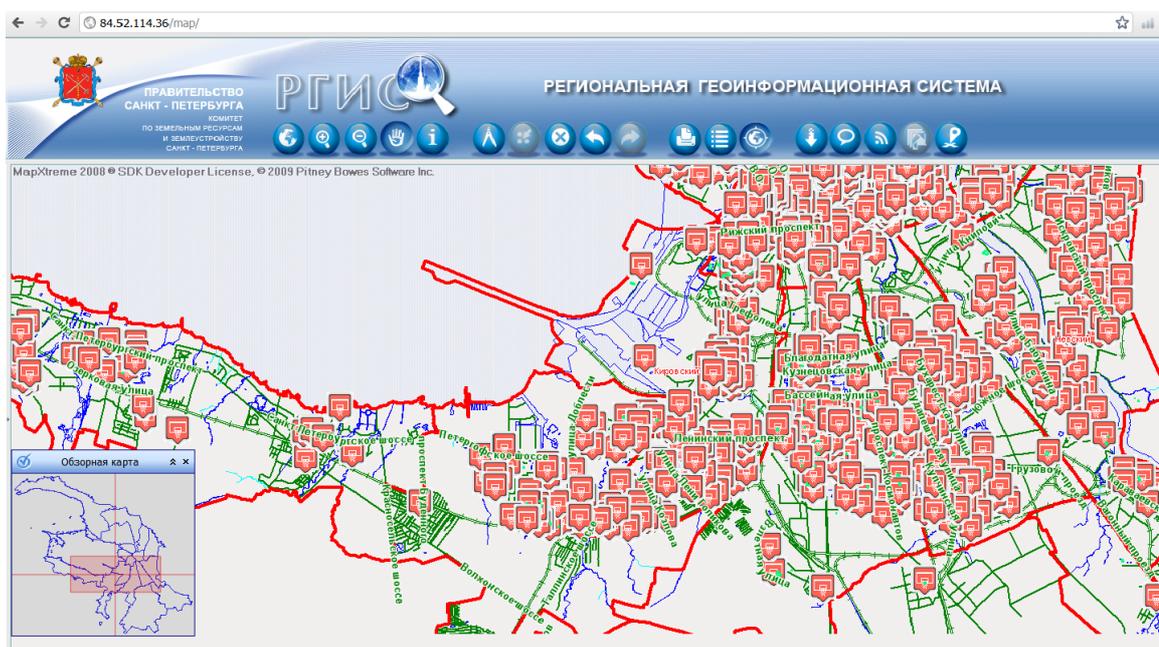


Рисунок 1. Объекты физкультуры и спорта. Южная часть С-Пб.

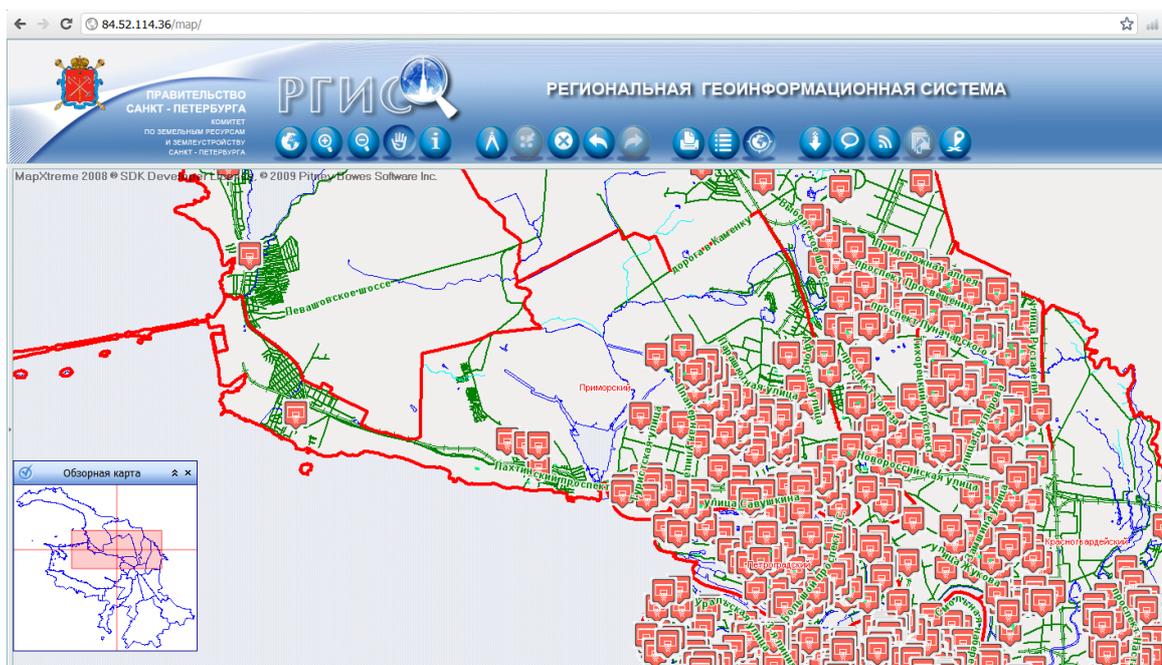


Рисунок 2. Объекты физкультуры и спорта. Северная часть С-Пб.

По документам Комитета по физической культуре и спорту общее число объектов спортивного назначения в Санкт-Петербурге составляет 5654 (таблица 2).

Таблица 2 - Спортивные объекты г. Санкт-Петербурга

Объекты спортивного назначения	Общее количество
Манежи	21
Крытые катки	11
Плоскостные спортивные сооружения	2311
Спортивные залы	1500
Нестандартные спортивные залы	375
Бассейны	145
Нестандартные чаши для плавания	8
Стрелковые тир	59
Другие спортооружения	987
Стадионы	25
Лыжные базы	20
Гребные базы и каналы	10
Велотреки	1
Мототреки	1

Продолжение таблицы 2 - Спортивные объекты г. Санкт-Петербурга

Дворцы спорта	12
Помещения спортивного назначения	157
Футбольные поля	8
Футбольные манежи	3
Итого:	5654

По окончанию нанесения всех объектов на слой была проведена выборка из таблицы объекты_кфкс по всем занесенным данным, результат выведен в список MapInfo (рисунок 3).

Название	Описание	Адрес	Длина	Ширин	Площадь	Примечание
МАНЕЖ	Спортивный манеж НГУ ФКСЗ им.Л.Ф.Лесгафта	ул.Декабристов, 36	31	11	341	Легкая атлетика, сектор д/прыжков в высоту
корт	СПб ГУ "ПКиО "Екатерингоф"	Лифляндская ул., 12	24	11	0	Теннис
Крытый каток	ТРЦ "Варшавский экспресс"	наб.Обводного канала, 118	45	11	495	Каток
Крытый каток	Каток СКА ЛенВО, крытый, с искусственным льдом	наб.р.Фонтанки, 112	45	20	900	Хоккей, фигурное катание
Спортивное сооружение	Стадион "Красный треугольник"	Лифляндская ул., 5	100	64	6 400	Спортдоро: бег, дорожки 4, сектор для прыжков в длину, площадка мии
Спортивное сооружение	НГУ ФКСЗ им.Л.Ф.Лесгафта	ул.Декабристов, 35 лит.В	91	46	4 186	Спортдоро: бег, дорожки 4
Спортивное сооружение	Государственный университет путей и сообщения	Садовая ул., 50-Б (Юсуповский сад)	35	18	630	Мини-футбол
Спортивное сооружение	Государственный университет путей и сообщения	Садовая ул., 50-Б (Юсуповский сад)	18	9	162	Волейбол
Спортивное сооружение	Государственный университет путей и сообщения (обще)	7-я Красноармейская ул., 12	30	18	540	Футбол
Спортивное сооружение	ГБОУ СО Вторая Санкт-Петербургская гимназия	Казанская ул., 27/13	15	20	300	Волейбол, стритбол
Спортивное сооружение	ГБОУ СО Вторая Санкт-Петербургская гимназия	Казанская ул., 48	18	9	162	спортплощадка
Спортивное сооружение	ГБОУ СОШ №234	Канонерская ул., 33	49	16	784	спортплощадка
Спортивное сооружение	ГБОУ СОШ №235	наб. р.Пряжки, 2-4-6	0	0	0	Мини-футбол, баскетбол, волейбол
Спортивное сооружение	ГБОУ СОШ №241	Римского-Корсакова пр., 4/38	20	16	320	Баскетбол
Спортивное сооружение	ГБОУ СОШ №241	Римского-Корсакова пр., 4/38	30	16	480	Футбол
Спортивное сооружение	ГБОУ СОШ №259	наб.Крюкова кан., 3	20	15	300	Футбол, стритбол
Спортивное сооружение	ГБОУ СОШ №260	Лермонтовский пр., 21	24	13	312	Футбол
Спортивное сооружение	ГБОУ СОШ №263	Старо-Петерофский пр., 33	40	20	800	Футбол
Спортивное сооружение	ГБОУ СОШ гимназия №272	8-я Красноармейская ул., 3	31	16	496	Футбол, стритбол
Спортивное сооружение	ГБОУ СОШ №278	Дровяная ул., 7а	25	15	375	Футбол
Спортивное сооружение	ГБОУ СОШ №281	Советский пер., 4/15	20	14	280	Футбол, стритбол
Спортивное сооружение	ГБОУ СОШ №286	10-я Красноармейская ул., 5	40	19	760	Футбол
Спортивное сооружение	ГБОУ СОШ №286	10-я Красноармейская ул., 5	20	12	240	Б/Б, в/Б, городки
Спортивное сооружение	ГБОУ СОШ №287	Бумажная ул., 5	40	20	800	Футбол, стритбол
Спортивное сооружение	ГБОУ СОШ №288	Курляндская ул., 43	30	30	900	Футбол
Спортивное сооружение	ГБОУ СОШ №307	Малодетскоеосельский пр., 23	73	34	2 482	Мини-футбол, ОФП
Спортивное сооружение	ГБОУ СОШ №317	Серпуховская ул., 39	30	16	480	Футбол
Спортивное сооружение	ГБОУ СОШ №522	Бронницкая ул., 34	26	22	572	Футбол
Спортивное сооружение	ГБОУ СОШ №564	наб.Обводного кан., 143	24	20	480	Волейбол, стритбол
Спортивное сооружение	ГБОУ СОШ №564	ул.Егорова, 24	33	12	396	Стритбол, футбол, волейбол
Строительное сооружение	ГБОУ СОШ №615	Нарвский пр., 6	11	6	66	Футбол
Спортивное сооружение	ГБОУ СОШ №624	8-я Красноармейская ул., 16	31	18	558	Футбол, стритбол
Спортивное сооружение	Детский дом №826	Писарева ул., 12	15	27	405	Футбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Лощманская ул., 4/Пряжки, 66	31	13	403	Футбол, стритбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Садовая ул., 125	22	14	308	ОФП
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	ул.Союза Печатников, 31	12	8	96	Стритбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Гороховая ул., 41	26	16	416	Футбол
Строительное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Б.Казачий пер., 10	30	20	600	Футбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Измайловский пр., 11	24	9	216	Стритбол, волейбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Московский пр., 79	29	37	1 073	Футбол
Строительное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Московский пр., 75	29	27	783	Футбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	наб.р.Фонтанки, 128	20	14	280	Баскетбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Курляндская ул., 16-18	30	15	450	Футбол
Спортивное сооружение	ГБОУ МУК "Центр-Комплекс" Адмиралтейского района	Бумажная ул., 22	33	14	462	Футбол

Рисунок 3. Выборка данных из таблицы объекты_кфкс

Разница между общим числом объектов по документам Комитета по физической культуре и спорту и результатам выборки составляет 2637 объектов спортивного назначения. В

пределах данной работы многие данные небыли нанесены на слой объекты_кфкс по следующим причинам:

- Во-первых, не было возможности определения точного местоположения объекта.
- Во-вторых, частично имеющаяся информация являлась юридически не достоверной или неправильно описанной.
- В-третьих, из-за отсутствия точной информации в адресе объектов, такой как корпус и/или литера.

В большинстве случаев эта проблема касается плоскостных сооружений, которые не были занесены в базу данных РГИС. Такие объекты необходимо находить вручную, с помощью «ортофотоплана» и в приближенном месте указывать их расположение, что является не правильным касательно юридически актуальной информации об объекте. Таким образом, мы сталкиваемся с проблемой геокодирования данных.

Для решения данных проблем необходимо установить:

- Единый формат хранения предоставляемого документа для геокодирования.
- Точный юридический адрес объекта.
- Стандартный набор информации об объекте.

С целью устранения вышеописанных проблем было разработано программное обеспечение, предоставляющее исполнительным органам государственной власти возможность выбирать и заносить юридически достоверную информацию об объектах, с возможностью сохранения данных в единый формат.

За период работы, по материалам, предоставленным Комитетом по физической культуре и спорту, была проведена большая работа, связанная с геокодированием объектов. Во время практики была создана база данных объектов спортивного назначения, которые обладают пространственной привязкой и имеют семантические характеристики. Разработано программное обеспечение для форматирования информации об объектах спортивного назначения. В результате проделанной работы можно сделать следующие выводы:

1. Созданный геоинформационный слой отображает юридически достоверную информацию исполнительных органов государственной власти о 3017 объектах спортивного назначения, расположенных в Санкт-Петербурге.
2. Состав и актуальность сведений о спортивных объектах делает РГИС эффективным инструментом для пользователей при принятии различных решений.
3. Тематический слой пополнит базу данных РГИС, содержащую более 120 слоев пространственных данных в единой государственной системе координат.

4. Геокодирование пространственных данных исполнительных органов государственной власти позволило определить явную проблему интеграции разнородных источников данных.

Объединение разнородных данных для последующего совместного анализа и обработки становится популярной темой исследования последнего десятилетия. Основной задачей технологии слияния данных является объединение данных из разных источников в интересах решения последующих содержательных задач: принятие управленческих решений, анализ состояния объектов, прогнозирование ситуаций и т.д. Интеграция разнородных данных в единую информационную среду обеспечивает возможность их комплексного анализа и позволяет получить качественно новые знания об объекте исследования [10].

1.3. Анализ прикладных систем и технологий обработки разнородных данных

Одной из актуальных проблем при использовании прикладных геосистем для доступа к разнородным массивам данных является оптимизация технологий доступа, хранения и обработки. В настоящее время большое внимание уделено именно процессу предоставления доступа к разнородным данным. Многие научные институты предлагают свои пути решения возникшей проблемы. Исследование показало, что разработанные новые модели в большей части нацелены на решение конкретных прикладных задач, поэтому в центре внимания все равно остается проблема усовершенствования и создания новой технологии доступа к данным [1]. Основной задачей разработки нового механизма доступа является объединение данных из разных источников в интересах решения последующих содержательных задач: принятие управленческих решений, классификация гетерогенных данных, определение состояния объектов, оценка ситуации и т.д.

В настоящее время интерес к работе с разнородными данными возрастает. В центре внимания находится решение различных ситуативных прикладных задач с использованием интеграции и обработки таких данных [24]. В практическом плане увеличивается число комплексных информационных решений, которые способны объединять, следующие типы ресурсов: геоинформационные ресурсы (картографические данные, спутниковые снимки, данные наблюдений и т. п.), электронные каталоги, изображения, аудио и видео материалы и другие ресурсы. Рассмотрим некоторые из функционирующих на данный момент систем обработки разнородных данных и других предложенных моделей.

Распределенная геоаналитическая система «Geo-Status». Концепция ситуационного центра губернатора Владимирской области. Основные преимущества системы:

- Решение широкого спектра информационно-аналитических задач с привязкой к пространственным данным.
- Полнофункциональная информационная система с развитыми механизмами работы с базой данных. Работа с базой данных реализована на платформе 1С.
- В качестве геосервера может использоваться как платный ArcGisServer, так и бесплатный GeoServer.
- Система многофункциональна, работает через Web, имеет развитые средства разделения доступа.
- В системе Geo-Status возможно оперативно, в реальном времени, получать необходимую информацию с визуальным отображением данных.
- Система позволяет в режиме реального времени, из любой точки обеспечить информационно-аналитическую поддержку процессов принятия управленческих решений.
- Распределенная геоаналитическая система (РГАС) создана, как единый информационный ресурс. Ключевым элементом системы является ситуационный центр.
- РГАС обеспечивает информационную и аналитическую поддержку принятия эффективных управленческих решений на основе объективной и многопараметрической информации.
- РГАС создана на платформе пространственных данных, что обеспечивает мониторинг приоритетных направлений развития и имеет широкую сферу применения.

На рисунке 4 представлена схема базы данных РГАС, в неё входят несколько баз данных, которые упорядочены по отраслям и сферам применения.

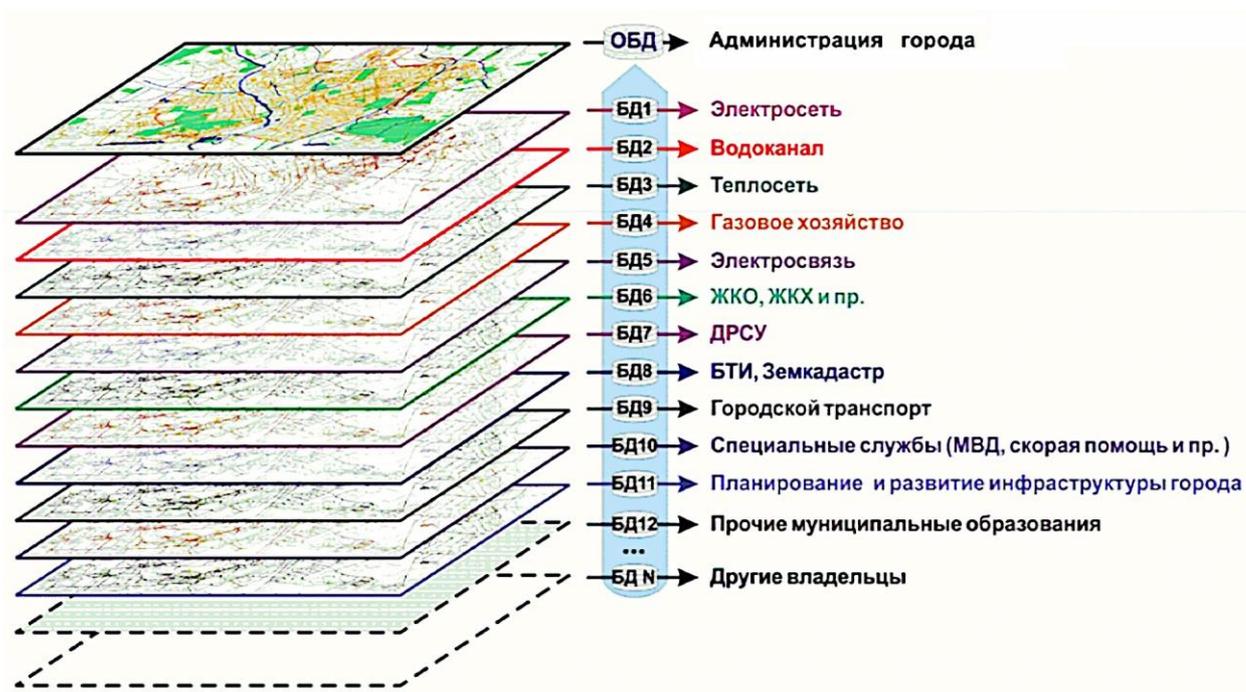


Рисунок 4. Слои представления данных в РГАС.

Для эффективного функционирования к системе РГАС предъявлены следующие требования:

- Автоматизированный процесс создания баз данных(сбор, хранение, анализ и передача информации);
- Хранение пространственных и атрибутивных данных для их дальнейшего анализа и обработки;
- Мониторинг в режиме реального времени (выполняется запрос о свойствах объектов и автоматизируется процесс анализа, сопоставляя множество параметров);
- Формирование и представление материалов для принятия управленческих решений.

Система реализована на 2 платформах. Первая платформа - 1С. Используется для организации работы с БД. За счет этого достигается максимальная концентрация функциональных возможностей при вводе, анализе и получении отчетности. Вторая платформа – геоинформационный сервис. Система поддерживает работу, как с платным, так и бесплатным сервером разработки «ESRI».

Система может работать, как через web-браузер, так и через специального тонкого клиента. Платформа 1С обеспечивает практически все современные методы межсистемного взаимодействия: com, soap, xml. Геоинформационный модуль базируется на открытой библиотеке «OpenLayers». Таким образом, возможна значительная адаптация функционала

работы с картой. Геоинформационный модуль может работать не только с локальными слоями, но и использовать сторонние по открытым протоколам WMS, WFS.

Организация интеграционных процессов реализовано с помощью программного комплекса SONICESB. Комплекс предназначен для организации и обеспечения информационного взаимодействия модулей ситуационного центра и взаимодействия с внешними по отношению к нему системами.

Особенности интеграции:

- Предоставление унифицированного интерфейса взаимодействия по общепринятым промышленным стандартам: SOAP, FTP, POP, JDBC.
- Интеллектуальная маршрутизация, настройка правил маршрутизации, основанных на атрибутах передачи данных, содержании данных и предопределенных настройках интеграционной платформы.
- Трансформация данных – предоставление данных в требуемом формате, обеспечение совместимости сопрягаемых систем и модулей за счет наличия специального интерфейса преобразования схем трансформации данных.
- Безопасное конфигурирование, управление и контроль функционирования всех сервисов трансформации данных.

Внедрение такой системы, предоставляет доступ к огромному объему данных, что способствует более эффективному принятию управленческих решений. Из описания системы можно сделать следующие интересные по данной теме выводы:

- Система РГАС имеет доступ к различным базам данных – в свою очередь по отношению друг к другу эти БД разнородны.
- Для соединения с данными используются общепринятые промышленные стандарты: SOAP, FTP, POP, JDBC.
- Для использования в системе РГАС сторонних данных, необходимо учитывать правила предоставления информации в требуемом формате.

Сервис-ориентированная геоинформационная система для обеспечения доступа к пространственным данным в задачах исследования природных экосистем. В Институте вычислительных технологий СО РАН разрабатывается сервис-ориентированная геоинформационная система для обеспечения доступа к пространственным данным. Система строится на основе каталога спутниковых данных Новосибирского научного центра СО РАН, который включает архивные данные со спутников серии LandSat по территории РФ за 1982–2002 гг., и с 2008 г. регулярно пополняется оперативными данными SPOT 4. С апреля 2010 г. к каталогу подключена система структурного восстановления оперативных данных,

поступающих с платформ TERRA/AQUA. Прием данных осуществляется Западно-Сибирским региональным центром приема и обработки данных в рамках соглашения о межведомственном сотрудничестве.

Система реализована в виде базового набора приложений, работающих в среде сервера приложений Tomcat. Подсистема пользовательских интерфейсов реализована с использованием технологий PHP/JavaScript. Доступ к системе реализован посредством модуля Central Authentication Service (CAS). Он позволяет организовать многоуровневую систему разграничения прав доступа с централизованной базой пользователей на основе LDAP-каталога Сибирского отделения РАН и реализовать практически индивидуальные настройки доступа к любому защищаемому ресурсу. Система состоит из следующих функциональных блоков (рисунок 5).

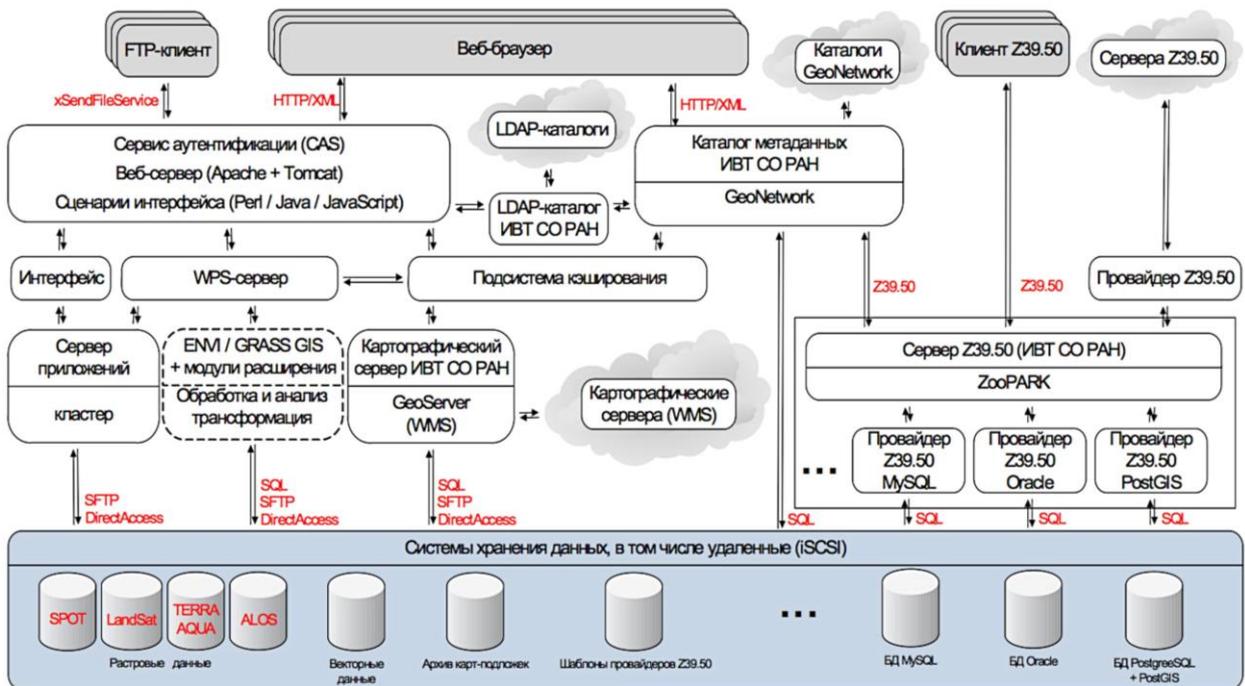


Рисунок 5. Структура картографической информационной системы.

Центральным блоком системы является подсистема картографических сервисов, реализованная на основе пакета GeoServer. Подсистема обеспечивает доступ к картографической информации, хранящейся в системе (базовые подложки, векторные слои, построенные по базам данных и др.). Для публикации динамических данных используется пакет UMN MapServer, который обеспечивает доступ к данным, формируемым в оперативном режиме, а также к пользовательским наборам данных.

Для тематической обработки данных в систему интегрирован комплекс программ, основанный на эффективных логико-вероятностных алгоритмах выбора информативных признаков и классификации.

Для расширения функциональности системы используется подсистема сервисов. На основе WPS-сервера deegree, распространяемого по лицензии GPL, разработан модуль для интеграции в систему алгоритмов обработки пространственных данных. Он осуществляет интерпретацию входных и выходных данных согласно спецификации протокола WPS и выполняет функции контейнера для неограниченного числа WPS-процессов. Архитектура модуля представлена на рисунке 6.

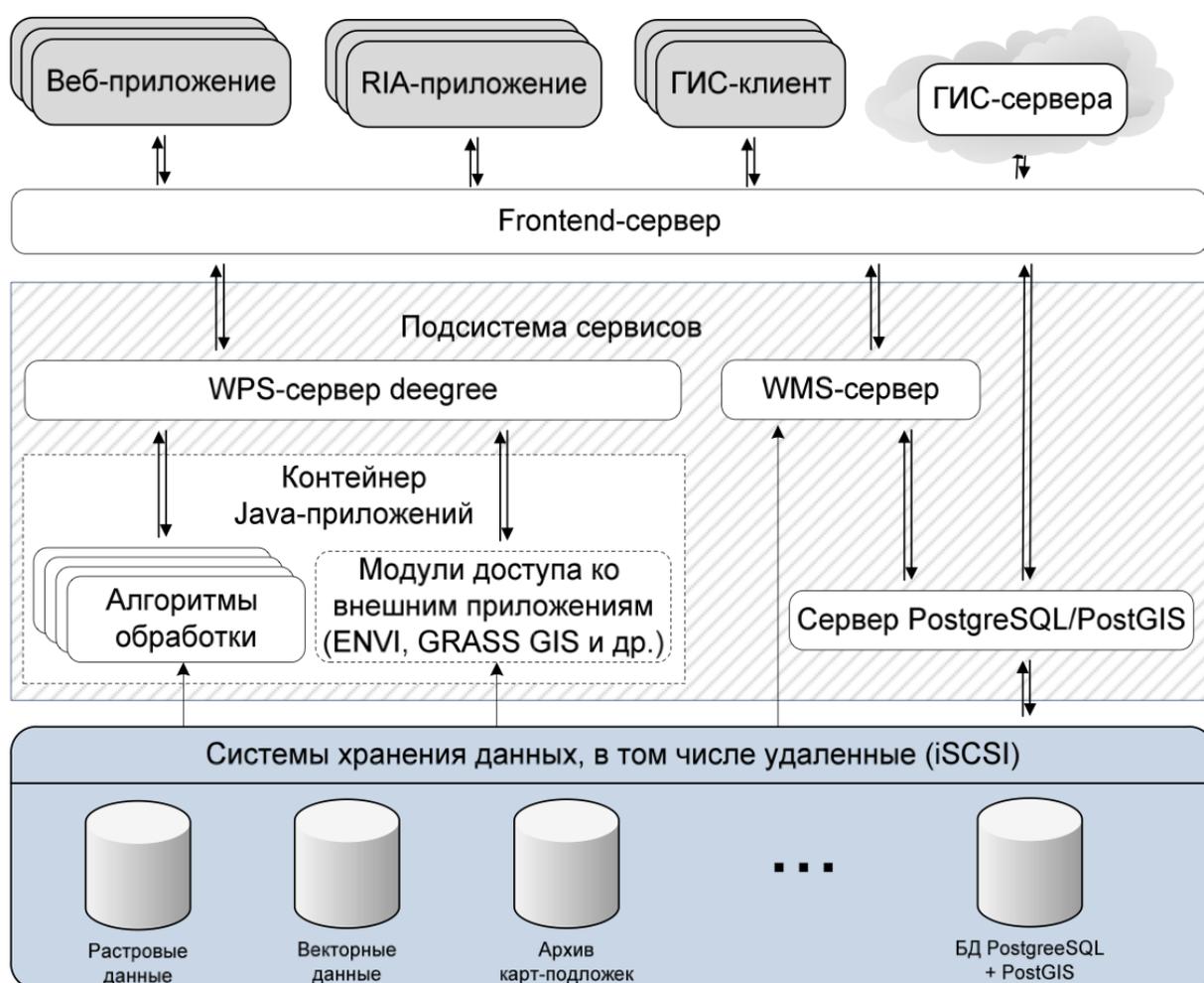


Рисунок 6. Структурная схема подсистемы сервисов.

Для обеспечения функционирования системы в распределенном режиме по протоколам доступа к метаданным и их представлению в нее интегрированы модули поддержки протокола Z39.50. Поисковая система позволяет не только находить данные по метаданным, но и выполнять комплексные запросы.

В настоящее время к системе подключено 26 институтов Сибирского отделения РАН. Система используется для выполнения междисциплинарных интеграционных проектов СО РАН. В систему внедрено несколько информационных ресурсов, разработанных сотрудниками Сибирского отделения РАН в рамках различных проектов и грантов и предоставляемых в виде веб-сервисов, в том числе:

- векторная карта растительности Западной Сибири и ландшафтная карта Иркутской области;
- векторная карта почв бореальной зоны Западно-Сибирской равнины и соответствующая ей карта растительности, содержащая 28 различных типов растительности (М 1 : 7 500 000);
- база данных по содержанию органического углерода в почвах Сибири;
- климатологические данные за период с 1989 по 2009 г. на ключевой участок с координатами: 53–70° с. ш., 59–93° в. д., подготовленные сотрудниками Института мониторинга климата и экологических систем СО РАН.

Разработан подход к интеграции разнородных пространственных данных, позволяющий осуществлять их комплексный анализ. Для этого все необходимые данные приводятся к единому пространственному разрешению с использованием регулярной сетки. Разработанный подход применялся для построения информационных моделей бореальных экосистем Западной и Восточной Сибири.

Если для интеграции спутниковых и картографических данных для минимальной функциональности информационной системы достаточно интеграции в рамках, например, каталога GeoNetwork, то для более широкого спектра информационных ресурсов функциональных возможностей этой системы становится явно недостаточно. Несмотря на то, что GeoNetwork при загрузке дополнительных схем данных и надлежащей настройки поддерживает каталогизацию и обеспечение доступа к ресурсам различного типа, способы управления ресурсами в этой системе оставляют желать лучшего. В этой системе отсутствует поддержка иерархических коллекций, включающих в том числе и разнородные информационные ресурсы, а также детализированное разграничение доступа к этим коллекциям на основе расширенных ролевых правил. Отсутствие этой функциональности в GeoNetwork существенно сужает рамки ее использования для коллективной работы по созданию тематической информационной системы [79].

В качестве основы системы, интегрирующей разнородные данные, была выбрана система DSpace. Информационная система DSpace обладает широкими возможностями по управлению цифровым контентом, но не содержит интерфейсов для работы с географическими

координатами. Учитывая, что DSpace широко используется для создания электронных библиотек, система была модифицирована для придания ей дополнительной функциональности. На рисунке 7 показана общая схема системы.

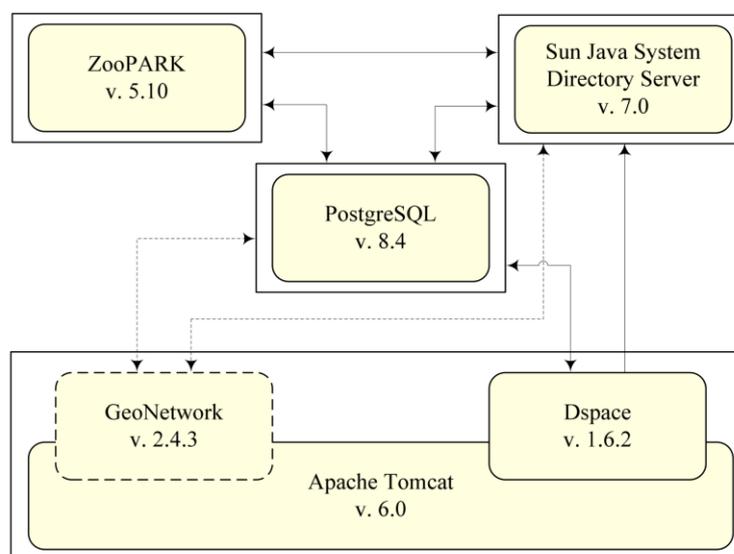


Рисунок 7. Структура серверного программного обеспечения информационной системы.

Наряду с компонентой, управляющей цифровым контентом (DSpace), система включает СУБД PostgreSQL для хранения метаданных, LDAP-сервер для обеспечения функции однократной аутентификации в корпоративном каталоге СО РАН, а также сервер ZooPARK для обеспечения географического поиска в ресурсах DSpace и обеспечения интеграции с другими информационными ресурсами, в том числе и с ресурсами ГИС.

Предложенная архитектура интеграции разнородных данных для задач исследования природных экосистем реализована в виде работающего прототипа информационной системы.

Интеграция гетерогенных источников данных на основе рекурсивной декомпозиции. Рассматривается модель интеграции данных, в которой должны поддерживаться унифицированное представление разнородных источников данных, управление ограничениями целостности, управление выполнением операций манипулирования данными и запросов, согласование данных из разных источников, возможность расширения и настройки на новые источники данных [113]. Предлагаемый подход к интеграции основан на рекурсивной декомпозиции источников данных, при которой каждый источник данных последовательно разбивается на атомарные элементы данных, причем на каждом уровне рекурсивной вложенности данные и их описания представляются единообразно. Такая модель позволяет осуществлять интеграцию различных источников данных на любом уровне

посредством задания связей между произвольными элементами схемы, ограничений целостности и допустимых операций. Разработанная модель представления источников данных, которая позволяет осуществлять многоуровневую интеграцию гетерогенных источников данных в единое информационное пространство, обеспечивает поддержку ограничений целостности на любом уровне интеграции источников данных, а также поддержку структурных и ассоциативных связей между источниками данных на любом уровне интеграции. Обеспечивается возможность динамического изменения схемы данных, а также расширяемость системы за счет новых моделей данных.

Система интегрирует разнородные источники данных в единую систему, с общей схемой для реализации единого механизма управления данными, независимого от моделей данных отдельных источников. В основе модели данных – рекурсивное представление источников данных. Данный способ представления позволяет моделировать произвольно сложные источники данных, сводя их к все более простым элементам. Каждый источник данных (на самом верхнем уровне вложенности существует один источник – сама схема) представляется в виде набора пар (атрибут, источник данных), который разделяет данный источник на составные источники. Данные текущего уровня помещаются в атрибуты пар. Самый нижний уровень вложенности соответствует неделимому элементу данных. Промежуточные уровни источников данных представляют собой метаданные. Очевидно, что при таком подходе к моделированию обработка данных и метаданных осуществляется единообразно. Каждому источнику данных ставится в соответствие набор связей с другими источниками (любого уровня), набор ограничений целостности и набор допустимых операций [104, 31, 80].

Архитектура системы интеграции гетерогенных источников данных является многослойной, в ней можно выделить три уровня (рисунок 8):

- уровень данных (I);
- уровень унифицированного представления данных (II);
- уровень логического представления данных (III).

На нижнем уровне системы располагаются источники данных и драйверы доступа к ним. Источники данных могут представлять собой реляционные, сетевые, иерархические, объектно-ориентированные, объектно-реляционные базы данных, электронные таблицы, XML-файлы, потоки данных, неструктурированные источники. Для доступа к источникам данных могут использоваться драйверы ODBC, провайдеры OLE DB, ADO.NET, либо иные специфические драйверы, которые поддерживают интерфейс выполнения манипуляций над источником данных, которые от него требуются в данной системе.

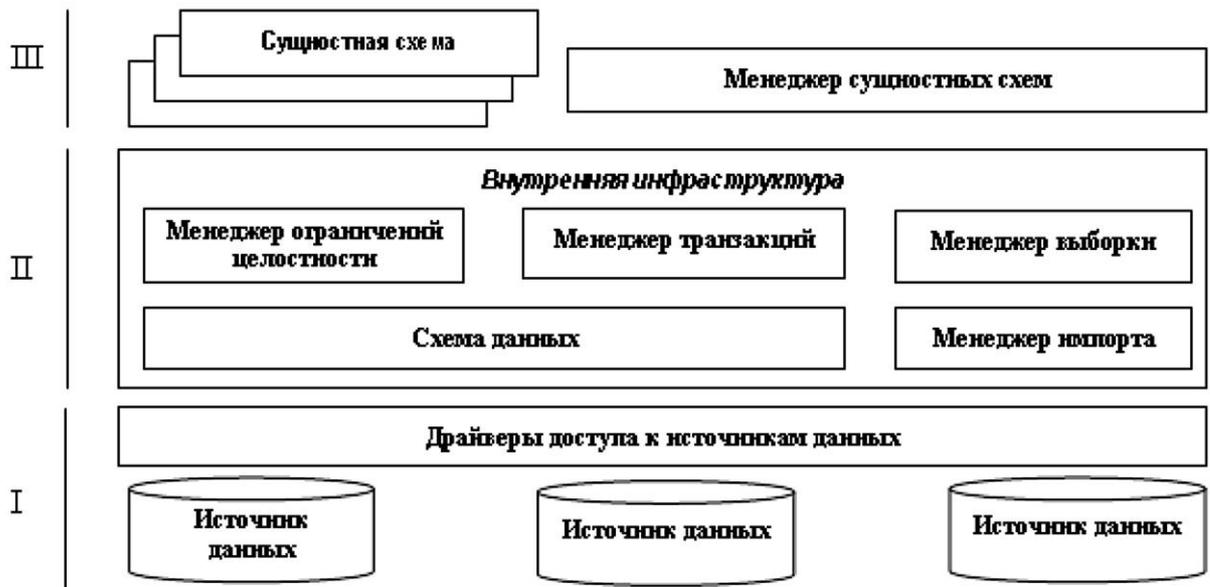


Рисунок 8. Архитектура системы интеграции данных.

На уровне внутренней инфраструктуры реализуется физическое представление схемы интеграции ИД, средства по поддержанию схемы в согласованном состоянии и выполнения базовых операций. Кроме этого, на данном уровне находится компонент, отвечающий за добавление в схему новых ИД.

Схема данных представляет собой внутреннюю рекурсивную структуру моделей источников данных в соответствии с описанной выше математической моделью.

Менеджер транзакций осуществляет выполнение глобальных транзакций в системе. Перед выполнением транзакции в объекте сохраняется информация о состоянии источников данных, которые будет охватывать транзакция, до ее начала. Это необходимо для осуществления процедуры отката в случае, если будет выполнено условие отката транзакции. На глобальном уровне транзакции выполняются последовательно и не могут пересекаться.

Менеджер ограничений целостности контролирует выполнение ограничений целостности во время выполнения операций глобальных транзакций, а также периодически с некоторым интервалом времени. Выполнение глобальных транзакций может состоять из произвольных операций манипулирования источниками данных, что может привести к рассогласованию данных в силу нарушения установленных ограничений целостности. Данный компонент осуществляет проверку после выполнения каждой базовой операции над источником данных на наличие нарушений ограничений целостности. Если нарушение было зафиксировано, то осуществляется откат операции и, по возможности, всей транзакции. Периодическая проверка выполнения ограничений целостности нужна на случай, если в источниках данных произошли локальные транзакции под управлением самих источников, что

привело к нарушению ограничений целостности, установленных в данной схеме. В случае обнаружения такого нарушения пользователю(администратору) сообщается об этом и работа с этим ИД становится невозможной до тех пор, как согласованность данных на глобальном уровне не будет восстановлена.

Менеджер выборки данных отвечает за соединение источников данных и представление данных в виде, отвечающем потребностям пользователя. На вход компонент получает запросы, сформулированные в терминах сущностной схемы, а на выходе выдает результат выборки данных из ИД.

Менеджер импорта осуществляет импорт в систему схем данных определенного типа. В основе работы компонента лежат алгоритмы декомпозиции источников данных. Расширение спектра поддерживаемых системой моделей данных требует добавления средств интерпретации данных моделей в этот компонент.

Внешнее представление – это логическое представление интегрированной схемы, отражающее сущности и взаимосвязи конкретных предметных областей.

Схема сущностей обеспечивает работу с системой в терминах предметной области. Компонент обеспечивает высокоуровневое представление источников данных, позволяющее решать задачи предметной области, а не тратить ресурсы на интерпретацию схемы интеграции. Построение подобной схемы сущностей, по сути, означает определение отображения схемы источников данных на предметно- ориентированную схему сущностей.

Менеджер схем сущностей позволяет создавать и редактировать новые схемы сущностей и устанавливать отображение между источниками данных и сущностями.

Предложенная модель представления источников данных, которая позволяет осуществлять многоуровневую интеграцию гетерогенных источников данных в единое информационное пространство, обеспечивает поддержку ограничений целостности на любом уровне интеграции источников данных; поддержку структурных и ассоциативных связей между источниками данных на любом уровне интеграции [39, 40].

Преимущества модели:

1. Возможность интеграции данных на любом уровне декомпозиции;
2. Возможность динамического изменения схемы данных;
3. Расширяемость системы за счет новых моделей данных;
4. Универсальное представление связей, ограничений целостности и операций над данными.

Технологическая платформа для распределенных гетерогенных систем сбора и обработки данных. В данной работе рассматривается именно разработка платформы для

систем сбора данных и принципов ее построения без привязки к конкретной предметной области, однако точкой отсчета является система сбора гидрологических данных, на примере которой предпринимается попытка выделения общих черт таких систем и основных проблем при их реализации [65].

Требованиями к такой платформе являются:

- организация сбора гидрометеорологических и гидрохимических данных из различных источников. Датчики, сбор с которых ведется в реальном времени. Архивы, синхронизация с которыми осуществляется по запросу пользователя. И ручной ввод данных, сбор которых не может быть организован автоматически.
- организация единого распределенного хранилища данных.
- организация многопользовательского удалённого доступа к собранным данным. Потребителями данных могут быть как средства визуализации и мониторинга в реальном режиме времени, так и различного рода программное обеспечение для проведения расчетов.

Общая структура системы представлена на рисунке 9, структура системы несколько уточняет классическую трехуровневую систему управления и сбора данных и содержит уровень объектов мониторинга, уровень сбора данных, уровень хранения и подготовки данных, сервисный уровень для поддержания клиентских предложений.

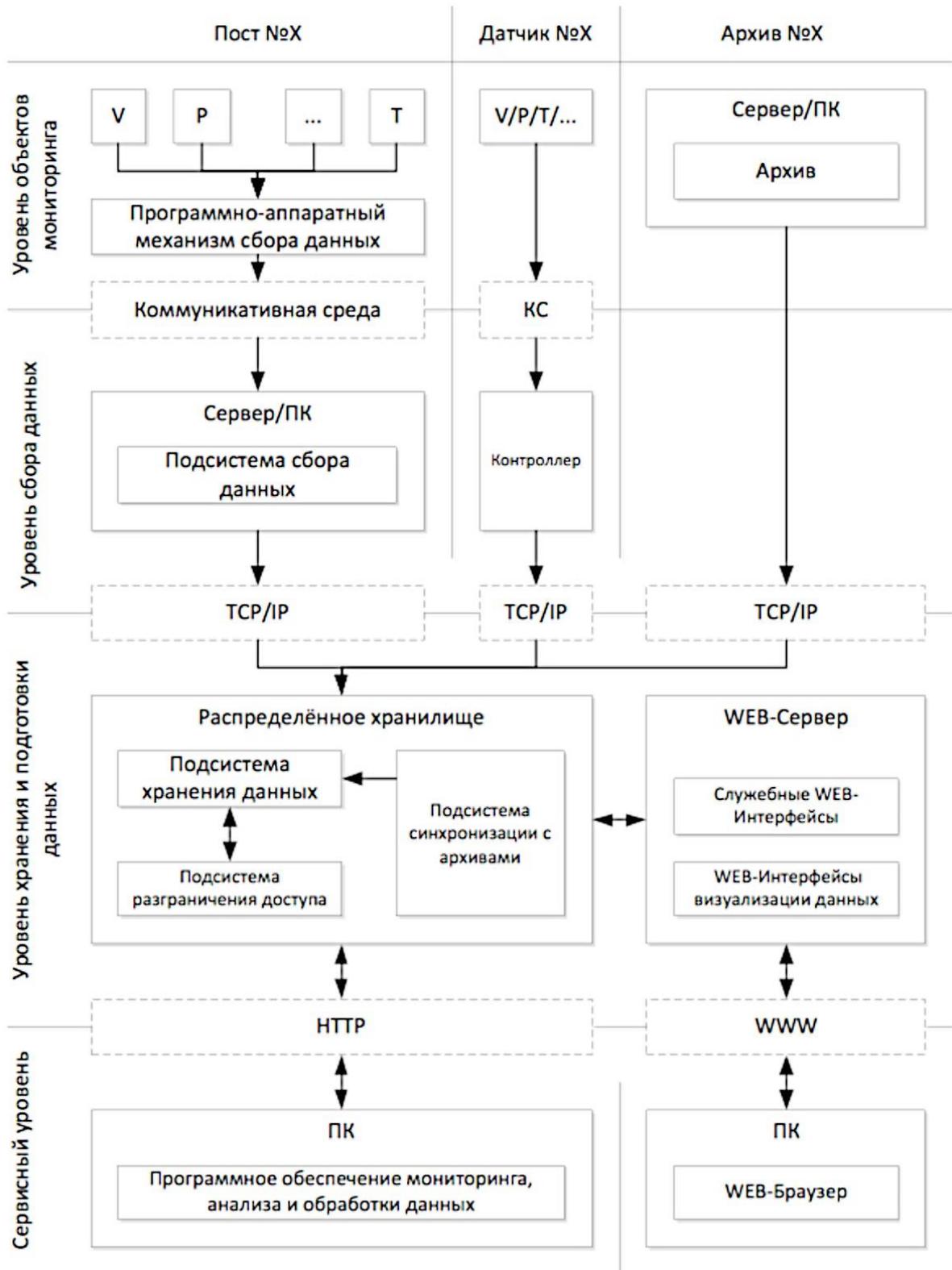
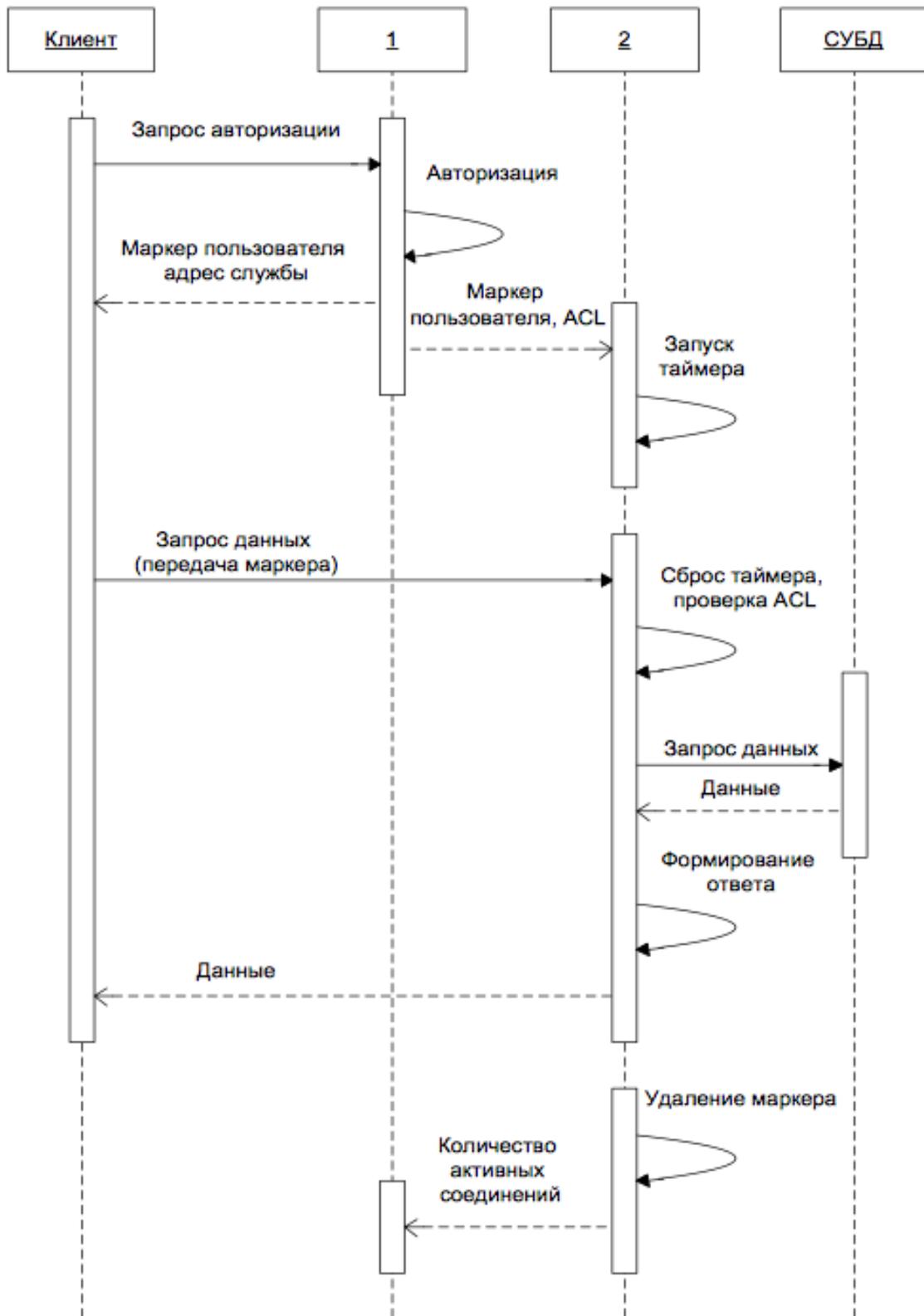


Рисунок 9. Архитектура системы

Распределенное хранилище подразумевает распределение нагрузки между дублирующими друг друга службами. Алгоритм получения данных и распределения нагрузки представлен в виде диаграммы последовательности UML и показан на рисунке 10.



1. Служба распределения нагрузки и авторизации
2. Служба разграничения доступа и обработки данных

Рисунок 10. Распределение нагрузки и разграничение доступа

Клиент отправляет запрос на авторизацию службе авторизации и распределения нагрузки. В случае успешной авторизации он получает адрес службы обработки данных и маркер, по которому остальные службы смогут его идентифицировать. Одновременно с этим служба обработки данных и разграничения доступа имеющая условно наименьшее (смотри далее) количество активных соединений, получает тот же маркер и список прав доступа пользователя и запускает таймер для проверки активности пользователя.

Далее клиент формирует запрос на данные и передает его по полученному адресу вместе со своим маркером. В случае наличия у службы маркера переданного маркера, клиент считается авторизованным и таймер его активности сбрасывается [105]. В случае отсутствия запросов с данным маркером в течение некоторого времени, сеанс считается завершенным маркер, и маркер удаляется, а служба авторизации уведомляется о количестве активных клиентов.

При получении запроса с определенным маркером служба распределения нагрузки проверят список прав доступа, запрашивает данные из БД и формирует ответ в соответствии с запрашиваемым форматом данных и правами пользователя. Далее запросы обрабатываются до тех пор, пока маркер пользователя не будет удален.

Развертывание и регистрация в системе дополнительных служб обработки данных и распределения доступа позволит масштабировать систему в целом. Каждая из таких служб может находиться на выделенном сервере.

Следует отметить, что равномерное распределение клиентов по службам не всегда эффективно. Такой алгоритм распределения нагрузки уместен только в том случае, если все службы имеют одинаковую производительность. Что справедливо только в одном случае – все службы запущены на машинах с одинаковой производительностью (или на одной машине) и имеют одинаковую конфигурацию. В остальных же случаях нужно вести речь об условном весе каждого клиента и условном весе службы. Вес клиента характеризует сложность и частоту отправляемых им запросов, а вес службы характеризует ее производительность относительно некоторой эталонной службы.

Достоинствами разрабатываемой платформы является обеспечение требований масштабируемости и надежности на уровне архитектуры и модели вычислений, а также инвариантность к предметной области и виду собираемых данных до определенного уровня абстракции. Среди недостатков работы следует отметить недостаточную проработанность механизма сопряжения системы в целом с различными видами датчиков на низком уровне.

Проводя анализ представленных систем и моделей, можно сделать очевидный вывод, что раздельно существуют системы аналитической обработки и оперативного управления

данными, системы управления документами и прочие системы. Различные задачи требуют использования различных моделей представления данных.

Предложенные общие структуры моделей интегрирования разнородных данных могут быть взяты за основу при создании подобных моделей в других областях применения с использованием геопространственных разнородных данных. Но ни одна из рассмотренных моделей не может в полной мере удовлетворить требования для создания геоинформационной системы с использованием разнородных пространственных данных, так как существуют определенные особенности в обработки таких данных и связанных с ними представлений [30, 34, 46].

1.4. Обзор программных инструментов для доступа к обработке баз данных

Механизмы доступа к базам данных снижают сложность обмена информацией с базами, однако интерпретация результатов их работы также достаточно трудоемка. Можно выделить несколько субъектов, участвующих в движении информации между базой данных и приложением:

1. Интерфейсная часть приложения или его программная часть, манипулирующая информацией, хранимой в базе данных;
2. Компоненты, обеспечивающие связь приложения с механизмом доступа к базе данных;
3. Механизм доступа к базе данных;
4. База данных.

На рисунке 11 представлена схема движения информации между приложением и базой данных. Из рисунка видно, что между клиентским приложением и СУБД не существует прямой связи, в следствие дополнительно встраиваются необходимые программные модули, позволяющие клиентскому приложению получать доступ к БД. Такие модули называются механизмами доступа к данным.

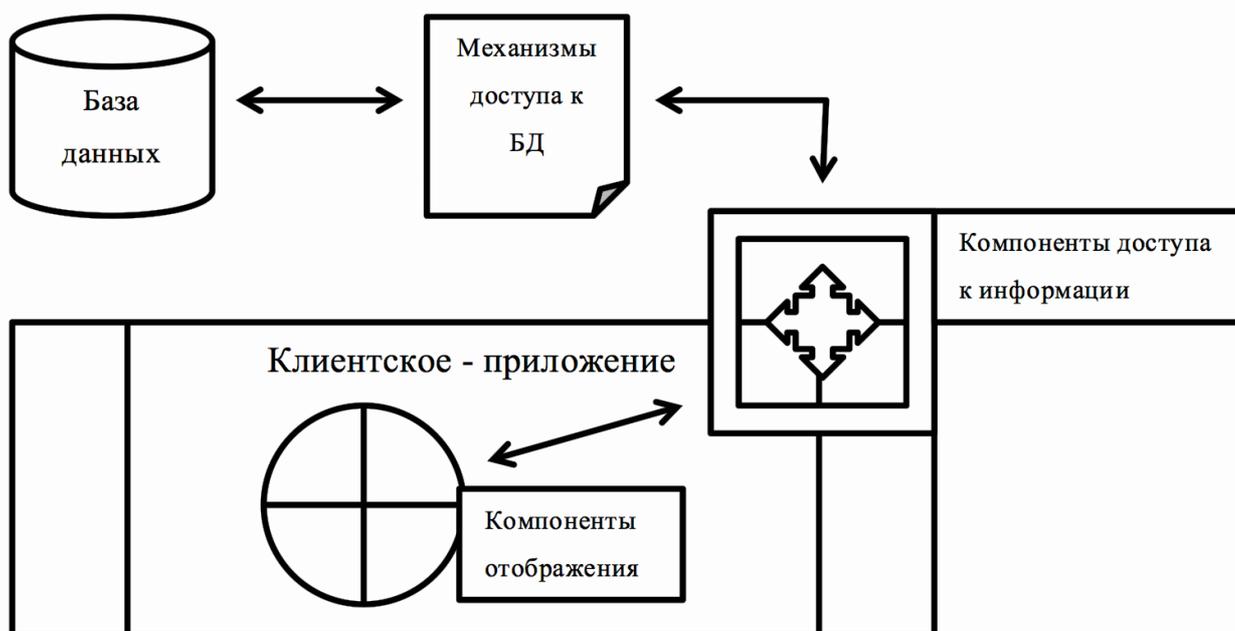


Рисунок 11. Схема движения информации от базы данных к приложению.

Они выполняют функцию получения данных из базы с учетом особенностей хранения или правил обмена, так же реализуют интерфейс для доступа к данным на логическом уровне, независимый от выбранного типа базы данных [90]. Таких механизмов разработано множество, причем большинство из них в некоторой степени универсальны и предназначены для работы с разными видами баз данных, а для доступа к широко распространенным базам данных обычно можно выбрать один из нескольких механизмов.

Первоначально преобладали способы, основанные на API библиотек СУБД, входящих в состав клиентского программного обеспечения, устанавливаемого на компьютерах пользователей. Как правило, клиентское программное обеспечение включало в себя и собственную среду разработки прикладного программного обеспечения. Это приводило к тому, что замена СУБД требовала переписывания значительной части кода клиентского приложения. Прикладной программный интерфейс (API – application programming interface) представляет собой набор различных функций, констант, классов, форматов запросов, вызываемых из клиентского приложения [37]. Более простым языком API – это готовый программный код. API предназначен для того, чтобы облегчить задачу написания приложения, благодаря использованию шаблонов кода.

На сегодняшний день самый известный сервис code.google.com, предоставляет около полусотни разнообразных API решений. Безусловно, в идеале хотелось бы, чтобы все подключаемые системы имели единообразное универсальное API и можно было бы использовать один код для всех систем. Главной особенностью такой системы управления был

бы единый интерфейс доступа ко всем данным. Как результат - в настоящее время главенствующими становятся универсальные механизмы доступа к данным.

Универсальный механизм доступа к данным обеспечивает возможность использования одного и того же интерфейса для доступа к разным типам СУБД. Универсальные механизмы доступа к данным обычно реализованы в виде библиотек и дополнительных модулей (драйверов или провайдеров). Библиотеки содержат определенный стандартный набор классов, методов, параметров, и обеспечивают стандартный интерфейс доступа к данным. Дополнительные модули реализуют непосредственное обращение к функциям клиентского API конкретных СУБД [106]. Причем, эти дополнительные модули, устанавливаются исходя из текущей потребности.

Универсальные механизмы доступа к данным обычно реализованы в виде библиотек и модулей, называемых драйверами или провайдерами, содержащих стандартный набор функций или классов, реализованных на основе функций клиентского API конкретных СУБД. Среди универсальных механизмов доступа к данным наиболее распространены в настоящее время следующие программные интерфейсы:

- ODBC - Open Database Connectivity.
- OLE DB - Object Linking and Embedding Database.
- ADO - ActiveX Data Objects.
- BDE - Borland Database Engine.
- JDBC - Java Database Connectivity.
- ADO.NET - ActiveX Data Objects технологии NET.

ODBC. Наиболее распространенный программный интерфейс, обеспечивающий доступ к данным. В начале 1990 г. существовало несколько поставщиков баз данных, каждый из которых имел собственный интерфейс. Если приложению было необходимо общаться с несколькими источниками данных, для взаимодействия с каждой из баз данных было необходимо написать свой код. Для решения возникшей проблемы Microsoft и ряд других компаний создали стандартный интерфейс для получения и отправки источникам данных различных типов. Этот интерфейс был назван Open Database Connectivity, или открытый механизм взаимодействия с базами данных.

Программный интерфейс разработан в сотрудничестве с Simba Technologies на основе спецификаций Call Level Interface (CLI) — программный стандарт, закрепленный в документе ISO/IEC 9075-3:2003. CLI описывает как именно программа должна отправлять SQL-запросы к

СУБД и как именно возвращенный набор записей должен быть последовательно обработан приложением.

Технология ODBC, предназначена для унификации доступа к информации, хранящейся в базах данных различных видов. ODBC состоит из набора драйверов, осуществляющих операции обмена с определенными базами данных, и менеджера драйверов, осуществляющего передачу запросов от приложения к драйверу и передачу информации от драйвера к приложению.

Этот интерфейс очень удобен: одно приложение может обращаться к различным базам данных через общий набор команд. ODBC дает возможность приложениям обращаться к данным из различных систем управления базами данных. Он обеспечивает максимально широкие возможности для взаимодействия. Интерфейс Microsoft Open Database Connectivity признан в качестве промышленного стандарта и является компонентом архитектуры *Microsoft WOSA* (Windows Open Services Architecture - открытая архитектура служб среды Windows). Как показано на рисунке 12, менеджер драйверов является промежуточным звеном между приложением и базами данных.

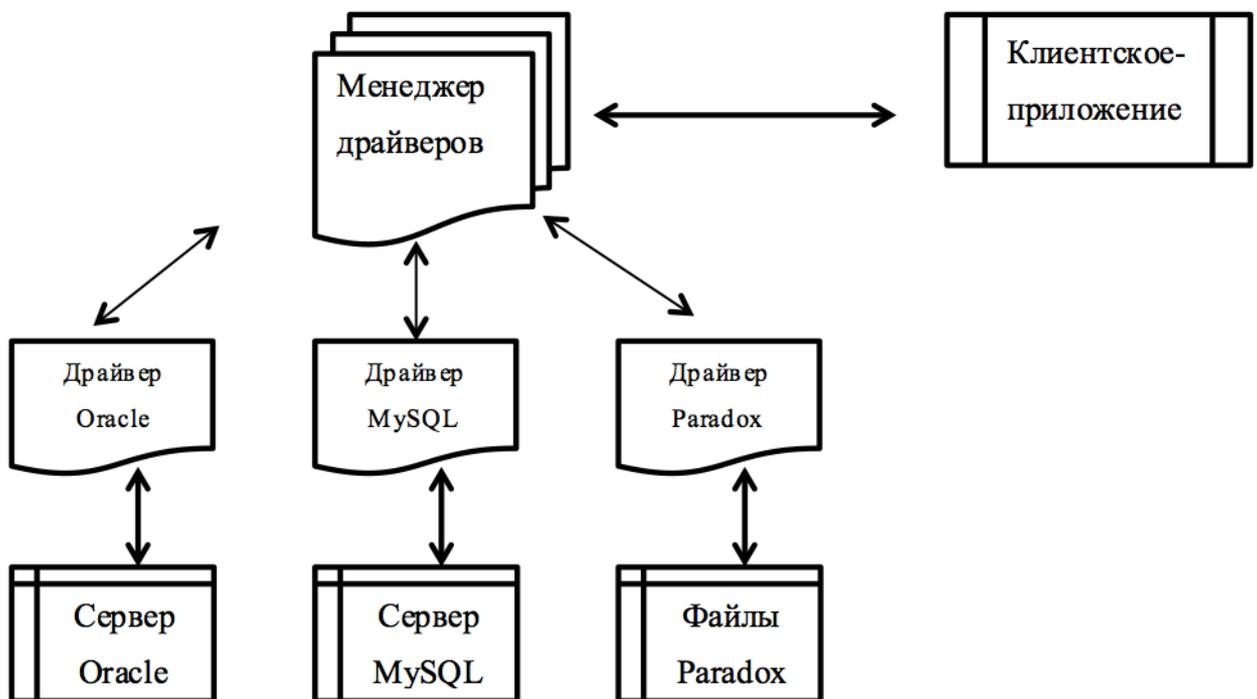


Рисунок 12. Архитектура ODBC

Интерфейс ODBC содержит набор функций, который управляет каждым инструментом базы данных. Если приложению нужно сменить используемую базу, разработчик просто

заменяет один драйвер другим, и приложение может работать как обычно, без необходимости модификации кода программы. ODBC использует низкоуровневый интерфейс, поэтому программисты на С и С++ могут задействовать все преимущества технологии ODBC.

Для получения и изменения данных используется язык запросов SQL, вне зависимости от того, поддерживается ли он базой данных, к которой обращается приложение. Если база данных не поддерживает язык SQL, то доступ к ней не отличается от доступа к БД, поддерживающим SQL. В этом и заключается унификация доступа к базам данных системой ODBC – приложение указывает название драйвера, который должен использоваться для подключения к базе данных, и передает запрос, в котором описан состав требуемой информации. Далее механизм ODBC выполняет все необходимые операции по получению информации, скрывая от приложения специфику работы с конкретной базой данных. Доступ приложений к ODBC осуществляется через API-функции, реализованные в динамических библиотеках.

OLE DB. Механизм работы с самыми разнообразными источниками данных на базе COM-интерфейса. OLE DB определяет набор COM-интерфейсов (Component Object Model), включающих различные сервисы однотипного доступа к различным данным (в том числе и к не реляционным БД, например, к папкам систем электронной почты или просто файлам), обеспечивая при этом поддержку работы с наборами данных и иерархическими наборами записей, подключенными непостоянно к сети. Для доступа к БД требуется установка OLE DB провайдера для СУБД (DLL СУБД загружается при выполнении приложения в его адресное пространство). Кроме того, фирмой разработан специальный провайдер (Microsoft OLE DB Provider for ODBC Drivers), который может работать не через API клиентской части СУБД, а через интерфейс ODBC API.

Характерной особенностью механизма OLE DB является то, что данные, возвращаемые OLE DB приложению, представляют собой не просто массив информации, а COM-объект, обладающий, помимо самой информации, методами управления этой информацией (например, фильтрация и сортировка). Логическая схема доступа к данным с помощью OLE DB представлена на рисунке 13.

Приложение, обращается к поставщику данных, который перенаправляет запрос от приложения к базе данных, либо обрабатывает запрос самостоятельно. Таким образом, поставщик данных аналогичен драйверам ODBC и BDE. Механизм OLE DB имеет доступ к ODBC, поддерживая, таким образом, большое количество драйверов, реализованных для ODBC.

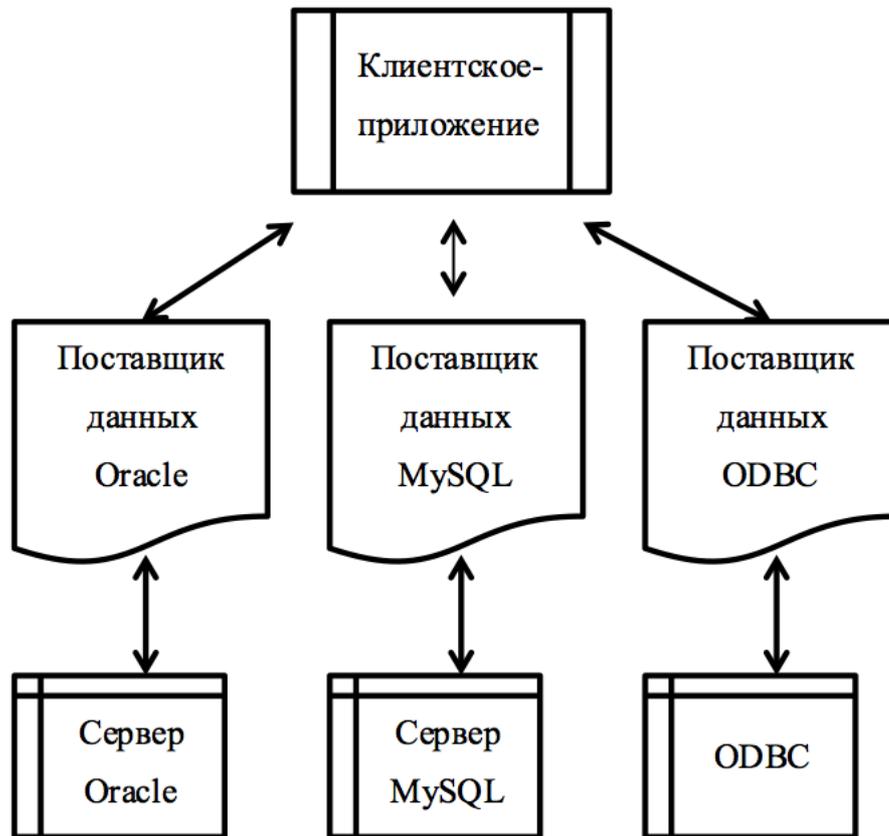


Рисунок 13. Движение информации между приложением и БД при использовании OLE DB

Недостатком OLE DB (и соответственно технологии ADO, построенной на OLE DB) является более низкое быстродействие по сравнению с механизмами ODBC и BDE. Это связано с использованием COM-объектов, применение которых нагружает операционную систему.

ADO. Своеобразная надстройка над OLE DB (использует библиотеки OLE DB) - дополнительный набор библиотек, содержащих COM-объекты, реализующие интерфейс доступа к данным. Этот набор библиотек первоначально включал две объектные технологии Microsoft: Data Access Objects (DAO) и Remote Data Objects (RDO) - два различных механизма доступа к локальным и удаленным базам данных соответственно.

Как ответ на потребность создания единой технологии, обеспечения единого подхода при работе с БД и единого интерфейса для доступа к локальным и удаленными данным, появилась технология ADO. ADO является более дружественной оболочкой базовой технологии OLE DB и позволяет работать с любыми базами данных.

Технология ADO (от англ. ActiveX Database Objects – механизм доступа к базам данных через объекты ActiveX). Обеспечивает удобный и надежный доступ к данным, хотя и несколько более медленный, чем в технологиях BDE и dbExpress. Хорошо подходит для работы с

системами управления базами данных (СУБД) от фирмы Microsoft (MS Access, MS SQL Server), поскольку не требует дополнительных компонентов (библиотек, драйверов), так как они уже есть на компьютере пользователя. Схема доступа к информации с помощью ADO представлена на рисунке 14.

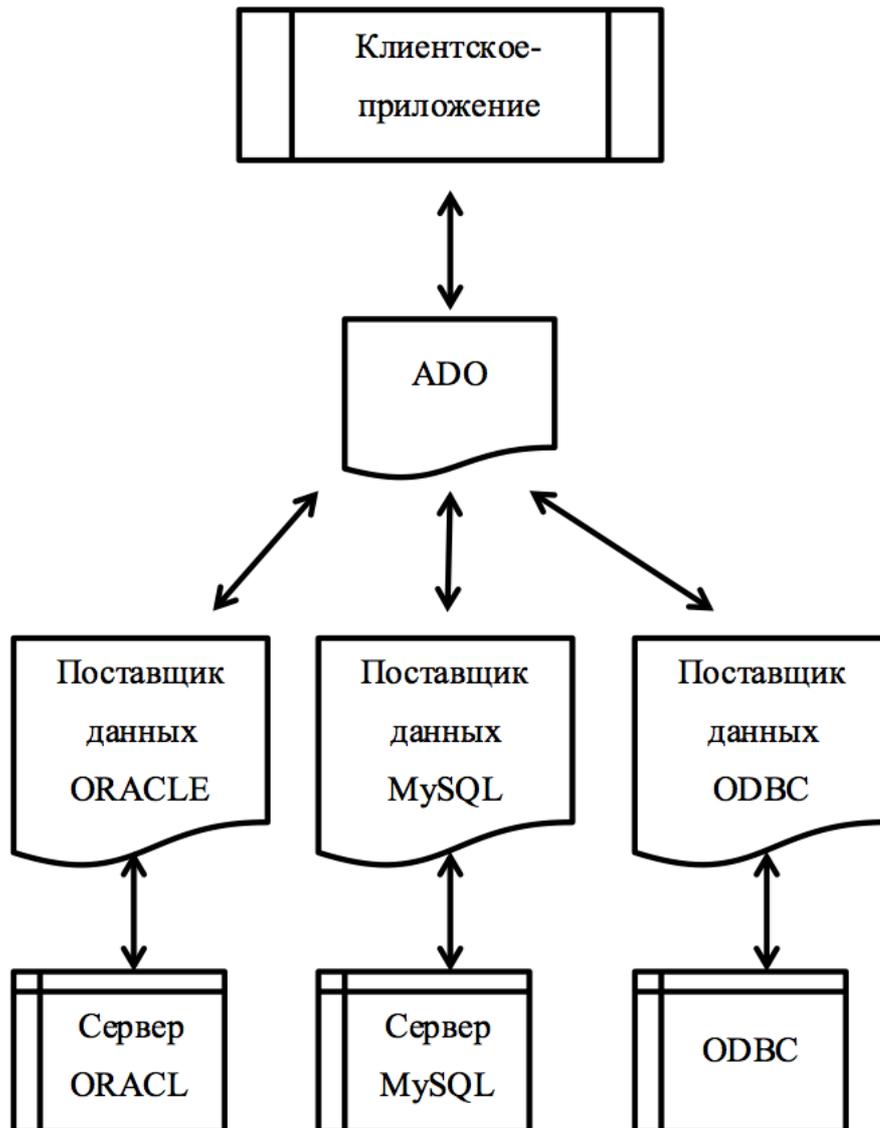


Рисунок 14. Движение информации между приложением и БД при использовании ADO.

Механизм ADO предоставляет несколько основных COM-объектов, используемых для получения и управления информацией (имеются дополнительные COM-объекты, расширяющие функциональность ADO). Непосредственный доступ из Delphi возможен только к ADO. Помимо компонентов для доступа к ADO, Delphi имеет компоненты для получения

информации из баз данных с помощью механизма BDE фирмы Borland, а также компоненты db Express, осуществляющие прямой доступ к различным видам баз данных.

BDE. Механизм доступа к базам данных фирмы Borland. Базовая технология доступа к БД является аналогом ODBC и имеет схожую с ней архитектуру. Универсальный механизм доступа к данным, базирующийся на двух группах библиотек-драйверов (SQL Links - для серверных СУБД и ODBS Links - для серверных и автономных СУБД). Эти библиотеки позволили применять стандартный набор функций для доступа к данным dBase, Access, FoxPro, ODBC-источников, а также большинства серверных СУБД [107]. BDE поддержан на уровне компонент в визуальных средах разработки фирмы Borland. Реализация механизма позволяет функциям приложения обращаться к функциям клиентского API или ODBC API, а также непосредственно манипулировать файлами данных некоторых СУБД. Архитектурная схема механизма представлена на рисунке 15.

Для доступа к базе данных с помощью BDE на компьютере должны быть установлены библиотеки BDE общего назначения (обычно устанавливаются вместе со средой разработки и затем включаются в .exe файл), а также BDE-драйвер для данной СУБД.

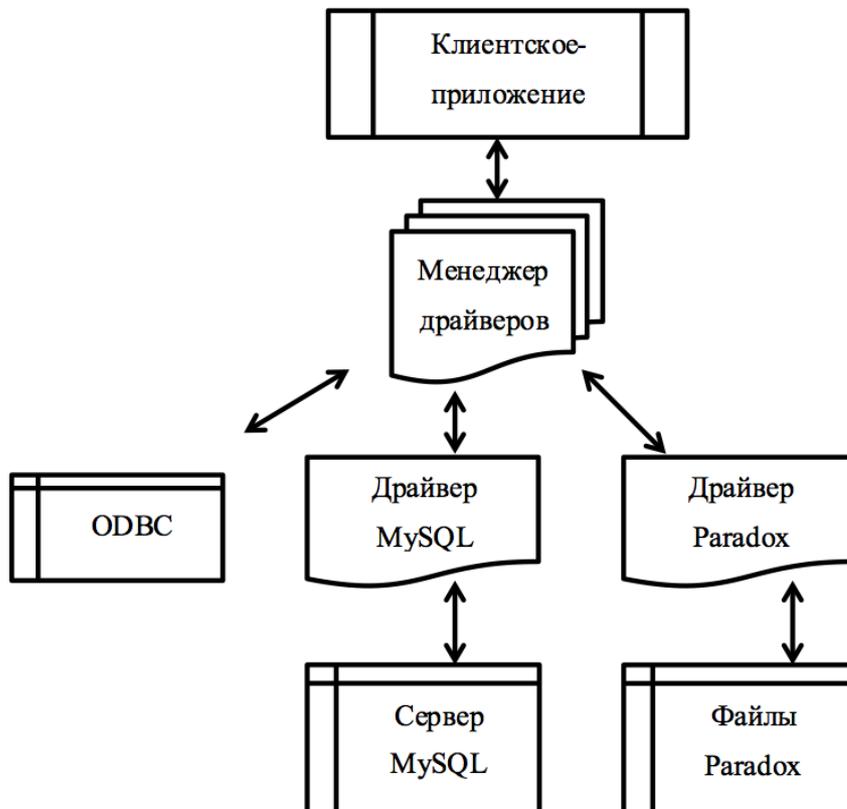


Рисунок 15. Движение информации между приложением и БД при использовании механизма BDE.

Существенным недостатком использования этой технологии является достаточно трудоемкий процесс развертывания программы работы с базой данных [58]. Подводя итог обзора механизмов доступа к гетерогенным данным можно сказать следующее - существует несколько способов доступа к данным из средств разработки и клиентских приложений, которые условно можно разделить на две категории:

1. Использование клиентского API (или клиентских СОМ-объектов).
2. Применение универсальных механизмов доступа к данным.

При использовании API следует понимать, что замена СУБД потребует переписывания значительной части кода клиентского приложения. В свою очередь следует подчеркнуть, что универсальные методы так же не лишены недостатков: снижение производительности, поставка соответствующих драйверов и их настройка и т. п.. Нельзя забывать и о том, что они значительно увеличивают производительность труда программиста, так как отпадает необходимость в изучении специфических интерфейсов и специфических средств разработки. Механизмы ODBC, OLE DB и ADO фирмы Microsoft представляют собой по существу промышленные стандарты [88].

В результате проведенного анализа можно сделать вывод, что ни одна из вышеописанных технологий доступа не может в полной мере удовлетворить требованиям прикладных задач использования распределённых вычислительных систем обработки разнородных геоданных, хотя некоторые из рассматриваемых технологий применялись в разработке геоинформационных систем в совокупном качестве и сложных форматах представления.

На этом фоне ставится следующая **научная задача** - разработка информационных технологий интеграции разнородных геоинформационных данных для обеспечения принятия решения по управлению территорией в различных сферах производственной деятельности. Данную технологию можно рассматривать как основу для реализации системы, поддерживающей оперативное управление разнородными данными и их аналитическую обработку.

В работе необходимо предусмотреть новый подход к обработке разнородных данных. Необходимо обеспечить интегрированное представление данных, независимо от того, где эти данные фактически расположены и в каком формате представлены. Следующим шагом в исследовании будет моделирование новой архитектуры обработки и управления разнородными распределёнными данными, на основе подобных и существующих систем.

1.5. Разработка модели обработки разнородных данных для решения прикладных задач

Во многих задачах устойчивого развития территорий, необходимо одновременно обрабатывать аэрокосмические данные, данные наземных измерений, разнообразную текстовую информацию и т.д.. Это накладывает существенные особенности на все процессы обработки данных и применения прикладных инструментов, используемых для этого.

Разработка технологии и программной реализации консолидации разнородной информации из различных источников требует получения ответов на следующие вопросы:

1. Каким образом может быть решена проблема слияния данных, полученных из распределенных источников, содержащих разнородные данные;
2. Как может быть решена проблема идентификации данных об одном и том же объекте, представленном фрагментами в различных БД;
3. Каким образом можно справиться с разнообразием структур представления данных в различных источниках;
4. Какой должна быть архитектура системы обработки данных.

Использование геопространственных данных связано с необходимостью постоянно интегрировать эти данные с геопространственными данными другой природы. При этом большая часть проблем связана с реализацией механизмов интегрирования разнородных данных [82].

Решение задач интегрирования разнородной геопространственной информации в различных предметных областях требует рассмотрения задачи разработки новой модели доступа к данным. Использование различных механизмов доступа к данным позволит упростить доступ к разнородной геопространственной информации. В постановке задачи обработки данных можно выделить основную задачу, имеющую большое значение - развитие модели доступа к распределенным разнородным пространственным данным [15].

Моделирование можно назвать одним из основных методов исследования сложных систем, который используется для принятия управленческих решений в различных сферах науки и техники. Идея моделирования состоит в замене реального объекта некоторой моделью, прообразом этого объекта. При этом у исследователя появляется возможность проводить эксперименты, которые были бы невозможны над реальным масштабным объектом. Так работа не с самим объектом или явлением, процессом, а его моделью происходит достаточно быстро и без существенных материальных затрат на его исследование в различных ситуациях. Однако вычислительные эксперименты с моделью системы позволяют изучить её поведение в достаточном объеме, что не позволяют теоретические подходы.

Модель подготовки разнородных данных необходимо разработать с учетом выявленных проблем интеграции разнородных данных в исследуемых моделях и системах. Обработка разнородной информации в единой системе позволит использовать различные данные в совокупности для производимых вычислений и операций, что будет способствовать получению результатов в виде расширенной статистической информации для поставленной задачи. Необходимо внести в разработку новой модели обработки разнородных данных следующие основные требования к функциональным характеристикам модели:

1. Модель должна обеспечивать контроль и обработку входящей разнородной информации.
2. Модель должна иметь собственный алгоритм нахождения драйверов доступа к данным.
3. Модель должна поддерживать возможность получения данных из разнородных источников (различных баз данных, документов, сайтов и др.).
4. Модель должна быть легко масштабируема относительно подключения различных источников.
5. Модель должна поддерживать возможность выполнения операций расширенных запросов (возможность работы по средствам реляционной алгебры, SQL).
6. Модель должна обладать возможностью определения формата входящей информации.

Модель подготовки разнородных данных необходимо разработать с учетом выявленных проблем в исследуемых моделях и системах. Обработка разнородной информации в единой системе позволит использовать различные данные в совокупности для производимых вычислений и операций, что будет способствовать получению результатов в виде расширенной статистической информации для поставленной задачи [20].

Подходы, основанные на обработке данных из источников, предоставляющих доступ к информации по какому-либо распространенному интерфейсу, сужают диапазон поддерживаемых информационных ресурсов до баз данных, предоставляющих доступ по поддерживаемому интерфейсу. Решением данной проблемы может быть создание надстройки над базами данных, обеспечивающей централизованный доступ и управление информацией из уже существующих и функционирующих источников [108].

Разработка инструментов данного типа связана с такими проблемами, как передача подзапросов источникам, выбор алгоритмов для манипулирования данными, организация обработки данных из различных источников. Процесс обработки запросов к различным

источникам информации может быть логически разделён на следующие стадии: анализ запроса, оптимизация с целью повышение производительности и выполнение [71, 72].

Процесс обработки запросов к различным источникам информации может быть логически разделён на следующие стадии: анализ запроса, оптимизация с целью повышение производительности и выполнение. Таким образом, система обработки и управления информацией из различных источников может быть разделена на ряд следующих функциональных составляющих.

В первую очередь, можно выделить языковые инструменты для описания нечетких запросов. Данная компонента должна преобразовать языковое представление запроса в древовидную структуру, содержащую операции из расширенной реляционной алгебры [10].

Следующей составляющей является оптимизатор запросов, который получает древовидное представление запроса от языковой компоненты, и производит оптимизации (не меняя тип представления запроса) с целью повышения производительности, основываясь на моделях стоимости исполнения различных операций для каждого исполнителя или адаптера источника.

Еще одной компонентой системы должно быть ядро исполнения запросов, которое получает оптимизированный запрос от оптимизатора, собирает необходимые данные из источников, и производит над ними операции, согласно плану выполнения запроса, полученного от оптимизатора.

В рамках основной цели научной работы, для уменьшения трудоемкости подключения новых источников данных, наиболее оптимальным решением является возможность поддержки любого из широко известных интерфейсов доступа к данным, такой подход позволит обрабатывать данные не только из структурированных источников, поддерживающих свой собственный язык запросов и определённые модели представления данных, но и из неструктурированных источников [109].

Глубокое рассмотрение процесса интеграции и взаимодействия разнородных данных между собой, выявило ряд проблем, основными из которых являются:

1. Отсутствие единого прикладного программного обеспечения для обработки гетерогенных геопространственных данных, или его наличие, целью которого является решение ряда типизированных задач, настроенного на конкретную сферу применения;
2. Отсутствие четко сформулированных алгоритмов обработки и взаимосвязи гетерогенных данных;
3. Частичная несовместимость форматов геопространственных данных;
4. Разнородность геопространственных данных;

5. Некорректное формулирование или полное отсутствие метаданных - описание наборов геопространственных данных.

На основе приведенных функциональных требований, модель можно разделить на три основные составляющие:

1. Планировщик, который получает план выполнения запроса от внешней компоненты системы (оптимизатора, или же непосредственно от языковых инструментов). На основе полученного плана определяет, какие компоненты ядра исполнителя необходимы для выполнения данного запроса, составляет для каждой из них задание, в которое входит подзапрос для данной конкретной компоненты, выделенный из общего плана выполнения запроса и некоторая метаинформация (например, адреса пересылки данных, адреса оповещения оповещений и др.).
2. Исполнители операций расширенной реляционной алгебры, организованные в древовидную структуру. Каждый исполнитель может поддерживать различные операции, алгоритмы их выполнения, используя различные технологии (например, средствами существующих баз данных, таких как PostgreSQL, AsterixDB, и др.) Следует отметить, что для добавления новых операций, достаточно лишь реализовать специальный интерфейс исполнителя, таким образом, достигается масштабируемость относительно поддержки новых операций и различных алгоритмов их выполнения.
3. Механизмы доступа к данным, которые, в свою очередь предоставляют интерфейс для извлечения данных из различных источников. Таким образом, достигается масштабируемость в отношении поддержки новых источников данных (то есть для извлечения данных из новой базы, или же типа документов, достаточно реализовать для него адаптер).

Следует отметить, что некоторые источники (использующие базы данных), имеют возможности для выполнения операций из плана выполнения, своими средствами, и могут так же реализовывать интерфейс исполнителя запросов [74, 75].

Для реализации концепций касающихся представления информации и операций для их обработки требуется детальное построение инструментов для обработки разнородной информации. Предполагается декомпозиция инструмента обработки данных на такие функциональные составляющие, как модуль запросов, модуль обработки, модуль интеграции и ядро исполнения системы – планировщик (рисунок 16).

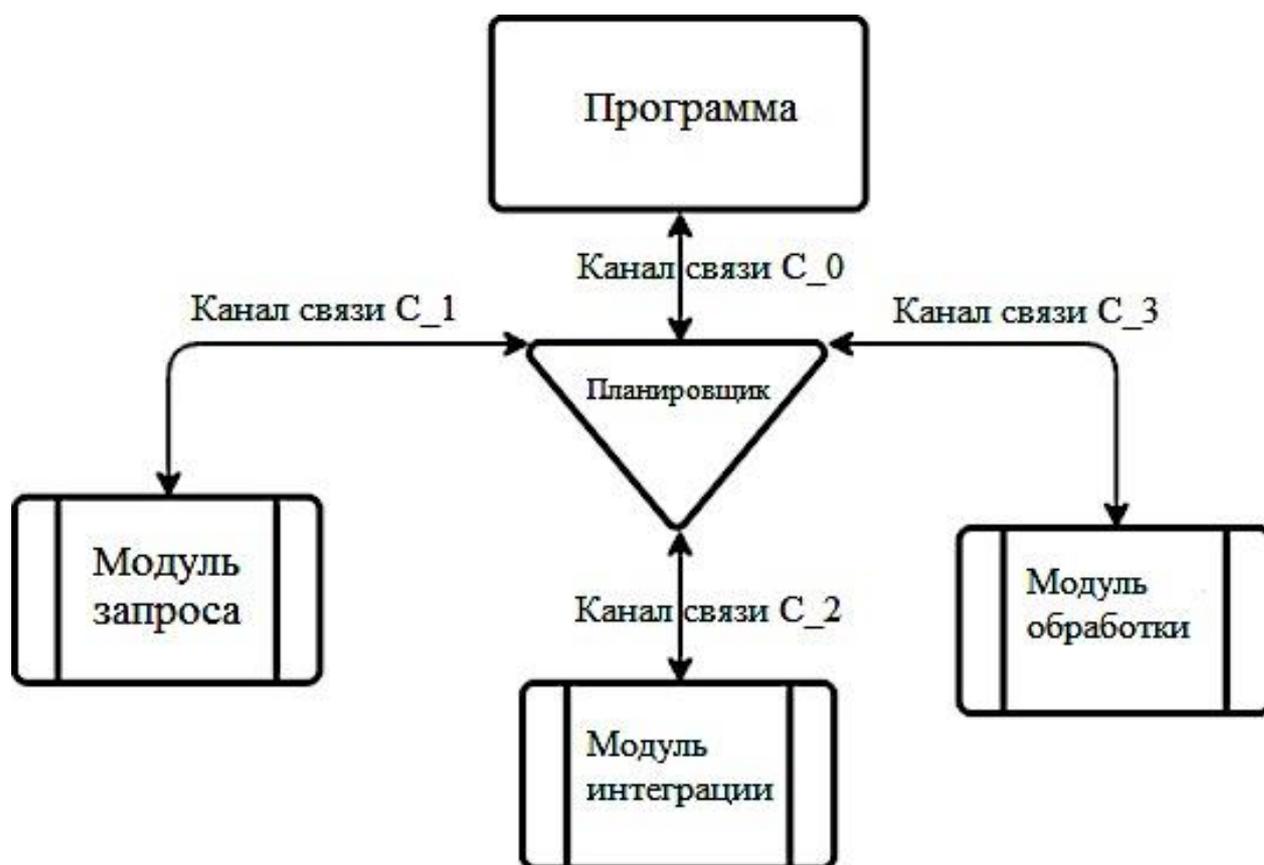


Рисунок 16. Схема декомпозиции модуля системы обработки данных

Выполнение запроса начинается, когда планировщику поступает план выполнения запросов. Планировщик формирует задания для каждого нового источника и рассылает данные задания по модулям системы. Далее, происходит процесс обработки информации согласно полученному заданию. В результате интерактивного взаимодействия с системой, осуществляется подключение новых данных, модуль определяет процедуру обработки разнородных данных, которая лучше всего отвечает характеру конкретной задачи интегрирования и типам используемых данных. В результате интегрирования информация передается в канал обработки данных, вызываются необходимые для его реализации сервисы и выполняется процедура интегрирования [21].

Общая схема модели обработки разнородных данных представлена на рисунке 17.

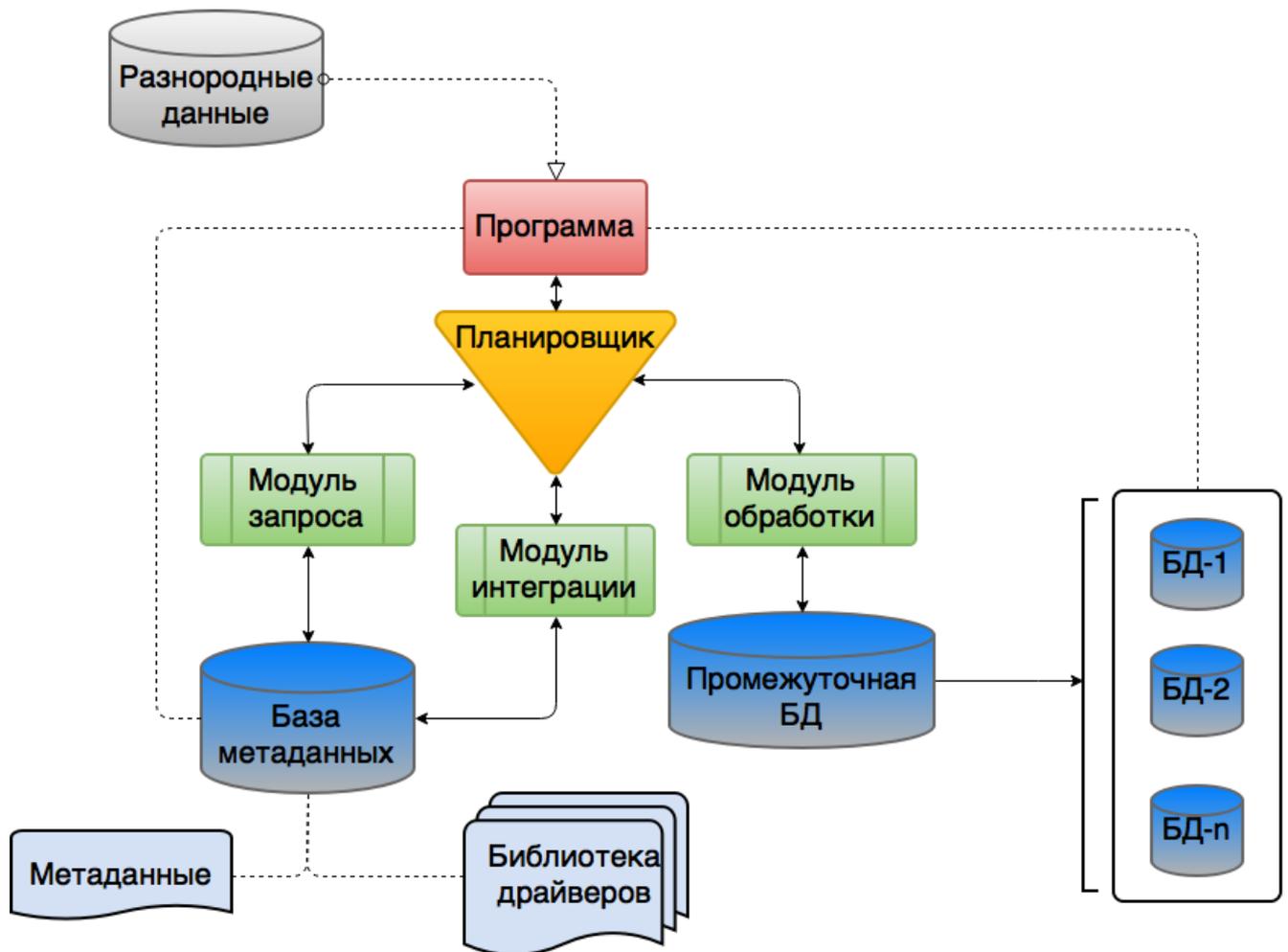


Рисунок 17. Общая схема модели обработки разнородных данных

Детальное представление схем модулей запросов, интеграции и обработки разнородных данных, а так же некоторые алгоритмы оптимизации определенных операций, которые были использованы в рамках данной научной работы представлены на рисунках 18-22.

В новой архитектуре модуль запросов представляет собой взаимосвязь базы метаданных, которая хранит драйверы доступа к разнородным данным с модулем интеграции разнородных данных. Задача модуля запросов сравнить расширение подгружаемого файла с расширениями в таблице.

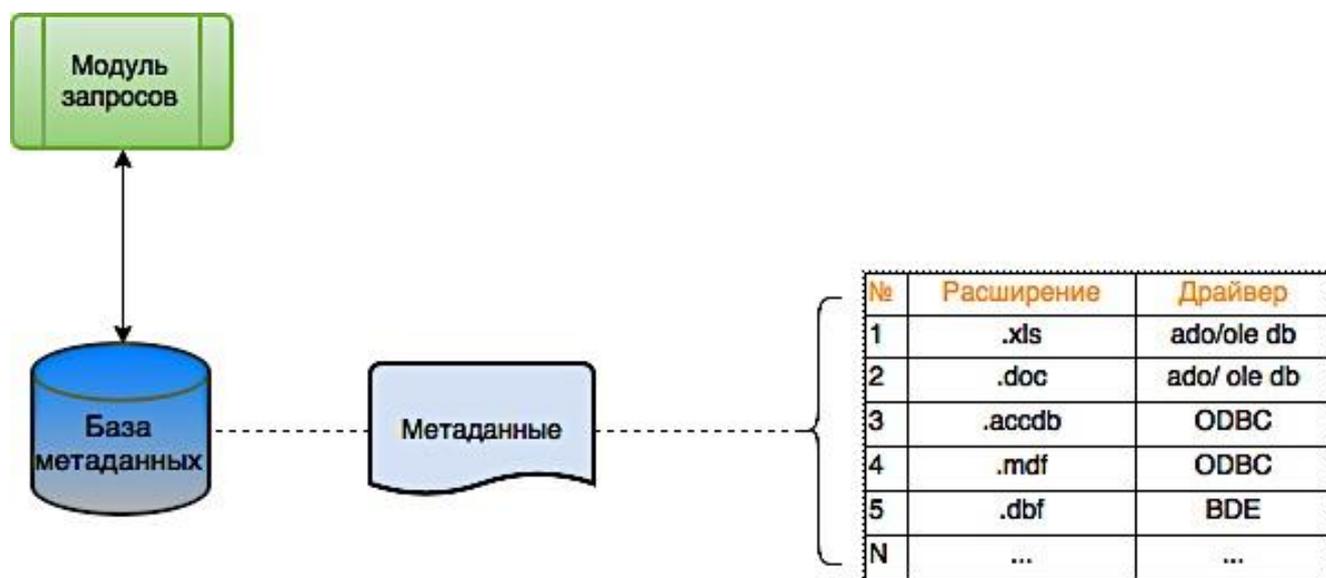


Рисунок 18. Архитектура модуля запроса

Если в таблице нет похожего (точного совпадения) расширения, тогда модуль отправляет данную информацию в планировщик. Планировщик оповещает систему сообщением («Формат не распознан!», «Найти расширение в интернете?»). Система продолжает работать в штатном режиме. Если в таблице есть похожее (точное совпадение) расширение, тогда модуль отправляет данную информацию в планировщик (сохраняет формат в память). Планировщик оповещает систему сообщением («Формат распознан!»). Планировщик отправляет информацию в модуль интеграции.

Для обеспечения доступа к базам данных используются различные технологии, набор библиотек, посредством которых происходит взаимодействие с базой данных. Для предоставления доступа к неопознанному заранее формату данных приложение-клиент, которое работает с базами, должно иметь в своем составе ряд библиотек, позволяющих подключиться к загружаемой базе.

Ниже представлена схема модели интеграции разнородных данных (рисунок 19). Процедура подключение нового источника выполняется экспертом. Вся дальнейшая работа системы (запросы, загрузка данных в хранилище) производится с учетом алгоритма обработки данных.

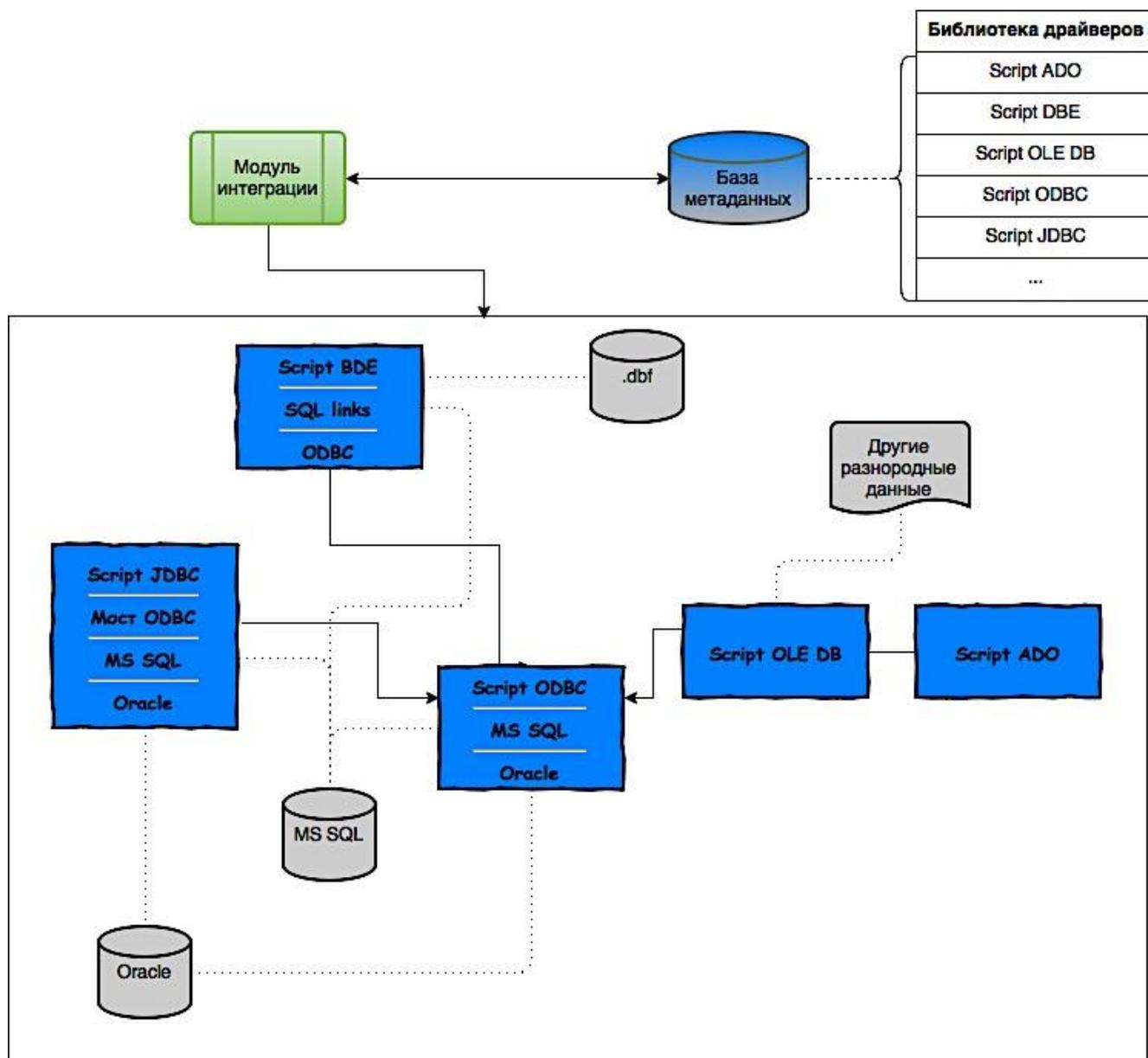


Рисунок 19. Архитектура модуля интеграции.

Особенность данного модуля интеграции заключается в сведении известных драйверов доступа к базам данных в единую автоматизированную информационную систему интегрирования разнородных источников данных. Основная идея такого подхода позволяет минимизировать участие человека в процедуре интеграции разнородных данных. Каждому расширению соответствует свой интерфейс обработки данных.

Если драйвер не найден, модуль отправляет информацию в планировщик, планировщик оповещает систему сообщением («Искомый драйвер не найден»), система работает в штатном режиме. Если драйвер найден, модуль интеграции запускает алгоритм соединения с данными. Если неудачное соединение, модуль отправляет данную информацию в планировщик

(«Соединение не установлено!»). Если удачное соединение, модуль отправляет планировщику информацию («Подключение данных»).

Архитектура модуля обработки данных показана на рисунке 20.

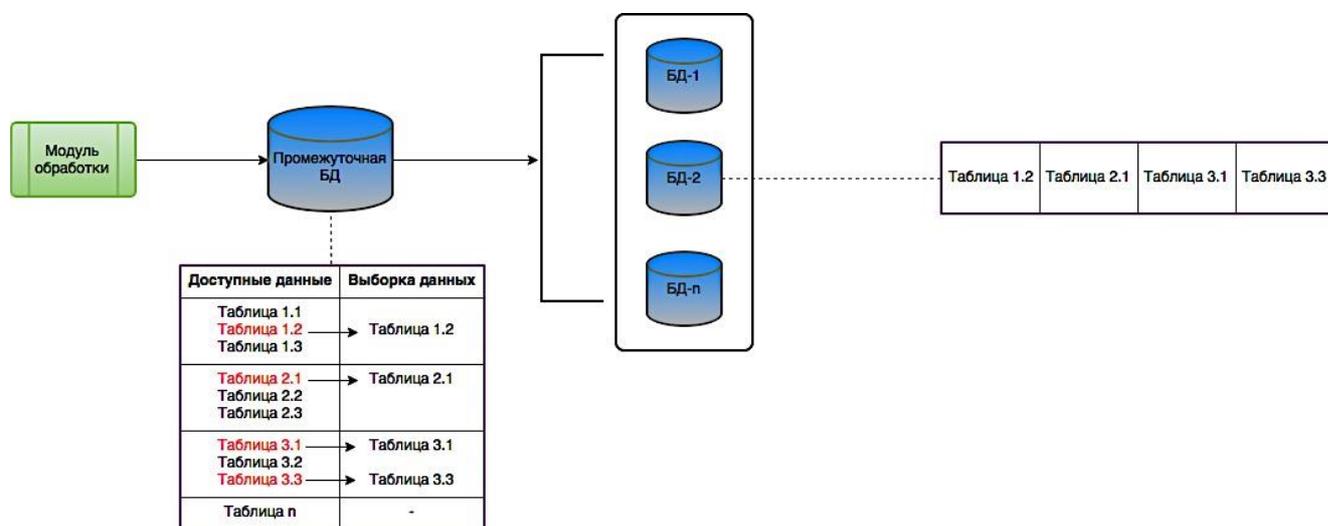


Рисунок 20. Архитектура модуля обработки данных.

Модуль обработки позволяет выбрать из всех имеющихся разнородных массивов данных только необходимые для дальнейшей обработки поля таблицы данных. Автоматизация процесса обработки разнородных данных позволяет получить новую базу данных. Выборка данных осуществляется по заранее заданным оператором признакам.

Скорость обработки данных зависит от следующего ряда возможных причин:

1. Формат разнородных данных.
2. Организация (структура) данных.
3. Аппаратное обеспечение (процессор, память, диски, сеть).
4. Усложненный алгоритм обработки данных (количество транзакций и запросов к базе данных).
5. Объем данных.
6. Конфигурация компьютера.
7. Свободное место на диске.
8. Программное окружение (операционная система и программные продукты).
9. Канал связи — технические средства, позволяющие осуществлять передачу данных на расстоянии.
10. Скорость передачи информации (скорость информационного потока) — количество информации, передаваемое за единицу времени.

11. Пропускная способность канала — максимальная скорость передачи информации по каналу связи в единицу времени.

В исследовании не учитываются все вышеописанные причины снижения скорости обработки данных. Учет производится только исходя из объема обрабатываемой информации и пропускной способности канала.

Общая схема передачи информации включает в себя отправителя информации, канал передачи информации и получателя информации. Основной характеристикой каналов передачи информации является их пропускная способность. Объем переданной информации V вычисляется по формуле (1.1):

$$V = qt, \quad (1.1)$$

где:

V – объем переданной информации;

q – пропускная способность канала (в битах в секунду или подобных единицах), пропускная способность канала равна количеству информации, которое может передаваться по нему в единицу времени.

t – время передачи информации (формула общего вида 1.2).

$$t = \frac{V}{q}, \quad (1.2)$$

В ряде случаев возможно наличие, как нескольких каналов связи, так и одного.

Первый случай, когда имеются множественные каналы обработки информации (1.3).

$$t_{\text{общ}} = \max_{i=1}^m \frac{V_i}{q_i}, \quad (1.3)$$

где:

V_i – источники информации (специфический формат);

q_i – множественные каналы связи;

$t_{\text{общ}}$ – общее время прохождения информации.

Второй случай, когда имеется один канал обработки информации (1.4).

$$t_{\text{общ}} = \frac{\sum_{i=1}^m V_i}{q}, \quad (1.4)$$

где:

V_i – источники информации (специфический формат);

q – канал связи;

$t_{\text{общ}}$ – общее время прохождения информации.

Чем больше обрабатывается данных из разнородных источников, тем длиннее становится очередь прохождения информации через канал, одновременно с этим увеличивается и время обработки данных.

Весь объем разнородных данных (массив информации) можно вычислить по следующей формуле (1.5):

$$V_i = \sum_{j=1}^n \alpha_{ji}, \quad (1.5)$$

где:

n – длина поля/записи (элемент матрицы);

$\sum_{j=1}^n \alpha_{ji}$ - сумма всех элементов массива данных;

V_i — объем переданной информации, n -го источника данных.

Для дальнейшего исследования нет необходимости использовать весь объем фактической информации V_i , n -го источника данных. Нам необходима только искомая информация для решения прикладной задачи. В следствии чего необходимо произвести фильтрацию данных.

Для удаления информации необходимо произвести множественную операцию по исключению некоторых рядов данных. Первоначально ставиться задача определения признака, по которому будут удалены ненужные для дальнейшего исследования ряды данных.

Математическая обработка данных задается следующим условием (1.6, 1.7):

$$t_{\text{общ}}^* = \frac{\sum_{i=1}^m V_i^*}{q}, \quad (1.6)$$

$$V_i^* = \min \sum_{i=1}^n \alpha_{ij}, \quad (1.7)$$

при условии, что $\alpha_{ij} = \beta_{ij}$,

где:

V_i^* – минимальное значение на участке;

α_{ij} – признаки искомой информация;

β_{ij} – критерий определения признаков искомой информации.

Этапы фильтрации данных:

1. Задается условие/критерий определения признаков искомой информации - β_{ij} . Критерий β_{ij} задается оператором системы заранее, исходя из установленных целей прикладной задачи. В дальнейшем, критерий будет заложен в базу знаний для подстановки в схожих вариантах обработки разнородных данных, при решении ряда прикладных задач.
2. Производится фильтрация данных. Фильтрация данных осуществляется по заранее установленному условию/признаку β_{ij} .
 - если $\alpha_{ij} \neq \beta_{ij}$, то условие считается не выполненным и элементам таблицы присваивается нулевой идентификатор (в результате операции сопоставления отдельные элементы $\alpha_{ij} = 0$), соответственно данные элементы будут удалены/исключены из таблицы;
 - если $\alpha_{ij} = \beta_{ij}$, то условие считается выполненным и данные элементы будут определены как искомые значения и продолжат дальнейшую обработку данных.

При условии, что объем данных будет уменьшен, в зависимости от выбранной информации для дальнейшего исследования, то время обработки информации сократиться. Расчет производится по формуле (1.8):

$$t = \frac{\sum_{i=1}^n V_i}{q}, \quad (1.8)$$

В результате автоматизации поток информации сокращается. Отобрав только необходимые для дальнейшей обработки поля таблицы данных, происходит сокращение объема обрабатываемой информации, что позволит значительно сократить время обработки данных (рисунок 21).

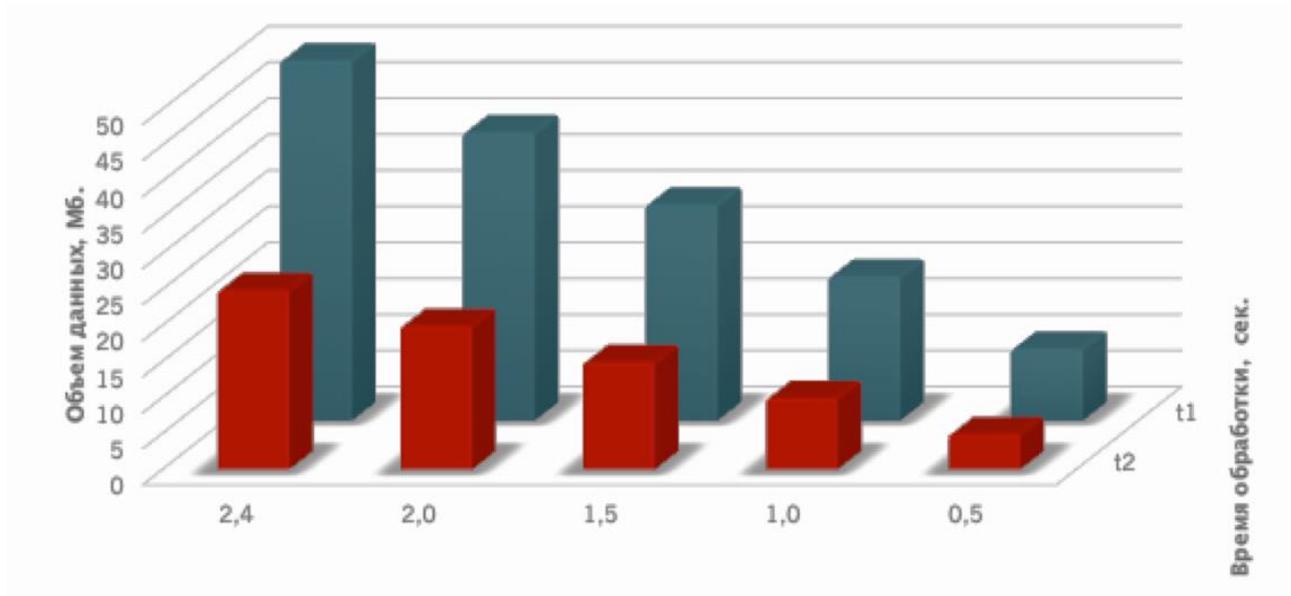


Рисунок 21. Сокращение объема обрабатываемой информации

В результате проведенного исследования по сокращению объема обрабатываемой информации, был сформирован алгоритм обработки с учетом фильтрации данных (рисунок 22).

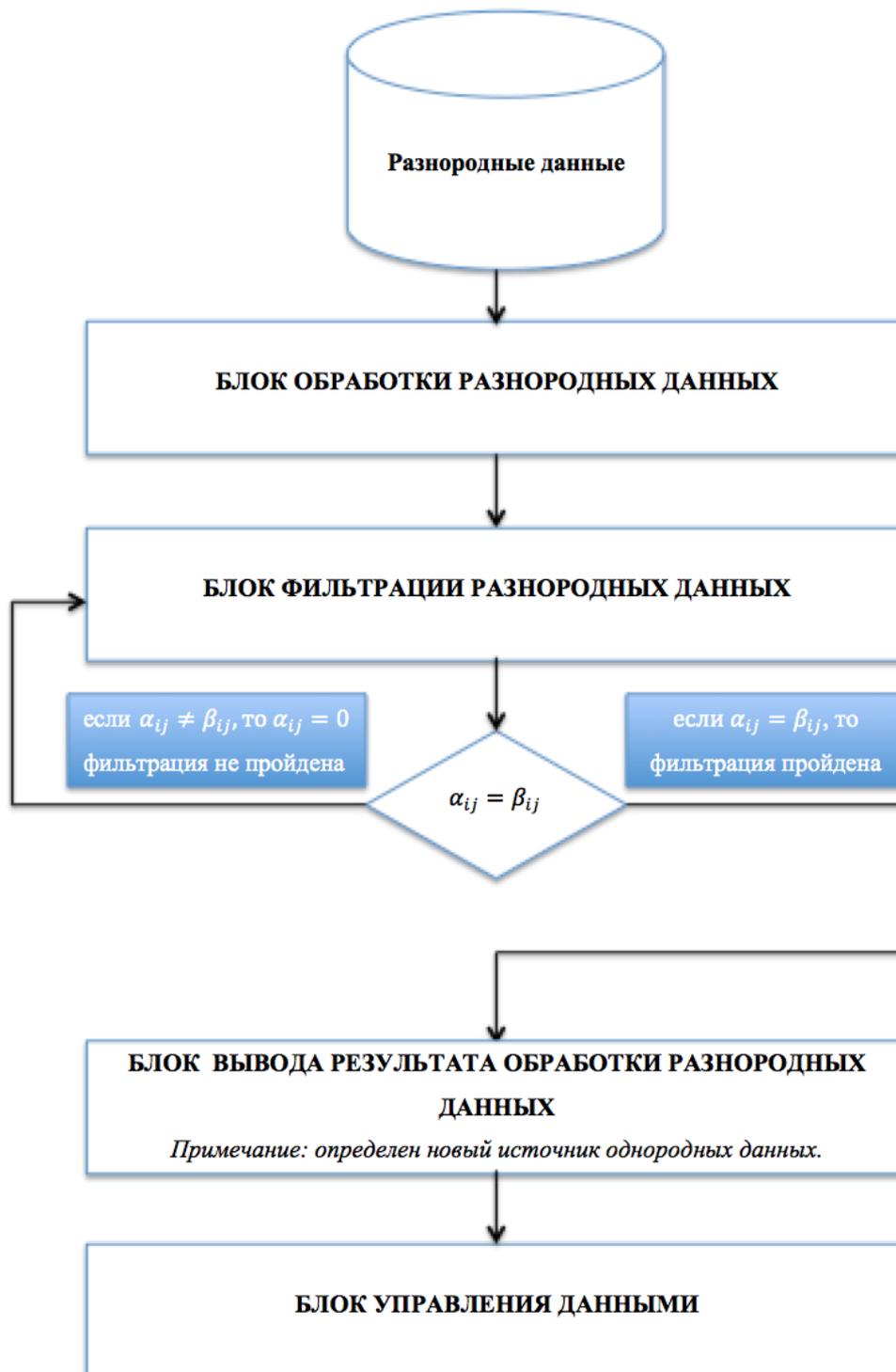


Рисунок 22. Алгоритм обработки разнородных данных

После завершения выполнения алгоритма интегрированные данные становятся доступными пользователю системы для дальнейшего комплексного анализа информации.

Для использования усовершенствованной модели для управления разнородными данными необходимо разработать специализированное программное обеспечение,

позволяющее получать расширенную информацию и проводить более тщательный анализ полученных статистических данных.

Вывод по главе

Решение задач интегрирования гетерогенной геопространственной информации в различных предметных областях потребует адаптации предложенной модели, как с точки зрения используемых данных, так и с точки зрения функциональности сервисов интегрирования. Предложенная в данной работе общая структура модели интегрирования может быть взята за основу при создании подобных моделей в других областях применения пространственных данных, а также позволит упростить доступ к разнородной геопространственной информации.

В главе были рассмотрены современные проблемы создания геоинформационной системы, осложненной разнородностью данных, так же представлена разработанная модель обработки разнородных данных, охватывающая все слои представления информации. Разработаны принципы и средства, способные эффективно синхронизировать доступ к пространственным геоданным в различных форматах [81, 110].

- Модель описывает оптимальное решение, которое позволяет поддерживать широко известные и часто используемые интерфейсы доступа к данным, такой подход позволит обрабатывать данные не только из структурированных источников, поддерживающих свой собственный язык запросов и определённые модели представления данных, но и из неструктурированных источников.
- Модель обработки разнородных данных, позволит сократить время обработки геоданных, за счет сокращения обрабатываемой информации путем встроенных процедур поиска критериев искомым данным.
- Разработанная модель, позволяет эффективно использовать доступ к пространственным геоданным в различных форматах без непосредственного участия всех данных.
- Предложенная в данной работе общая структура модели интегрирования может быть взята за основу при создании подобных моделей в других областях применения пространственных данных, а также позволит упростить доступ к разнородной геопространственной информации.

Данная модель рассматривается как основа для реализации системы, поддерживающей оперативное управление разнородными данными и их аналитическую обработку.

2. Разработка методики управления разнородными пространственно-распределенными данными в геоинформационной системе на основе метеорологической информации

Экономическая и социальная деятельность человека находится в постоянной взаимосвязи с природными, в том числе гидрометеорологическими условиями. В течение последних десятилетий во всех странах мира, в том числе и в России, отмечаются нарастающие экономические и социальные потери от неблагоприятных проявлений погоды и климата [14]. Современное производство становится все более подвержено влиянию погодных условий и это влияние фактически эквивалентно повышению средних и предельных издержек производства [103], что приводит к повышению себестоимости продукции и снижению прибыли.

Поскольку прогнозы погоды объективно обладают ограниченной надежностью, возникают риски принятия неточных управленческих решений, имеющих нежелательные экономические последствия. Большинство экономических субъектов не только не владеют правилами принятия решений в условиях погодной (частичной) неопределенности, но и в большинстве случаев, не располагают необходимыми для этой цели данными [96].

Эффективное использование информации об ожидаемом состоянии погодных условий позволяет адаптировать хозяйственную деятельность к неблагоприятным проявлениям гидрометеорологической среды и тем самым снижать потери от неблагоприятных и опасных условий погоды. Оптимальная хозяйственная деятельность должна осуществляться с обязательным использованием программного инструментария, учитывающего неопределенность реализации информационных гидрометеорологических ресурсов.

Для эффективного принятия управленческих решений в области ресурсосбережения требуется привлечение большого массива разнородной пространственно-распределенной метеорологической информации, что существенно осложняет обработку данных, проведение анализа и дальнейшее предоставление вариантов оптимального управленческого решения.

В рамках исследования диссертационной работы будет произведена обработка разнородных массивов пространственных метеорологических данных с использованием модели обработки разнородных данных, что позволит в свою очередь автоматизировать процесс по средствам предлагаемых в работе решений и произвести комплексный анализ разнородной информации.

2.1. Алгоритм управления разнородными метеорологическими данными для подготовки варианта оптимального управленческого решения

В настоящее время прогностическая гидрометеорологическая информация может и должна применяться в процессах оперативного предупреждения, управления и предотвращения метеорологических потерь. Несмотря на то, что гидрометеорологическая служба имеет возможность предоставлять потребителям (лицам, принимающим экономические решения на основе информации об ожидаемом состоянии атмосферы) прогностическую информацию в виде любого числа градаций, отражающих влияние метеорологических условий на хозяйственную (производственную) деятельность, большинство из них игнорируют сложную погодозависимость и используют прогнозы в виде простой альтернативы действий: «применять меры защиты, если ожидается опасная погода» или «выполнять производственные операции, если ожидается благоприятная погода». Данный (альтернативный) регламент хотя и позволяет получить некоторые экономические результаты, но соответствует нижнему пределу потенциальной полезности краткосрочных гидрометеорологических прогнозов.

Необходимо использовать научный подход в управлении экономико-метеорологическими рисками. Оценить риски возможно, если имеется два вида информации: информация о событии (неблагоприятном условии погоды) которое вызывает риск, а также информация об экономических последствиях от возможных проявлений риска.

Современная теория принятия индивидуальных решений построена на модели ожидаемой полезности, в которой одним из ключевых компонентов является присутствие случайного фактора (неопределенности относительно будущих событий). Погодохозяйственное решение, учитывающее влияние погодных факторов на экономическую деятельность, содержит неопределенность в отношении реализации состояния погоды, на которую оно было ориентировано. Потребитель должен располагать набором погодохозяйственных решений (действий), дифференцированных в зависимости от интенсивности I и продолжительности μ воздействующего метеорологического фактора.

Анализ погодных условий, оказывающих влияние на производственную деятельность, позволяет ранжировать степень их опасности и определить пороговые значения метеорологических величин, при которых необходима дифференциация защитных мер. Полагается, что в случае правильного прогноза ожидаемого уровня опасности погодных условий и применения мер защиты, размер предотвращаемого ущерба должен быть больше издержек на проведение защитных мероприятий.

Применение многофазового погодо-хозяйственного регламента [93, 96] позволит пользователю гидрометеорологической информации эффективно управлять экономико-метеорологическими рисками и предоставит выбор оперативных решений с учетом вероятности осуществления ожидаемых погодных условий [16, 17].

В ходе исследования была поставлена задача разработки методики автоматизации подготовки варианта управленческого решения на основе управления разнородными пространственно-распределенными метеорологическими данными. На рисунке 23 изображена общая схема управления метеорологическими данными для составления вариантов оптимального управленческого решения.

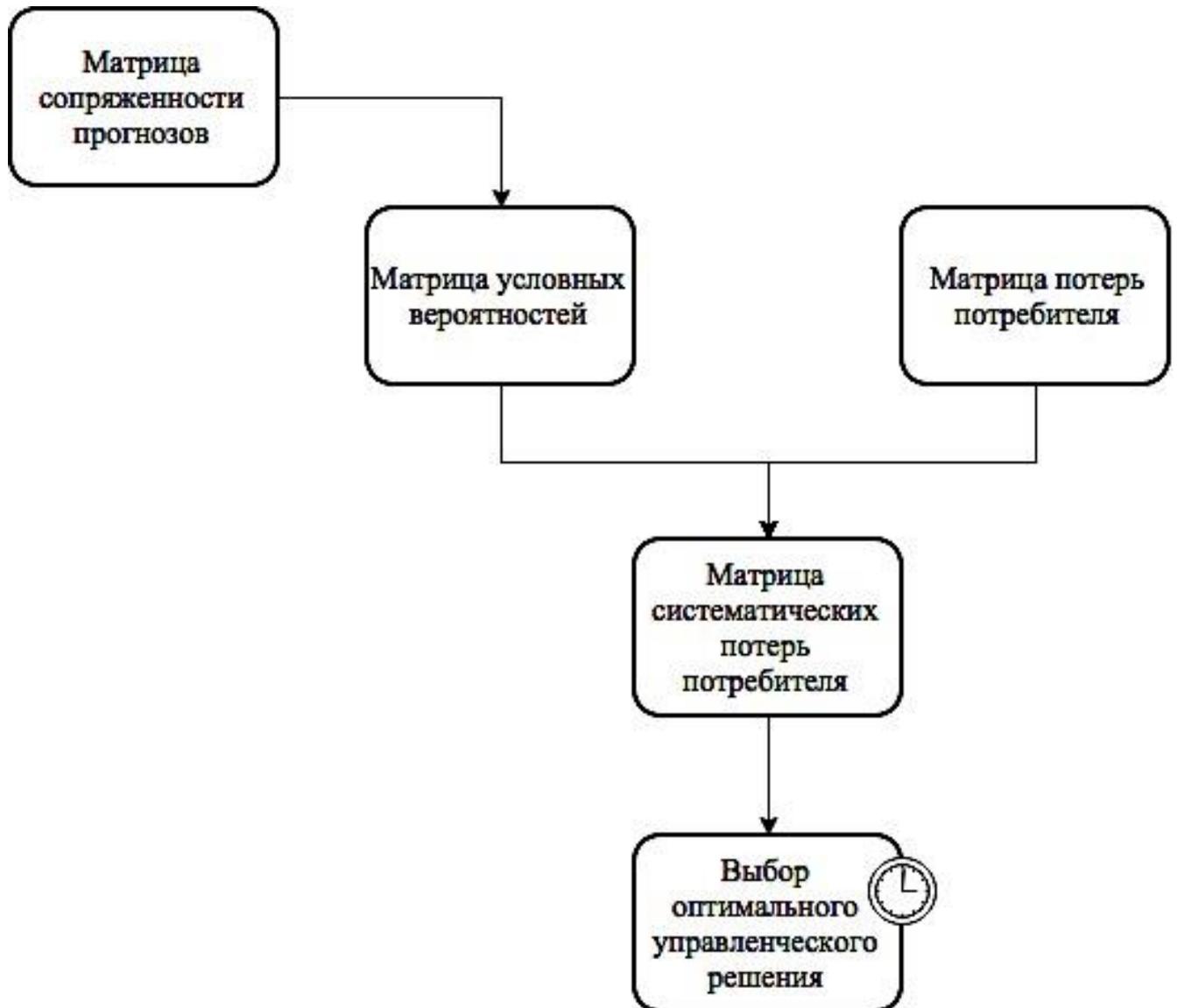


Рисунок 23. Погодо-хозяйственный регламент

Данный регламент требует привлечения метеорологических данных прогноза погоды и фактических значений погодных условий. Данный вид информации зачастую хранится в различных базах данных, что затрудняет обработку и проведение комплексного анализа данных. Схема выбора оптимального управленческого решения дополнена разнородными пространственно-распределёнными метеорологическими данными (рисунок 24) [85].

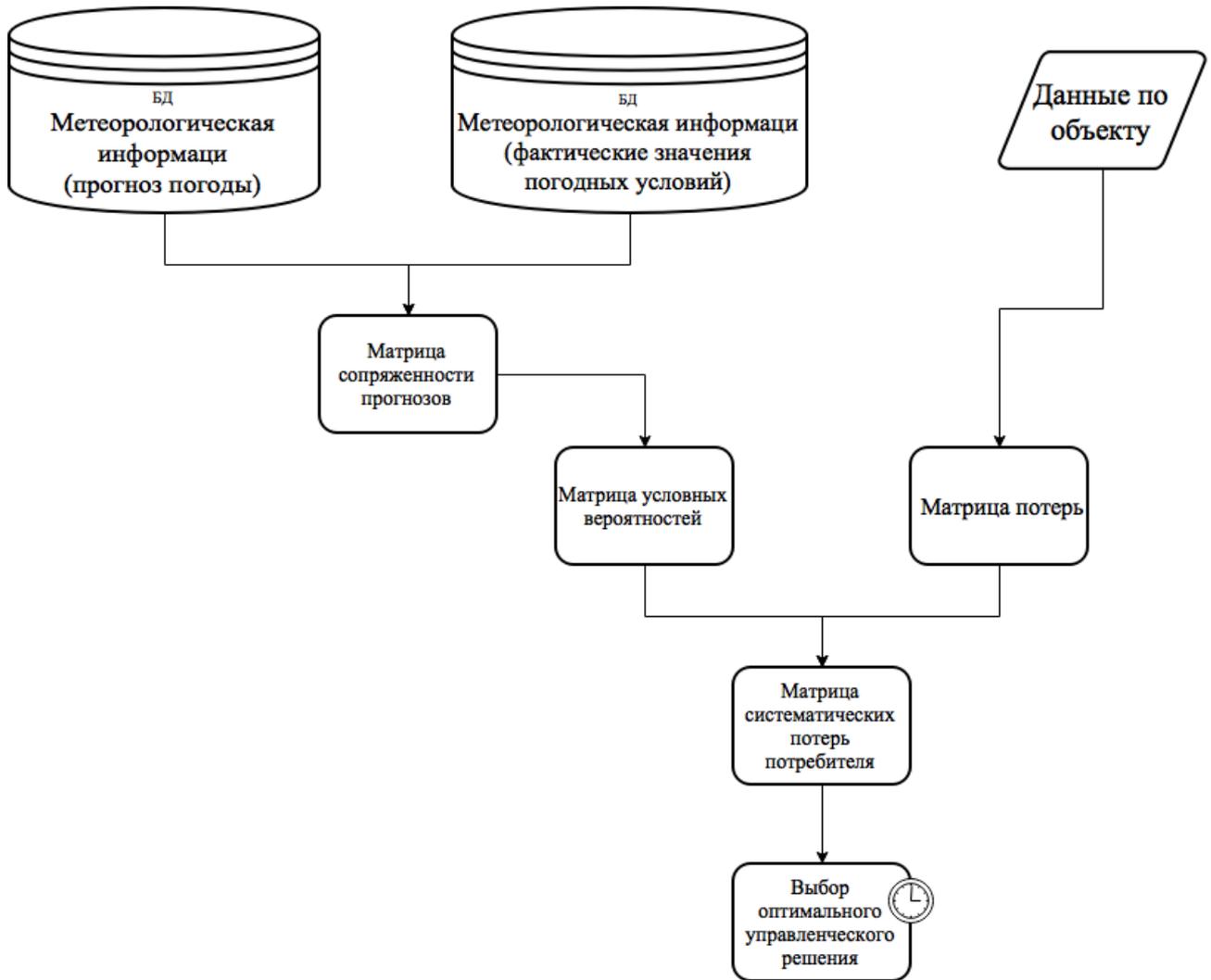


Рисунок 24. Схема управления разнородными метеорологическими данными

Ниже представлено описание основных этапов оптимизации использования метеорологической информации.

ШАГ 1: Описание погодозависимости

В соответствии со спецификой погодозависимости экономической деятельности, пользователем выбираются значения погодных условий, при которых производственные процессы отличаются от процессов, проводимых при благоприятных условиях погоды.

1.1 Пользователем выбирается вид влияющих метеорологических условий (осадки, температура воздуха, ветер, видимость и т.д.)

1.2 Пользователем задается дискретность метеорологической информации, используемой при принятии решений :

- количество градаций метеорологических условий (n);
- минимальные и максимальные значения метеорологического параметра в пределах каждой градации (ширина градации).

1.3 Пользователем вводятся значения издержек (доходов), обусловленных ожиданием и реализацией погодных условий. (примечание: возможно подключение отдельных модулей, позволяющих определить функцию погодозависимости пользователя - информацию для п. 1.1, 1.2, 1.3)

Введенные значения преобразуются в матрицу потерь (доходов). Общий вид матрицы потерь приведен в таблице 3.

Таблица 3 – Многофазовая матрица потерь $\|s_{ij}\|$

Фактическая погода, Φ_i	Действия потребителя d_j , в ожидании погоды Π_j						
	d_1	d_2	d_3	...	d_j	...	d_m
Φ_1	s_{11}	s_{12}	s_{13}	...	s_{1j}	...	s_{1m}
Φ_2	s_{21}	s_{22}	s_{23}	...	s_{2j}	...	s_{2m}
Φ_3	s_{31}	s_{32}	s_{33}	...	s_{3j}	...	s_{3m}
...
Φ_i	s_{i1}	s_{i2}	s_{i3}	...	s_{ij}	...	s_{im}
...
Φ_n	s_{n1}	s_{n2}	s_{n3}	...	s_{nj}	...	s_{nm}

По главной диагонали расположены потери потребителя при правильном прогнозе (прогнозируемая интенсивность и продолжительность гидрометеорологических условий совпадает с фактически осуществившейся). Слева от главной диагонали расположены потери соответствующие ошибкам–пропускам (интенсивность и продолжительность фактически наблюдавшихся гидрометеорологических условий больше чем ожидаемых). Справа от главной диагонали - потери соответствующие ошибкам страховкам (наблюдались более благоприятные условия погоды, чем прогнозировались).

ШАГ 2: Составление матрицы сопряженности метеорологических прогнозов

Примечание: программа содержит 2 базы данных с метеорологической информацией.

- База данных «Фактические значения гидрометеорологических величин и явлений погоды» - массив фактических значений гидрометеорологических условий, с привязкой к территории и времени.
- База данных «Методические прогнозы гидрометеорологических условий» - содержит информацию о прогнозируемых значениях гидрометеорологических величин

2.1 Пользователь выбирает:

- населенный пункт, в котором расположен производственный объект;
- период, за который оценивается качество прогностической информации (с ... по....).

2.2 По параметрам, указанным в пункте 1.2 (выбранное количество и заданная ширина градаций метеорологических условий) из элементов баз данных «Прогнозы гидрометеорологических условий» и «Методические прогнозы гидрометеорологических условий» составляется матрица сопряженности метеорологических прогнозов (методических прогнозов).

Общий вид матрицы сопряженности приведен в таблице 4.

Таблица 4 - Многофазовая матрица сопряженности прогнозируемых P_j и фактических Φ_i условий погоды

Фактическая погода, Φ_i	Прогноз, P_j						$\sum_{j=1}^m n_j$
	P_1	P_2	P_3	...	P_j	P_m	
Φ_1	n_{11}	n_{12}	n_{13}	...	n_{1j}	n_{1m}	n_{10}
Φ_2	n_{21}	n_{22}	n_{23}	...	n_{2j}	n_{2m}	n_{20}

Продолжение таблицы 4 - Многофазовая матрица сопряженности прогнозируемых Π_j и фактических Φ_i условий погоды

Φ_3	n_{31}	n_{32}	n_{33}	...	n_{3j}	n_{3m}	n_{30}
...
Φ_i	n_{i1}	n_{i2}	n_{i3}	...	n_{ij}	n_{im}	n_{i0}
$\sum_{i=1}^n n_i$	n_{01}	n_{02}	n_{03}	...	n_{0j}	n_{0m}	N

где

n_{i0} - фактически наблюдавшееся число случаев выбранной фазы погоды Φ_i ;

n_{0j} - число прогнозов выбранной фазы погоды Π_j ;

n_{ij} - число сопряжений прогноза Π_j и фактической погоды Φ_i в пределах заданной градации;

N - общее число прогнозов.

2.3 На основании результатов матрицы сопряженности в пункте 2.2 проводится проверка наличия связи между исследуемыми признаками с использованием критерия независимости Пирсона (формула 2.1):

$$\chi^2 = \sum_{i=1}^n \sum_{j=1}^m \frac{(n_{ij} - n_{ij}^{сл})^2}{n_{ij}^{сл}}, \quad (2.1)$$

где

χ^2 - критерий независимости Пирсона;

n_{ij} - частоты матрицы сопряженности прогнозов;

$n_{ij}^{сл} = \frac{n_{i0} \cdot n_{0j}}{N}$ - распределение частот случайного прогноза, характеризующее независимость случайных величин, принятых в качестве нулевой гипотезы h_0 (независимость Π_j и Φ_i).

Этот критерий позволяет дать оценку статистической значимости выводов, получаемых по выборочным таблицам сопряженности.

Примечание:

- при положительном исходе проверки, проведенной в пункте 2.3, пользователь переходит к следующему этапу;

- при отрицательном исходе проверки, проведенной в пункте 2.3, пользователь направляется к пункту 2.1 для уточнения выбранных данных.

ШАГ 3 : Выбор оптимального управленческого решения

3.1 На основании пункта 2.2. (матрица сопряженности метеорологических прогнозов) рассчитываются условные вероятности q_{ij} осуществления фазы погоды Φ_i при известном тексте прогноза Π_j (т.е. вероятность, с которой можно ожидать неблагоприятное явление (условие) погоды, если об этом явлении уже имеется предварительная информация в виде прогноза):

$$q_{ij} = \frac{\left(\frac{n_{i0} \cdot n_{ij}}{N \cdot n_{i0}}\right)}{\frac{n_{0j}}{N}} = \frac{n_{ij}}{n_{0j}}, \quad (2.2)$$

где

$n_{ij}, n_{i0}, n_{0j}, N$ - частоты матрицы сопряженности прогнозов.

Общий вид матрицы условных вероятностей $\|q_{ij}\|$ приведен в таблице 5.

Таблица 5 - матрица условных вероятностей $\|q_{ij}\|$

Фактическая погода, Φ_i	Прогноз, Π_j						
	Π_1	Π_2	Π_3	...	Π_j	...	Π_m
q_{1j}	q_{11}	q_{12}	q_{13}	...	q_{1j}	...	q_{1m}
q_{2j}	q_{21}	q_{22}	q_{23}	...	q_{2j}	...	q_{2m}
q_{3j}	q_{31}	q_{32}	q_{33}	...	q_{3j}	...	q_{3m}
...
q_{ij}	q_{i1}	q_{i2}	q_{i3}	...	q_{ij}	...	q_{im}
...
q_{nj}	q_{n1}	q_{n2}	q_{n3}	...	q_{nj}	...	q_{nm}
$\sum_{i=1}^n q_i$	1	1	1	1	1	1	1

3.2 На основании матрицы потерь $\|s_{ij}\|$ и матрицы условных вероятностей $\|q_{ij}\|$ разрабатывается матрица систематических потерь $\|R_{kj}\|$, элементы которой представляют собой средние потери потребителя, при условии, что получив данный текст прогноза Π_j , он вправе выбрать ориентацию d_k на любую иную погоду Φ_i . Элементы матрицы систематических потерь $\|R_{kj}\|$ определяются согласно формуле:

$$\bar{R}_{kj} = \sum_{i=1}^{n=m} s_{ki} \cdot q_{ij}, \quad (2.3)$$

где s_{ki} – элементы матрицы потерь $\|s_{ij}\|$, отвечающие k -му решению (действию потребителя) в ожидании погоды Φ_i .

Величина \bar{R}_{kj} рассматривается как критерий выбора оптимального решения. Общий вид матрицы систематических потерь приведен в таблице 6.

Таблица 6 - Матрица систематических потерь потребителя \bar{R}_{kj}

Действия потребителя, d_k	Прогноз, Π_j				
	Π_1	Π_2	Π_3	...	Π_m
d_1	\bar{R}_{11}	\bar{R}_{12}	\bar{R}_{13}	...	\bar{R}_{1m}
d_2	\bar{R}_{21}	\bar{R}_{22}	\bar{R}_{23}	...	\bar{R}_{2m}
d_3	\bar{R}_{31}	\bar{R}_{32}	\bar{R}_{33}	...	\bar{R}_{3m}
...
d_k	\bar{R}_{k1}	\bar{R}_{k2}	\bar{R}_{k3}	...	\bar{R}_{km}

Матрица систематических потерь позволяет пользователю осуществлять выбор оперативных решений с учетом вероятности осуществления ожидаемых погодных условий.

3.3 Пользователь вводит прогноз метеорологических условий, и получает рекомендацию по выбору оптимального оперативного решения $d_{\text{копт}}$ с учетом неопределенности реализации текста прогноза. При известном прогнозе уровня опасности погодных условий Π_j оптимальное решение $d_{\text{копт}}$ обеспечивает минимум \bar{R}_{kj} :

$$d_{\text{копт}} \Leftrightarrow \min \bar{R}_{kj} (\Pi_j), \quad (2.4)$$

Шаг 4 : Оценка экономической полезности метеорологических прогнозов

4.1 Пользователь выбирает:

- период, за который необходимо оценить экономическую полезность метеорологической информации;
- вводит величину стоимости единицы прогностической информации ($Z_{пп}$ или C).

4.2 За выбранный период разрабатываются матрицы сопряженности методических и инерционных прогнозов погоды (согласно пункту 2.2, для разработки матрицы сопряженности инерционных прогнозов используется только база данных фактических значений метеорологических условий).

4.3 Разрабатываются матрицы систематических потерь потребителя при использовании методических и инерционных прогнозов (выполняются операции, предусмотренные пунктами 3.1, 3.2)

4.4 Определяются средние байесовские потери при условии доверия текстам методических и инерционных прогнозов (формула 2.5):

$$\bar{R} = \sum_{j=1}^m p_{0j} \bar{R}_{k=j}(\Pi_j), \quad (2.5)$$

4.5 Определяются средние байесовские потери при условии выбора оптимальных решений (только по данным методического прогноза), (формула 2.6):

$$\bar{R}_{\text{опт}} = \sum_{j=1}^m p_{0j} \min \bar{R}_{kj}(\Pi_j), \quad (2.6)$$

4.6 Оценка экономического эффекта пользователя проводится по следующим формулам:

- Экономический эффект, получаемый за счет преимущества методических прогнозов относительно инерционных (формула 2.7):

$$\mathcal{E} = N(\bar{R}_{\text{ин}} - \bar{R}_{\text{мет}} - Z_{\text{пп}}), \quad (2.7)$$

- Прирост экономического эффекта, получаемый за счет оптимального использования методических прогнозов (формула 2.8):

$$\mathcal{E} = N(\bar{R}_{\text{мет}} - \bar{R}_{\text{опт}} - Z_{\text{пп}}), \quad (2.8)$$

- Народохозяйственный экономический эффект (формула 2.9):

$$\mathcal{E} = \beta \cdot N(\bar{R}_{\text{ин}} - \bar{R}_{\text{мет}} - C), \quad (2.9)$$

где

\mathcal{E} - экономический эффект;

β - коэффициент долевого участия прогностического подразделения в получении экономического эффекта использования прогнозов;

$Z_{\text{ин}}$ - стоимость единицы прогностической информации;

N - общее число прогнозов за период;

$\bar{R}_{\text{ин}} - \bar{R}_{\text{м}}$ - снижение потерь потребителя за счет использования методических прогнозов.

На основе этапов оптимизации использования метеорологической информации разработан алгоритм автоматизации подготовки варианта управленческого решения с использованием модели обработки разнородных пространственно-распределенных данных (рисунок 25).

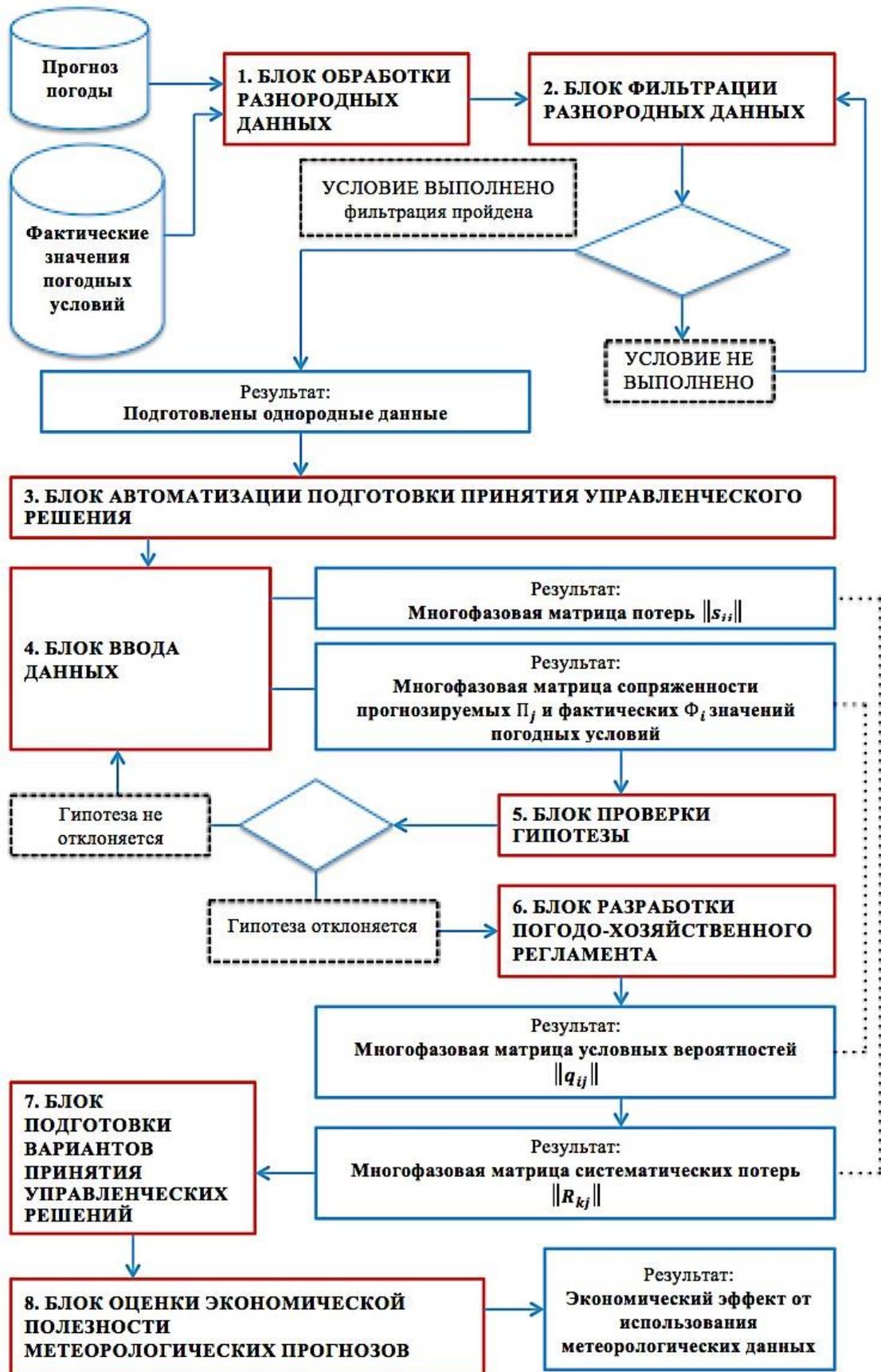


Рисунок 25. Алгоритм автоматизации управления данными для поддержки принятия решений

Алгоритм позволяет использовать автоматизацию обработки разнородных метеорологических данных для подготовки варианта управленческого решения по управлению территорией.

Прогностическая информация, как и фактические значения погодных условий зачастую не доступны потребителю в единообразном виде, данные могут быть представлены в различных форматах, использовать различные структуры построения данных, содержать различные наборы данных.

Исходя из определенных проблем при обработке разнородных данных целесообразно использовать представленный алгоритм решений для поддержки потребителей гидрометеорологической информации в доступном виде, оптимальным вариантом использования данного алгоритма является его включение в разработку геоинформационной системы поддержки принятия управленческих решений, что позволит автоматизировать подготовку варианта оптимального управленческого решения с учетом обработки разнородных пространственно-распределенных метеорологических данных.

Данный алгоритм можно применять к различным прикладным задачам, требующим учет метеорологической информации. В ходе исследований была поставлена задача апробации данной методики управления метеорологическими данными для решения задачи оптимизации отпуска тепла на теплоэлектроцентралях (ТЭЦ).

2.2. Использование методики обработки разнородной метеорологической информации на примере оптимизации отпуска тепла ТЭЦ

Климат в целом и его реализация в погодных условиях существенно определяют различия погодозависимости теплоэнергетики. Теплоэнергетика как специфическая отрасль экономики и особая инфраструктура хозяйственной деятельности зависит от погоды-климатического фактора. Учет таких информационных природных ресурсов как прогноз температуры воздуха и скорости ветра занимает в теплоэнергетике ведущую роль. Адаптация под ожидаемую погоду определяет эффективность теплоисточника с точки зрения ресурс потребления, а значит и ресурсосбережения.

В энергетическом балансе страны расходы топлива на теплоэлектроцентралях постоянно растут. Ускоренное развитие промышленности в будущем потребует еще больше энергоресурсов. Следовательно, реализация энергосберегающей политики в системе теплоснабжения страны может дать ощутимый экономический эффект. Экономия ресурсов

ТЭЦ зависит в большей части от прогнозов погоды или иной метеорологической информации. Внедрение научных методов использования метеорологической информации и, прежде всего, прогнозов погоды позволяет значительно снизить издержки в экономике страны за счет влияния погодных условий, получить наибольший эффект при минимальных затратах [32].

В системе централизованного теплоснабжения большое распространение получили теплоэлектроцентрали (ТЭЦ) - предприятия по комбинированной выработке электроэнергии и тепла. Технологически ТЭЦ ориентированы на приоритет электроснабжения, попутно производимое тепло востребовано в большей степени в холодный период года. Влияние метеорологических факторов на производство тепловой энергии прослеживается на всех этапах функционирования ТЭЦ и сказывается непосредственно на задании режима работы теплоисточника, на колебаниях спроса на тепловую энергию, а также на условия эксплуатации и содержания тепловых сетей.

Использование в теплоэнергетике метеорологических прогнозов, рассматривается, прежде всего, как привлечение метеорологических (погодных) ресурсов в целях минимизации тепловых потерь, а значит – экономии топливных ресурсов. В работе приведены методические основы оценки экономической полезности прогнозов температуры в теплоэнергетике.

Реализация методических прогнозов решает ряд производственных задач, в частности - регулировать рабочую мощность теплоисточника, поддерживать необходимую температуру теплоносителя, но главное обеспечить оптимальную экономию топлива, все это достигается за счет выбора оптимального управленческого решения.

Это означает, что, с одной стороны, постоянно адаптироваться к ожидаемым, нередко значительным колебаниям температуры воздуха и скорости ветра и обеспечить температуру в жилых помещениях не ниже $18\text{ }^{\circ}\text{C}$, а с другой - не израсходовать лишнего топлива при ожидаемых понижениях наружной температуры. Оптимизация использования прогнозов температуры воздуха предполагает минимизацию потерь, которые возможны в непрерывном функционировании ТЭЦ. Эти потери, связанные как с перерасходом топлива (ситуации избыточного теплоснабжения), так и в случаях недостаточного теплоснабжения.

Значительная неоднородность климатических и погодных условий проявляется в выборе известного в теплоэнергетике показателя расчетной температуры наружного воздуха t_p . Его значение используется при определении максимальной нагрузки теплоисточника Q_p^t в расчетах теплопередачи ограждающих конструкций и в решении ряда других задач.

Для задания соответствующей нагрузки теплоисточника, устанавливается необходимое (в соответствии с прогнозом температуры наружного воздуха) значение температуры теплоносителя.

Наблюдаемые отклонения от температурного графика вызваны такими факторами, как неточность контроля расхода тепла в теплосистеме, а также возможными ошибками в прогностической информации, поступающей на ТЭЦ. Ориентируясь на ожидаемую среднесуточную температуру воздуха, устанавливается такой тепловой режим выдачи прямой воды t' , поступающей от ТЭЦ в систему теплоснабжения, при котором потребители должны быть обеспечены необходимым теплом. В действительности идеальная подстройка невозможна в связи с большими расстояниями в системе передачи (транспортировки) тепла [45, 67, 73, 76].

Для транспортировки исходного теплового состояния теплоносителя требуются многие часы работы теплоисточника (и тепловой сети), чтобы донести тепло до «крайних» потребителей. Тепловая сеть любой ТЭЦ, обеспечивающая теплом крупный район города, может иметь протяженность десятки километров. Чем больше расстояние трубопроводной передачи тепла, тем больше времени требуется для его «прокачки» в теплосети. Поэтому в крупных городах, а в мегаполисах в особенности, как правило, используются прогнозы среднесуточной температуры воздуха $\bar{t}_{пр}$.

Конечно, ожидаемые значения температуры воздуха $\bar{t}_{пр}$, на которые должен ориентироваться диспетчер ТЭЦ, не всегда будут совпадать с фактической температурой воздуха. Регулирование отопления сопровождается ошибками, вызывающими перерасход или недодачу тепла, что в свою очередь является следствием ошибочности оперативных прогнозов среднесуточной температуры воздуха, на которые потребитель постоянно ориентируется. Следует заметить, что правильный прогноз резких изменений температуры воздуха позволяет избежать крупных потерь перерасхода или недодачи тепла [86, 87].

2.2.1 Разработка функции потерь для теплоэнергетики с использованием модели управления разнородными данными

Отмечено, что у потребителя (ТЭЦ) имеется возможность задания теплового режима в зависимости от масштабов тепловой сети на отрезки времени 6, 12, или 24 часа. При этом, температура прямой воды ($t'_в$), поступающей в систему теплоснабжения, задается в соответствии с температурным графиком ТЭЦ, отражающим расход тепла Q (Гкал/ч) в зависимости от ожидаемой температуры наружного воздуха (t). Необходимый при этом расход тепла определяется следующим образом:

$$Q = Q_p^T \frac{t_n - t}{t_n - t_p}, \quad (2.10)$$

где Q_p^T - мощность теплоисточника (ТЭЦ) – тепловая нагрузка, отвечающая расчетной температуре воздуха t_p ;

t_n - температура воздуха, необходимая в отапливаемых зданиях и принимаемая равной 18 °С;

t_p - расчетная температура наружного воздуха (средняя из восьми самых холодных зим за последний 50-летний период наблюдений в данном пункте; возможно использование формулы Чаплина:

$$t_p = 0.4t_{cx} + 0.6t_{ам}, \quad (2.11)$$

где

t_{cx} - средняя температура воздуха за самый холодный месяц;

$t_{ам}$ - абсолютный минимум температуры воздуха.

Поскольку диспетчер ориентируется на прогнозируемую среднесуточную температуру воздуха ($\bar{t}_{пр}$), то очевидно, что возможны отклонения прогнозируемого значения от фактически осуществившегося ($\bar{t}_{ф}$). Отсюда следуют две реальные ситуации.

В случае $\Delta t_{пр} = (t_{пр} - t_{ф}) < 0$ (прогнозируется более низкая температура чем фактически наблюдается) в сеть подается лишнее тепло – избыточное теплоснабжение, т. е. наблюдается перерасход топлива и тепловой энергии.

Тогда:

$$Q_{п} = \Delta Q_{п} = Q_{пр} - Q_{р} > 0, \quad (2.12)$$

где

$Q_{пр} = Q_{ф}$ фактически заданное диспетчером количество тепла, согласно прогноза температуры воздуха;

$Q_{р}$ - необходимое количество тепла, отвечающее последующей фактической температуре воздуха.

Если $\Delta t_p > 0$, т.е. $t_\phi < t_{пр}$ отмечается недостаточное теплоснабжение.

Иначе:

$$Q_H = \Delta Q_H = Q_{пр} - Q_p < 0, \quad (2.13)$$

Последствия: материальные потери в сфере производства, перерасход электроэнергии и природного газа на обогрев и потери рабочего времени, связанные с простудными заболеваниями.

По своей природе величина ΔQ носит случайный характер и может анализироваться с позиции известных законов статического распределения. С учетом плотности вероятности $\varphi(\Delta Q)$ функции интегрального распределения можно записать так:

$$Q_{п} = \int_{\Delta Q_0}^{\Delta Q_{п}^{\max}} \Delta Q_{п} \varphi(\Delta Q_{п}) d(\Delta Q_{п}), \quad (2.14)$$

$$Q_H = \int_{\Delta Q_0}^{\Delta Q_H^{\max}} \Delta Q_H \varphi(\Delta Q_H) d(\Delta Q_H), \quad (2.15)$$

Интегрируя выражения (2.14) и (2.15) с учетом плотности вероятности и ошибок прогнозирования получим функции тепловых потерь.

$$Q_{п}(\Delta t_{пр}) = \int_{\Delta Q_0}^{\infty} \frac{\Delta Q_{п}}{\sqrt{2\pi}\sigma_{\Delta Q}} \exp\left[-\frac{1}{2}\left(\frac{\Delta Q_{п} + \Delta Q_H}{\sigma_{\Delta Q}}\right)^2\right] d(\Delta Q_{п}) = \frac{\sigma_{\Delta Q}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\Delta Q_0 + \frac{Q_p^t}{t_{п} - t_p} \Delta t_{пр}}{\sigma_{\Delta Q}}\right)^2\right] - \frac{Q_p^t}{t_{п} - t_p} \Delta t_{пр} \left[\frac{1}{2} - \Phi\left(\frac{\Delta Q_0 + \frac{Q_p^t}{t_{п} - t_p} \Delta t_{пр}}{\sigma_{\Delta Q}}\right)\right], \quad (2.16)$$

$$Q_H(\Delta t_{пр}) = \int_{\Delta Q_0}^{-\infty} \frac{\Delta Q_H}{\sqrt{2\pi}\sigma_{\Delta Q}} \exp\left[-\frac{1}{2}\left(\frac{\Delta Q_H + \Delta Q_{пр}}{\sigma_{\Delta Q}}\right)^2\right] d(\Delta Q_H) = \frac{\sigma_{\Delta Q}}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\Delta Q_0 - \frac{Q_p^t}{t_{п} - t_p} \Delta t_{пр}}{\sigma_{\Delta Q}}\right)^2\right] + \frac{Q_p^t}{t_{п} - t_p} \Delta t_{пр} \left[\frac{1}{2} - \Phi\left(\frac{\Delta Q_0 - \frac{Q_p^t}{t_{п} - t_p} \Delta t_{пр}}{\sigma_{\Delta Q}}\right)\right], \quad (2.17)$$

где $\sigma_{\Delta Q}$ - среднее квадратичное отклонение расхода тепла $\Delta Q(Q_H, Q_{п})$;

Среднее квадратичное отклонение расхода тепла рассчитывается следующим образом:

$$\sigma_{\Delta Q} = 0.155 Q_P^T \frac{t_n - \bar{t}_{or}}{t_n - t_p}, \quad (2.18)$$

ΔQ_0 - погрешность контроля расхода тепла измерительными приборами;

Ошибка контроля тепла рассчитывается по формуле вида:

$$\Delta Q_0 = \frac{2 Q_P^T \delta t_B}{\Delta t_{Bпр}}, \quad (2.19)$$

$\Phi(t)$ - табулированная функция интеграла вероятности.

Структура формул (2.16) и (2.17) отражает две функциональные компоненты избыточного $Q_{п}$ и недостаточного $Q_{н}$ расхода тепла. Первая из них - экспоненциальная, вторая - линейная, что позволяет выполнить расчеты $Q_{п}$ и $Q_{н}$ при малых и больших ошибках прогноза $|\Delta t_{пр}| \geq 1.8 \sigma_{\Delta Q} (t_n - t_p) / Q_p^t$

В качестве предельной величины ошибок $|\Delta t_{пр}|$, определяющей области малых и больших значений $\Delta t_{пр}$, используется формула:

$$|\Delta t_{пр}| = \frac{3 \sigma_{\Delta Q}}{Q_p^t} (t_n - t_p), \quad (2.20)$$

Если отмечается, что $\Delta t_{пр} > |\Delta t_{пр}|$, то используются линейные оценки $Q_{п}$ и $Q_{н}$.

Иначе

$$Q_{п} = - \frac{Q_p^t}{t_n - t_p} \Delta t_{пр} \quad (2.21)$$

и

$$Q_{н} = - \frac{Q_p^t}{t_n - t_p} \Delta t_{пр} \quad (2.22)$$

При малых ошибках ($\Delta t_{пр} < |\Delta t_{пр}|$) расчеты выполняются согласно формул

$$Q_{п} = \sigma_{\Delta Q} \widetilde{Q}_{п}[\lambda, \theta_{пр}] \quad (2.23)$$

$$Q_{н} = \sigma_{\Delta Q} \widetilde{Q}_{н}[\lambda, \theta_{пр}] \quad (2.24)$$

Функции (2.23) и (2.24) табулированы, где

$$\lambda = \frac{\Delta Q_0}{\sigma_Q}, \quad (2.25)$$

и относительная ошибка прогноза

$$\Theta_{\text{пр}} = \frac{\Delta t_{\text{пр}}}{\sigma_{\Delta Q}(t_{\text{п}} - \bar{t}_{\text{от}})} \quad (2.27)$$

при этом $\bar{t}_{\text{от}}$ - средняя температура воздуха за отопительный период.

Для расчета тепловых затрат $\tilde{Q}_{\text{п}}$ и $\tilde{Q}_{\text{н}}$, что приведено в (2.23) и (2.24), необходимо предварительно установить функции вида:

$$\tilde{Q}_{\text{п}} = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} (\lambda + \Theta_{\text{пр}})^2 \right] - \Theta_{\text{пр}} \left[\frac{1}{2} - \Phi(\lambda + \Theta_{\text{пр}}) \right] \quad (2.28)$$

и

$$\tilde{Q}_{\text{н}} = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} (\lambda - \Theta_{\text{пр}})^2 \right] + \Theta_{\text{пр}} \left[\frac{1}{2} - \Phi(\lambda - \Theta_{\text{пр}}) \right], \quad (2.29)$$

здесь $\Theta_{\text{пр}}$ - относительная ошибка прогноза среднесуточной температуры воздуха, определяемая по формуле:

$$\Theta_{\text{пр}} = \frac{Q_{\text{п}}^T}{\sigma_{\Delta Q}} \frac{\Delta t_{\text{пр}}}{t_{\text{п}} - t_{\text{р}}} \quad (2.30)$$

Величина $\tilde{Q}_{\text{п}}$ и $\tilde{Q}_{\text{н}}$ находятся в зависимости от λ и $\Theta_{\text{пр}}$ по таблице 7 - значения $\tilde{Q}_{\text{п}}$ и $\tilde{Q}_{\text{н}}$ в зависимости от относительной ошибки прогноза $\Theta_{\text{пр}}$ и параметра λ . Безразмерный параметр λ рассчитывается по формуле:

$$\lambda = \frac{\Delta Q_0}{\sigma_{\Delta Q}}, \quad (2.31)$$

Таблица 7 – Относительные ошибки прогноза

$\tilde{Q}_{п,н}$	λ	Относительная ошибка прогноза $\Theta_{пр}$								
		-3,0	-2,0	-1,0	-0,5	0	0,5	1,0	2,0	3,0
$\tilde{Q}_п$	0,1	3,0	2,01	1,08	0,70	0,40	0,16	0,08	0,01	0
	0,5	3,0	2,00	1,04	0,65	0,35	0,15	0,06	0,01	0
	1,0	2,99	1,93	0,90	0,51	0,24	0,10	0,03	0	0
	1,5	2,99	1,74	0,66	0,32	0,13	0,04	0,01	0	0
	2,0	2,77	1,40	0,40	0,16	0,05	0,01	0	0	0
$\tilde{Q}_{п,н}$	λ	Относительная ошибка прогноза $\Theta_{пр}$								
		-3,0	-2,0	-1,0	-0,5	0	0,5	1,0	2,0	3,0
$\tilde{Q}_н$	0,1	0	0,01	0,08	0,16	0,40	0,70	1,08	2,01	3,0
	0,5	0	0,01	0,06	0,15	0,35	0,65	1,04	2,00	3,0
	1,0	0	0	0,03	0,10	0,24	0,51	0,90	1,93	2,99
	1,5	0	0	0,01	0,04	0,13	0,32	0,66	1,74	2,99
	2,0	0	0	0	0,01	0,05	0,16	0,40	1,40	2,77

Исходные данные, на основании которых выполняется исследование экономической полезности прогнозов температуры воздуха для теплоэнергетики Санкт-Петербурга, включают информацию двух видов: метеорологическую (фактические и прогностические сведения о температуре воздуха у поверхности земли) и теплотехническую (теплотехнические характеристики ТЭЦ Санкт-Петербурга). Исходные данные по теплоисточникам Санкт-Петербурга на 2016 г. представлены в таблице 8.

Таблица 8 - Станции ТЭЦ Санкт-Петербурга

№ п/п	Станция	Год ввода в эксплуатацию	Электрическая мощность	Тепловая мощность, Гкал/ч	Выработка электроэнергии, млн кВт	Выработка тепловой энергии, Гкал
1	Автовская ТЭЦ	1956	321 МВт	1849	1097,508	3 000 404
2	Василеостровская ТЭЦ	1932	135 МВт	1113	707,3	1 663 365
3	Выборгская ТЭЦ	1954	250,5 МВт	1056	620,029	1 087 868
4	Каскад Вуоксинских ГЭС	1949	240 МВт	-	1270,027	-
5	Каскад Ладозских ГЭС	2010	345 МВт	-	1168,197	-
6	Нарвская ГЭС	1955	124,8 МВт	-	517,551	-
7	Первомайская ТЭЦ	1957	524 МВт	1419	1774,257	1 433 235
8	Правобережная ТЭЦ	2006	643 МВт	1283	1985,609	1 937 640
9	Северная ТЭЦ	1975	500 МВт	1188	1704,713	2 732 306

Продолжение таблицы 8 - Станции ТЭЦ Санкт-Петербурга

10	Центральная ТЭЦ	1999	55 МВт	1346	144,516	2 161 804
11	Южная ТЭЦ	1978	1207 МВт	2353	3912,229	3 645 321

Исследование экономической полезности прогнозов температуры воздуха в дальнейшем будет выполняться на примере исходных данных Правобережной ТЭЦ. Правобережная ТЭЦ была построена как новая станция, заменившая собой первенца плана ГОЭЛРО, ТЭЦ-5 Красный Октябрь. Обеспечивает электрической и тепловой энергией промышленные предприятия, жилые и общественные здания Невского и Красногвардейского районов Санкт-Петербурга с населением свыше 700 тыс. человек.

Краткая информация по Правобережной ТЭЦ:

- Установленная электрическая мощность — 643,0 МВт;
- Установленная тепловая мощность — 1283,0 Гкал/ч;
- Выработка электроэнергии в 2015 году — 1985,609 млн кВтч;
- Отпуск тепловой энергии в 2015 году — 1 937 640 Гкал;
- Основное топливо — газ, резервное — мазут.

Для расчета функций тепловых и стоимостных потерь, необходимых для оценки экономической полезности прогнозов температуры воздуха для ТЭЦ, использовались следующие параметры:

1. Мощность теплоисточника, $Q_p^T = 1283$ Гкал/час
2. Расчетная температура воздуха t_p (климатическая характеристика, в соответствии с которой проектируется тепловая мощность ТЭЦ, данные на основе свода правил СП 131.13330.2012 "СНИП 23-01-99*. строительная климатология" актуализированная редакция СНИП 23-01-99*), $t_p = - 26$ °С
3. Температура воздуха внутри отапливаемых помещений, $t_{п} = 18$ °С
4. Средняя температура воздуха за отопительный сезон, $\bar{t}_{от} = - 3.6$ °С
5. Абсолютная погрешность измерения температуры воды, $\delta t_{в} = 1$ °С
6. Средняя разность температур воды прямой t' (поступающей от ТЭЦ в систему теплоснабжения) и обратной t'' (возвращаемой в ТЭЦ) $\Delta t_{врг}$, $t' = 120$ °С, $t'' = 50$ °С, $\Delta t_{врг} = 70$ °С
7. Стоимость отпускаемого на ТЭЦ тепла, $c_{п} = 0,12$ тыс. руб.
8. Масштаб потерь в расчете на единицу недоданной тепловой энергии $c_{п}$. Принято $c_{п} = 3 \div 5 c_{п}$

9. Период прогноза температуры воздуха $\tau = 24$ ч.
10. Среднее квадратичное отклонение расхода тепла $\sigma_{\Delta Q} = 97,62$ Гкал/час.
11. Ошибка контроля тепла $\Delta Q_0 = 36,66$ Гкал/час.
12. Относительная ошибка прогноза для малых ошибок будет определяться величиной $\Theta_{\text{пр}} = 0,299\Delta t_{\text{пр}}$.
13. Параметр $\lambda = 0,38$
14. Ошибка прогноза $|\Delta t_{\text{пр}}| = 10,04$ °С.

Для малых ошибок прогноза находим относительную ошибку $\Theta_{\text{пр}}$ при выбранных значениях $\Delta t_{\text{пр}}$:

$\Delta t_{\text{пр}}$	$\Delta t_{\text{пр}} < 0$			$\Delta t_{\text{пр}} = 0$	$\Delta t_{\text{пр}} > 0$		
	-9	-6	-3	0	3	6	9
$\Theta_{\text{пр}}$	-2,69	-1,79	-0,90	0,00	0,90	1,79	2,69

Тепловые потери $Q_{\text{п}}$ и $Q_{\text{н}}$ при малых ошибках прогноза определяются на основании предварительно установленных тепловых затрат $\tilde{Q}_{\text{п}}$ и $\tilde{Q}_{\text{н}}$, которые в зависимости от λ и $\Theta_{\text{пр}}$ находятся по таблице 5.

Для Правобережной ТЭЦ по таблице 5, при $\lambda = 0,38$ находим:

$\Delta t_{\text{пр}}$	-9	-6	-3	0	3	6	9
$\Theta_{\text{пр}}$	-2,69	-1,79	-0,90	0,00	0,90	1,79	2,69
$\tilde{Q}_{\text{п}}$	2,689	1,805	0,972	0,366	0,084	0,022	0,003
$\tilde{Q}_{\text{н}}$	0,003	0,022	0,084	0,366	0,972	1,805	2,689

При больших ошибках $\Delta t_{\text{пр}}$ значения тепловых потерь $Q_{\text{п}}$ и $Q_{\text{н}}$ устанавливаются по формулам (2.21) и (2.22) (линейная зависимость $Q_{\text{п}}$ и $Q_{\text{н}}$ от $\Delta t_{\text{пр}}$). Для Правобережной ТЭЦ $Q_{\text{п}} = -29,159 \Delta t_{\text{пр}}$ и $Q_{\text{н}} = 29,159 \Delta t_{\text{пр}}$. Соответственно находим при больших ошибках для Правобережной ТЭЦ:

$\Delta t_{\text{пр}}$	-15	-12	-9...+9	12	15
$Q_{\text{п}}$	437,386	349,909			
$Q_{\text{н}}$				349,909	437,386

Далее устанавливается *функция стоимостных потерь*. Для этого определяются стоимостные потери в виде:

$$s_{\text{п}} = \tau c_{\text{п}} Q_{\text{п}} - \text{при перерасходе тепла (тыс. руб.)}, \quad (2.32)$$

$$s_{\text{н}} = \tau c_{\text{н}} Q_{\text{н}} - \text{при недодаче тепла (тыс. руб.)}, \quad (2.33)$$

где $s_{\text{п}}$ и $s_{\text{н}}$ - суммарные стоимостные потери, $\tau = 24$ часа - при использовании суточных прогнозов,

$c_{\text{п}}$ - стоимость единицы выработки тепла на ТЭЦ (стоимость 1 Гкал),

$c_{\text{н}}$ - масштаб потерь при недодаче тепла. Это те потери, которые могут быть на производстве, при недополучении тепла в коммунальном хозяйстве, что вызывает расход электрической энергии и бытового газа на отопление, включая другие издержки.

Значение $s_{\text{п}}$ и $s_{\text{н}}$ находим при известном $c_{\text{п}}$ и $c_{\text{н}}$. Для Правобережной ТЭЦ $c_{\text{п}} = 1,2$ тыс. руб., $c_{\text{н}} = 1,2 \cdot 3 = 3,6$ тыс. руб. Величины $s_{\text{п}}$, $s_{\text{н}}$ и $s(\Delta t_{\text{пр}}) = s_{\text{п}} + s_{\text{н}}$ выражаем в тысячах рублей.

Анализ функции стоимостных потерь показывает, что минимум потерь приходится на ошибку примерно $\Delta t_{\text{пр}} = -3$ °С. Это значение ошибки рассматривается как *оптимизационная поправка* $\delta t_{\text{оп}}$ к ожидаемой среднесуточной температуре воздуха. Потребитель (диспетчер), учитывая соотношение штрафов за счет разного знака ошибок, должен понизить ожидаемую по прогнозу температуру воздуха на величину $\delta t_{\text{оп}}$. Все полученные значения ($\Theta_{\text{пр}}$, $\tilde{Q}_{\text{п}}$, $Q_{\text{п}}$, $s_{\text{п}}$, $\tilde{Q}_{\text{н}}$, $Q_{\text{н}}$, $s_{\text{н}}$) вносятся в таблицу 9, что отвечает расчету функции потерь $s(\Delta t_{\text{пр}})$, которая получает численное выражение.

Таблица 9 - Расчет тепловых и стоимостных потерь (тыс. руб.) для Правобережной ТЭЦ

параметры	ошибка прогноза ($\Delta t_{\text{пр}}$)										
	Перерасход тепла					0,0	Недодача тепла				
	-15,0	-12,0	-9,0	-6,0	-3,0		3,0	6,0	9,0	12,0	15,0
$\Theta_{\text{пр}}$	нет	нет	-2,7	-1,8	-0,9	0,0	нет	нет	нет	нет	нет
$\tilde{Q}_{\text{п}}$	нет	нет	2,7	1,8	1,0	0,4	0,1	0,0	0,0	0,0	0,0
$Q_{\text{п}}$	437,4	349,9	262,5	176,3	94,9	35,7	8,2	2,1	0,3	0,0	0,0
$s_{\text{п}}$	12596,7	10077,4	7560,8	5076,3	2732,9	1027,8	237,0	61,0	8,8	0,0	0,0
$\Theta_{\text{пр}}$	нет	нет	нет	нет	нет	0,0	0,9	1,8	2,7	нет	нет

Продолжение таблицы 9 - Расчет тепловых и стоимостных потерь (тыс. руб.) для
Правобережной ТЭЦ

параметры	ошибка прогноза ($\Delta t_{пр}$)										
	Перерасход тепла						Недодача тепла				
	-15,0	-12,0	-9,0	-6,0	-3,0	0,0	3,0	6,0	9,0	12,0	15,0
\tilde{Q}_H	0,0	0,0	0,0	0,0	0,1	0,4	1,0	1,8	2,7	нет	нет
Q_H	0,0	0,0	0,3	2,1	8,2	35,7	94,9	176,3	262,5	349,9	437,4
s_H	0,0	0,0	26,3	182,9	711,0	3083,4	8198,8	15228,8	22682,3	30232,1	37790,2
$s(\Delta t_{пр})$	12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2

Известно, что перерасход тепловой энергии - избыточная выработка тепла $Q_{п}$ на ТЭЦ – рассматривается как прямой расход топливных ресурсов, т.е. природного газа. При известной себестоимости выработки тепла $c_{п}$ устанавливается общий расход на единичный прогноз [18].

Если на ТЭЦ вырабатывается недостаточное количество тепловой энергии $Q_{н}$, не обеспечивающее нормальное теплоснабжение, то устанавливаются связанные с этим потери и принимаются равными $c_{н}=3c_{п}$. Исходя из этих условий и периода действия прогноза τ суммарные тепловые издержки, выраженные в стоимостной форме и записываются так:

$$S(\Delta t_{пр}) = \tau [c_{п}Q_{п} + c_{н}Q_{н}], \quad (2.34)$$

где

$s(\Delta t_{пр})$ - функция стоимостных потерь, т.е. потери зависимые от ошибки прогноза ($\Delta t_{пр}$);

$c_{п}$ – стоимость выработки единицы тепловой энергии, руб;

$c_{н}$ – стоимость недодачи тепловой энергии, руб;

$Q_{п}$ - избыточное теплоснабжение, Гкал/ч;

$Q_{н}$ – недостаточное теплоснабжение, Гкал/ч.

Расчет перерасхода $Q_{п}$ или недодачи $Q_{н}$ тепловой энергии осуществляется по формулам (2.16) и (2.17) как на участке линейного, так и экспоненциального распределения Q в зависимости от возможных ошибок прогнозирования $\Delta t_{пр}$.

На основании таблицы 6 составляется матрица потерь для Правобережной ТЭЦ, при известных градациях прогноза температуры воздуха (Приложение 5). Разработанная матрица потерь используется затем в течение многих лет как консервативная экономическая характеристика ТЭЦ.

Матрица потерь позволяет получить следующую полезную информацию:

- Описывает экономические последствия использования прогнозов температуры воздуха;
- Является необходимым условием для выбора оптимального управленческого решения (является исходной информацией для разработки матрицы систематических потерь).

2.2.2. Разработка погодо-хозяйственного регламента с использованием разнородных метеорологических данных

Для достоверной статистической оценки необходимо исследование метеорологических сведений в течение нескольких отопительных периодов. В рамках исследования выполняется разработка матриц сопряженности прогностических и фактических значений температуры воздуха.

Согласно известным научным разработкам [86, 87] в целях оптимизации выработки тепловой энергии используется прогноз среднесуточной температуры воздуха в данном пункте. По выбранному пункту рассчитываются прогностические $\bar{t}_{\text{пр}}$ ($\bar{t}_{\text{пр}} = 0.5(\bar{t}_{\text{пр.ночь}} + \bar{t}_{\text{пр.день}})$) и фактические $\bar{t}_{\text{ф}}$ ($\bar{t}_{\text{ф}} = 0.5(\bar{t}_{\text{ф.ночь}} + \bar{t}_{\text{ф.день}})$) среднесуточные значения температуры воздуха. На этом основании разрабатываются матрицы сопряженности методических и инерционных прогнозов.

Прогностические $t_{\text{пр}}^{-\text{ин}}$ и фактические $t_{\text{ф}}^{-\text{ин}}$ значения для матрицы сопряженности инерционных прогнозов устанавливаются только по фактическим материалам. Методика разработки представлена в работах [83]. На основании имеющихся прогностических данных, рассматривались прогнозы температуры воздуха для Правобережной ТЭЦ за отопительный сезон 2013-2014 г., число суточных прогнозов $N=212$.

Общий вид матрицы сопряженности многофазовых прогнозов с учетом обработки разнородных данных, представлен на рисунке 26.



Рисунок 26. Составление матрицы сопряженности многофазовых прогнозов с использованием модели управления разнородными данными

При разработке матрицы сопряженности необходимо, прежде всего, решить задачу определения числа градаций k в матрице сопряженности.

Для этого используются следующие формулы:

$$k_1 \leq 5 * \lg N \quad (2.35)$$

$$k_2 \leq \sqrt{N} \quad (2.36)$$

$$\begin{array}{l} N \dots 50 \ 100 \ 500 \\ k_3 \dots 8 \ 10 \ 13 \end{array} \quad (2.37)$$

Используя прогностические данные по Санкт-Петербургу за отопительные периоды 2012 – 2013 гг. находим:

Годы	N	k_1	k_2	k_3	\bar{k}
2012 – 2013	212	12	15	11	13

Заметим, что число градаций и их ширина должна определяться с учетом специфики зависимости теплоэнергетики от метеорологических условий. В данном исследовании принято число градаций $k=14$, однако в теплые зимы число градаций будет естественным образом уменьшено.

Ширина градации l может быть задана из условий прогностической градации или требований потребителя. Возможно использование формулы:

$$l = \frac{A_{max} - A_{min}}{k}, \quad (2.38)$$

где A_{max}, A_{min} - экстремальные значения температуры воздуха (при $A_{max} = 8 \text{ } ^\circ\text{C}$);
 k - число градаций.

В работе исследовались матрицы сопряженности прогнозов следующего вида (таблице 10):

Таблица 10 - Матрица сопряженности прогнозов среднесуточной температуры воздуха

Фактическая погода, Φ_i	Прогноз, Π_j						$\sum_{j=1}^m n_j$
	Π_1	Π_2	Π_3	...	Π_{13}	Π_{14}	
	-30...-28	-27...-25	-24...-22		+6...+8	9 и более	
Φ_1 -30...-28	n_{11}	n_{12}	n_{13}	...	n_{19}	n_{110}	n_{10}
Φ_2 -27...-25	n_{21}	n_{22}	n_{23}	...	n_{29}	n_{210}	n_{20}
Φ_3 -24...-22	n_{31}	n_{32}	n_{33}	...	n_{39}	n_{310}	n_{30}
...
Φ_{13} +6...+8	n_{91}	n_{92}	n_{93}	...	n_{99}	n_{910}	n_{90}
Φ_{14} 9 и более	n_{101}	n_{102}	n_{103}	...	n_{109}	n_{110}	n_{100}
$\sum_{i=1}^n n_i$	n_{01}	n_{02}	n_{03}	...	n_{09}	n_{010}	N

При известном распределении градаций $t(\Pi_1) < t(\Pi_m)$ и соответственно $t(\Phi_1) < t(\Phi_m)$ ошибочные прогнозы распределяются следующим образом: в правом верхнем углу - ошибки-пропуски (прогнозировалась более высокая температура, а была более низкая), в левом нижнем углу – ошибки-страховки [86, 87].

На основании данных, полученных в ходе исследования, были разработаны матрицы сопряженности методических и инерционных прогнозов температуры воздуха за отопительный сезон 2013-2014 г, результаты представлены в Приложении 2.

Для того чтобы установить существует ли связь между исследуемыми признаками используется *критерий независимости Пирсона* (формула 2.1). Этот критерий позволяет дать оценку статистической значимости выводов, получаемых по выборочным таблицам сопряженности.

Теоретическое (табличное) значение $\chi_{\alpha, \nu}^2$ устанавливается исходя из числа степеней свободы ν в таблице сопряженности и заданного уровня значимости α (выбор вероятности ошибки) (рисунок 27):

$$\nu = (\psi_1 - 1)(\psi_2 - 1), \quad (2.39)$$

где

ν - число степеней свободы;

ψ_1 - число строк;

ψ_2 - число столбцов.

Уровень значимости α может быть задан в пределах (от 0.001 до 0.05). данное число будет указывать на вероятность ошибки в существовании связи между признаками. так, при $\alpha = 0.05$ в 5 % случаев возможна независимость признаков – подтверждение нулевой гипотезы h_0 . Для проверки нулевой гипотезы используется правило:

- если $\chi^2 > \chi_{\alpha, \nu}^2$, то гипотеза о независимости признаков Π_j и Φ_i , отклоняется – между признаками существует взаимозависимость;
- если $\chi^2 \leq \chi_{\alpha, \nu}^2$, то гипотеза о независимости не отклоняется – связь между признаками носит случайный характер, признается случайной.

Критерий независимости Пирсона χ^2 позволяет лишь установить наличие или отсутствие связи между признаками Π_j и Φ_i , однако не измеряет эту связь, т.е. не является мерой связи.

Значения χ^2 , соответствующие значениям $P(\chi^2)$ и числам степеней свободы ν

ν	P										
	0,99	0,95	0,90	0,50	0,25	0,10	0,05	0,025	0,01	0,005	0,001
1	0,032	0,024	0,02	0,46	1,32	2,71	3,84	5,02	6,63	7,88	10,8
2	0,02	0,10	0,21	1,39	2,77	4,61	5,99	7,38	9,21	10,6	13,8
3	0,12	0,35	0,58	2,37	4,11	6,25	7,81	9,35	11,3	12,8	16,3
4	0,30	0,71	1,06	3,36	5,39	7,78	9,49	11,1	13,3	14,9	18,5
5	0,55	1,15	1,61	4,35	6,63	9,24	11,1	12,8	15,1	16,7	20,5
6	0,87	1,64	2,20	5,35	7,84	10,6	12,6	14,4	16,8	18,5	22,5
7	1,24	2,17	2,83	6,35	9,04	12,0	14,1	16,0	18,5	20,3	24,3
8	1,65	2,73	3,49	7,34	10,2	13,4	15,5	17,5	20,1	22,0	26,1
9	2,09	3,33	4,17	8,34	11,4	14,7	16,9	19,0	21,7	23,6	27,9
10	2,56	3,94	4,87	9,34	12,5	16,0	18,3	20,5	23,2	25,2	29,6
11	3,05	4,57	5,58	10,3	13,7	17,3	19,7	21,9	24,7	26,8	31,3
12	3,57	5,23	6,30	11,3	14,8	18,5	21,0	23,3	26,2	28,3	32,9
13	4,11	5,89	7,04	12,3	16,0	19,8	22,4	24,7	27,7	29,8	34,5
14	4,66	6,57	7,79	13,3	17,1	21,1	23,7	26,1	29,1	31,3	36,1
15	5,23	7,26	8,55	14,3	18,2	22,3	25,0	27,5	30,6	32,8	37,7
16	5,81	7,96	9,31	15,3	19,4	23,5	26,3	28,8	32,0	34,3	39,3
17	6,41	8,67	10,1	16,3	20,5	24,8	27,6	30,2	33,4	35,7	40,8
18	7,01	9,39	10,9	17,3	21,6	26,0	28,9	31,5	34,8	37,2	42,3
19	7,63	10,1	11,7	18,3	22,7	27,2	30,1	32,9	36,2	38,6	43,8
20	8,26	10,9	12,4	19,3	23,8	28,4	31,4	34,2	37,6	40,0	45,3
21	8,90	11,6	13,2	20,3	24,9	29,6	32,7	35,5	38,9	41,4	46,8
22	9,54	12,3	14,0	21,3	26,0	30,8	33,9	36,8	40,3	42,8	48,3
23	10,2	13,1	14,8	22,3	27,1	32,0	35,2	38,1	41,6	44,2	49,7
24	10,9	13,8	15,7	23,3	28,2	33,2	36,4	39,4	43,0	45,6	51,2
25	11,5	14,6	16,5	24,3	29,3	34,4	37,7	40,6	44,3	46,9	52,6
26	12,2	15,4	17,3	25,3	30,4	35,6	38,9	41,9	45,6	48,3	54,1
27	12,9	16,2	18,1	26,3	31,5	36,7	40,1	43,2	47,0	49,6	55,5
28	13,6	16,9	18,9	27,3	32,6	37,9	41,3	44,5	48,3	51,0	56,9
29	14,3	17,7	19,8	28,3	33,7	39,1	42,6	45,7	49,6	52,3	58,3
30	15,0	18,5	20,6	29,3	34,8	40,3	43,8	47,0	50,9	53,7	59,7

Рисунок 27. Табличное значение $\chi_{\alpha, \nu}^2$

В данной таблице уровень значимости α обозначен как уровень вероятности ошибки P .

На основании матрицы сопряженности метеорологических прогнозов рассчитываются условные вероятности q_{ij} осуществления фазы погоды Φ_i при известном тексте прогноза Π_j . Результат расчетов представлен в Приложении 5. На рисунке 28 изображен алгоритм обработки разнородных метеорологических данных и получение матрицы условных вероятностей.



Рисунок 28. Расчет матрицы условных вероятностей на основе обработки разнородных метеорологических данных

Учитывая специфику работы теплоисточника и постоянную адаптацию теплоносителя под ожидаемую среднесуточную температуру воздуха выше были рассмотрены особенности разработки функции потерь данного потребителя (теплоисточника) в виде функции стоимостных потерь:

- Представление функции потерь в дискретном, матричном виде порядка $n \times m$.
- Особое значение в исследовании приобретает задача выбора критерия оптимизации решений.

При известных вероятностных характеристиках прогноза целесообразно использовать байесовский подход к выбору оптимальных решений и стратегий.

С учетом условных вероятностей осуществления текстов прогнозов (градаций температуры воздуха) q_{ij} и элементов матрицы потерь s_{ij} определяются средние потери потребителя при условии сохранения его права, получив данный текст прогноза Π_j , выбирать ориентацию d_k на любую иную температуру воздуха. Такие потери называются систематическими и определяются по формуле:

$$\bar{R}_{kj} = \sum_{k=1}^{n=m} s_{kj} \cdot q_{ij}, \quad (2.48)$$

где s_{kj} – элементы матрицы потерь $\|s_{ij}\|$, отвечающие k -му решению (действию потребителя) в ожидании температуры воздуха Φ_i .

Величина \bar{R}_{kj} рассматривается как критерий выбора лучшего, оптимального решения. На этом основании рассчитывается матрица систематических потерь потребителя $\|\bar{R}_{kj}\|$, отражающая потери теплоисточника по всему множеству действий диспетчера при ориентации на данный прогноз Π_j . Общий вид матрицы систематических потерь приведен в таблице 11.

Таблица 11 - Матрица систематических потерь потребителя \bar{R}_{kj}

Действия потребителя, d_k	Прогноз, Π_j				
	Π_1	Π_2	Π_3	...	Π_m
d_1	\bar{R}_{11}	\bar{R}_{12}	\bar{R}_{13}	...	\bar{R}_{1m}
d_2	\bar{R}_{21}	\bar{R}_{22}	\bar{R}_{23}	...	\bar{R}_{2m}
d_3	\bar{R}_{31}	\bar{R}_{32}	\bar{R}_{33}	...	\bar{R}_{3m}
...
d_k	\bar{R}_{k1}	\bar{R}_{k2}	\bar{R}_{k3}	...	\bar{R}_{km}

Оптимальное решение отвечает условию:

$$\bar{R}(d_k, \Pi_j)_{\text{опт}} = \min_{d_k} \bar{R}_{kj}, \quad (2.49)$$

Матрица систематических потерь выступает в качестве регламента выбора оптимальных оперативных решений при задании температуры теплоносителю.

Исследования, выполненные в работах позволяют рассматривать средние (в статистическом смысле) потери как показатель меры погодозависимости производственной деятельности теплоисточника при использовании им прогностических ресурсов. В качестве базовых рассматриваются средние потери в условиях реализации методических прогнозов при задании режима работы ТЭЦ.

На множестве значений \bar{R}_{kj} в матрице систематических потерь определяются средние потери за отопительный период при использовании методических прогнозов и стратегии постоянного доверия им.

$$\bar{R}_M = \sum_{j=1}^m p_{0j} \bar{R}_{k=j}(\Pi_j), \quad (2.50)$$

При оптимальной стратегии, что соответствует выбору оптимального решения $d_k = d_{k_{\text{опт}}}$, средние потери есть величина:

$$\bar{R}_{\text{мо}} = \sum_{j=1}^m p_{0j} \min \bar{R}_{kj}(\Pi_j), \quad (2.51)$$

В результате исследования была построена матрица систематических потерь для Правобережной ТЭЦ (Приложение 6).

- Матрица систематических потерь является регламентом выбора диспетчером оптимального задания режима работы ТЭЦ.
- Матрица систематических потерь позволяет определить ежегодную полезность прогнозов по завершению отопительного периода.
- Матрица систематических потерь может стать новым видом информационного продукта или разновидностью импакт прогноза, который учитывает неопределенность реализации текста прогноза.

2.2.3. Оценка экономической полезности прогнозов температуры воздуха для теплоэнергетики

Возможность экономически выгодной реализации прогнозов содержится в байесовском подходе к оценке результативности использования прогнозов в хозяйственной практике. В основе такого подхода лежит оценка средних потерь (средних в статистическом смысле), имеющих вероятностное содержание. На этом основании, во-первых, устанавливается более верная ориентация производственных операций (так или иначе связанных с погодой), а также предпринимаются необходимые меры защиты (вплоть до кардинальных). В результате оценивается экономическая полезность – экономический эффект и экономическая эффективность использования прогнозов погоды.

Весь спектр прогнозируемых (Π_i) и фактических (Φ_j) значений метеорологических элементов и явлений погоды можно представить в виде последовательного непрерывного ряда значений или ряда градаций, или фаз погоды. Матрица сопряженности $\Pi_i \sim \Phi_j$ позволяет рассчитать вероятности (p_{ij}) осуществления каждой фазы. Тем самым все категорические прогнозы переводятся в вероятностную форму, что необходимо потребителю для последующего выбора оптимальной хозяйственной стратегии.

Допуская возможность использования инерционных прогнозов, определяются средние потери при известной уже матрице потерь и матрице сопряженности инерционных прогнозов

$$\bar{R}_{ин} = \sum_{j=1}^m p_{0j}^{ин} \bar{R}_{k=j}(\Pi_j), \quad (2.52)$$

Приведенные здесь формализации (формулы с (2.48) по (2.52)) положены в основу оценки экономической полезности прогнозов температуры воздуха для теплоэнергетики Санкт-Петербурга. Показателями экономической полезности прогнозов выступает экономический эффект и экономическая эффективность.

Экономический эффект использования оперативных методических прогнозов есть *сбереженные* потребителем *материальные ценности* (получаемых за счет преимущества методических прогнозов относительно инерционных) и *за вычетом затрат на получение прогностической информации* в Росгидромете.

В работе рассматривается оценка экономического эффекта на основании формулы 2.9.

Оценка экономического эффекта устанавливается как при условии доверия прогнозам, так и при возможности выбора оптимальных решений. Наряду с оценкой экономического эффекта выполняются расчеты экономической эффективности по формуле вида:

$$P = \frac{\mathcal{E}}{N \cdot Z_{\text{ин}}}, \quad (2.53)$$

где

P – экономическая эффективность;

\mathcal{E} - экономический эффект;

N - общее число прогнозов за отопительный период;

$Z_{\text{ин}}$ - стоимость единицы прогностической информации в данном УГМС.

Экономическая эффективность отражает удельный экономический эффект – величина безразмерная, позволяющая установить величину сэкономленных материальных ценностей (средств, ресурсов и т.п.) данного потребителя, приходящихся на один рубль затрат на разработку прогнозов в иной отрасли – в гидрометслужбе.

Полученные в ходе исследования результаты позволяют определить средние потери за отопительный период при использовании методических (\bar{R}_m) и инерционных ($\bar{R}_{\text{ин}}$) прогнозов. *В итоге определяется экономический эффект (\mathcal{E}) и экономическая эффективность (P) использования конкретным ТЭЦ прогнозов среднесуточной температуры воздуха.*

Оценка экономического эффекта оперативных метеорологических прогнозов может дать достоверные результаты только в том случае, если выполняются следующие условия:

1. Поскольку осуществление любого прогноза носит вероятностный характер, оценивать экономический эффект единичного прогноза допустимо лишь при условии, что для полученной величины экономического эффекта рассчитывается процент его достоверности. При ином подходе полученный экономический эффект может оказаться случайным.

2. Для оценки экономического эффекта необходимо выбирать достаточно продолжительный период времени (месяц, сезон), что позволяет считать выбранный непрерывный ряд прогнозов статистически обеспеченным.

3. На основании выдаваемого потребителю прогностического материала устанавливается начальный уровень отсчета полезности прогнозов – базовый вариант. Это могут быть инерционные, случайные или климатологические прогнозы. Привлечение стандартного прогноза в качестве базового варианта позволяет установить преимущество оперативных методических прогнозов и является условием их дальнейшего совершенствования.

4. Учитывая экономические последствия воздействия погоды на потребителя в виде тех или иных потерь, в качестве основной характеристики экономического эффекта рассматривается уменьшение средних (статистических) потерь ΔR .

5. Обязательным условием оценки экономического эффекта оперативных методических прогнозов является учет произведенных на них затрат Z_{III} .

В работе были выполнены исследования по данным Правобережной ТЭЦ Санкт-Петербурга, рассчитаны функции (матрицы) метеорологических потерь и произведена оценка экономического эффекта и экономической эффективности использования ТЭЦ прогнозов средней температуры воздуха за отопительные периоды 2013-2014 гг.

Расчеты выполнены в два этапа. На первом этапе были разработаны матрицы сопряженности многофазовых методических прогнозов температуры воздуха. На втором - установлены систематические и общие потери ТЭЦ при использовании методических и инерционных прогнозов. Далее, по рассчитанным общим средним потерям и выбранной стратегии поведения потребителя, рассчитываем экономический эффект при использовании методических (\mathcal{E}_m) и инерционных прогнозов температуры воздуха, также - экономический эффект оптимального использования методических прогнозов (\mathcal{E}_{mo}), экономическую эффективность (P) использования методических прогнозов.

Оценка экономического эффекта устанавливается как при условии доверия прогнозам, так и при возможности выбора оптимальных решений. Полученные в ходе исследования результаты позволяют определить средние потери за отопительный период при использовании методических (\bar{R}_m) и инерционных ($\bar{R}_{ин}$) прогнозов [86, 87].

Результаты расчетов показателей экономической полезности прогнозов среднесуточной температуры воздуха для Правобережной ТЭЦ представлены в таблице 12.

Таблица 12 - Оценка показателей экономической полезности прогнозов среднесуточной температуры воздуха (Правобережной ТЭЦ, 2013-2014 год), в рублях

Показатели	Использование прогнозов в режиме доверия	Оптимальное использование прогнозов
Экономический эффект	103150,5	167733,1
Экономическая эффективность	486,5	791,1

В современных условиях развития топливно-энергетического комплекса придаётся большое значение стратегии энергосбережения и проблеме эффективного использования нефтегазовых топливных ресурсов с учетом метеорологической информации.

Оптимизация использования метеорологической информации способствует снижению потерь от неблагоприятных и опасных условий погоды, а так же способствует повышению эффективного принятия управленческих решений по управлению территорией.

Методика оптимизации управленческих решений с использованием метеорологической информации позволяет не только оценить экономическую полезность прогнозов температуры воздуха для ТЭЦ, но и установить оптимальные для данной ТЭЦ условия их реализации, что заметно повышает экономический эффект этих прогнозов. Таким образом, решается целевая задача погодозависимого потребителя - минимизация экономико-метеорологических рисков при принятии хозяйственных решений на основе информации об ожидаемом состоянии гидрометеорологической среды.

Автоматизация многофазового погодо-хозяйственного регламента, позволяющего потребителю подстраиваться под ожидаемые погодные условия с минимальными метеорологическими издержками, создает ряд предпосылок к разработке геоинформационной системы поддержки принятия управленческих решений:

- автоматизация управления разнородных метеорологических данных;
- оптимизация использования метеорологической информации;
- учет неопределенности реализации метеорологического прогноза;
- снижение потерь от неблагоприятных и опасных условий погоды;
- адаптация хозяйственной деятельности к неблагоприятным проявлениям гидрометеорологической среды;
- минимизация затрат на поиск, обработку, проведение анализа и исследований, составление вариантов решений;
- расширение предлагаемых вариантов управленческих решений;
- минимизация участия человека в процессе подготовки вариантов управленческих решений.

3. Разработка геоинформационной системы поддержки принятия управленческих решений

На основе модели управления разнородными данными разрабатывается программное решение, позволяющее потребителю оценить экономические преимущества оптимального использования метеорологической информации в производственном процессе. При реализации программного решения использовалась модель обработки метеорологических данных применительно к задаче регулирования отпуска тепла на основании ожидаемой температуры воздуха. В настоящее время особое внимание уделяется решению таких прикладных задач именно с помощью информационных технологий. Проведя анализ существующих в открытом доступе программных продуктов по решению данной проблемы, можно сделать явный вывод, что на данный момент не достаточно программных решений в области импакт составляющих прогноза.

В настоящее время экономические агенты, принимающие решения в соответствии с ожидаемой погодой, могут использовать прогнозы, приобретенные в рамках специализированного гидрометеорологического обеспечения в подразделениях Росгидромета, либо в иных, в том числе, коммерческих прогностических организациях. Не исключается также использование прогностической информации, размещенной на интернет-ресурсах в открытом доступе. Кроме того, потребитель вправе использовать ориентацию на иные прогнозы, разработка которых не требует применения физически обоснованных методов прогнозирования. В качестве таких прогнозов могут выступать инерционные прогнозы, в основе которых лежит текущая погода, или случайные прогнозы, когда значение метеорологической величины выбирается случайным образом из всей области ее распределения. Таким образом, потребители сталкиваются с проблемой поиска источника информации, который наилучшим образом обеспечивает достижение ожидаемых экономических результатов – снижение издержек по метеорологическим причинам.

Целесообразно создать такой информационный продукт, который позволит потребителям гидрометеорологической информации получить доступ к использованию программных ГИС-решений, средству анализа информации и получению актуальных сведений, а так же предоставит им возможность выбора стратегически важного решения по управлению территорией в единой системе.

3.1. Методика проектирования геоинформационной системы поддержки принятия решений на основе модели обработки разнородных данных

В двадцать первом веке информационные технологии могут предоставить широкий набор способов создания как информационной, так и геоинформационной системы. Выбор происходит исходя из требований пользователей и области применения специализированного программного обеспечения.

К примеру ГИС создаются различными организациями, реализующими функционал системы под конкретные прикладные задачи. Интерфейс каждой системы индивидуален, системы различаются так же и по набору геоданных, функционалу, назначению системы, картографическим основам, методу анализа и выводу результатов, механизму обработки данных, сервисам системы, архитектуре системы, механизму интеграции данных, целевой установке и другим различным факторам.

Исследуя в работе геоинформационные системы обработки разнородных пространственных данных для эффективного решения комплексных задач по управлению территориями целесообразно использовать ГИС поддержки принятия решений (ГИС ППР, рисунок 29) [70].

В процессе принятия управленческих решений лицу, принимающему решение (ЛПР), приходится учитывать большое количество показателей, критериев, факторов, влияющих на поставленную в задаче цель. Принятие решения, как правило, есть выбор из списка возможных альтернатив. Принять «правильное» решение значит выбрать такую альтернативу из числа возможных, в которой с учетом всех разнообразных критериев, факторов и требований будет оптимизирована общая ценность, то есть она в максимальной степени будет способствовать достижению поставленной цели. Кроме того, практически в любых управленческих задачах существуют разного рода неопределенности, связанные с противоречивостью критериев, неполнотой знаний о проблеме, невозможностью количественного измерения тех или иных факторов и показателей [35, 89, 91].

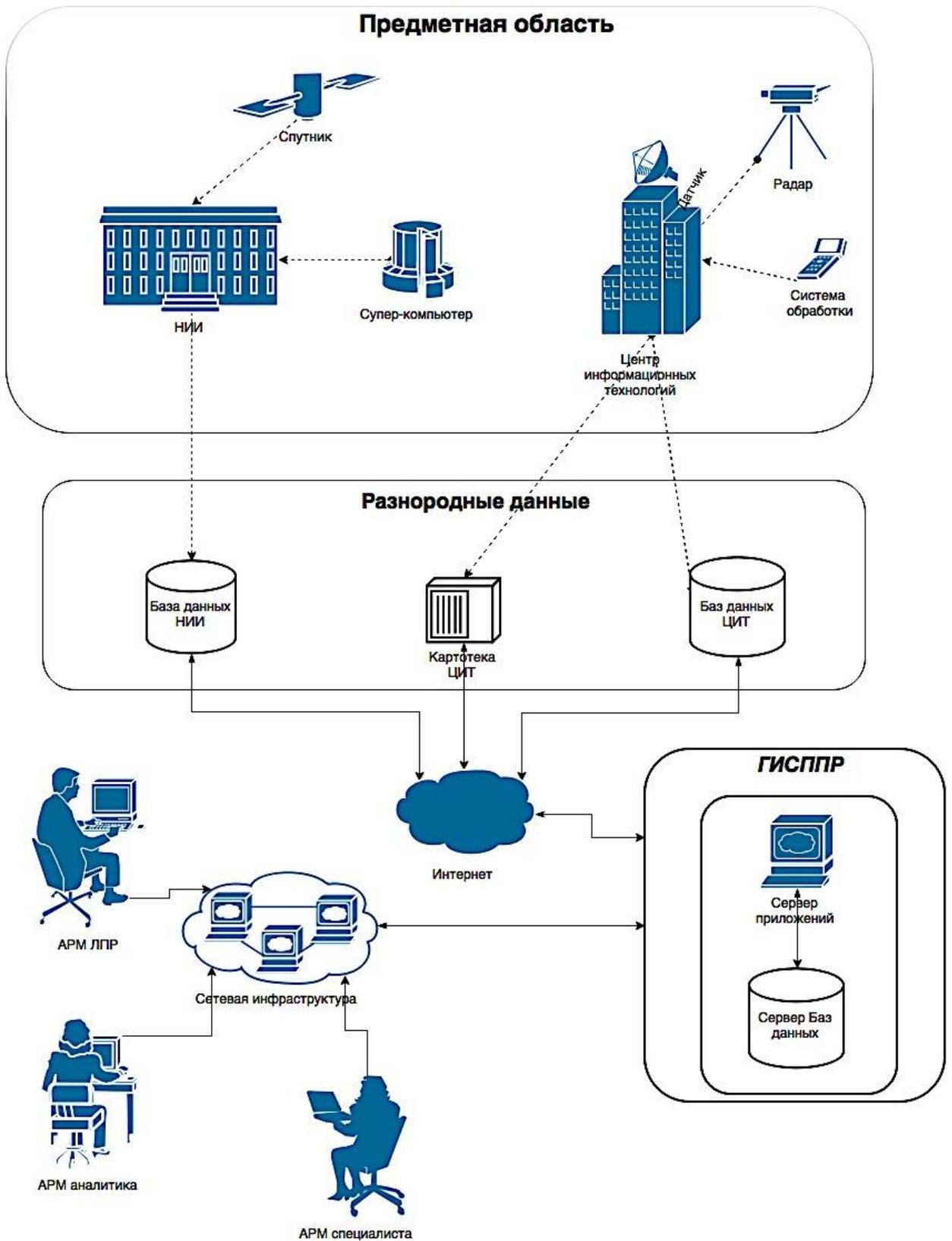


Рисунок 29. Схема потоков разнородной информации при работе с ГИСПП

Этапы работы с ГИСПП (над прикладной задачей) можно упрощенно представить в виде следующей последовательности (рисунок 30):

1. Ввод и предварительная обработка данных в ГИС,
2. Увязка баз картографических и фактографических данных.
3. Структурирование, анализ и интерпретация данных в ГИС с использованием встроенных инструментов обработки данных и средств анализа.
4. Формализация задачи в блоке СППР с использованием обработанных ГИС данных и пользовательских экспертных оценок и предпочтений.
5. Обработка данных средствами СППР (моделирование, оценка, ранжирование и т.д.)
6. Анализ и моделирование полученных СППР результатов
7. Принятие окончательного решения и выдача управляющих воздействий на основе анализа результатов ГИС ППР [68].

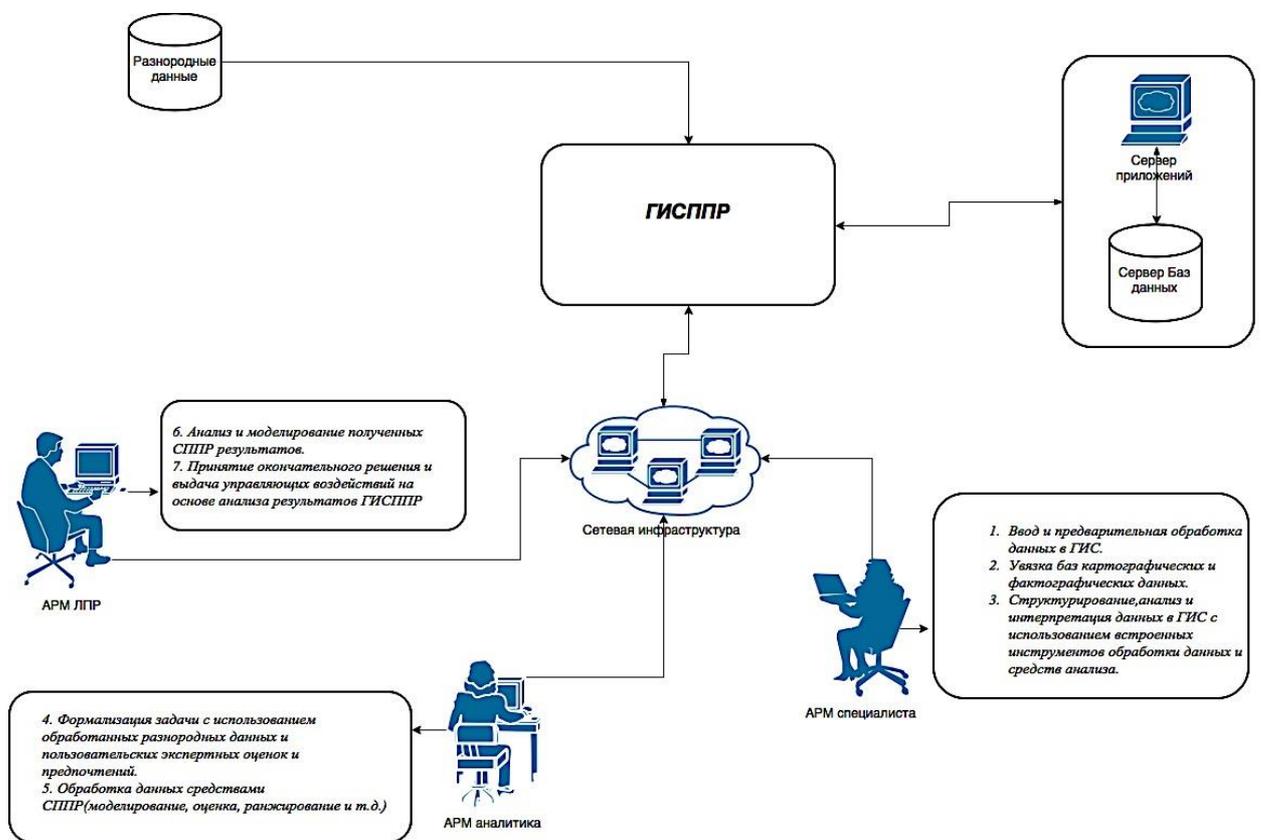


Рисунок 30. Этапы работы с ГИС ППР

Расчет в ГИС СППР формирует обобщенные приоритеты - величины, определяющие степень предпочтительности элемента относительно цели решаемой проблемы. Средства графического анализа СППР позволяют провести ранжирование альтернативы по обобщенным приоритетам. Кроме того, можно провести анализ по локальным приоритетам, чтобы выяснить из каких составляющих получился данный показатель, какой критерий получил наибольший

или наименьший вес и т.д. Есть средства определения чувствительности полученного результата к изменению входных данных и экспертных суждений.

Результирующие данные СППР могут быть сохранены, что с легкостью позволяет их использовать для интерпретации в ГИС. Например, мы можем по этим данным построить шкалу и интерпретировать полученные результаты средствами ГИС.

ГИСППР позволит решать широкий круг задач, к примеру:

- Анализ и планирование развития территории;
- Анализ и управление ресурсами;
- Оценка и планирование состояния окружающей среды;

Применение вышеперечисленных особенностей ГИСППР свидетельствует о преимуществах данной системы как в точности анализа, так и в удобстве, простоте и эффективности работы в процессе принятия решений.

Проектирование геоинформационной системы поддержки принятия решений предусматривает использование разработчиками определенной технологии проектирования, которую можно охарактеризовать как совокупность методологии и средств проектирования ИС, а также методов и средств его организации. В её основе лежит технологический процесс, определяющий действия, их последовательность, а также требуемый состав исполнителей, средств и ресурсов. Проектирование ИС можно разделить на ряд взаимосвязанных, последовательно или параллельно выполняющихся, цепочек действий. Они, в свою очередь, делятся на проектировочные операции, которые формируют или редактируют результаты проектирования. Таким образом технология проектирования выполняется на основе того или иного метода по средствам выполнения последовательных операций, в результате чего становится понятным не только что должно быть сделано, но и как, кем и в какой последовательности [42, 4].

Требования к подобным картографическим системам плохо формализованы и не сформированы в общую концепцию проектирования подобных систем, так же проектирование системы не стандартизировано. Все это доставляет неудобства разработчикам геоинформационных систем.

Применение ГИС для решения различных задач обуславливает разные подходы к процессу проектирования таких систем. В связи с отсутствием ГОСТ на разработку геоинформационной системы в работе представлена одна из возможных методик проектирования геоинформационной системы поддержки принятия решений на основе модели обработки разнородных данных.

Цель методики - поэтапная разработка геоинформационной системы для поддержки принятия решений с использованием механизмов интеграции разнородных пространственных данных и типовых ГИС-решений.

Геоинформационные технологии призваны автоматизировать многие трудоёмкие операции, ранее требовавшие больших временных, энергетических, психологических и других затрат от человека. Однако разные этапы технологической цепочки поддаются большей или меньшей автоматизации, что в значительной степени может зависеть от правильной постановки исходных задач.

Проектирование информационной системы в соответствии с государственными стандартами 19-й и 34-й серии можно свести к следующим стадиям (рисунок 31). Стандартизация позволяет не только взаимодействовать с другими информационными системами, но и использовать создаваемые технологические решения в других проектах и информационных системах [5, 6, 62].

Прежде чем начать проектирование геоинформационной системы необходимо определиться с объектом проектирования. Под объектом проектирования понимают отдельные элементы или комплексы функциональных и обеспечивающих частей. Функциональными элементами принято считать задачи или комплексы задач и функции управления. Обеспечивающая часть состоит из элементов или комплексов информационного, программного, технического и других видов обеспечения системы. Для геоинформационной системы управления территориями, объектом проектирования служат метеорологические, гидрологические и иные параметры, оказывающие влияние на хозяйственную деятельность [19].

Стадии	Этапы работ
1. Формирование требований к АИС	1.1. Обследование объекта и обоснование необходимости создания АИС. 1.2. Формирование требований пользователя к АИС. 1.3. Оформление отчета о выполненной работе и заявки на разработку АИС (тактико-технического задания).
2. Разработка концепции АИС	2.1. Изучение объекта. 2.2. Проведение необходимых научно-исследовательских работ. 2.3. Разработка концепции АИС и выбор варианта концепции АИС, удовлетворяющего требованиям пользователя. 2.4. Оформление отчета о выполненной работе.
3. Техническое задание	3.1. Разработка и утверждение технического задания на создание АИС.
4. Эскизный проект	4.1. Разработка предварительных проектных решений по системе и ее частям. 4.2. Разработка документации на АИС и ее части.
5. Технический проект	5.1. Разработка проектных решений по системе и ее частям. 5.2. Разработка документации на АИС и ее части. 5.3. Разработка и оформление документации на поставку изделий для комплектования АИС и (или) технических требований (технических заданий) на их разработку. 5.4. Разработка заданий на проектирование в смежных частях проекта объекта автоматизации.
6. Рабочая документация	6.1. Разработка рабочей документации на систему и ее части. 6.2. Разработка или адаптация программ.
7. Ввод в действие	7.1. Подготовка объекта автоматизации к вводу АИС в действие. 7.2. Подготовка персонала. 7.3. Комплектация АИС поставляемыми изделиями (программными и техническими средствами, программно-техническими комплексами, информационными изделиями). 7.4. Строительно-монтажные работы. 7.5. Пусконаладочные работы. 7.6. Проведение предварительных испытаний. 7.7. Проведение опытной эксплуатации. 7.8. Проведение приемочных испытаний.
8. Сопровождение АИС	8.1. Выполнение работ в соответствии с гарантийными обязательствами. 8.2. Послегарантийное обслуживание.

Рисунок 31. Этапы проектирования информационной системы в соответствии с ГОСТ

Помимо объекта проектирования существует ещё и субъект проектирования информационной системы, который может состоять из коллективов специалистов, осуществляющих проектную деятельность, и заказчика требуемой ИС. Субъектом

проектирования в данной системе являются муниципальные органы власти, а также крупные промышленные компании, заинтересованные в проведении подобных исследованиях. Масштаб проектируемой системы определяет количество проектного коллектива и время, отведенное на создание автоматизированной системы [1-3, 7-9]. В случае создания геоинформационной системы управления территориями на основе распределенных разнородных данных, существует необходимость в участии нескольких проектных коллективов на местах распределения баз данных и создании головной организации, которая бы координировала деятельность команд-исполнителей на местах.

Основу же технологии проектирования составляет методология проектирования, она определяется, как набор методов, реализующих некоторые концепции и принципы проектирования информационной системы. Сложность построения подобной системы состоит в применении различных технологий создания предметных информационных систем.

В геоинформационной системе управления территориями на основе разнородных баз данных необходимо сочетать геоинформационные технологии и технологии создания систем управления, что неразрывно влечёт за собой создание методики проектирования подобных систем. Изначально в проект должны быть заложены технологии, обеспечивающие процесс функционирования системы:

1. Нарращивание, изменение и распределение автоматизированных функций системы, конфигурирование системы;
2. Возможность автоматического обмена информацией между объектами как внутри системы, так и с внешними объектами;
3. Возможность динамической перестройки системы;
4. Автоматическую модернизацию системы, её компонентов при утрате актуальности информации в базах данных;
5. Взаимозаменяемость сервисов, для бесперебойной работы системы;
6. Защита системы от несанкционированного доступа, изменения и подмены данных;
7. Обеспечение взаимосвязи разноформатных, гетерогенных баз данных.

Применение подобных технологий требует дополнительного формального описания предметной области:

1. Описание задач автоматизации и их информационного обеспечения;
2. Описание структуры и объема информации, которая участвует в информационном обмене задач, объектов;
3. Описание требуемой структуры и объема информации при создании персональных баз данных пользователя;

4. Описание полномочий пользователя при работе с функциями и информацией.

Важным этапом разработки ГИС является обоснование выбора моделей данных и ее представления, что определяет требования к СУБД в создаваемой системе. Поскольку система предполагает работу с большими объемами записей, то СУБД должна обеспечивать быстрое действие и удобство работы, а также должна удовлетворять не только текущим, но и будущим потребностям.

Приведем критерии, предъявляемые реализуемой системе.

1. В созданной системе должна использоваться реляционная структура базы данных, построенная на взаимоотношении совокупности взаимосвязанных таблиц. Каждая из групп должна хранить информацию о конкретном картографическом слое и об объектах данного слоя, включая геоданные и атрибуты.
2. СУБД должна обладать масштабируемостью, мобильностью, распределенностью, а также обеспечивать поддержку широкого круга сетевых протоколов для работы в многопользовательском режиме.
3. СУБД должна иметь средства разработки WEB-приложений и поддержку широко известных языков программирования;
4. СУБД должна поддерживать различные программные и аппаратные платформы, так же оптимизировать ресурсы для максимальной производительности.
5. СУБД должна быть хорошо документирована и иметь поддержку производителем.

Для организации разработки подобной системы может быть введён в состав информационного обеспечения объектов предметных областей и их структуры, входящих в систему разнородных баз данных, представленное в виде множеств иерархически упорядоченных реляционных таблиц — информационных ресурсов. Что предусматривает создание информационного хранилища с описанием атрибутов и их принадлежности к каждой разнородной базе данных. Описание помогает быстрее отправлять запрос к содержащей необходимую информацию базе данных. Тем самым задается информация, необходимая для построения информационной системы данного вида. При этом реализуются задание различных мета свойств информации в базе данных [36, 51].

Обеспечение целостности данных является важнейшей задачей при проектировании и эксплуатации систем обработки данных (СОД). Проблема целостности состоит в обеспечении правильности данных в базе данных в любой момент времени». Целостность – актуальность и непротиворечивость информации, ее защищенность от разрушения и несанкционированного изменения.

Целостность является одним из аспектов информационной безопасности наряду с доступностью – возможностью с приемлемыми затратами получить требуемую информационную услугу, и конфиденциальностью – защитой от несанкционированного прочтения. Целостность данных – неотъемлемое свойство базы данных, и ее обеспечение является важнейшей задачей проектирования БД. Целостность данных описывается набором специальных предложений, называемых ограничениями целостности. Ограничения целостности представляют собой утверждения о допустимых значениях отдельных информационных единиц и связях между ними. Эти ограничения определяются в большинстве случаев особенностями предметной области. При выполнении операций над БД проверяется выполнение ограничений целостности.

Различают логическую и физическую целостность БД. Логическая целостность – состояние БД, характеризующееся отсутствием нарушений ограничений целостности, присущих логической модели данных (т.е. неявных ограничений), и явных ограничений, заданных декларативным или процедурным путем. Физическая целостность – отсутствие нарушений спецификаций схемы хранения, а также физических разрушений данных на носителе.

Для контроля целостности БД применяется также механизм триггеров. Триггер – это действие, которое активизируется при наступлении указанного события (вставки, удаления, обновления записи). Триггеры специфицируются в схеме базы данных. Более широким понятием по отношению к триггеру является понятие хранимая процедура. Хранимые процедуры описывают фрагменты логики приложения, хранятся и исполняются на сервере, что позволяет улучшать характеристики производительности.

- Транзакции. Транзакция представляет собой законченную совокупность действий над БД, которая переводит ее из одного целостного в логическом смысле состояния в другое. К транзакциям предъявляется набор требований, известный под названием ACID (Atomicity, Consistency, Isolation, Durability). Эти требования вытекают из определения транзакции.
- Атомарность. Транзакция представляет собой некоторый набор законченных действий. Система обеспечивает их выполнение по принципу «все или ничего» - либо выполняются все действия, тогда транзакция «фиксируется»; либо, если возможность выполнить все действия отсутствует, например, в случае сбоев, транзакция «откатывается» назад, а БД остается в исходном состоянии.
- Согласованность. Предполагается, что в результате выполнения транзакции система переходит из одного корректного состояния в другое.

- **Изолированность.** При выполнении транзакции данные могут временно находиться в несогласованном состоянии. Такие данные не должны быть видны другим транзакциям, пока изменения не будут завершены (т.е. пока вес модификации не будут формально зафиксированы).
- **Долговечность.** Если транзакция зафиксирована, то ее результаты должны быть долговечными. Состояния всех объектов сохраняются даже в случае аппаратных или системных сбоев.

Таблицы проектируемой БД рекомендуется семантически разделить на три типа:

- таблицы с данными (геоданные с их атрибутивная часть);
- таблицы со словарями (различные словари, которые предназначены для приведения исходных разнородных данных к единому виду, применяются для фильтрации и классификации данных);
- служебные таблицы (используются с целью формализации отношений между объектами системы).

При введении понятия информационного ресурса необходимо обосновать состав, структуру, свойства и связи информационных ресурсов, которые адекватно описывают предметную область каждой отдельно взятой системы и из которой состоит геоинформационная система.

Следующими действиями при проектировании подобных систем должно являться:

1. Описание распределения ресурсов по местам сбора данных;
2. Описание обмена информацией между объектами автоматизации и с внешними системами в виде ресурсов;
3. Описание доступа пользователей к ресурсам, что позволяет продемонстрировать характер взаимодействия пользователя с системой;
4. Создание под хранение информационных ресурсов реляционной схемы базы данных.

Данные виды описаний можно рассматривать в контексте задач инфологического моделирования предметной области разрабатываемой системы. Для проведения такого моделирования существуют различные математические модели, методы и методики, в том числе и аппарат моделей сущность-связь или аппарат объектно-ориентированного анализа. В связи с этим разработаны и повсеместно применяются CASE-средства, которые могут обеспечить работу с данными математическими аппаратами. Так топология связей сущность-связь позволяет вычислить группировки сущностей, связанные между собой, которые рассматриваются в контексте единого целого как объекты предметной области. Такую информацию можно получить из анализа родительских и дочерних связей.

Немаловажным остается тот факт, что за основу геоинформационной системы управления территориями взята именно функциональность и наглядность геоинформационных систем. Следовательно, основное место в методике проектирования данных систем занимают технологии создания геопространственного поля.

Часто для публикации ГИС-систем в интернете используются готовые решения, обладающие наполнением и представляющие собой готовые геопространственные поля (картографические основы). На рынке присутствуют следующие картографические основы: GoogleMaps, OpenStreetMap, Яндекс-карты, kosmosnimki.ru, Карты@mail.ru, Рамблер-карты. Выбор геопространственного поля предлагается осуществлять по следующим параметрам:

- Актуальность карт. Должны использоваться только актуальные карты. Дороги, реки, здания, пешеходные дорожки и другие объекты, отображаемые на карте, должны соответствовать действительному представлению этих объектов;
- Информационная наполненность. Помимо графического отображения, картографическая основа должна позволить получить краткую семантическую справку по каждому из объектов;
- Отсутствие «избыточности» наполнения. Визуально картографическая основа должна иметь такую топографическую точность, которая бы удобно воспринималась как схематическое изображение объекта и имела минимум погрешностей. Не должна использоваться карта, наполненная избыточной информацией, то есть на карте должны присутствовать только те объекты, которые имеют значение при перемещении;
- Наличие возможностей API (application programming interface). Геопространственное поле должно обладать возможностями API (интерфейс прикладного программирования) для использования во внешних программных продуктах.

Требования предъявляются так же и к распределенности и разнородности баз данных, которые можно предъявить к созданию подобной геоинформационной системы заключаются в разделении системы на ядро и надстройку, при этом ядро должно быть универсально для выполнения любых функций ГИС.

Любое проектирование начинается с технических и организационных вопросов, которые перерастают в системное проектирование, параллельно с которым выполняется проектирование программного обеспечения (рисунок 32).

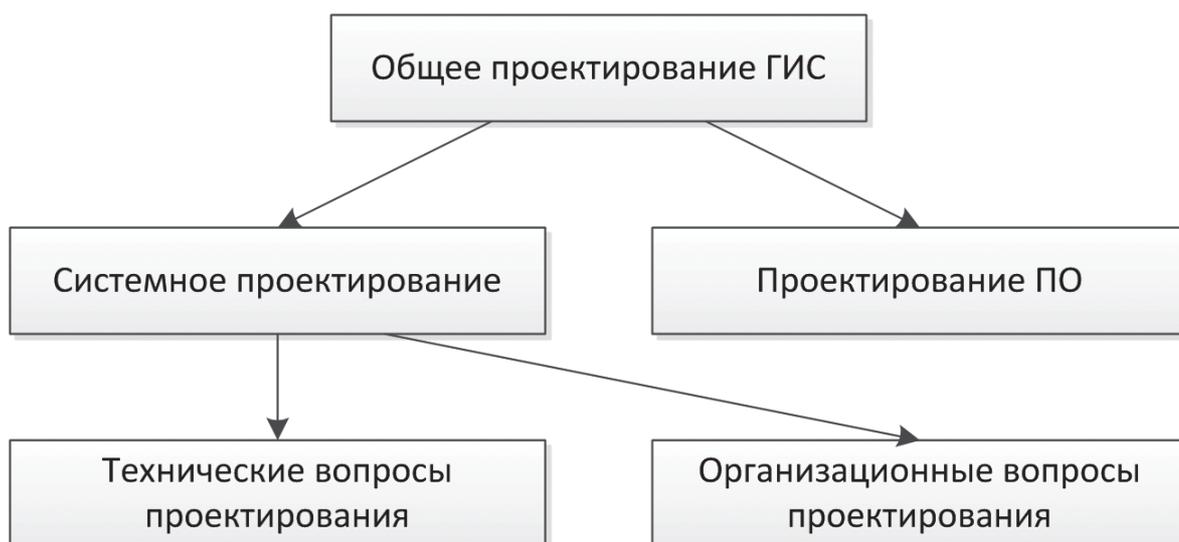


Рисунок 32. Универсальные этапы проектирования геоинформационных систем

Техническое проектирование, при реализации геоинформационной системы управления территорией на основе разнородных баз данных должно содержать в себе следующие функции:

1. Отладка работы системы по предъявляемым требованиям;
2. Законченность оформления требований к содержанию и функционированию системы по отношению к пользователю;
3. Утверждение форматов данных, с которыми будет работать система, из чего вытекает анализ входящих в систему гетерогенных баз данных;
4. Формирование требований к системе, таких как гибкость и надежность;
5. Выбор программного обеспечения осуществления проектируемой системы;
6. Структуризация потоков информации в информационной системе.

После определения функций в техническом проектировании производится моделирование. Строятся следующие модели:

- Модель потребностей данных, где отображается взаимосвязь наличия данных и возможностей проведения анализа;
- Модель потребностей приложений, где система рассматривается как набор функций, которые она может выполнить.

При организационном проектировании разработчик должен предложить следующие решения для возникающих перед ним задач:

1. Предложения по превращению проектируемой системы в рентабельную;
2. Расчёт стоимости получаемой с помощью системы информации;
3. Предложения по упрощению адаптации геоинформационной системы;
4. Предложения как удешевить и упростить поддержку ГИС.

Особенности проектирования подобной системы заключаются в формировании сложного интерфейса работы с данными, в неизвестности окончательного набора функций разрабатываемой системы до окончания разработки, в обеспечении расширяемости системы, в обеспечении поддержки модульности и скриптового языка программирования, в больших размерах проектируемой системы [38].

Для уменьшения объема данных необходимых для проведения единовременного анализа можно произвести фрагментацию данных по следующим параметрам:

1. Фрагментация по координатам;
2. Фрагментация по наиболее часто используемым объектам.

Также можно предложить ограничение доступа к определенным зонам карт. Проектирование геоинформационной системы управления территориями должно состоять как минимум из четырёх основных шагов (рисунок 33).



Рисунок 33. Шаги проектирования геоинформационной системы управления на основе гетерогенных баз данных

Начальное представление заключается в описании разрабатываемой системы, анализе осуществимости системы, возможно ли вообще создать такую геоинформационную систему с

требуемыми параметрами. Сюда входят описание не только предметной области системы, но и описание предметных областей включенных в неё гетерогенных баз данных, а также описанные выше информационные ресурсы системы, организованные в виде реляционной базы данных.

Концептуальное представление основано на определении требований и проектировании виртуальной интегрированной базы данных. Распределенность и разнородность баз данных включающихся в геоинформационную систему предусматривает место объединения и приведения к общему формату получаемой информации. Виртуальная интегрированная база данных позволяет создать на основе разноформатных локальных баз данных обобщенную, приведенную к единому формату виртуальную БД, которая может быть представлена в виде реляционной для упрощённого подключения ГИС-приложения.

Детальное представление системы заключается в отображении разрабатываемого проекта на конкретный ГИС-пакет, если используются готовые программные продукты геоинформационных систем. Или создание собственного приложения для отображения представленной информации в виртуальной базе данных и реализации необходимых функций.

Реализация системы предусматривает подключение базы данных к ГИС-приложению и программирование всех предусмотренных для выполнения функций системы.

Поэтапное проектирование позволяет избежать ошибок и предусмотреть ситуации прекращения функционирования подобной системы [53-55, 63, 64].

Возможно выделить пять основных этапов процесса проектирования ГИС.

1. Анализ системы поддержки принятия решений. Процесс начинается с определения всех типов решений, для принятия которых требуется информация. Должны быть учтены потребности каждого уровня и функциональной сферы.
2. Анализ информационных требований. Определяется, какой тип информации нужен для принятия каждого решения.
3. Агрегирование решений, т.е. группировка задач, в которых для принятия решений требуется одна и та же или значительно перекрывающаяся информация.
4. Проектирование процесса обработки информации. На данном этапе разрабатывается реальная система сбора, хранения, передачи и модификации информации. Должны быть учтены возможности персонала по использованию вычислительной техники.
5. Проектирование и контроль над системой. Важнейший этап – это создание и воплощение системы. Оценивается работоспособность системы с разных позиций, при необходимости осуществляется корректировка. Любая система будет иметь недостатки, и поэтому её необходимо делать гибкой и приспособляемой.

Прежде всего, это формулирование требований к используемым информационным продуктам и выходным материалам, получаемым в результате обработки. Сюда можно отнести требования к распечатке карт, таблиц, списков, документов; к поиску документов и т.д. В результате должен быть создан документ с условным названием «Общий список входных данных».

Следующий шаг – определение приоритетов, очередности создания и основных параметров (территориального охвата, функционального охвата и объёма данных) создаваемой системы. Далее устанавливают требования к используемым данным с учётом максимальных возможностей их применения.

Итак, мы выбрали метод, которым будем руководствоваться при проектировании геоинформационной системы.

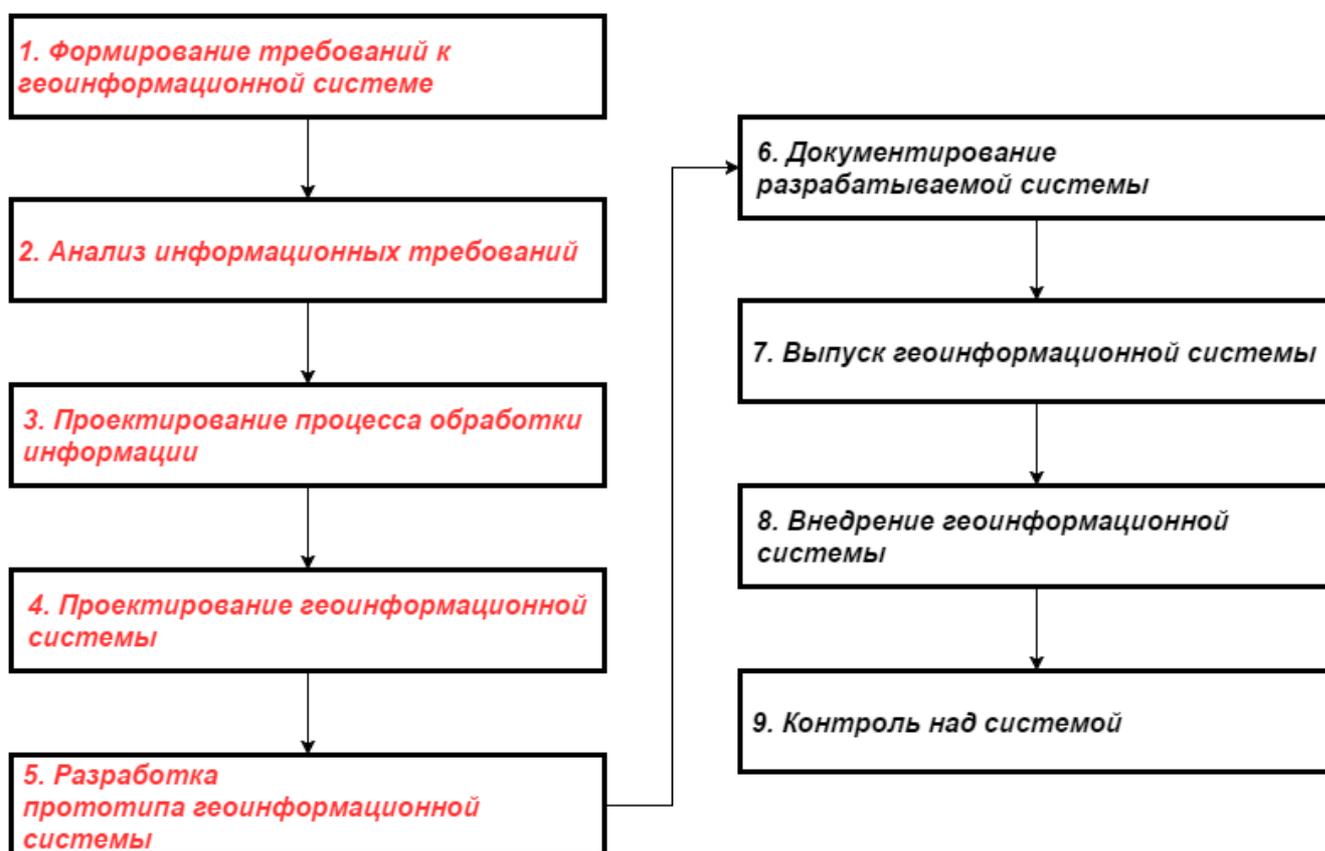


Рисунок 34. Этапы проектирования ГИС

Теперь нам необходимо спланировать комплекс работ по созданию нашей системы в соответствии с типовыми этапами разработки ГИС, краткая характеристика которых приведена в таблице 13.

Таблица 13 - Этапы проектирования ГИС

№	Наименование этапа	Описание
1	Формирование требований к геоинформационной системе.	<p>1.1. Формируются требования пользователя к системе.</p> <p>1.2. Определяются основные задачи геоинформационной системы.</p> <p>1.3. Определяется цель создания геоинформационной системы (ГИС).</p> <p>1.4. Проводится декомпозиция задач по модулям и определяются функции с помощью которых решаются эти задачи.</p> <p>1.5. Проводится описание функций на языке производственных требований (описание процессов предметной области).</p> <p>Результат:</p> <ul style="list-style-type: none"> • Обоснование необходимости создания геоинформационной системы. • Концептуальная модель ГИС, состоящая из описания предметной области, основной цели, функций и задач системы.
2	Анализ информационных требований.	<p>2.1. Изучение объекта, территории.</p> <p>2.2. Проведение научных исследований.</p> <p>2.3. Определяются типы решений, для принятия которых требуется информация (должны быть учтены потребности каждого уровня и функциональной сферы).</p> <p>2.4. Определяется, какой тип информации необходим для принятия каждого типа решения.</p> <p>2.5. Формулируются требования к входным источникам информации, с учетом разнородности данных.</p> <p>2.6. Проводится описание функций на языке функциональных требований (описание обрабатываемых данных).</p> <p>Результат:</p> <p>Концептуальная модель ГИС, состоящая из описания типов решений, ресурсов и потоков разнородных данных.</p>

Продолжение таблицы 13 - Этапы проектирования ГИС

3	Проектирование процесса обработки информации.	<p>3.1. Формулируются требования к выходным данным для поддержки принятия решений.</p> <p>3.2. Разрабатывается система сбора, хранения, передачи и модификации информации, с учетом разнородности данных.</p> <p>3.3. Разрабатываются таблицы данных, словарей, служебных таблиц.</p> <p>3.4. Подключается система обработки разнородных данных (включая надстройки в зависимости от используемых данных).</p> <p>3.5. Проводится описание функций на языке технических требований (аппаратное, программное, лингвистическое обеспечение ГИС).</p> <p>Результат:</p> <ul style="list-style-type: none"> • Концептуальная модель ГИС, состоящая из описания требований и ограничений к технической реализации системы. • Аппаратно-технический состав создаваемой ГИС.
4	Проектирование геоинформационной системы.	<p>4.1. Формулируются требования к компонентам системы.</p> <p>4.2. Проектируется архитектура ГИС.</p> <p>4.3. Проектируется логическая схема ГИС.</p> <p>4.4. Проектируется геопространственное поле.</p> <p>4.5. Проектируется интерфейс системы.</p> <p>Результат:</p> <p>Техническое задание на разработку геоинформационной системы</p>
5	Разработка прототипа геоинформационной системы.	<p>5.1. Выбирается лингвистическое обеспечение (среда разработки - инструментарий)</p> <p>5.2. Разработанная логическая схема воплощается в реальные объекты, при этом логические схемы реализуются в виде объектов базы данных, а функциональные схемы - в пользовательские формы и приложения.</p> <p>5.3. Проводится разработка программного обеспечения.</p> <p>5.4. Проводится тестирование системы.</p> <p>Результат:</p> <p>Разработан работоспособный прототип ГИС.</p>

Продолжение таблицы 13 - Этапы проектирования ГИС

6	Документирование разрабатываемой системы.	<p>6. Выпуск сопроводительной документации:</p> <ul style="list-style-type: none"> • Руководство разработчика системы • Руководство пользователя системы • Техническая документация • Другие виды сопроводительных документов <p>Результат: Комплект документации: разработчика, администратора, пользователя геоинформационной системы.</p>
7	Выпуск геоинформационной системы.	<p>7.1. Реализация рабочей версии системы.</p> <p>7.2. Осуществляется корректировка: информационного, аппаратного, программного и методического обеспечения.</p> <p>Результат:</p> <ul style="list-style-type: none"> • Рабочая версия ГИС. • Оптимальный пакет сопроводительных документов ГИС.
8	Внедрение геоинформационной системы	<p>8.1. Ввод рабочей версии геоинформационной системы в эксплуатацию.</p> <p>8.2. Обучение персонала работе с системой.</p> <p>Результат: Внедрение рабочей версии геоинформационной системы в работу.</p>
9	Контроль над системой.	<p>9. Сопровождение геоинформационной системы.</p> <p>Результат:</p> <ul style="list-style-type: none"> • Поддержание актуальной версии системы • Выявление и устранение недочетов системы

Методика проектирования геоинформационной системы поддержки принятия решений для управления территориями на основе распределенных разнородных баз данных позволит решить задачи построения любой геоинформационной системы на основе разнородных и распределенных баз данных, не зависимо от выбранной территории, что существенно продвигает в развитии управленческую деятельность в регионах. Представленная методика создания информационного ресурса на основе ГИС-технологий, отражает степень разработанности вопроса.

3.2. Верификация модели обработки и управления разнородными данными на примере задачи оптимизации регулирования отпуска тепла на ТЭЦ в геоинформационной системе

Во второй главе была рассмотрена методика оптимизации использования метеорологической информации для предоставления оптимального управленческого варианта решения на примере оптимизации отпуска тепла на ТЭЦ. Автоматизация методических прогнозов решает ряд производственных задач, в частности - регулировки рабочей мощности теплоисточника, а так же поддержания необходимой температуры теплоносителя, но главное обеспечение оптимальной экономии топлива, все это достигается за счет выбора оптимального управленческого решения.

Реализация данной методики в геоинформационной системе требует формирования технического задания на разработку системы. В связи с тем, что стандарт на разработку ГИС еще не введен в эксплуатацию и не присутствует в общем доступе, то в качестве регламента разработки будет использована предложенная в работе методика проектирования геоинформационной системы для поддержки принятия решений.

Исходя из предложенной методики ниже представлены основные этапы формирования необходимой информации для разработки ГИС, этапы 6-9 в диссертационной работе не рассматривались.

1 Этап: Формирование требований к геоинформационной системе.

1.1 Требования пользователя к системе. Система должна:

- Упростить получение варианта управленческого решения;
- Обрабатывать метеорологическую информацию;
- Производить анализ метеорологической информации;
- Производить экономические расчеты исходя из выбора решения;
- Обладать интуитивно понятным интерфейсом;
- Выполнять поставленные задачи.

1.2 Задачи выполняемые ГИС:

- Автоматизация обработки разнородных данных;
- Обработка разнородных данных с использованием методики управления данными для поддержки принятия управленческих решений;

- Создание единой базы данных, как источника актуальных и достоверных сведений обо всех пространственных данных, что способствует повышению точности принимаемых решений;
- Обеспечения актуальной и комплексной информацией о пространственных объектах из единого источника для оперативного исследования, оценки и обоснования управленческих решений;
- Повышение инвестиционной привлекательности системы за счет представления актуальной и достоверной комплексной информации о пространственных объектах;
- Создание условий, обеспечивающих доступ к обобщенным пространственным данным и их эффективное использование;
- Упрощение процедур, связанных с получением и обработкой разнородных данных через Интернет.

1.3 Целью создания ГИС является - создание эффективного средства информационной поддержки повышения процедуры принятия решений, обеспечиваемой функциями анализа пространственных данных и представлением результатов научных исследований в наглядном и удобном для восприятия виде.

1.4 Декомпозиция задач по модулям:

- Модуль обработки разнородных данных;
- Модуль автоматизации подготовки варианта управленческого решения;
- Модуль ввода данных;
- Модуль вывода данных;
- Модуль администрирования.

Ряд этапов уже был частично рассмотрен в контексте диссертационной работы, в частности анализ информационных требований представлен во второй главе. Далее в работе будут рассмотрены остальные этапы проектирования геоинформационной системы поддержки принятия управленческих решений на основе обработки разнородных данных.

3.2.1 Формирование базы разнородных метеорологических данных по территории Санкт-Петербурга

Наличие баз статистических данных является необходимым условием обеспечения точности любого расчета. Формирование единой актуальной упорядоченной системы хранения статистических данных является основой для дальнейшей обработки и исследования. Информационные данные для геокодирования были предоставлены в разнородных форматах:

1. База данных (массив) по фактической температуре воздуха (ВНИИГМИ-МЦД). Формат данных – txt.
2. База данных прогнозов (Северо-Западное УГМС). Формат – xls.

Используя полученные разнородные данные с учетом предложенной в работе методики управления метеорологическими данными применительно к задаче оптимизации регулирования отпуска тепла на ТЭЦ, схема управления будет иметь следующий вид (Рисунок 35).

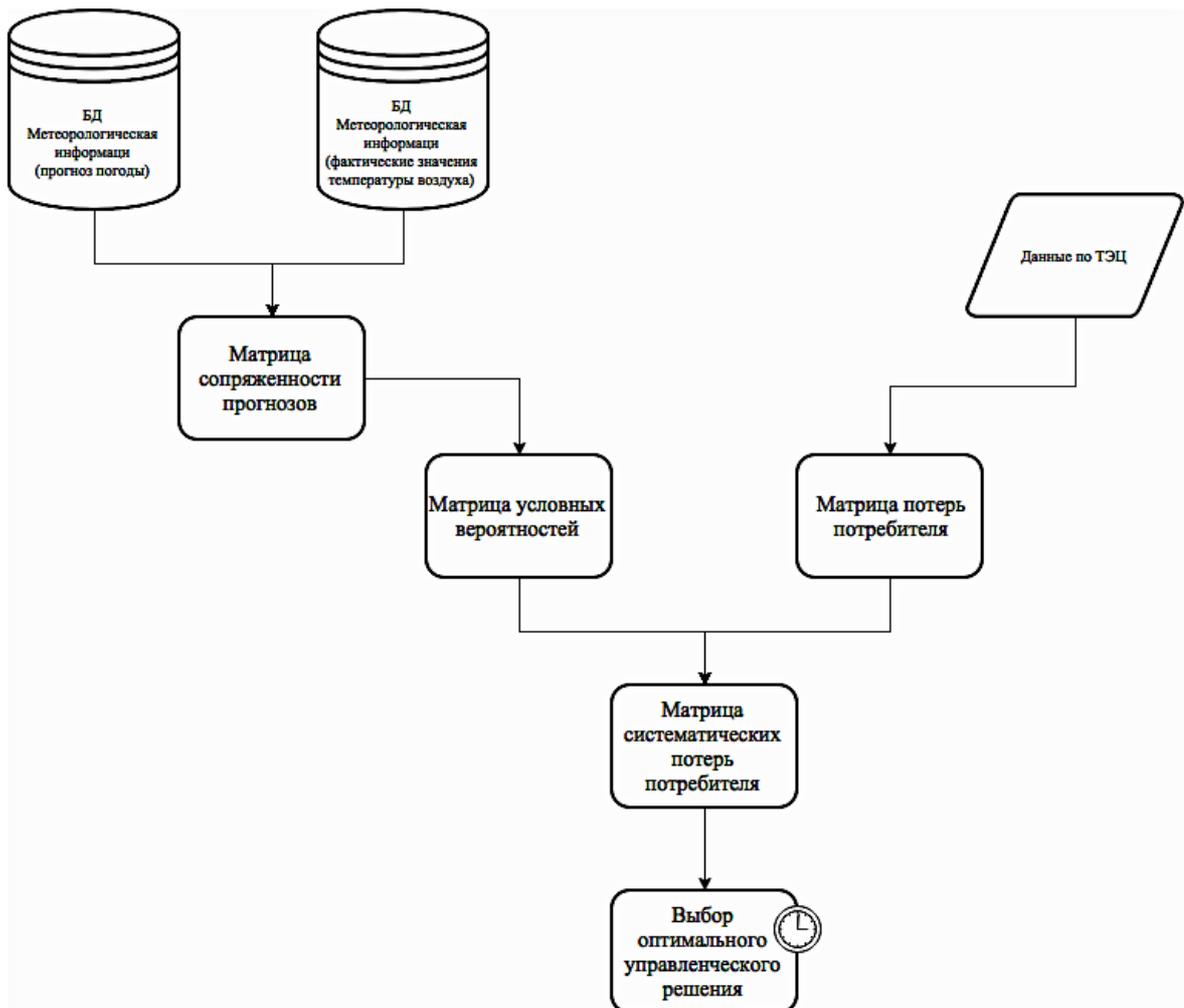


Рисунок 35. Схема управления разнородными метеорологическими данными

Исходя из поставленной задачи выборка данных будет осуществляться со станции - Санкт-Петербург (рисунок 36). На сервере ВНИИГМИ-МЦД доступ к массиву данных, выборке данных по интересующим пользователя станциям, их просмотр и копирование обеспечивается специализированной технологией (<http://meteo.ru/it/178-aisori>). Авторы канд. физ.-мат. наук В.М. Веселов и канд. техн. наук И.Р. Прибыльская. Перейдя по ссылке и пройдя регистрацию в специализированном ПО, необходимо выбрать станцию.

Выбор данных

English Раздел БД Сутки Источник данных TTTR - Температура и осадки Описание

Подготовьте список станций

Выберите станцию (станции)
Сортировать по: Индексу Названию
Выделить все

20046	Им. Э. Т. Кренкеля, ГМО	Россия
20069	Визе	Россия
20087	Голомянный	Россия
20107	Баренцбург	Россия
20289	Русский	Россия
20292	Им. Е. К. Федорова, ГМО	Россия
20476	Стерлегова	Россия
20667	Им. М. В. Попова	Россия
20674	Диксон	Россия
20744	Малые Кармакулы	Россия
20891	Хатанга	Россия
20946	Им. Е. К. Федорова	Россия
20982	Волочанка	Россия
21432	Котельный	Россия
21611	Терпий-Тумса	Россия
21647	Шалаурова	Россия
21802	Саскылах	Россия
21824	Тикси	Россия
21908	Джалинда	Россия
21921	Кюсюр	Россия

Выбранные станции (1/600)
Сортировать по: Индексу Названию

26063	Санкт-Петербург	Россия
-------	-----------------	--------

Справка Выход OK Очистить

Рисунок 36. Выборка данных с сервера ВНИИГМИ-МЦД

В таблице 14 представлены координаты метеорологической станции в г. Санкт-Петербург, где производились замеры показателей. Высота над уровнем моря: 11.2 м.

Таблица 14 - Координаты метеорологической станции

Индекс ВМО	Наименование станции	Координаты станции		Начало наблюдений	Примечание
		широта	долгота		
26063	Санкт-Петербург	59° 58'	30° 18'	1834	Перенос в 07.1933 г. на 5км к ССЗ.

Определение условий отбора данных представлено на рисунках 37-38.

Подготовка запроса

Сейчас работают 2

Список элементов

Индекс ВМО
 Год
 Месяц
 День
 Общий признак качества температур
 Минимальная температура воздуха
 Средняя температура воздуха
 Максимальная температура воздуха
 Количество осадков

Условия отбора данных

Имя признака	От (>=)	До (<=)
Год	2000	2015
Месяц	1	2
День		

Использовать разделитель ";"

Рисунок 37. Выборка данных с сервера ВНИИГМИ-МЦД

Выбор данных

English
Раздел БД
Источник данных

Подготовьте список станций

Выберите станцию (станции)
 Сортировать по: Индексу Названию

20046	Им.Э.Т.Кренкеля, ГМО	Россия
20069	Визе	Россия
20087	Голомянный	Россия
20107	Баренцбург	Россия
20289	Русский	Россия
20292	Им.Е.К.Федорова, ГМО	Россия
20476	Стерлегова	Россия
20667	Им.М.В.Попова	Россия
20674	Диксон	Россия
20744	Малые Кармакулы	Россия
20891	Хатанга	Россия
20946	Им.Е.К.Федорова	Россия
20982	Волочанка	Россия
21432	Котельный	Россия
21611	Терпий-Тумса	Россия
21647	Шалаурова	Россия
21802	Саскылах	Россия
21824	Тикси	Россия
21908	Джалинда	Россия
21921	Кюсюр	Россия

Выбранные станции (1/600)
 Сортировать по: Индексу Названию

26063	Санкт-Петербург	Россия
-------	-----------------	--------

Рисунок 38. Выборка данных с сервера ВНИИГМИ-МЦД

Результат критериев выборки статистических данных представлен на рисунке 39.

Результат выборки (показан фрагмент, полный объем 49 kB (упаковано в 8 kB) можно получить при нажатии кнопки "Открыть" или "Упаковать и сохранить")

Открыть Упаковать и сохранить Обозначения

Новые данные Новый запрос Выход Справка Комментарий Комментарий(95)

1	2	3	4	5	6	Q	7	Q	8	Q	9	D	Q
26063;2001;	1;	6;0;	-2.5;0;	-1.0;0;	0.7;0;	0.0;2;0							
26063;2001;	1;	7;0;	-2.6;0;	-0.7;0;	1.6;0;	7.1;0;0							
26063;2001;	1;	8;0;	-0.7;0;	0.9;0;	2.4;0;	1.2;0;0							
26063;2001;	1;	9;0;	0.4;0;	1.1;0;	2.1;0;	1.4;0;0							
26063;2001;	1;	10;0;	-1.2;0;	-0.5;0;	0.4;0;	0.7;0;0							
26063;2001;	1;	11;0;	-2.5;0;	-1.1;0;	-0.4;0;	7.9;0;0							
26063;2001;	1;	12;0;	-4.4;0;	-2.8;0;	-1.2;0;	1.5;0;0							
26063;2001;	1;	13;0;	-7.4;0;	-5.5;0;	-2.8;0;	0.0;3;0							
26063;2001;	1;	14;0;	-4.6;0;	-1.9;0;	0.4;0;	0.4;0;0							
26063;2001;	1;	15;0;	-0.8;0;	0.1;0;	0.6;0;	0.0;3;0							
26063;2001;	1;	16;0;	0.0;0;	0.4;0;	0.9;0;	0.0;2;0							
26063;2001;	1;	17;0;	-2.0;0;	-0.8;0;	0.5;0;	0.0;2;0							
26063;2001;	1;	18;0;	-3.3;0;	-2.3;0;	-0.5;0;	0.0;3;0							
26063;2001;	1;	19;0;	-2.0;0;	-1.4;0;	-0.9;0;	0.0;3;0							
26063;2001;	1;	20;0;	-2.2;0;	-1.2;0;	-0.1;0;	0.0;3;0							
26063;2001;	1;	21;0;	-10.3;0;	-7.6;0;	-1.8;0;	0.0;2;0							
26063;2001;	1;	22;0;	-10.3;0;	-8.7;0;	-6.9;0;	0.0;2;0							
26063;2001;	1;	23;0;	-12.3;0;	-11.0;0;	-9.0;0;	0.2;0;0							
26063;2001;	1;	24;0;	-11.8;0;	-10.2;0;	-8.0;0;	0.0;3;0							
26063;2001;	1;	25;0;	-8.2;0;	-6.1;0;	-3.6;0;	0.0;2;0							
26063;2001;	1;	26;0;	-4.7;0;	-2.4;0;	0.6;0;	1.6;0;0							
26063;2001;	1;	27;0;	-1.3;0;	-0.3;0;	0.5;0;	0.5;0;0							
26063;2001;	1;	28;0;	-0.5;0;	1.0;0;	1.9;0;	0.0;3;0							
26063;2001;	1;	29;0;	-0.5;0;	0.4;0;	1.4;0;	0.0;3;0							
26063;2001;	1;	30;0;	-2.6;0;	-0.8;0;	0.1;0;	1.5;0;0							
26063;2001;	1;	31;0;	-5.0;0;	-3.8;0;	-2.2;0;	0.7;0;0							
26063;2001;	2;	1;0;	-6.7;0;	-4.4;0;	-3.1;0;	0.2;0;0							

Рисунок 39. Интерфейс обработки запроса

Формат представления данных по запросу выборки данных с сервера ВНИИГМИ-МЦД по умолчанию txt (текстовый). При сохранении выборки данных пользователь получает два файла формата txt: файл с данными и файл с метаданными, описывающий формат записи в файлах данных.

Перечень станций России для новой версии архива составлен на основании списка станций Росгидромета, включенных в глобальную сеть наблюдений за климатом (утвержденного Руководителем Росгидромета 25 марта 2004г.) и списка метеорологических станций Росгидромета, подготовленного в Главной Геофизической Обсерватории им. А.И. Воейкова (исп. Зав. ОМРЭИ ГГО В.И.Кондратюк). Список станции и информация по ним содержится в наборе «Каталог станций».

Для некоторых станций информация заканчивается более ранними годами, так как:

- Станции закрыты (как на территории России, так на территории независимых государств, бывших республик СССР);
- Данные по станциям не представлены для подготовки “Метеорологического ежемесячника станций стран содружества независимых государств, часть 1 “Ежедневные данные””

Данные по станциям, находящимся на территории независимых государств – бывших республик СССР, публикуются в соответствии с решениями Исполкома межгосударственного совета по метеорологии стран СНГ об утверждении списка станций для международного обмена (Алма-Аты, 1993 год).

Массив состоит из 600 файлов данных в формате ASCII с именами вида IIII.dat, где III – синоптический индекс станции (индекс ВМО). Записи в файлах данных упорядочены по возрастанию ключевых элементов: год, месяц, день. Описание формата записи приведено в таблице 15.

Таблица 15 - Формат записи в файлах данных

Номер поля	Позиция	Длина Поля	Наименование поля	Примечание
1	1-5	5	Индекс ВМО станции	Фиксировано для файла
	6	1	Пробел	
2	7-10	4	Год	
	11		Пробел	
3	12-13	2	Месяц	
	14	1	Пробел	
4	15-16	2	День	
	17	1	Пробел	
5	18	1	TFLAG - групповой признак качества для показателей температуры воздуха	См. таблицу 3.1.1.2
	19	1	Пробел	
6	20-24	5	TMIN-минимальная температура воздуха за сутки	В градусах Цельсия с точностью 0.1 градуса
	25	1	Пробел	
7	26	1	QTMIN- признак качества для TMIN	См. таблицу 3.1.1.3
	27	1	Пробел	
8	28-32	5	TMEAN - среднесуточная температура воздуха	В градусах Цельсия с точностью 0.1 градуса
	33	1	Пробел	
9	34	1	QTMEAN - признак качества для TMEAN	См. таблицу 3.1.1.3
	35	1	Пробел	
10	36-40	5	TMAX - максимальная температура воздуха за сутки	В градусах Цельсия с точностью 0.1 градуса
	41	1	Пробел	
11	42	1	QTMAX - признак качества для TMAX	См. таблицу 3.1.1.3
	43	1	Пробел	
12	44-48	5	R - суточная сумма осадков	В миллиметрах с точностью 0.1 мм
	49	1	Пробел	

Продолжение таблицы 15 - Формат записи в файлах данных

13	50	1	CR - дополнительный признак к R	См. таблицу 3.1.1.4
	51	1	Пробел	
14	52	1	QR - признак качества для R	См. таблицу 3.1.1.5
	53-54	2	Символ конца строки '\n'	

TFLAG - групповой признак качества для показателей температуры воздуха введен в состав записи с целью проинформировать пользователей о наличии ошибочных данных в архивных источниках, на основании которых осуществлялось формирование массива (таблицы 16-18).

Таблица 16 - Значения TFLAG (группового признака качества для характеристик температуры воздуха)

TFLAG	Кодируемая ситуация			
	QTMIN	QTMEAN	QTMAX	Условия
0	0	9	9	Без условий
	9	0	9	
	9	9	0	
	0	0	9	TMIN<TMEAN
	0	9	0	TMIN< TMAX
	9	0	0	TMEAN < TMAX
	0	0	0	TMIN < TMEAN < TMAX
1	0	0	9	Если нарушено хотя бы одно из соотношений: TMIN<TMEAN TMEAN<TMAX TMIN<TMAX
	0	9	0	
	9	0	0	
	0	0	0	
9	9	9	9	-

Таблица 17 - Значения QTMIN, QTMEAN, QTMAX, QR

Q- флаги	Значения Q-флагов
0	Значение достоверно
1	Значение не согласуется с данными архива срочных наблюдений
9	Значение забраковано или наблюдения не проводились.

CR	
0	измеренное количество осадков 0,1 мм и более
1	осадки измерены за несколько дней
2	измерения осадков производились, но осадков не было (R = 0)
3	наблюдались только следы осадков (< 0,1 мм) R = 0
9	значение забраковано или наблюдения не проводились.

Фактические значения температуры воздуха за рассматриваемый период получены по данным ВНИИГМИ-МЦД. Продолжительность отопительного сезона для Санкт-Петербурга определялась с октября по апрель. Исходные значения (прогностические $\bar{t}_{пр}$ и фактические $\bar{t}_{ф}$) среднесуточной температуры воздуха по Санкт-Петербургу предварительно были табулированы и представлены в таблице 19.

Таблица 19 – Прогностические и фактические значения среднесуточной температуры воздуха Санкт-Петербурга за отопительный сезон 2013-2014 года.

Число		Октябрь	Ноябрь	Декабрь	Январь	Февраль	Март	Апрель	Май
1	прогноз	4	6	-1	0	-11	1	0	6
	Факт	4	7	-1	2	-8	0	2	6
2	прогноз	5	7	-4	-2	-6	1	1	6
	Факт	5	8	-5	-2	-6	2	-1	4
3	прогноз	7	8	-2	-3	-6	2	1	6
	Факт	9	8	-1	-1	-4	3	3	6
4	прогноз	8	7	0	-1	-2	2	0	6
	Факт	8	6	0	1	-2	2	1	7
5	прогноз	10	6	-2	1	-3	2	3	5
	Факт	10	7	-2	3	-4	2	3	7
6	прогноз	9	5	0	0	-6	2	3	5
	Факт	9	7	1	1	-5	1	3	5
7	прогноз	10	5	-2	2	-6	3	3	6
	Факт	11	6	0	3	-5	3	3	5
8	прогноз	11	5	-4	3	-2	3	3	7
	Факт	12	6	-2	3	0	4	2	5
9	прогноз	12	6	-7	4	0	3	1	11
	Факт	13	6	-7	4	3	4	2	11
10	прогноз	12	6	-6	2	1	5	2	15
	Факт	12	6	-8	4	2	6	2	16
11	прогноз	10	4	-1	-4	1	3	4	12
	Факт	10	6	-2	-5	2	3	6	12
12	прогноз	8	2	4	-6	1	3	5	13
	Факт	8	3	3	-6	1	3	7	11

Продолжение таблицы 19 – Прогностические и фактические значения среднесуточной температуры воздуха Санкт-Петербурга за отопительный сезон 2013-2014 года.

13	прогноз	9	5	3	-8	2	5	7	13
	Факт	10	6	5	-8	2	4	8	16
14	прогноз	5	4	-4	-9	1	4	7	10
	Факт	4	5	-3	-10	2	3	8	9
15	прогноз	7	3	-5	-10	1	2	7	8
	Факт	7	4	-4	-9	1	2	8	9
16	прогноз	4	6	-2	-11	0	-1	6	9
	Факт	5	6	0	-8	1	-1	6	10
17	прогноз	2	6	3	-11	1	-3	7	13
	факт	4	8	5	-10	2	-4	7	14
18	прогноз	4	4	1	-10	2	-5	10	17
	факт	6	3	2	-12	1	-4	11	18
19	прогноз	2	3	1	-12	0	-4	11	23
	факт	3	3	1	-13	0	-1	10	24
20	прогноз	2	4	0	-14	-2	-3	11	25
	факт	4	5	2	-13	-1	-3	10	25
21	прогноз	2	5	1	-14	-3	3	12	22
	факт	2	5	3	-9	-2	4	14	23
22	прогноз	1	5	2	-11	-1	6	11	17
	факт	1	6	4	-10	1	6	13	18
23	прогноз	4	6	3	-13	2	4	5	20
	факт	4	6	5	-18	3	4	8	21
24	прогноз	9	4	3	-14	3	5	5	22
	факт	10	4	2	-14	6	5	5	23
25	прогноз	8	-1	4	-11	3	5	8	23
	факт	9	-1	4	-11	3	4	8	23
26	прогноз	7	-2	2	-10	1	3	9	20
	факт	7	-3	3	-8	2	4	11	20
27	прогноз	9	1	3	-12	1	4	10	12
	факт	9	1	3	-11	2	3	11	7
28	прогноз	11	3	3	-16	0	2	11	8
	факт	11	4	4	-12	2	3	11	9
29	прогноз	10	-1	4	-17	0	1	11	12
	факт	10	1	4	-15	0	3	13	11
30	прогноз	9	-4	3	-15	0	2	9	14
	факт	9	-4	4	-17	0	4	8	14
31	прогноз	6	0	2	-16	0	1	0	15
	факт	6	0	3	-14	0	1	0	16

В результате исследования выполненного коллективом авторов (д.т.н., профессор Истомин Е.П., к.т.н. Соколов А.Г., к.г.н. Фокичева А.А., Степанов С.Ю., Петров Я.А.) было получено свидетельство о государственной регистрации базы данных «База данных метеорологических параметров» №2016620986.

3.2.2 Разработка геоинформационной системы с использованием разнородных пространственно-распределенных метеорологических данных

На этапе проектирования геоинформационной системы была реализована модель, которая обеспечивает наглядное представление возможных вариантов использования разрабатываемой системы поддержки принятия решений, а так же представляет структуру системы, связи между структурными классами и логику действия программы. Конечная модель системы разработана в виде группы диаграмм на унифицированном языке визуального моделирования UML [74].

Существует достаточно много CASE-инструментов моделирования и проектирования систем и баз данных (не только с помощью UML). В данном учебном пособии для примера моделирования системы выбран программный инструмент моделирования StarUML. Данная программная платформа имеет свободную лицензию и доступна для установки с официального сайта StarUML. StarUML поддерживает одиннадцать различных типов диаграмм, принятых в нотации UML 2.0, а также подход MDA (модельно-настраиваемая архитектура), предлагает настройку параметров пользователя для адаптации среды разработки, поддерживает расширения, предоставляет различного рода модули, расширяющие возможности StarUML [50, 59].

Построение основных диаграмм будет достаточным для понимания структуры системы. Поведение системы (т.е. функциональность, которую она обеспечивает) описывают с помощью функциональной модели (рисунок 40-41), которая отображает системные прецеденты (use cases, случаи использования), системное окружение (действующих лиц, актеров, actors) и связи между ними (use cases diagrams).

- Актером (действующее лицо, actor) называется любой объект, субъект или система, взаимодействующая с моделируемой системой извне.
- Вариант использования (прецедент, use case) — описание множества последовательности действий (включая варианты), выполняемых системой для того, чтобы актер мог получить определенный результат
- Связи и взаимоотношения, существующие между элементами модели, в UML описываются с помощью отношений, изображаемых на диаграммах.

В основном на диаграммах прецедентов используются следующие типы отношений:

- ассоциации (association relationship);
- включения (include relationship);
- расширения (extend relationship);

- обобщения (generalization relationship).

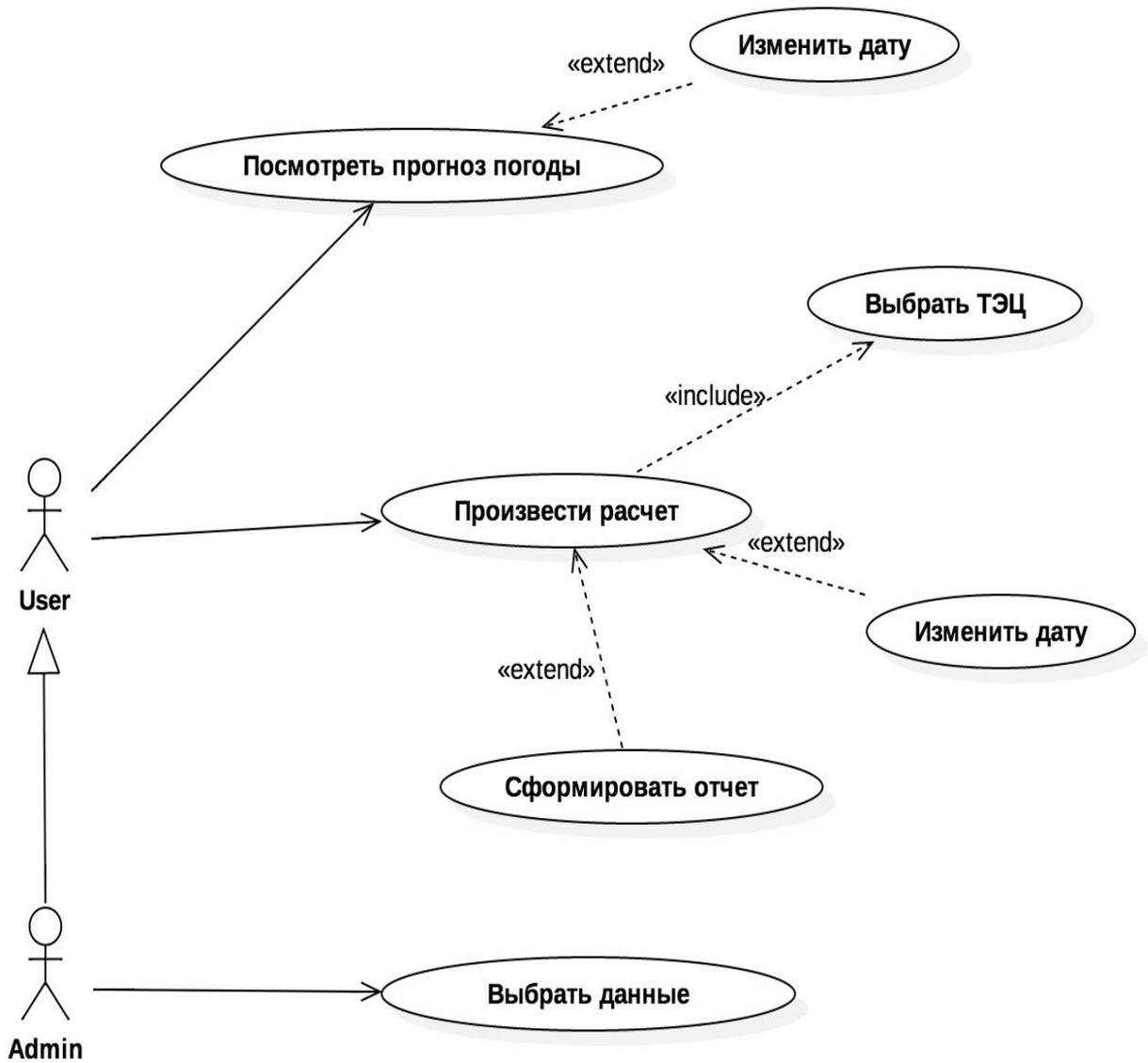


Рисунок 40. Диаграмма вариантов использования системы



Рисунок 41. Диаграмма вариантов использования, поясняющая прецедент «Выбрать данные»

Для реального описания системы потребуются более специфические данные, которые отражены в потоке событий. Потоки событий уточняют или детализируют последовательность действий, совершаемых системой при выполнении ее вариантов использования, а также описывают логику переходов через варианты использования. Диаграмма деятельности обеспечивает способ моделирования потока событий и отражает последовательность выполнения операций (рисунок 42).

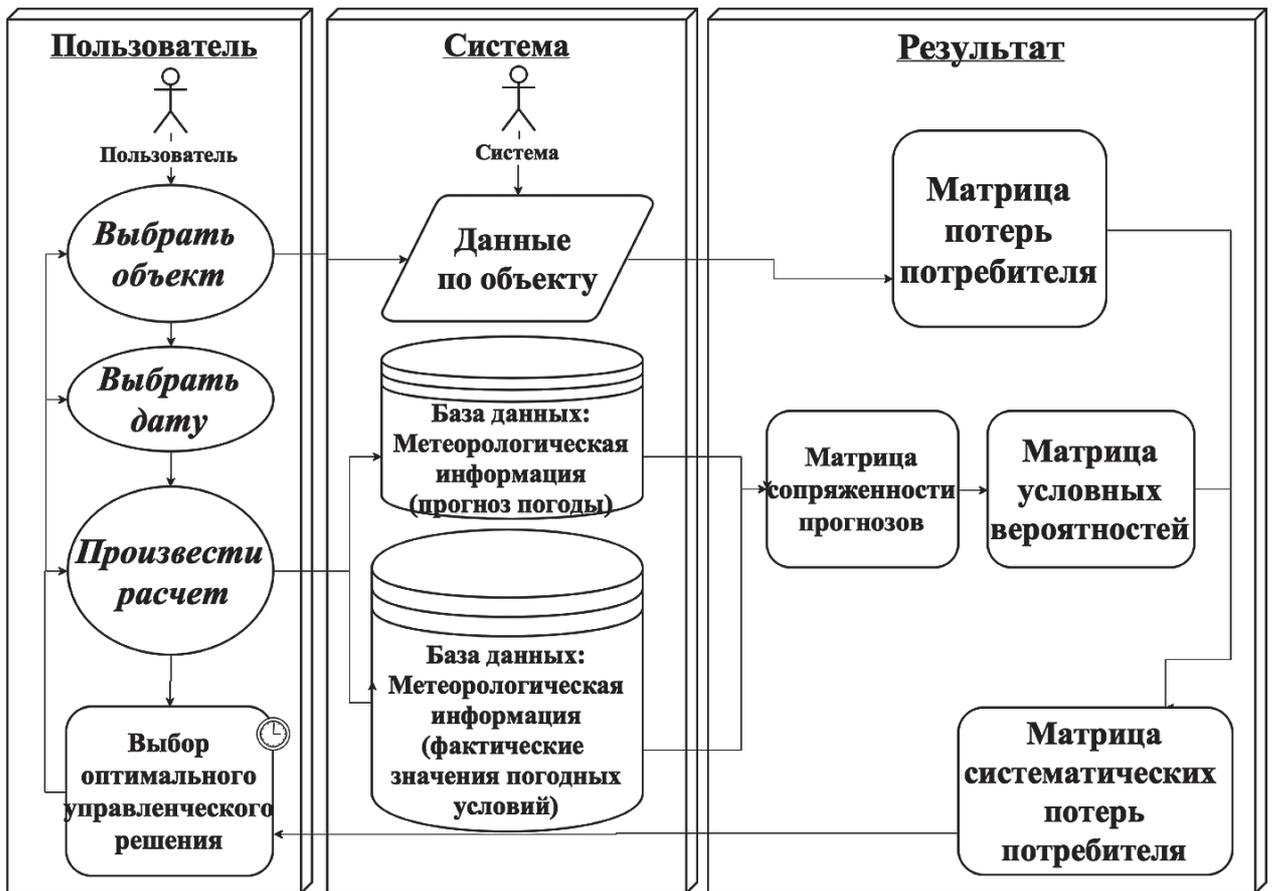


Рисунок 42. Схема управления разнородными метеорологическими данными применительно к задаче оптимизации регулирования отпуска тепла на ТЭЦ

Набор диаграмм деятельности является совокупностью алгоритмов реализации выполняемых системой операций. Каждая диаграмма акцентирует внимание на последовательности выполнения всех операций, которые в целом приводят к получению желаемого результата. Последовательность событий определена в строгом порядке, управление данными производится при непосредственном участии пользователя. Система определяет входные данные и используя алгоритм обработки выводит результат.

На рисунке 43 отражено пространственное размещение всех компонентов, участвующих в процессе работы системы. В процессе выбора СУБД были рассмотрены следующие продукты: MySQL, PostgreSQL, CA Ingres, Firebird. Наиболее подходящей СУБД явилась MySQL.

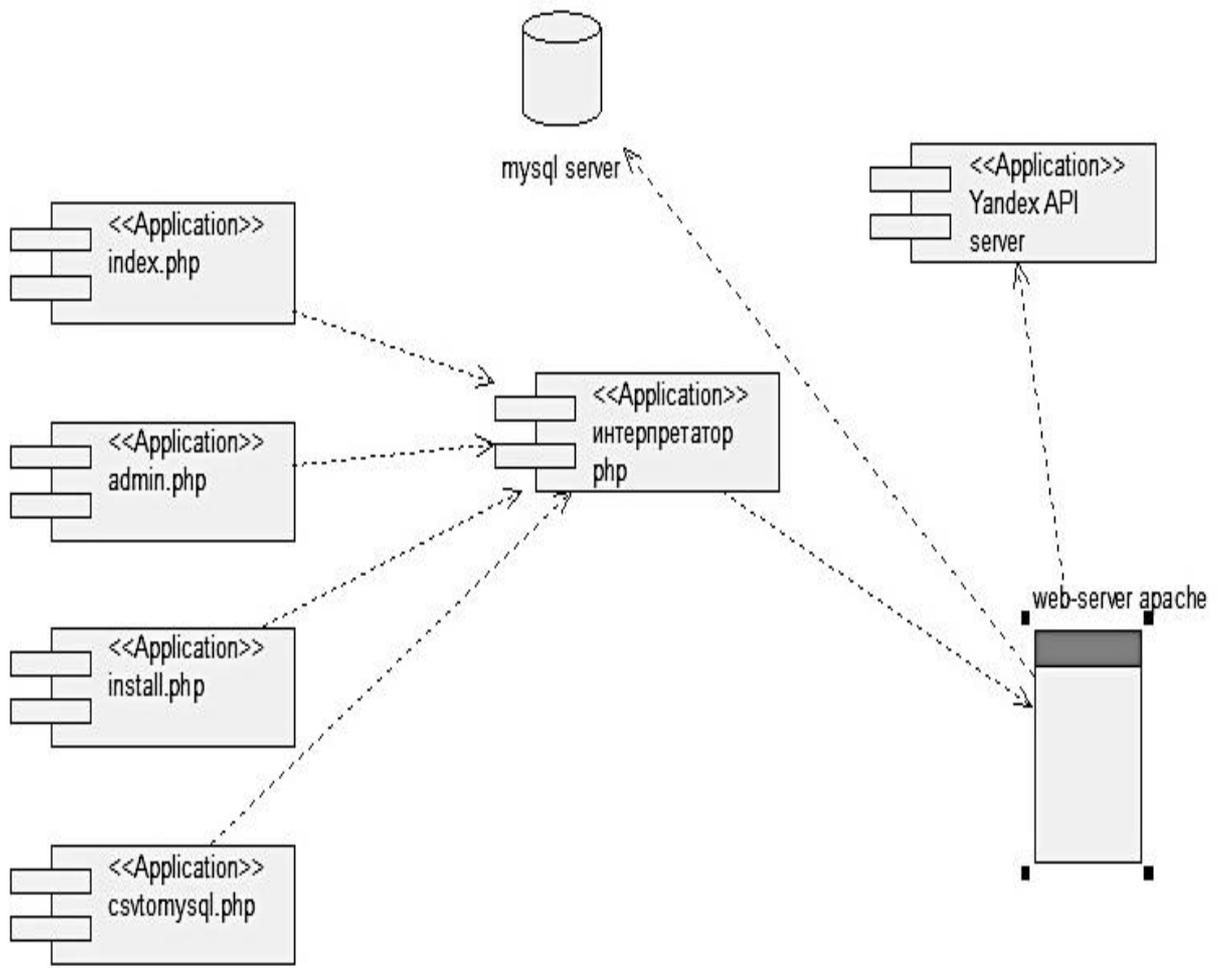


Рисунок 43. Диаграмма компонентов

Проектируемая система имеет клиент-серверную архитектуру, в которой в роли сервера выступает сервер базы данных MySQL. Тип операционной системы для использования сервера может быть любым совместимым с MySQL. Для корректной работы клиентской многопользовательской части приложения необходимы интернет-браузер и доступ в интернет. Модульность программной реализации обеспечивается системными средствами PHP, клиентское приложение является объектной оболочкой над реляционной базой данных. В общем виде архитектура системы представлена на рисунке 44 [78].



Рисунок 44. Архитектура системы

В разрабатываемой системе можно выделить пять основных компонентов.

- СУБД. Структурированные архивы пространственно-привязанных картографических и семантических данных, снабженные соответствующими метаданными, хранящиеся в БД.
- Картографическая основа. Геопроостранственное поле, отвечающее за интерактивное представление картографических данных, используя сервисы API.
- CMS. Система управления контентом, реализующая логику приложения, связь с картографическими сервисами API и обеспечивающая работу с хранилищем данных.
- Приложения. Набор программных модулей, реализованных на языке PHP.
- Web-Клиент. Графический интерфейс пользователя, реализованный в виде Web-приложения.

Для взаимодействия с системой на клиентском уровне пользователь оперирует web-браузером, который является «тонким» клиентом, установленным на рабочей станции.

Средство отображения данных и инструмент управления ГИС представляют собой интерфейс эксплуатации и администрирования системы, предоставляющий пользователю возможность в интуитивно понятной форме сформировать задание на обработку пространственно-распределенных данных (рисунок 45) [54, 55, 60].

СИСТЕМА УПРАВЛЕНИЯ ПРОСТРАНСТВЕННЫМИ ДАННЫМИ

[Главная](#) [Меню](#) [Карта](#) [О сервисе](#) [Справка](#)

Авторизация

Имя пользователя:
введите логин

Пароль:
введите пароль

Запомнить меня

Пользователь системы:
Админ

Территория исследования

Управление данными

Выберите вид влияющих метеорологических условий	<input type="text" value="Температура воздуха"/>	Выберите период, за который оценивается качество прогностической информации	<input type="text" value="С"/>
			<input type="text" value="дд. мм. гггг, -- : --"/>
			<input type="text" value="По"/>
			<input type="text" value="дд. мм. гггг, -- : --"/>
Дискретность метеорологической информации: * количество градаций метеорологических условий (n)	<input type="text" value="введите значение"/>	Ширина градации:	<input type="text" value="* минимальное значение"/>
			<input type="text" value="введите значение"/>
* Значения издержек (доходов), обусловленных ожиданием и реализацией погодных условий	<input type="text" value="введите значение"/>	* Выберите населенный пункт, в котором расположен производственный объект	<input type="text" value="введите населенный п"/>
<input type="button" value="Произвести расчёт"/>			

Рисунок 45. Web-интерфейс геоинформационной системы управления пространственными данными

Система прежде всего, должна быть удобна с точки зрения разработчика и интуитивно понятна со стороны пользователя. Интерфейс разработан с помощью языков HTML, PHP и JavaScript, проектирование элементов которого произведено с помощью возможностей InstantCMS. Модуль графического представления, расположенный на сервере приложений (уровень бизнес логики), позволяет объединить располагающиеся на различных серверах картографические материалы и картографическую основу «Google-карты» в единое геопространственное поле. Для реализации функциональности использован инструментарий среды «Google-карты», позволяющий объединить картографические данные, расположенные на различных серверах в различных форматах посредством инструментария API и языка JavaScript.

Данные, передаваемые от CMS (системы управления контентом) на рабочую станцию, конечному пользователю обрабатываются web-браузером и визуализируются в виде интерактивного онлайн-приложения [23].

Задача, сформированная пользователем системы, а также параметры визуализации результатов передаются на сервер приложений в виде запроса к CMS, принимающему и возвращающему sql-запросы. Запрос содержит указания на обрабатываемые характеристики, пространственные и временные границы интересующей области (широта, долгота), тип выбранной карты и выбранную территорию. На сервере приложений CMS передает данные в блок приложений, где модули вычислений обрабатывают запрос пользователя, взаимодействуя с модулем доступа к данным.

Каждый модуль имеет доступ к архивам данных СУБД через специальную библиотеку функций. Модуль доступа к данным обеспечивает поиск, чтение и выборку данных из архивов. В процессе вычислений используется накопленная в базе данных информация, после чего передается в модуль графического представления. Модуль графического представления использует полученную информацию и, объединив картографические материалы, из базы данных и из среды «Google-карты», визуализирует данные, запрошенные пользователем. Далее при поддержке Instant CMS (платформа публикации интернет-сайтов) результат отправляется пользователю системы [21, 27, 77].

Геоинформационная система позволяет собирать, хранить, отображать, редактировать и обрабатывать объединенные картографические разнородные пространственно-распределенные данные. Система поддержки принятия решений протестирована в интернет-браузерах Internet Explorer, Google Chrome, Safari. В настоящее время пользователю системы доступна информация по одному тематическому слою – слой объектов теплоэлектростанций Санкт-Петербурга. Состав и актуальность информационных данных делает ГИС эффективным инструментом для пользователей при принятии управленческих решений. ГИС позволит решать сложные задачи географического анализа [18].

Для основных потребителей гидрометеорологической информации ГИС поддержки принятия решений по оптимизации отпуска тепла на ТЭЦ станет каждодневным рабочим инструментом.

Дальнейшее развитие системы предусматривает:

- Разработку документации системы (руководство пользователя, руководство программиста);
- Внедрение дополнительных модулей по обработке разнородных данных;
- Разработку новых тематических слоев;
- Добавление новых разнородных пространственных данных.

В работе используются информационные ресурсы, которые в свою очередь являются одними из видов экономических ресурсов. В последние годы значительно повысилась их ценность, т.к. скорость обработки, накопления и обмена информации вырастает многократно.

Система позволяет может быть оценена как численными методами, так и без их учета. Экономический эффект от использования геоинформационной системы управления разнородными пространственно-распределенными данными можно оценить по следующим критериям:

- Сокращение времени обработки информации;

- Автоматизация обработки разнородной информации;
- Минимизация участия человека в процессе обработки данных;
- Автоматизация подготовки выбора поддержки принятия решений.

Вывод по главе

В третьей главе был проведен анализ проектирования информационных систем. На основе государственного стандарта была разработана методика проектирования геоинформационной системы для поддержки принятия управленческих решений. Методика проектирования геоинформационной системы была модифицирована внедрением в часть основных этапов проектирования системы существенным дополнением - использование разнородной информации. Модифицированная методика проектирования позволяет реализовать методы и подходы создания геоинформационной системы поддержки принятия решений для управления территориями на основе разнородных данных.

Для апробации предложенной в работе модели был осуществлен сбор и обработка разнородной информации по определенным параметрам. Было проведено проектирование системы с использованием методологии объектно-ориентированного анализа, выраженная по средствам визуального языка проектирования UML и программного продукта STAR UML [57]. В результате чего были построены концептуальная, статическая и физическая модель системы и представлены в виде диаграмм вариантов использования, компонентов и развертывания. На стадии проектирования были уточнены и скорректированы решения реализации геоинформационной системы, продуманы функциональные возможности системы относительно пользователей и обслуживающего персонала, скорректировано физическое представление систем и размещение отдельных компонентов [29].

В ходе исследований была спроектирована геоинформационная система для поддержки принятия управленческого решения по управлению территорией применительно к задаче оптимизации отпуска тепла ТЭЦ. Данная система разработана на основе предложенной в работе модели обработки разнородных данных, так же система использует разработанный алгоритм управления метеорологическими данными, что позволяет потребителям гидрометеорологической информации получить сведения помогающие принять эффективное управленческое решение.

ЗАКЛЮЧЕНИЕ

В работе были получены следующие основные результаты:

1. Модель обработки и управления разнородной геопространственной информацией, основанная на технологиях доступа к данным как локально, так и через глобальную сеть Интернет. Разработаны принципы эффективной синхронизации разнородных пространственных данных. Приведен алгоритм модуля обработки данных, который позволяет получить доступ к любым источникам информации, необходимым для принятия управленческого решения. Представлена схема управления разнородными данными;
2. Методика управления данными на примере регулирования отпуска тепла в зависимости от температуры окружающей среды. Методика применена в новом качестве, её результат может стать новым видом информационного продукта, который учитывает неопределенности реализации текста прогноза и позволяет оптимизировать управление ресурсами. Рассмотренный метод использования метеорологической информации позволяет обеспечить принятие оптимальных управленческих решений в сфере теплоэнергетики, обеспечивающий минимизацию потерь по метеорологическим причинам;
3. Автором предложена модификация методики проектирования геоинформационной системы с использованием модели обработки разнородной геопространственной информации, с учетом требований к системе поддержки принятия решений. Автором были предложены требования к создаваемой системе управления данными;
4. Геоинформационная система поддержки принятия решений на основе модели управления разнородными геоданными применительно к задачам управления территориями с использованием метеорологических данных. Решена научно-техническая задача, имеющая существенное значение для улучшения процесса управления теплоисточниками. Представлена схема обработки разнородной метеорологической информации в геоинформационной системе для получения практического результата. Реализация геоинформационной системы позволяет в автоматическом режиме поддерживать актуальными базы данных, что способствует повышению точности принимаемых решений.

Проводимые в диссертационной работе исследования по обработке разнородных пространственных данных позволили определить явную проблему интеграции разнородных

источников данных. В ходе исследования был произведен анализ опыта российских и зарубежных организаций в области разработки систем обработки разнородных данных, был выявлен ряд положительных и отрицательных сторон каждой системы. Автором проведен широкий анализ механизмов, моделей и систем обработки распределенной пространственной информации с использованием информационных технологий и выявлено, что ни один представленный инструмент не может полностью удовлетворить требованиям обработки разнородных данных при создании геоинформационных систем.

Предложенные в работе методы и модели позволили автоматизировать обработку разнородных данных, что способствовало увеличению скорости обработки данных и минимизации участия человека в данном процессе. Исследования показали, что обработка разнородных данных позволит проводить комплексный анализ различных массивов информации. Предложенные автором научные решения позволяют обрабатывать большое количество разнородных источников информации.

В работе была проведена автоматизация многофазового погодо-хозяйственного регламента с использованием метеорологических разнородных данных, что позволит потребителям гидрометеорологической информации подстраиваться под ожидаемые погодные условия, а так же предоставит возможность адаптации хозяйственной деятельности к неблагоприятным проявлениям гидрометеорологической среды, используя учет неопределенности реализации метеорологического прогноза.

Полученные в ходе работы результаты позволят эффективно использовать ресурсы в интересах экономического развития территории, в частности:

1. Обеспечения органов государственной власти актуальной и комплексной информацией для оперативного исследования, оценки и обоснования управленческих решений;
2. Повышения эффективности процедуры принятия решений, обеспечиваемой функциями анализа пространственных данных и представлением результатов научных исследований в наглядном и удобном для восприятия виде.

Предложенные в работе результаты могут быть использованы при решении ряда других прикладных задач по управлению территорией.

Список сокращений и условных обозначений.

АИС – автоматизированная информационная система

АСУ – автоматизированная система управления

АПК – автоматизированный программный комплекс

АРМ – автоматизированное рабочее место

БД – база данных

ГИС – геоинформационная система

ГИСППР – геоинформационная система поддержки принятия решений

ГБД – гетерогенная база данных

ИС – информационная система

ЛВС – локальная вычислительная сеть

ЛПР – лицо принимающее решение

ПО – программное обеспечение

РГИС – региональная геоинформационная система

РИС – распределенная информационная система

РБД – распределенная база данных

РГГМУ – Российский государственный гидрометеорологический университет

СУБД – система управления базами данных

СПИСОК ЛИТЕРАТУРЫ

1. ISO/IEC 10746-2:1996, Information technology — Open Distributed Processing — Reference Model: Foundations.3.2.5: *knowledge that is exchangeable amongst users about things, facts, concepts, and so on, in a universe of discourse*
2. ISO/IEC 2382-1:1993, Information technology — Vocabulary — Part 1: Fundamental terms.01.01.01: *knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning*
3. ГОСТ 34.003-90 Информационная технология. Комплекс стандартов на автоматизированные системы. Термины и определения.
4. ГОСТ 34.201-89 Информационная технология. Комплекс стандартов на автоматизированные системы. Виды, комплектность и обозначение документов при создании автоматизированных систем.
5. ГОСТ 34.321-96 Информационные технологии. Система стандартов по базам данных. Эталонная модель управления данными/ п. 2.1
6. ГОСТ 34.321-96 Информационные технологии. Система стандартов по базам данных. Издания. – М.:2001. – 24 с.
7. ГОСТ 34.601-90 Информационная технология. Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Стадии создания.
8. ГОСТ 34.602-89 Информационная технология. Комплекс стандартов на автоматизированные системы. Техническое задание на создание автоматизированной системы.
9. ГОСТ 34.603-92 Информационная технология. Виды испытаний автоматизированных систем.
10. Акальцов В.П. Базы данных. В 2-х кн. Книга 2. Распределенные и удаленные базы данных: учебник. – М.: ИД «ФОРУМ»: ИНФА – М, 2009. – 272 с.
11. Алешин Л.И. Информационные технологии: Учебное пособие/ Л.И. Алешин. – М.: Литера, 2008. – 424 с. (серия «современная библиотека», выпуск 35).
12. Арсеньев Б.П. Интеграция распределенных баз данных [Текст] : учебное пособие / Б. П. Арсеньев, С. А. Яковлев. - СПб. : Лань, 2001. - 461 с.
13. Б.П. Арсеньев, Интегрированные распределенные базы данных/ Б.П. Арсеньев, С.А.Яковлев. – М.: Лань, 2001. – 464 с.
14. Бедрицкий А.И., Коршунов А.А., Хандожко Л.А., Шаймарданов М.З. Показатели влияния погодных условий на экономику: региональное распределение экономических

- потерь и экономической выгоды при использовании гидрометеорологической информации и продукции // Метеорология и гидрология. 1999. № 3. С. 5-17.
15. Белодед Б.В., Глоба Л.С., Терновой М.Ю., Шпогина Е.С., Единый интерфейс доступа к гетерогенным базам данных. // Материалы 21-й международной крымской конференции «СВЧ-техника и телекоммуникационные технологии» - Севастополь, 2011. – с. 47-50.
16. Бескид П.П., Богданов П. А., Зоринова Е.М., Истомин Е.П. Колбина О.Н., Крылов А.Г., Миранков В.А., Слесарева Л.С., Степанов С. Ю., Соколов А.Г., Тимофеева А.Г., Фокичева А.А., Цыбин А.Р., Шишкин А.Д. Управление рисками устойчивого развития объектов и территорий в пространственном аспекте: монография / под общ. ред. Е.П. Истомина. - СПб.: РГГМУ, 2016. – 250 с. // ISBN 978-5-86813-447-0, УДК 338.245/355.07, ББК 65.29
17. Бескид П.П., Зоринова Е.М., Истомин Е.П. Колбина О.Н., Крылов А.Г., Слесарева Л.С., Степанов С. Ю., Соколов А.Г., Тимофеева А.Г., Фокичева А.А., Цыбин А.Р. Разработка основ методологии геоинформационного управления объектами и территориями // монография - СПб.: РГГМУ, 2016 – 286 с. // ISBN 978-5-86813-448-7, УДК 338.245/355.07, ББК 65.29
18. Бугаевский Л.М., Цветков В.Я. Геоинформационные системы: Учебное пособие для вузов. М.: «Златоуст», 2000. - 222 с.
19. Бурковский В.Л., Дорофеев А.Н. Система автоматизированного управления метрологическим обеспечением машиностроительного предприятия// Системы управления и информационные технологии: Сб. науч. тр. Воронеж, 2000. – С.79-86.
20. Бурковский В.Л., Дорофеев А.Н., Назаров В.Н. Интеграция гетерогенных баз данных в системы принятия решений // Теория активных систем: Сб. науч. тр. Межд. науч. – практ. конф. Москва, 2001. Том 1. – С.83-84.
21. Бурковский В.Л., Дорофеев А.Н., Семенин С.В., Моделирование и алгоритмизация управления гетерогенными базами данных в распределенных информационных системах. Воронеж. Гос.техн.университет, 2003. – 136 с.
22. Бурковский В.Л., Рыков С.А., Голиков А.А., Управление гетерогенными распределенными объектами информационных систем реального времени. Воронеж. гос. техн. ун-т. - Воронеж : [б. и.], 2012. - 193 с.
23. Воробьев В.И., Хандожко Л.А. Метеорологический фактор ресурсосбережения в теплоэнергетике Санкт-Петербурга. – Ученые записки. Научно-теоретический журнал, вып.1. – СПб.: изд. РГГМУ, 2005, с. 163 – 175.

24. Гарсия-Молина Г., Ульман Д., Уидом Д. Системы баз данных. Полный курс: Пер. с англ.: Уч. пос. – М.: Издательский дом «Вильямс», 2003. – 1088 с.: ил.
25. Геоинформатика: учеб. для студ. вузов / Е.Г. Капралов, А.В. Кошкарев, В.С. Тикунов и др под ред. В.С. Тикунова. – М.: Издательский центр «Академия», 2005. – 480с.
26. Глебова Н. ГИС для управления городами и территориями // ArcReview, 2006. - № 3(38).
27. Дворецкий С.И. Моделирование систем: учебник для студ. высш. учеб. заведений/ С.И. Дворецкий, Ю.Л. Муромцев, В.А. Погонин, А.Г. Схиртладзе. – М.: Издательский центр «Академия», 2009. – 320 с.
28. Диго С.М., Базы данных: проектирование и использование: Учебник. – М.: Финансы и статистика, 2005. – 592 с.
29. Дорофеев А.Н., Бурковский В.Л., Анализ сериализуемости глобальных транзакций в распределенной системе мульти-БД // Современные проблемы информатизации в непромышленной сфере и экономике: Сб. науч. тр. VI Международной электронной науч. конф. Воронеж, 2001. – С.89-90.
30. Дорофеев А.Н., Бурковский В.Л., Применение протокола атомарной фиксации в системе мульти-БД // Промышленная автоматика: Межвузовский сборник научных трудов. Воронеж, 2001. – С. 30.
31. Дулин С.К., Духин С.В. Множественно-реляционная модель как средство обеспечения гомоморфизма геоинформационной системы. – М.: ВЦ РАН, 2006. – 30 с.
32. Духин С.В. Формализация геоданных на основе множественно-реляционной модели // Системы и средства информатики. Специальный выпуск «Математические модели и методы информатики, стохастические технологии и системы». – М.: ИПИ РАН, 2005. С. 253–269.
33. Емельянова Н.З., Партыка Т.Л., Попов И.И. Основы построения автоматизированных информационных систем: Учебное пособие. – М.: ФОРУМ: ИНФРА-М, 2007. – 416 с.
34. Зайченко В.Ю. Геоинформатика как самостоятельная наука и отдельная научная дисциплина // Геоинформатика. – №3. – 2009. – С. 57-61.
35. Ивакин А.Я., Шкодырев В.П. Интеллектуальные геоинформационные системы и управление пространственными процессами. Учебное пособие - СПб.: Политехнический университет, 2012.
36. Истомин Е.П. Сетевые методы и модели распределения автоматизированных систем: диссертация на соискание докт. техн. наук: 05.13.06/ Истомин Евгений Петрович. – СПб., 1998. – 263 с.

37. Калиниченко Л.А. Методы и средства интеграции неоднородных баз данных. – М.: Наука, 1983. – 423 с.: ил.
38. Капралов Е.Г., Геоинформатика / Е.Г. Капралов. – М.: Издательский центр «Академия», 2008. – 384 с.
39. Карпик А.П. Методологические и технологические основы геоинформационного обеспечения территорий: монография. – Новосибирск: СГГА, 2004. – 260с.
40. Кашников А., Лядова Л., Интеграция гетерогенных источников данных на основе рекурсивной декомпозиции. // International Journal “Informational Technologies & Knowledge”. Vol.5, №3, 2011. – с. 274-284.
41. Когаловский М.Р. Очерк истории отечественных технологий баз данных // Открытые системы. 2002. №1.
42. Когаловский М.Р. Энциклопедия технологий баз данных. М.: «Финансы и статистика», 2002, 800с.
43. Когаловский М.Р., Методы интеграции данных в информационных системах // Институт проблем рынка РАН. – Москва, 2010.
44. Колбина О.Н. Информационная система прогноза рисков наводнений в Санкт-Петербурге./ Истомин Е.П., Петров Я.А., Слесарева Л.С./ Информационные технологии системы: управление, экономика, транспорт, право: Сб. науч.тр./Вып. 10-СПб: ООО «Андреевский издательский дом» - 2013г.,156 с.
45. Колбина О.Н. Обзор технологических методов доступа к базам данных распределенных ГИС / Зоринова Е.М., Истомин Е.П., Колбина О.Н., Степанов С.Ю. //Информационные технологии и системы: управление, экономика, транспорт, право:Сб. тр. Международной практической конференции «Инфогео-2014»/Вып. 3(14) – СПб: ООО «Андреевский издательский дом» - 2014 г., 116 с.
46. Колбина О.Н. Разработка и развитие методов, моделей и систем геоинформационного управления пространственно-распределенными объектами: отчет о НИР/ Истомин Е.П., Соколов А.Г., Колбина О.Н./ – Санкт-Петербург: Российский государственный гидрометеорологический университет, 2013. – 102 с. 16.
47. Колбина О.Н. Современные и теоретические аспекты управления распределенными базами данных. / Истомин Е.П., Колбина О.Н./ Информационные технологии системы: управление, экономика, транспорт, право: Сб. науч. тр./ Вып. 1-СПб: ООО «Андреевский издательский дом» - 2011 г.,105 с.

48. Колбина О.Н. Феномен геоинформационного управления и принципы его реализации / Истомин Е.П., Кирсанов С.А., Соколов А.Г., Колбина О.Н. / – Вестник СПбГУ, серия 7, выпуск 4, СПбГУ, 2014.
49. Коннолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация и сопровождение. Теория и практика, 2-е изд.: Пер. с англ.: Уч. пос. – М.: Издательский дом «Вильямс», 2000. – 1120 с.:ил. – Парал. тит. англ. 28. Bright M. W., Hurson A. R., Pakzad S.H. A Taxonomy and current issues in multidatabase systems. Computer, pages 50-60, March 1992.
50. Коновалова Н.В., Геоинформационные системы управления территориями. Региональный уровень // Геодезия и картография. – 2016. – № 2. – С. 54-57.
51. Коровин Е.Н. Методология прогнозирования и оптимального управления территориально распределенными социально-экономическими системами на основе трансформации информации и многовариантного моделирования: Дис. на соиск. уч. ст. д-ра техн. наук – Воронеж: Воронежский гос. техн. ун-т., – 2005. – 356 с.
52. Костенко В.А., Информационные технологии. Основы теории и практики программирования в Internet. В 3 ч. Ч.1. Распределенные системы, динамические библиотеки и СОМ – объекты, связь в распределенных системах, сеть Internet, электронная почта: учебное пособие/ В.А. Костенко. – Орел: Академия ФСО России, 2009. – 244 с.
53. Кошкарев А.В. Понятия и термины геоинформатики и ее окружения: учебно-справочное пособие. – М.: ИГЕМ РАН, 2000. – 76с.
54. Кудашев Е.Б., Марков С.Ю., Попов М.А., Использование гетерогенной пространственной информации при решении задач устойчивого развития территорий // Журнал электронные библиотеки. – Т. 14, Вып. 3., 2011.
55. Лайкин В.И., Упоров Г.А. Геоинформатика: учебное пособие / Лайкин В.И., Упоров Г.А. – Ком-сомольск-на-Амуре: Изд-во АмГПУ, 2010. – 162 с.
56. Лурье И.К. Основы геоинформационного картографирования: учеб. пособие. – М.: Изд-во Моск. ун-та, 2000. – 143 с.
57. Нотация и семантика языка UML[Электронный ресурс] //Национальный открытый университет «Интуит» - Режим доступа: <http://www.intuit.ru/> (Дата обращения 15.06.14).
58. Оззу М.Т., Валдуриз П. Распределенные и параллельные системы баз данных // Системы управления базами данных. 1996. No4.

59. Пат. No2014621110 Российская Федерация, База данных Ice_Concentration [Текст]/Истомин Е.П., Колбина О.Н., Степанов С.Ю., Миранков В.А.; заявитель и патентообладатель РГГМУ. – Заявл. No 2014620482, опубл. 07.08.14.
60. Плешков В.Г., Ребрий А.В., Степанова Я.В. Инфраструктура пространственных данных – архитектура, ориентированная на сервисы // Геодезия и картография. – 2005. – №5. – С. 50-52.
61. Распределенные системы. Принципы и парадигмы / Э. Таненбаум, М.ван Стеен. – СПб.: Питер, 2003. – 877 с.: ил. – (Серия «Класика computerscience»).
62. РД 50-34.698-90 Автоматизированные системы. Требования к содержанию документов.
63. Розенберг И.Н., Духин С.В. Автоматизированная система ведения геоинформационной базы данных, увязанная с параметрами работы и развития ОАО «РЖД» //Труды девятой научно-практической конференции «Информационные технологии в железнодорожном транспорте». «ИНФОТРАНС-2004», 6-9 октября, Санкт-Петербург, 2004. С. 12 – 16.
64. Ролланд Ф. Основные концепции баз данных: Пер. с англ.: Уч. пос. – М.: Издательский дом «Вильямс», 2002. – 256 с.: ил.
65. Саймон А. Р. Стратегические технологии баз данных: менеджмент на 2000 год. Пер. с англ./ Под ред. и с предисл. М.Р. Когаловского. – М.: Финансы и статистика, 1999. – 480 с.: ил.
66. Самарский А.А. Математическое моделирование / А.А. Самарский, А.П. Михайлов. – М.: Физматлит, 2001. – 320 с.
67. Семакин И.Г. Информационные системы и модели. Элективный курс: Учебное пособие/ И.Г. Семакин, Е.К.Хеннер. – М.: БИНОМ. Лаборатория знаний, 2005. – 303 с.
68. Середович В.А., Геоинформационные системы (назначение, функции, классификация) [Текст]: монография/ В.А. Середович, В.Н. Ключнеченко, Н.В. Тимофеева. – Новосибирск: СГГА, 2008. – 192 с.
69. Слесарева Л.С., Истомин Е.П. Применение стохастических моделей для прогнозирования рисков в геосистемах. – Ученые записки РГГМУ, вып. 17, 2011
70. Слесарева Л.С. Развитие методов геомоделирования и оценки рисков в геосистемах природного характера (на примере наводнений): диссертация на соискание канд. техн. наук: 25.00.35/ Слесарева Людмила Сергеевна. – СПб., 2013. – 174 с.
71. Советов Б.Я. Моделирование систем/ Б.Я. Советов, С.А. Яковлев. – М.: Высш.шк., 2005. – 343 с.
72. Степанов С.Ю. Анализ моделей и систем обработки гетерогенных данных для использования в прикладных ГИС/ Истомин Е.П., Колбина О.Н., Степанов С.Ю.,

- Сидоренко А.Ю. // Научный вестник. 2015. N 4(6). с. 53-62 (ISSN 2411-1872, Св-во рег. СМИ ЭЛ № ФС 77-55459).
73. Степанов С.Ю. Анализ статистических данных космических измерений с приполярной территории для разработки стохастической модели оценки георисков/ Истомин Е.П., Сидоренко А.Ю.// Информационные технологии системы: управление, экономика, транспорт, право: Сб.науч.тр./Вып. 2 (11) -СПб: ООО «Андреевский издательный дом» - 2013г.,245 с.
74. Степанов С.Ю. Геоинформационные системы для предупреждения чрезвычайных ситуаций РФ и США / Степанов С.Б., Кошкарев Я.А., Пинчук Р.Б. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб. тр. Международной научно-практической конференции «Инфогео 2015»/Вып. 2 (16)/ Под ред. д.т.н., проф. Истомина Е.П., д.т.н. – СПб.: ООО «Андреевский издательский дом» - 2015 г., 214 с.
75. Степанов С.Ю. ГИС. Обзор программного обеспечения/ Гришин Н.М., Колбина О.Н., Соловьева А.Д, Степанов С.Ю. //Информационные технологии и системы: управление, экономика, транспорт, право: Сб. тр. Международной практической конференции «Инфогео-2014»/Вып. 3(14) – СПб: ООО «Андреевский издательский дом» - 2014 г., 116 с.
76. Степанов С.Ю. Методика проектирования геоинформационной системы управления территориями Заполярья на основе распределенных гетерогенных баз данных / Истомин Е.П., Колбина О.Н., Степанов С.Ю. // Ученые записки Российского государственного гидрометеорологического университета № 39. Научно- теоретический журнал. – СПб.: РГГМУ, 2015. – 250 с., ISSN 2074-2762
77. Степанов С.Ю. Механизмы моделей и систем интеграции разнородных геопространственных данных/ Истомин Е.П., Колбина О.Н., Степанов С.Ю., Сидоренко А.Ю. // Наука и современность. 2015. N 1(3). с. 89-97 (ISSN 2411-2127, Св-во рег. СМИ ЭЛ № ФС 77-55458)
78. Степанов С.Ю. Модели обработки гетерогенных данных с помощью ГИС программ для решения задач оценки вероятностных событий/ Истомин Е.П., Колбина О.Н., Степанов С.Ю., Сидоренко А.Ю. // Наука, образование, общество. 2015. N 4(6). с. 20-27. (ISSN 2411-2224, Св-во рег. СМИ ЭЛ № ФС 77-56325)
79. Степанов С.Ю. Прогнозирование гидрометеорологической обстановки западного сектора Северного морского пути с применением линейного фильтра Калмана-Бьюси/Сидоренко А.Ю., Новиков В.В., Истомин Е.П., Степанов С.Ю.//

- Информационные технологии и системы: управление, экономика, транспорт, право: Сб. тр. Международной практической конференции «Инфогео-2014»/Вып. 3(14) – СПб: ООО «Андреевский издательский дом» - 2014 г., 116 с.
80. Степанов С.Ю. Разработка модели доступа и технологий обработки гетерогенных баз данных для использования в прикладных ГИС/ Истомин Е.П., Колбина О.Н., Степанов С.Ю. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб. научных трудов./Вып. 1 (15) / Под ред. д.т.н., проф. Истомина Е.П. – СПб: ООО «Андреевский издательский дом» - 2015 г., 66 с.
81. Степанов С.Ю. Реализация модели подготовки гетерогенных данных в автоматизированной системе / Степанов С.Ю., Петров Я.А. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб. тр. Международной научно-практической конференции «Инфогео 2015»/Вып. 2 (16)/ Под ред. д.т.н., проф. Истомина Е.П., д.т.н. – СПб.: ООО «Андреевский издательский дом» - 2015 г., 214 с.
82. Степанов С.Ю. Сложная информационная система прогнозирования рисков с применением фильтра Калмана – Бьюси / Истомин Е.П., Новиков В.В., Сидоренко А.Ю., Колбина О.Н., Степанов С.Ю. // – Ученые записки РГГМУ, выпуск 36, РГГМУ, 2014.
83. Степанов С.Ю. Сравнительный анализ открытых геоинформационных систем/Информационные технологии системы: управление, экономика, транспорт, право: Сб.науч.тр./Вып. 10-СПб: ООО «Андреевский издательский дом» - 2013г.,156 с.
84. Степанов С.Ю. Технический анализ механизмов доступа к гетерогенным данным для использования в прикладных ГИС/ Истомин Е.П., Колбина О.Н., Степанов С.Ю. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб. научных трудов./Вып. 1 (15) / Под ред. д.т.н., проф. Истомина Е.П. – СПб: ООО «Андреевский издательский дом» - 2015 г., 66 с.
85. Степанов С.Ю., Фокичева А.А. Разработка модели управления пространственно-распределенными разнородными данными для поддержки принятия решений / Ученые записки Российского государственного гидрометеорологического университета № 45. Научно- теоретический журнал. – СПб.: РГГМУ, 2016. – 264 с., ISSN 2074-2762
86. Ташкин А.О., Семенов С.П. Методика разработки геоинформационной системы для маломобильных граждан // Современные проблемы науки и образования. – 2014. – № 1.
87. Теленик С. Ф., Жук С. В., Лыско В. Т., Ефремов К. В. Каталогизация и интеграция разнородных информационных ресурсов // Молодой ученый. — 2013. — №5. — С. 176-179.

88. Тикунов В.С. Основы геоинформатики: В 2кн. Кн1: учебное пособие для вузов – М.: Академия, 2004. – 352с.
89. Тикунов В.С. Основы геоинформатики: В 2кн. Кн2: учебное пособие для вузов – М.: Академия, 2004. – 480с.
90. Томилин В.В., Нориевская Г.М. Использование ГИС в муниципальном управлении // Практика муниципального управления, 2007. - №7
91. Трифонова Т.А., Мищенко Н.В. Геоинформационные системы и дистанционное зондирование в экологических исследованиях – М.: УМО РФ, 2005. – 349с.
92. Филиппов Д.Н., Фортыхина Е.А., Фокин В.С. Введение в геоинформационные системы. Учебное пособие. М.:РГОТУПС, 2007 г. - 163 с.
93. Фокичева А.А. Многофазовый погодо-хозяйственный регламент как инструмент снижения экономико-метеорологических рисков / Межвузовский сборник научных трудов: Информационные технологии и системы: управление, экономика, транспорт, право. – СПб: ООО «Андреевский издательский дом». – 2014. – № 1 (12). – С. 88 – 97.
94. Фокичева А.А., Истомин Е.П. О некоторых вопросах управления территорией с учетом нестабильности климатической системы / Информационные технологии и системы: управление, экономика, транспорт, право: Сб. тр. Международной научно-практической конференции «Инфогео 2014». – СПб.: ООО «Андреевский издательский дом». – 2014. – Вып. 3 (14). – С.62–66.
95. Фуфаев Э.В., Разработка и эксплуатация удаленных баз данных: учебник для студ. сред. проф. образования/ Э.В. Фуфаев, Д.Э. Фуфаев.-М. : Издательский центр «Академия», 2008. – 256 с.
96. Хандожко Л.А. Оптимальные погодо-хозяйственные решения .–СПб.: Изд. РГГМУ, 1999. – 161 с.
97. Хандожко Л.А. Экономическая метеорология. - СПб.: Гидрометеиздат, 2005. – 490 с.
98. Хандожко Л.А., Цепелев В.Ю., Колесов А.М., Фокичева А.А. Проблема оптимального использования метеорологических ресурсов в условиях нестабильной климатической системы. – Сборник трудов международной школы-конференции «Изменение климата и окружающая среда». – СПб.: изд. Гранд, 2005, с. 334-342.
99. Харрингтон Джен, Разработка баз данных: Пер. с англ. – М.: ДМК Пресс, 2005. – 272 с.
100. Черкашин А.К., Китов А.Д., Бычков И.В., Геоинформационная система управления территорией / А.К.Черкашин, А.Д.Китов, И.В.Бычков и др. Иркутск: Издательство Института географии СО РАН, 2002.- 151 с.
101. Черняк Л., Интеграция данных: синтаксис и семантика. // Открытые системы. 2009, №10.

102. Шаститко А.Е., Плаксин С.М. Концептуальные подходы к моделированию результатов гидрометеорологического обеспечения экономической деятельности. – М.: изд. МАКС Пресс, 2009, 86 с
103. Шаши Шекхар, Санжей Чаула Основы пространственных баз данных/ Пер. с англ. М.: КУДРИЦ-ОБРАЗ, 2004. – 336 с.
104. Щербинин Ю.Б. Нетрадиционные подходы к созданию геоинформационных систем управления муниципальными образованиями. – СНИБ "Эльбрус".
105. Яйли Е.А., Музалевский А.А. Риск: анализ, оценка, управление. СПб, изд. РГГМУ, 2005, 230 с.
106. Attaluri G.K., Bradshaw D.P., Coborn N. The CORDS Multidatabase Project. Technical Report. 1995.
107. Bernstein P.A., Hadzilacos V., Goodman N. Concurrency Control and Recovery in Database Systems, Addison-Wesley Series In Computer Science, Addison-Wesley, United States of America, 1987.
108. Breitbart Y. Multidatabase Interoperability, Sigmod Record, 19(3):53-60, September 1990.
109. Hadzilagos V.A. Theory of Reliability in Database Systems. Journal of the ACM, vol.35, No. 1, 1988, Pg.: 121-145.
110. InfoCity - виртуальный город компьютерной документации [Электронный ресурс]. – Режим доступа: http://www.infocity.kiev.ua/prog/delphi/content/delphi074_2.phtml.
111. Kent W., Solving Domain Mismatch and Schema Mismatch Problems with an Object-Oriented Database Programming Language. // Proceedings of the International Conference on Very Large Data Bases, 1991.
112. Razuvayev V.N., Apasova E.G., Martuganov R.A., Steurer P., Vose R., 1993. Daily Temperature and Precipitation Data for 223 U.S.S.R. Stations. ORNL/CDIAC, Numerical data package – 040, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
113. Zisman A., Kramer J. Towards Interoperability in Heterogeneous Database Systems. Imperial College Research Report No. Doc 95/11, December, 1995.

ПРИЛОЖЕНИЕ

Приложение 1. Программный код обработки матрицы сопряженности прогноза и фактических значений температуры воздуха.

```
SubMatrix()
```

```
Dim i As Integer ' ежедневные показатели температуры
```

```
Dim j As Integer ' месяц
```

```
Dim m As Integer ' прогноз
```

```
Dim n As Integer ' факт
```

```
Dim averSh As Worksheet
```

```
Dim matrSh As Worksheet
```

```
Set averSh = Worksheets("Среднее")
```

```
Set matrSh = Worksheets("Лист1")
```

```
For j = 3 To 8 ' 6 месяцев за 3 года
```

```
For i = 3 To 63 Step 2
```

```
    Select Case averSh.Cells(i, j)
```

```
        Case -30 To -28
```

```
            n = 3
```

```
        Case -27 To -25
```

```
            n = 4
```

```
        Case -24 To -22
```

```
            n = 5
```

```
        Case -21 To -19
```

```
            n = 6
```

```
        Case -18 To -16
```

```
            n = 7
```

Case -15 To -13

n = 8

Case -12 To -10

n = 9

Case -9 To -7

n = 10

Case -6 To -4

n = 11

Case -3 To -1

n = 12

Case 0 To 2

n = 13

Case 3 To 5

n = 14

Case 6 To 8

n = 15

Case Is >= 9

n = 16

End Select

Select Case averSh.Cells(i + 1, j)

Case -30 To -28

m = 6

Case -27 To -25

m = 7

Case -24 To -22

m = 8

Case -21 To -19

m = 9

Case -18 To -16

m = 10

Case -15 To -13

m = 11

Case -12 To -10

m = 12

Case -9 To -7

m = 13

Case -6 To -4

m = 14

Case -3 To -1

m = 15

Case 0 To 2

m = 16

Case 3 To 5

m = 17

Case 6 To 8

m = 18

Case Is >= 9

m = 19

End Select

matrSh.Cells(m, n) = matrSh.Cells(m, n) + 1

matrSh.Cells(m, n).Interior.ColorIndex = 23

Next i

Next j

End Sub

Приложение 2. Матрица сопряженности прогнозов и фактической среднесуточной температуры воздуха за отопительный сезон 2013-2014 г.

t														
	-30...-28	-27...-25	-24...-22	-21...-19	-18...-16	-15...-13	-12...-10	-9 ...-7	-6 ...-4	-3...-1	0...+2	+3...+5	+6...+8	+9 и более
-30...-28	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-27...-25	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-24...-22	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-21...-19	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-18...-16	0	0	0	0	0	2	0	0	0	0	0	0	0	0
-15...-13	0	0	0	0	2	2	1	0	0	0	0	0	0	0
-12...-10	0	0	0	0	1	0	6	1	0	0	0	0	0	0
-9 ...-7	0	0	0	0	0	1	4	2	1	0	0	0	0	0
-6 ...-4	0	0	0	0	0	0	0	0	10	3	0	0	0	0
-3 ... -1	0	0	0	0	0	0	0	0	3	13	1	0	0	0
0 ... +2	0	0	0	0	0	0	0	0	0	6	35	3	0	0
+3 ... +5	0	0	0	0	0	0	0	0	0	0	19	43	3	0
+6 ... +8	0	0	0	0	0	0	0	0	0	0	0	13	27	1
+9 и более	0	0	0	0	0	0	0	0	0	0	0	0	3	14
Σ	0	0	0	0	3	5	11	3	14	22	55	59	33	15

Приложение 3. Программный код обработки матрицы условных вероятностей прогноза и фактических значений температуры воздуха.

```
Option Explicit
```

```
Sub Matrix()
```

```
Dim i As Integer ' ежедневные показатели температуры
```

```
Dim j As Integer ' месяц
```

```
Dim m As Integer ' прогноз
```

```
Dim n As Integer ' факт
```

```
Dim averSh As Worksheet
```

```
Dim matrSh As Worksheet
```

```
Set averSh = Worksheets("—реднее")
```

```
Set matrSh = Worksheets("Тист1")
```

```
For j = 3 To 8 ' 6 месяцев за 3 года
```

```
    For i = 3 To 63 Step 2
```

```
        Select Case averSh.Cells(i, j)
```

```
            Case -30 To -28
```

```
                n = 3
```

```
            Case -27 To -25
```

```
                n = 4
```

```
            Case -24 To -22
```

```
                n = 5
```

```
            Case -21 To -19
```

```
                n = 6
```

```
            Case -18 To -16
```

```
                n = 7
```

```
            Case -15 To -13
```

```
                n = 8
```

Case -12 To -10

n = 9

Case -9 To -7

n = 10

Case -6 To -4

n = 11

Case -3 To -1

n = 12

Case 0 To 2

n = 13

Case 3 To 5

n = 14

Case 6 To 8

n = 15

Case Is >= 9

n = 16

End Select

Select Case averSh.Cells(i + 1, j)

Case -30 To -28

m = 6

Case -27 To -25

m = 7

Case -24 To -22

m = 8

Case -21 To -19

m = 9

Case -18 To -16

m = 10

Case -15 To -13

m = 11

Case -12 To -10

```

    m = 12
Case -9 To -7
    m = 13
Case -6 To -4
    m = 14
Case -3 To -1
    m = 15
Case 0 To 2
    m = 16
Case 3 To 5
    m = 17
Case 6 To 8
    m = 18
Case Is >= 9
    m = 19

End Select

    matrSh.Cells(m, n) = matrSh.Cells(m, n) + 1
matrSh.Cells(m, n).Interior.ColorIndex = 23

Next i
Next j

For i = 6 To 19 'считаем сумму по столбцам в матрице сопряженности
    For j = 3 To 16
        matrSh.Cells(20, j) = matrSh.Cells(20, j) + matrSh.Cells(i, j)
    Next j
Next i

'расчет матрицы условных вероятностей

For i = 6 To 19
    For j = 3 To 16

```

If matrSh.Cells(20, j) = 0 Then

 matrSh.Cells(i + 26, j) = 0

Else

 matrSh.Cells(i + 26, j) = matrSh.Cells(i, j) / matrSh.Cells(20, j)

End If

 matrSh.Cells(46, j) = matrSh.Cells(46, j) + matrSh.Cells(i + 26, j)

Next j

Next i

End Sub

Приложение 5. Матрица потерь (величина расчета измеряется в тыс. руб.)

t	-30	-27	-24	-21	-18	-15	-12	-9	-6	-3	0	3	6	9
	-28	-25	-22	-19	-16	-13	-10	-7	-4	-1	2	5	8	11
-28	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2								
-25	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2							
-22	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2						
-19	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2					
-16	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2				
-13	12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2			
-10		12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2		
-7			12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2	
-4				12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1	37790,2
-1					12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1	30232,1
2						12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8	22691,1
5							12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8	15289,8
8								12596,7	10077,4	7587,1	5259,2	3443,9	4111,3	8435,8
11									12596,7	10077,4	7587,1	5259,2	3443,9	4111,3

Приложение 6. Матрица систематических потерь (величина расчета измеряется в тыс. руб.)

t														
	-30...-28	-27...-25	-24...-22	-21...-19	-18...-16	-15...-13	-12...-10	-9...-7	-6...-4	-3...-1	0...+2	+3...+5	+6...+8	+9 и более
-30...-28	0,0	0,0	0,0	0,0	8397,8	9069,6	1145,2	0,0	0,0	0,0	0,0	0,0	0,0	0,0
-27...-25	0,0	0,0	0,0	0,0	10917,2	7065,8	7787,1	4198,9	0,0	0,0	0,0	0,0	0,0	0,0
-24...-22	0,0	0,0	0,0	0,0	8417,2	7657,9	10767,1	11756,9	899,8	0,0	0,0	0,0	0,0	0,0
-21...-19	0,0	0,0	0,0	0,0	6035,2	5496,7	8281,0	9247,3	9717,5	1717,7	0,0	0,0	0,0	0,0
-18...-16	0,0	0,0	0,0	0,0	4049,0	4539,5	5940,7	6811,1	10439,4	8817,7	229,0	0,0	0,0	0,0
-15...-13	0,0	0,0	0,0	0,0	3888,8	6070,7	4164,7	4654,1	7954,4	10424,9	8199,3	640,5	0,0	0,0
-12...-10	0,0	0,0	0,0	0,0	6994,3	10179,0	4261,7	3666,3	5628,4	7948,8	10902,4	9693,1	1145,2	0,0
-9...-7	0,0	0,0	0,0	0,0	13005,1	16014,6	7486,3	5552,8	3880,6	5646,5	8405,0	10505,9	11222,5	839,8
-6...-4	0,0	0,0	0,0	0,0	20224,0	22856,4	13470,3	10720,5	4277,1	4030,0	6030,4	8017,4	10080,0	12428,8
-3...-1	0,0	0,0	0,0	0,0	27718,5	30266,9	20685,3	17756,9	7998,7	4519,0	4083,1	5679,8	7601,8	9911,4
0...+2	0,0	0,0	0,0	0,0	35270,8	19654,3	28177,0	25204,8	14349,8	8191,0	3959,3	3877,8	5305,8	7431,9
+3...+5	0,0	0,0	0,0	0,0	12596,7	6046,4	31606,3	32751,5	21643,7	14429,8	7066,5	4184,1	3669,6	5138,2
+6...+8	0,0	0,0	0,0	0,0	0,0	7558,0	13741,9	25193,5	29156,1	21700,9	13056,6	7831,4	4443,7	3488,4
+9 и более	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	33471,3	29206,1	20271,4	14155,9	8665,7	4399,6

Приложение 7. Свидетельство о государственной регистрации базы данных «База данных метеорологических параметров»

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации базы данных
№ 2016620986
«База данных метеорологических параметров»

Правообладатель: *федеральное государственное бюджетное образовательное учреждение высшего образования «Российский государственный гидрометеорологический университет» (РГГМУ) (RU)*

Авторы: *Истомин Евгений Петрович (RU), Степанов Сергей Юрьевич (RU), Соколов Александр Геннадьевич (RU), Петров Ярослав Андреевич (RU), Фокичева Анна Алексеевна (RU)*

Заявка № **2016620547**
Дата поступления **28 апреля 2016 г.**
Дата государственной регистрации
в Реестре баз данных **20 июля 2016 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

 **Г.П. Ивлиев**



Приложение 8. Свидетельство о государственной регистрации базы данных «База данных
SeaIce version 1.0»



Приложение 9. Свидетельство о государственной регистрации базы данных «База данных Ice_Concentration»

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации базы данных

№ 2014621110

База данных Ice_Concentration

Правообладатель: *федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Российский государственный гидрометеорологический университет» (RU)*

Авторы: *Истомин Евгений Петрович (RU), Колбина Ольга Николаевна (RU), Степанов Сергей Юрьевич (RU), Миранков Валерий Александрович (RU)*

Заявка № **2014620482**

Дата поступления **17 апреля 2014 г.**

Дата государственной регистрации

в Реестре баз данных **07 августа 2014 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

Б.П. Симонов

