

На правах рукописи



Колбина Ольга Николаевна

Разработка геоинформационной системы оценки параметров климатических условий на основе распределенных гетерогенных баз данных

Специальность 25.00.35 – Геоинформатика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург - 2015

Работа выполнена на кафедре прикладной информатики федерального государственного бюджетного образовательного учреждения высшего профессионального образования «Российский государственный гидрометеорологический университет»

Научный руководитель

доктор технических наук, профессор
Истомин Евгений Петрович

Официальные оппоненты:

доктор технических наук, профессор
кафедры прикладных информационных
технологий Санкт-Петербургского
государственного университета
аэрокосмического приборостроения
Татарникова Татьяна Михайловна
кандидат технических наук, доцент
кафедры математики и программного
обеспечения военно-космической академии
им. А.Ф. Можайского
Котиков Павел Евгеньевич

Ведущая организация

Санкт-Петербургский государственный
электротехнический университет «ЛЭТИ»
им.В.И. Ульянова (Ленина) (СПбГЭТУ)

Защита состоится 17 марта 2015 г. на заседании диссертационного совета: Д 212.197.03 при Российском государственном гидрометеорологическом университете по адресу: 195196, Санкт-Петербург, пр. Металлистов, д. 3, ауд. 102 в 15 час. 30 мин.

С диссертацией можно ознакомиться в библиотеке Российского государственного гидрометеорологического университета или на сайте <http://www.rshu.ru/university/dissertations/>

Автореферат разослан 17 02 2015 г.

Ученый секретарь
диссертационного совета Д 212.197.03
д.г.н.



Е.С. Попова

Актуальность. В современном мире северные регионы России и Арктика занимают всё более значимые позиции среди остальных территорий РФ. Это связано, во-первых, с их ролью в формировании глобального климата и устойчивости состояния биосферы; во-вторых, наличием обширных запасов углеводородов на их территории и подводных месторождений Арктики; в-третьих, военно-стратегическим и транспортным потенциалом региона.

Одним из факторов, оказывающим влияние на социально-экономическое развитие Арктической зоны, являются экстремальные природно-климатические условия, включая низкие температуры воздуха, сильные ветры и наличие ледяного покрова на акватории арктических морей.

Для ведения любой хозяйственной деятельности в регионе будь то ведение разведки и добычи углеводородного сырья, или судоходство, или ведение рыболовного промысла необходима точная и своевременная информация о климатических условиях на период проведения работ, наличие которой обеспечивает безопасную и продуктивную работу.

Данные относительно высоты волн, скорости ветра, океанических течений, температуры играют важную роль при выборе методов перемещения судов и персонала с платформы на берег, выборе типов оборудования, необходимого для проведения запланированных работ, выборе конструкции эксплуатационных платформ, а также объектов инфраструктуры, способной выдержать установленный срок службы, и выборе стратегий ликвидации чрезвычайных ситуаций и многое другое.

Внедрение информационных технологий в деятельность организаций привело к развитию распределенных информационных систем (РИС), реализующих различные информационные среды на основе разнообразных баз данных. Это привело к появлению неоднородного информационного пространства, реализующего гетерогенные базы данных. К управлению такими распределенными системами предъявляются повышенные требования с точки зрения обеспечения целостности и непротиворечивости данных,

степени однородности информационного пространства. Проблема состоит в технологической сложности перехода от централизованного однородного управления данными к распределенной обработке данных в гетерогенной информационной среде, включающей автономные системы работы в локальной сети гетерогенных баз данных, не представляющие возможность контроля выполнения глобальных транзакций. Принципы создания распределенных геоинформационных систем на основе гетерогенных баз данных, в том числе, могут быть применены для успешного освоения арктических территорий РФ.

На сегодняшний день можно говорить о недостаточном развитии методов хранения и использования геоинформации в распределенных гетерогенных базах данных.

Вследствие чего подтверждается **актуальность проблемы** совершенствования методов хранения и использования геоинформационных данных в различных форматах и их пространственно обособленных системах, базах данных, решению которой посвящена диссертация. Методы включают в себя оперативное извлечение и обработку данных, которые хранятся в различных системах и форматах, с целью своевременной обработки и предоставления информации о состоянии характеристик окружающей среды, влияющих на ведение хозяйственной деятельности.

Степень разработанности проблемы. В процессе исследования были изучены труды отечественных и зарубежных ученых по рассматриваемой проблеме, таких как Фуфаев Э.В., Оззу М.Т., Валдуриц П., Гаскаров Д.В., Кутузов О.И., Диго С.М., Б.П. Арсеньев, С.А.Яковлев., Бурковский В.Л., Дорофеев А.Н., Семьин С.В., Акальцов В.П., Ивакин А.Я., Шкодырев В.П., Середович В.А. и других.

Проанализированы материалы научных организаций, нормативно-правовые документы, программы и концепции федеральных и региональных органов власти, связанные с развитием арктического региона и проблемами информационного обеспечения его освоения.

Проблематика хранения и использования геоданных в различных форматах и их пространственно обособленных системах и базах данных в России изучена недостаточно. Не в полной мере разработаны методы и технологии хранения и использования геоинформации на основе распределенных, гетерогенных баз данных и знаний. 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 обособленных данных различных форматов.

Предметом исследования выступает разработка технологий использования гетерогенных баз данных в распределенных географических информационных системах.

Цель диссертационной работы заключается в совершенствовании технологий хранения и использования геоинформации на основе распределенных гетерогенных баз данных.

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

- 1) обосновать требования к геоданным используемым в прикладных геоинформационных системах с элементами управления;
- 2) построить модель геоинформационной системы с учетом препроцессорной подготовки гетерогенных данных;
- 3) модифицировать методику проектирования информационно-управляющей системы с учетом распределенных гетерогенных баз данных;
- 4) разработать математическую модель обработки гетерогенных данных в рамках верификации геоинформационной системы.

Теоретические и методические основы исследования.

Теоретической основой исследования являются труды отечественных и зарубежных ученых в области проблематики хранения и использования

геоданных в различных форматах, распределенных информационных систем, аналитические обзоры, посвященные анализу технологий доступа к различным видам данных, нормативно-правовые документы, связанные с методикой построения информационных систем. Методической основой исследования является:

- обобщение и анализ существующего опыта создания распределенных информационных систем с гетерогенными базами данных;
- системный анализ и концептуальное моделирование;
- аналитические исследования;
- математическое моделирование;
- прогнозирование;
- метод геоинформационного моделирования.

Обоснованность и достоверность результатов исследования, выводов и рекомендаций обеспечивается:

– 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;

– использованием для достижения цели работы нормативных документов, программ, документов федеральных и региональных органов власти, касающихся методики построения информационных систем;

– применением принципов системного анализа и концептуального моделирование, 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) требования к геоданным, используемым в прикладных геоинформационных системах с элементами управления;
- 2) модель геоинформационной системы с препроцессорной подготовкой гетерогенных данных;
- 3) модифицированная методика проектирования информационно-управляющей системы на основе распределенных гетерогенных баз данных;
- 4) математическая модель обработки гетерогенных данных в рамках верификации геоинформационной системы.

Научная новизна состоит в том, что:

- были обоснованы требования к геоданным используемым в геоинформационных системах с элементами управления, что позволит сократить сроки подготовки данных;

- впервые представлена модель геоинформационной системы на основе гетерогенных распределенных данных, охватывающая все слои представления информации: данные, обработка, интерфейс. Разработаны принципы и средства, позволяющие эффективно размещать и синхронизировать доступ к пространственным геоданным в различных форматах без их конвертации. Для объединения и преобразования данных к желаемому для пользователя виду используется «виртуальный процессор данных». Предложенная модель позволяет сократить время доступа к разноформатным геоданным;

- в методике впервые объединены технологии создания информационно-управляющей и геоинформационной системы, осложненной гетерогенностью и распределенностью баз данных, входящих в её основу. В

связи с чем было предложено создание виртуальной интегрированной базы данных, а также описание информационных ресурсов системы в виде реляционной базы данных. В методике усовершенствована технология, позволяющая привлекать разноформатные данные;

-математическая модель, предложенная для верификации ГИС, впервые применяется в исследуемых системах, отягощенных гетерогенными БД. В геоинформационной системе проводится анализ не только поступающих локальных данных, но и привлекаются из удаленного сервера данные, преобразующиеся с помощью «блока подготовки данных» в необходимый формат для их последующего использования. Полученные таким образом данные средствами проверки статистических гипотез и алгоритмов, позволяющих решать задачи прогнозирования множества состояний при полной определенности исходных значений, могут быть использованы, в том числе, и для прогнозирования выбранных климатических параметров. Представленная геоинформационная система существенно отличается от имеющихся тем, что в основе моделирования лежат не физические, а математические вероятностные модели.

Теоретическая значимость результатов диссертации заключается в том, что в работе был обобщен опыт практического применения технологий сбора и обработки информации, телекоммуникационных систем распространения пространственно-временной геоинформации, построения распределенных и гетерогенных баз данных, научно-методического аппарата создания информационных систем. Предложены оригинальные концептуальные принципы взаимодействия разноформатных источников информации, баз данных, с прямым пользователем и на их основе разработаны новые структурные модели геоинформационных систем с использованием гетерогенных данных. Усовершенствована методика создания информационно-управляющей системы с учетом распределенных гетерогенных баз данных.

Практическая и научная значимость работы заключается в том, что решена научно-техническая задача, имеющая существенное значение для геомоделирования и системного анализа многоуровневой и разнородной информации: исследованы технологии сбора и обработки информации с использованием вычислительной техники, телекоммуникационных систем распространения пространственно-временной геоинформации, построение распределенных и гетерогенных баз данных, возможность совершенствования технологий хранения и использования геоинформации на основе распределенных гетерогенных баз данных по средствам применения новых методик и моделей построения ГИС, что соответствует п.п. 7, 9 области исследования паспорта специальности 25.00.35 «Геоинформатика».

Предложенные в работе оригинальные авторские разработки могут быть использованы на практике в реальных условиях создания специализированных геоинформационных систем, использующих данные, находящиеся в распределенных и различных по форматам хранения базах данных. Для этого в работе даются конкретные рекомендации по их разработке, выделены основные этапы проектирования, построены диаграммы адекватности модели. Основные выводы и положения диссертационного исследования были использованы в научно-исследовательских работах и учебном процессе.

Апробация работы:

1. Свидетельство о регистрации электронного ресурса «Программа расчета георисков и вероятности появления нагонного наводнения» №19308. 24.09.2013 г.
2. Свидетельство о регистрации базы данных «Ice_Concentration» №2014621110, дата регистрации 07.08.2014 года.
3. Победитель конкурса грантов для студентов вузов, расположенных на территории Санкт-Петербурга, аспирантов вузов, отраслевых и академических институтов, расположенных на территории Санкт-Петербурга в 2013 году.

4. Международная научно-практическая конференция «Инфогео 2013» секция «геоинформационные системы» (26-28 ноября 2013г.).
5. Международная научно-практическая конференция «Инфогео 2014» секция «геоинформационные системы» (6-8 октября 2014г.).
6. Участие в НИР №1223 «Разработка и развитие методов, моделей и систем геоинформационного управления пространственно-распределенными объектами», 2013 год.
7. Участие в НИР №74.20.56 «Разработка методических основ геоинформационного управления рисками развития рекреационных приморских территорий», 2014 год.

Публикации.

По материалам диссертации опубликовано 13 работ, в том числе 2 статьи в ведущих журналах и изданиях, рекомендованных ВАК.

Структура и объем диссертации. Диссертация состоит из введения, трех глав, заключения, списка использованной литературы, списка сокращений и условных обозначений, приложений. Объем работы составляет 150 листов, в том числе 50 рисунков, 4 таблицы и три приложения. Библиография диссертационной работы составляет 67 наименований.

Содержание работы

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***: «Модели и методы распределенной обработки геоданных» были **сформулированы требования к геоданным, используемым в прикладных геоинформационных системах с элементами управления.**

Наличие и возможности информационных сетей различного масштаба, от локальных до глобальных, позволяют реализовать в управлении

развитие пространственно-распределенных организационно-технических систем.

Поэтому современные факторные условия позволяют по-новому взглянуть на систему управления как сложную, распределенную в геопространстве систему взаимодействия различных субъектов и объектов управления – что и объясняет возникновение феномена геоинформационного управления.

Феномен геоинформационного управления в современных условиях приводит к необходимости рассматривать и реализовать систему управления деятельностью человека как многомерное пространственное явление, включающее определенным образом распределенные объективные и субъективные факторы, и реализуемое в геопространстве.

С точки зрения геоинформационного управления геопространство можно структурировать и выделить требования к представлению геоданных (рисунок 1).

Исходя из предложенной структуры, были сформулированы следующие требования к геоданным, используемым в прикладных геоинформационных системах с элементами управления:

1. Размещение любых объектов, конструкций происходит в принятой системе координат, например, трехмерное пространство - $x, y, z - \varphi, \lambda, h$, или другой принятой системе координат.
2. Обязательное размещение в географическом пространстве объективных и субъективных факторов, влияющих на принятие и реализацию управленческих решений, которые характеризуются:
 - 2.1. координатами размещения – например, координаты положения в трехмерном пространстве - φ, λ, h ;
 - 2.2. особенностями и параметрами факторов - природные, ресурсные, предпринимательские и др.;
 - 2.3. доступностью и затратами для их использования в целях управления и др.



Рисунок 1. – Структура пространства решений в геоинформационном управлении

3. Обязательное наличие данных о событиях, возникающих и влияющих на процесс управления, которые характеризуются:
 - 3.1. координатами (φ, λ, h),
 - 3.2. временем возникновения (t),
 - 3.3. масштабами (пространственными, временными),
 - 3.4. последствиями (уровнем воздействия на геоинформационную систему) и др.
4. Наличие данных о состояниях объектов управления, характеризующиеся некоторой совокупностью переменных (атрибутов). Каждое такое состояние можно представить вектором n -мерного пространства переменных состояния, включающим:

- 4.1. координаты размещения объектов в географическом пространстве (φ, λ, h);
- 4.2. взаимное влияние объектов - система отношений, связей;
- 4.3. свойства, атрибуты, в том числе время;
- 4.4. значимость для управления, например, вес для выбора альтернатив;
- 4.5. степень участия в событиях и др.

Проводится **анализ современных технологий доступа к данным, заключающийся в определении нехватки простого метода работы с данными разного формата.** В результате чего была выявлена необходимость в совершенствовании технологий хранения и использования геоинформации на основе распределенных гетерогенных баз данных и проведена классификация, которая в дальнейшем подтвердила актуальность работы.

Проанализированы телекоммуникационные системы распространения пространственно-временной геоинформации и выделены два подхода к созданию единой распределенной геоинформационной системы и объединению, тем самым, различных баз хранения информации.

Рассмотрены основные направления и существующие проблемы в реализации геоинформационных систем на основе гетерогенных баз данных.

Рассматриваются подходы к организации распределенных информационных систем на базе гетерогенных база данных, где была выявлена проблематичность и незавершенность данного вопроса. Применение гетерогенных баз данных в распределенной информационной системе (РИС) до сих пор остается малоизученным направлением. Разработчики программных продуктов только начинают решать существующие проблемы в уже созданных технологиях, но не рассматривают вариант создания более совершенной технологии построения распределенной ИС на базе гетерогенных распределенных баз данных, которая могла бы избежать возникновения рассматриваемых проблем. Таким образом, создание новой модели архитектуры для подобных систем на основе

уже существующих является следующим шагом в освоении технологий работы с разноформатными распределенными базами данных и её применением в создании геоинформационных систем.

Во второй главе «Разработка модели геоинформационной системы на основе гетерогенных баз данных» представлена разработанная модель геоинформационной системы с препроцессорной подготовкой гетерогенных данных.

Определены современные проблемы создания геоинформационной системы, осложненной гетерогенностью и распределенностью баз данных.

За несколько последних лет было предложено немало технологий использования гетерогенных и распределенных баз данных (CORBA, JAVA, EJB), но ни одна из них не может обеспечить сохранение автономности источников данных. Поэтому на данный момент широкую известность получили лишь подходы доступа к гетерогенной базе данных распределенной информационной системы только для чтения данных. Такое состояние вызвано простотой реализации этого подхода, так как операции чтения двух независимых транзакций не конфликтуют и, соответственно, не могут нарушить целостности и непротиворечивости данных в распределенной информационной системе. Развитие технологий и требования пользователей подталкивает разработчиков к созданию такой структуры информационной системы, где предусматривалась бы возможность доступа пользователей как локальных, так и глобальных, не только для чтения данных, но и для их изменения.

Были проанализированы существующие проблемы реализации модифицирующих глобальных транзакций и выделены достаточные условия для безотказного управления транзакциями в распределенной информационной системе на основе гетерогенных баз данных:

- все гетерогенные базы данных должны быть локально атомарны;

- в отдельно взятой гетерогенной базе данных не должно выполняться больше одной подтранзакции глобальной транзакции;
- распределенная информационная система на основе гетерогенных баз данных должна справляться с появлением глобальных тупиков.

В распределенной информационной системе на основе гетерогенных баз данных должно существовать условие автономности интегрируемых баз данных. Для сохранения этого условия необходим протокол управления модифицирующими глобальными транзакциями, в качестве критерия корректности которого, была выбрана глобальная сериализуемость. Для такого протокола должны быть характерны следующие признаки:

- 1) каждая локальная СУБД должна гарантировать локальную сериализуемость и безкаскадные истории;
- 2) локальные истории выполнения транзакций должны быть без тупиков.

Использование данного протокола управления транзакциями предусматривает определенную модель управления распределенной информационной системой на основе гетерогенных баз данных, обеспечивающую интеграцию разрозненных баз данных в единое информационное пространство (рисунок 2).

На представленном рисунке показана архитектура системы, основанная на отдельно взятой гетерогенной базе данных, к каждой из которых можно обратиться по средствам одних и тех же методов доступа. Модель успешно работает в многопользовательском режиме и предотвращает рассогласование данных, вызванное одновременной работой с одними и теми же данными, а также организывает программную архитектуру.

В ходе исследования определена структурная модель системы, способная усовершенствовать работу по использованию информации, учитывающая разноформатность и распределенность предоставленных данных (рисунок 3).

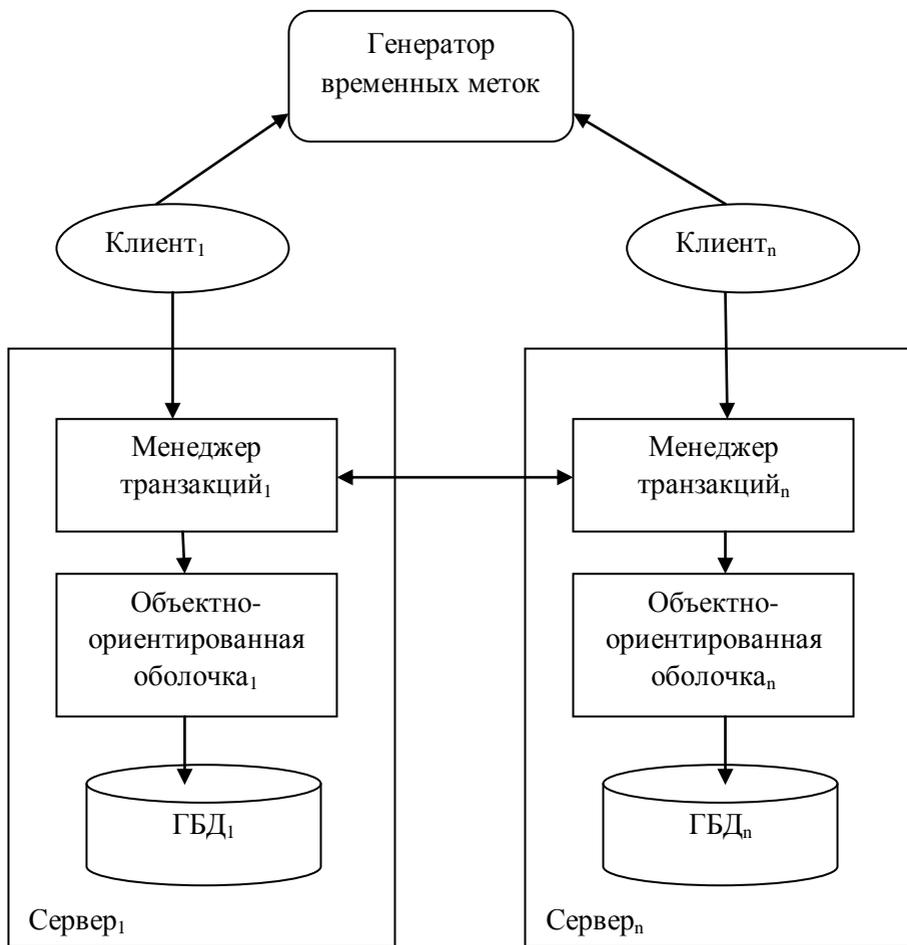


Рисунок 2 - Модель системы управления распределенной информационной системы на основе гетерогенных баз данных

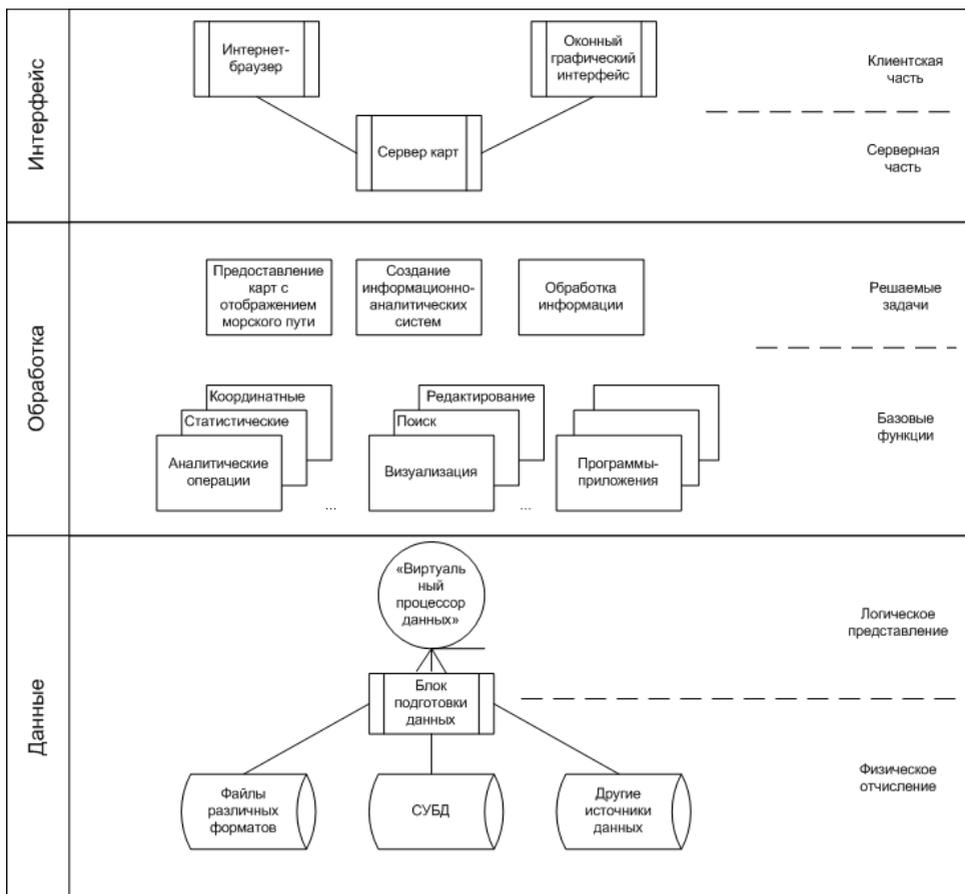


Рисунок 3 - Структурная модель геоинформационной распределенной системы на основе гетерогенных баз данных

На основании структурной модели формируется потоковая модель, которая представляет структуру прохождения информационных данных в рамках решаемой задачи (рисунок 4).

Клиент-серверная модель соединений, взявшая за основу принципы структурного программирования, а именно разделения процессов ввода-вывода и вычислительной обработки, в последнее время не справляется с

растущими потребностями разработчиков в области надежности и производительности систем.

Наиболее отвечающая всем требованиям равномерного распределения вычислительной нагрузки, обеспечения надежности и масштабируемости, является многоуровневая модель создания сложных программных систем. Она содержит три уровня представления данных «данные – обработка – интерфейс». Достоинствами данного представления являются независимость системы от количества программных компонентов, входящих в каждый уровень, свой набор протоколов и прикладных интерфейсов для каждого уровня, определяющий свою архитектуру.

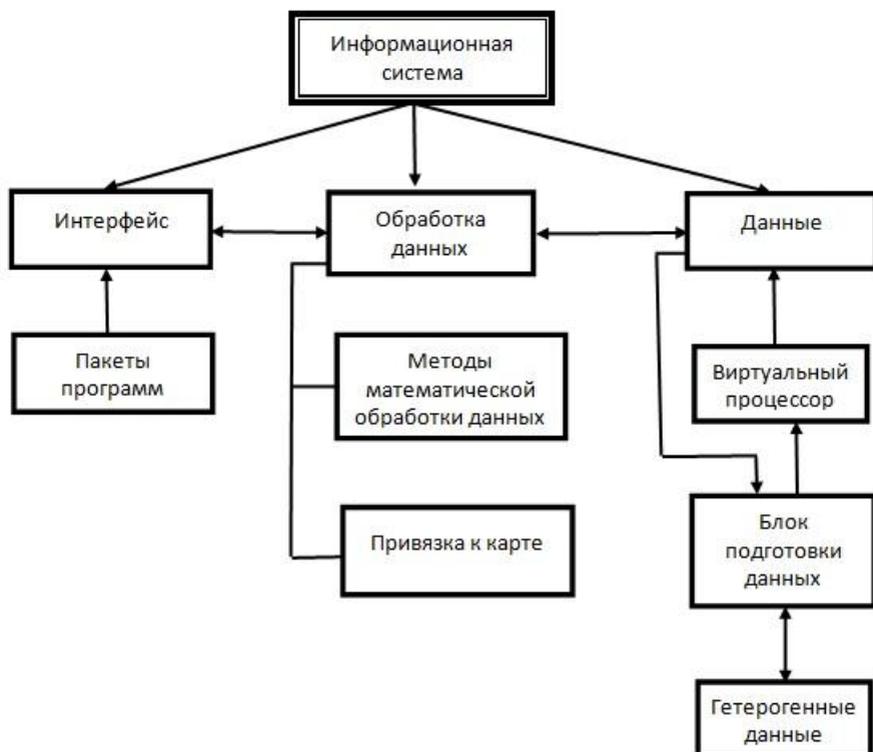


Рисунок 4 - Поточковая модель геоинформационной распределенной системы на основе гетерогенных баз данных

Во избежание трудностей доступа к данным было принято решение использовать механизм «виртуального процессора данных». Здесь он определяется как программно-технологический модуль, который позволяет единообразно получить доступ ко всем данным системы через стандартные интерфейсы (протоколы) и скрывающий при этом особенности размещения данных, структуры и форматы данных. При этом нет необходимости преобразования данных к формату системы, они остаются в исходном представлении и могут использоваться локально своими программными продуктами.

При более детальном рассмотрении «виртуального процессора данных» можно утверждать, что он состоит из набора не конкретизированных виртуальных баз данных и представляет собой некий виртуальный банк данных. Информационными источниками баз фактических состояний является цикл функционирования организационных подсистем сложного объекта управления, при рассмотрении геоинформационной системы как системы принятия управленческих решений. Информационными источниками баз плановых состояний и нормативных состояний могут являться входные документы, виртуальные результаты принятия управленческих решений, а также результаты решения регламентированных задач. В соответствии с регламентированными задачами и ситуациями в соответствующих базах гетерогенного банка данных текущего периода, а также, при необходимости, из баз предшествующих периодов, целесообразно формировать локальные по отношению к процессору базы данных.

Для решения задач управления территориями или в нерегламентированных ситуациях управления необходимо на основе виртуальной базы данных, гетерогенных баз данных и результатов обработки ситуаций формировать персональные базы для индивидуального пользователя данной системой. Соответственно, местоположение лица, принимающего решение будет критерием формирования персональной базы и обращения к ближайшему механизму «виртуального процессора данных».

В соответствии с временными параметрами технологии формирования персональной виртуальной базы данных сформирована схема

зависимостей между модулями подсистем представленных в «виртуальном процессоре данных» (рисунок 5).

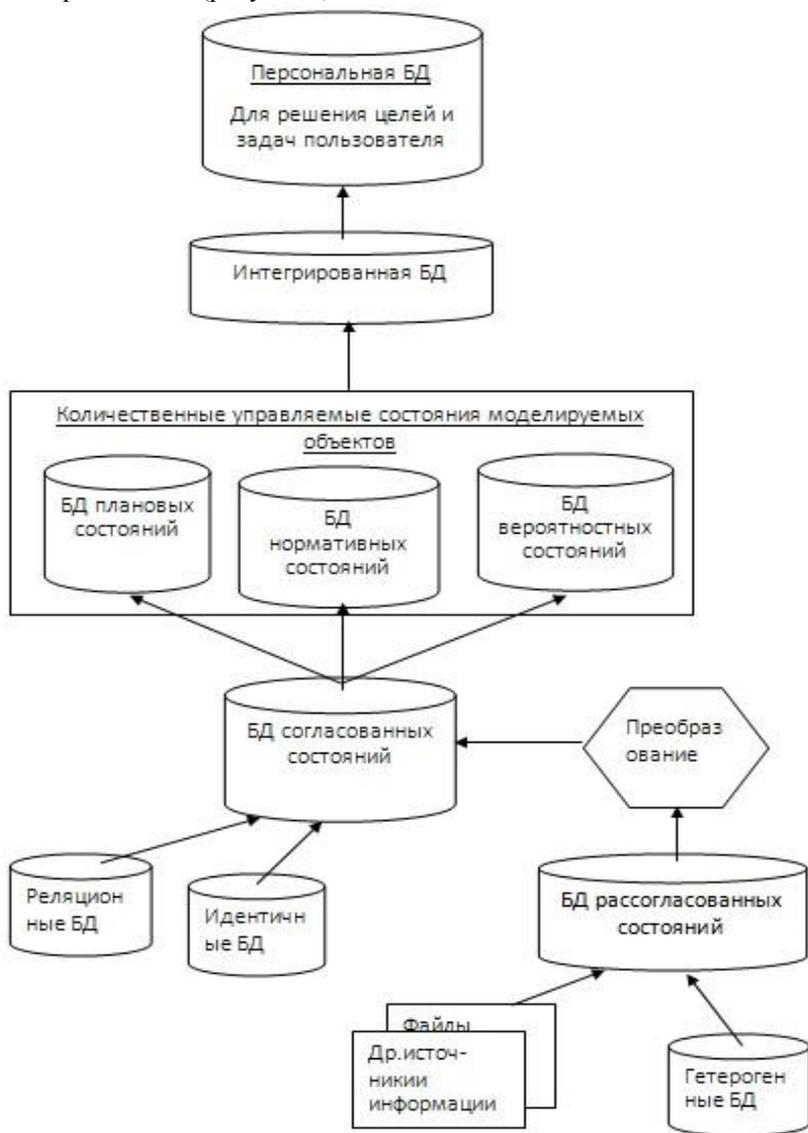


Рисунок 5 - Общая схема зависимостей между модулями подсистем данных в «виртуальном процессоре данных»

Представленные модели не могут напрямую использоваться при разработке ГИС, поэтому возникла задача разработки новой или модификации известной методики проектирования информационных систем.

Во второй главе также **представлена модифицированная методика проектирования информационно-управляющей системы на основе распределенных гетерогенных баз данных.**

За основу методики были взяты государственные стандарты 34-й и 19-ой серии, которые сводятся к этапам построения информационной системы.

Методика проектирования геоинформационной системы управления на основе гетерогенных баз данных заключается в объединении технологий создания информационно-управляющей и геоинформационной системы, при этом осложненной гетерогенностью и распределенностью баз данных, входящих в её основу. Изначально в проект должны быть заложены технологии, обеспечивающие процесс функционирования системы, из которых вытекают требования дополнительного формального описания предметной области.

Возможным решением для организации разработки подобной системы может быть ввод в состав информационного обеспечения объектов предметных областей и их структуры, входящих в систему гетерогенных 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 гетерогенной 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 метасвойств информации в базе данных.

Техническое проектирование, при реализации геоинформационной системы управления территориями на основе гетерогенных 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.

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.

Проектирование геоинформационной системы управления территориями должно состоять как минимум из четырёх основных шагов (рисунок 6).



Рисунок 6 - Шаги проектирования геоинформационной системы управления на основе гетерогенных баз данных

В ходе исследования выбрана спиральная модель проектирования для геоинформационной системы управления на основе гетерогенных баз данных. Именно она предусматривает углубление и последовательную конкретизацию деталей проекта и в результате выбирается обоснованный вариант, который удовлетворяет действительным требованиям заказчика и доводится до реализации.

В *третьей главе* «Информационная система обработки геоданных в Арктическом регионе РФ» представлена математическая модель

обработки гетерогенных данных в рамках верификации геоинформационной системы, которая позволяет одновременно использовать данные для оценок вероятностных характеристик двух параметров (силы ветра и плотности льда), выраженная в схеме взаимосвязи данных (рисунок 7).

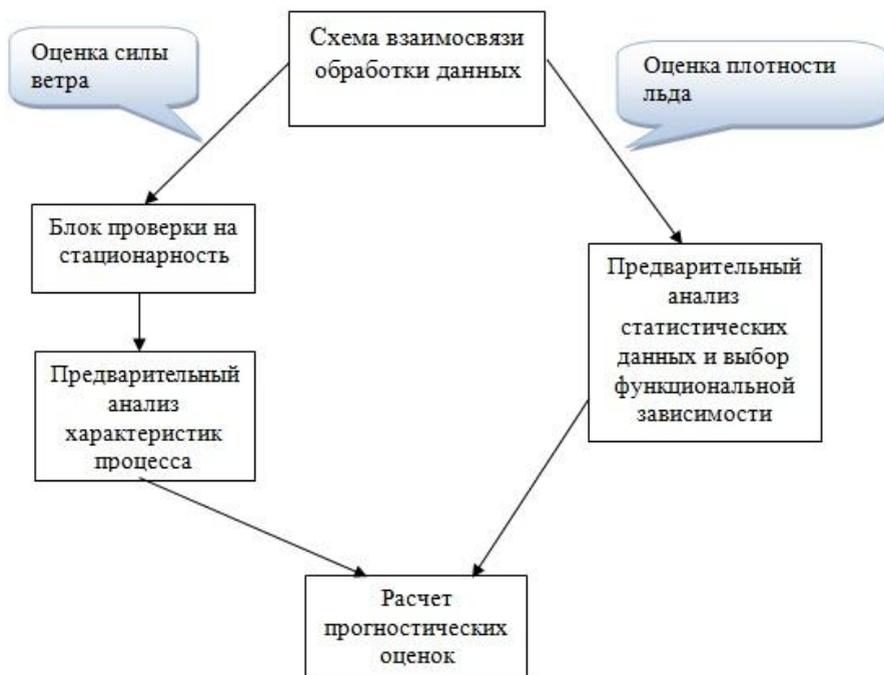


Рисунок 7 - Схема взаимосвязи данных в математической модели

При оценке вероятности наступления какого-либо события можно воспользоваться известными методами оценки вероятностных характеристик процессов. Один из таких методов был использован при моделировании оценки вероятности возникновения нагонных наводнений в городе Санкт-Петербург, который представлен временным рядом.

Возможность применения данного метода для оценки вероятности превышения силы ветра заданного порога на период времени Δt , зависит от удовлетворения следующим условиям:

- 1) величина значения исследуемого параметра климатического явления $x(t)$ должна быть описана в виде случайных функций $x(t) = F(t, \theta)$, где t -время, θ – вектор случайных параметров, влияющих на изменение $x(t)$;
- 2) функция распределения случайного процесса должна быть описана с достаточной достоверностью под нормальный закон распределения;
- 3) существуют такие временные интервалы, на которых $x(t)$ является стационарным случайным процессом.

В большинстве случаев геоинформационные данные могут быть представлены в виде временных рядов или иных случайных процессов.

В качестве случайного события при данной постановке задачи может быть рассмотрена задача превышения силы ветра одного из уровней описанного баллами, представленными в шкале Бофорта, в которой существует интервал скорости, на основании которого присваивается тот или иной бал. Чаще нас будет интересовать случай выхода за пределы 11 балла.

Построим случайную функцию, описывающую изменение силы ветра во времени $S(t)$ (рисунок 8).

При известном случайном процессе можно определить вероятность невыхода силы ветра за промежуток, не превышающий выбранный балл по шкале Бофорта по следующей формуле:

$$P_i(t_3) = P(S(t_3)/S(t_0) = S_0 \leq \delta_i), \quad (1)$$

где:

$P_i(t_3)$ – вероятность невыхода силы ветра за пределы i -го балла по шкале Бофорта,

δ_i - границы шкалы Бофорта на i -ом срезе,

S_0 - начальное значение силы ветра,

$S(t_3)$ -случайная функция изменения силы ветра на заданный момент времени t_3

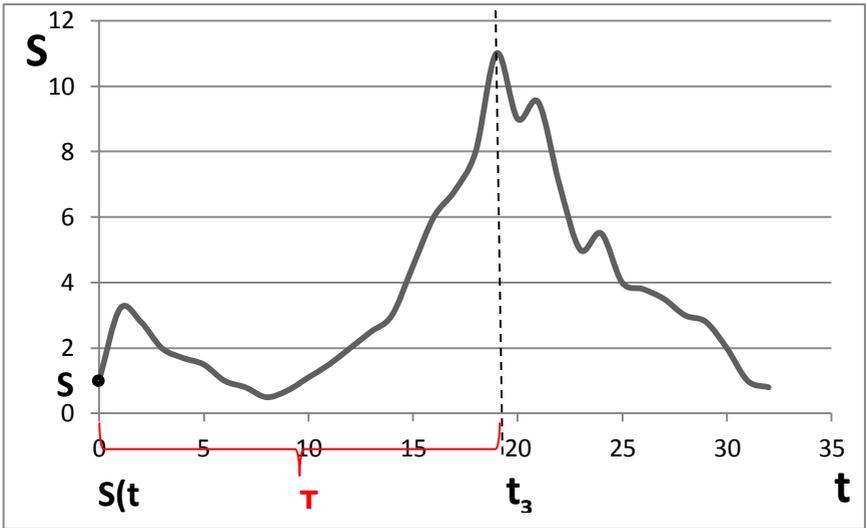


Рисунок 8 Случайная функция, описывающая изменение силы ветра во времени

$S(t_0)$ - случайная функция изменения силы ветра на начальный момент времени t_0

После соответствующих преобразований, используя ранее предложенные оценки, получим:

$$P_1(t_3) = \Phi\left(\frac{\delta_i - m_s - r_s(0, t_3) \cdot (s_0 - m_s)}{\sigma_s \sqrt{1 - r_s^2(0, t_3)}}\right) - \Phi\left(\frac{-m_s - r_s(0, t_3) \cdot (s_0 - m_s)}{\sigma_s \sqrt{1 - r_s^2(0, t_3)}}\right) \quad (2)$$

где: m_s – математическое ожидание случайного процесса;

σ_s - среднеквадратическое отклонение;

$r_s(0, t_3)$ – автокорреляционная функция.

Для предсказания плотности льда применяется алгоритм, позволяющий решать задачи прогнозирования состояний при полной определенности исходных значений множества, а именно оптимальный фильтр Калмана – Бьюси (рисунок 9).



Рисунок 9 - Алгоритм фильтра Кальмана-Бьюси

Алгоритм состоит из двух повторяющихся фаз: предсказание и корректировка. На первой рассчитывается предсказание состояния в следующий момент времени (с учетом неточности их измерения). На второй, новая информация корректирует предсказанное значение (также с учетом неточности и зашумленности этой информации).

Расчетные параметры алгоритма:

x_k^- - предсказание состояния системы в текущий момент времени;

x_{k-1} - состояние системы в прошлый момент времени;

u_{k-1} - управляющее воздействие в прошлый момент времени;

Fx_{k-1} - матрица перехода между состояниями;

Bu_{k-1} - матрица применения управляющего воздействия;

P_{k-1} - ошибка в прошлый момент времени;

Q - ковариация шума процесса;

H – матрица измерений, отображающая отношение измерений и состояний;

R – ковариация шума измерений;

z_k - измерение в текущий момент времени;

I – матрица идентичности.

Для проверки были использованы данные за выбранный период времени с 01.01.2012 по 24.08.2013 года в трех точках, взятых на протяжении Северного морского пути и расположенных в координатах (50; 80), (135; 85), (173; 95). Информация была представлена как CSV файл и преобразована в базу данных, которая впоследствии была запатентована. В результате был получен устойчивый прогноз, что и показано на рисунке 10.

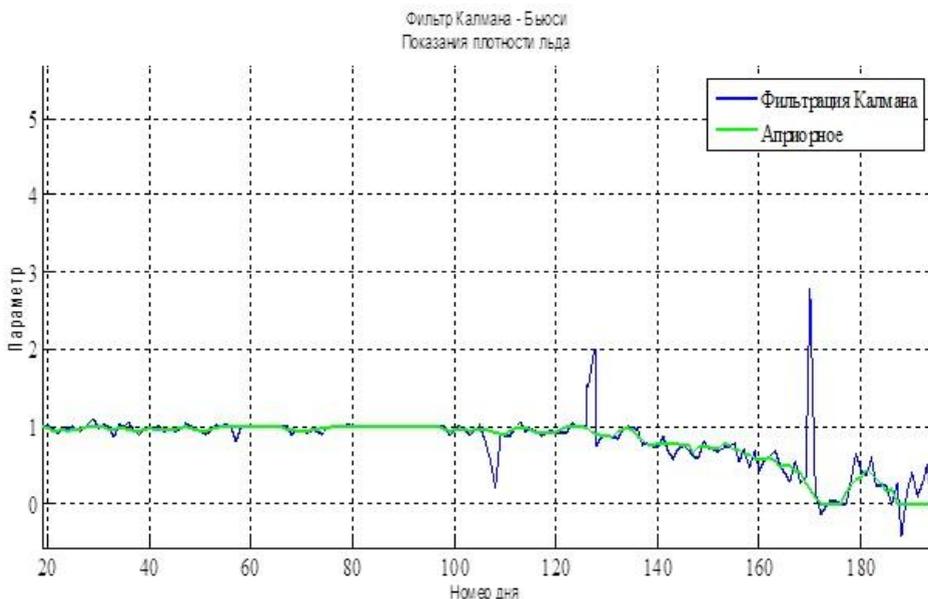


Рисунок 10 - Прогноз изменения концентрации плотности льда на протяжении Северного морского пути с помощью фильтра Калмана – Бьюси в одной из трех точек.

Приводятся результаты моделирования геоинформационной системы «Оценка климатических параметров для ведения хозяйственной деятельности на арктической территории РФ», подтверждающие **возможность реализации** выбранной модели геосистемы. При построении была использована методология объектно-ориентированного анализа и проектирования, выраженная по средствам визуального языка проектирования UML и программного продукта RationalRose 7.0. Построены концептуальная, статическая и физическая модель сложной системы, представленные в виде диаграмм вариантов использования, классов, 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 системы и размещение отдельных компонентов, подтверждающие адекватность выбранной модели геосистемы.

Основные результаты и выводы. Личный вклад автора.

1. Обоснованы требования к геоданным, используемым в прикладных геоинформационных системах с элементами управления, которые получены автором при выполнении НИР «Разработка и развитие методов, моделей и систем геоинформационного управления пространственно-распределенными объектами».
2. Автором лично проведена классификация современных информационных технологий в области хранения, передачи и обработки геоинформации с использованием вычислительной техники по критериям распределенности и гетерогенности данных и выявлено, что ни одна представленная технология не может удовлетворить требованиям создания распределённой вычислительной системы обработки геоданных с применением гетерогенных БД.
3. Построена модель геоинформационной системы с учетом препроцессорной подготовки гетерогенных данных, поддерживающая многопользовательский режим и основанная на технологиях доступа к данным как локально, так и через глобальную сеть Интернет. Разработаны принципы и средства, способные эффективно размещать и синхронизировать доступ к пространственным геоданным в различных форматах без явной конвертации. Для объединения и преобразования данных к желаемому виду пользователя автором было предложено использовать «виртуальный процессор данных». Как элемент виртуального процессора, «блок подготовки данных» в системе позволяет получить доступ к любым источникам информации, необходимым для принятия управленческого решения.

4. Модифицирована методика проектирования геоинформационной системы на основе распределенных гетерогенных баз данных, которая заключается в объединении технологий создания информационно-управляющей и геоинформационной системы, осложненная гетерогенностью и распределенностью баз данных, входящих в её основу. Автором были предложены: измененная схема проектирования, включающая новые шаги и функции.
5. Разработана математическая модель обработки гетерогенных данных в рамках верификации геоинформационной системы, которая позволяет одновременно обрабатывать данные по двум параметрам (силы ветра и плотности льда), с целью получения оценок вероятностных характеристик данных процессов. Автором лично была предложена схема взаимосвязи различных данных в математической модели.

В диссертации рассматривались данные из разрозненных источников. Всего было представлено 2 параметра, однако предложенные автором научные решения позволяют решать, не только монопараметрические задачи, а также многомерные задачи более чем для двух параметрах и на значительных территориях. Однако это не являлось целью исследования, но задача может быть предметом дальнейших исследований.

По теме диссертации опубликованы следующие работы:

в изданиях, рекомендованных ВАК:

1. 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.
2. Колбина О.Н. 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.

другие публикации:

3. Колбина О.Н. Современные и теоретические аспекты управления распределёнными базами данных/ Истомин Е.П., Колбина О.Н. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб.науч.тр./Вып. 1(9)-СПб.:ООО «Андреевский издательский дом» - 2011г.,105 с
4. Колбина О.Н. Применение распределённых баз данных в геоинформационных системах прогнозирования георисков / Истомин Е.П., Колбина О.Н., Зоринова Е.М. // Сборник трудов международной научно-практической конференции «Инфогео-2013», ООО «Андреевский издательский дом»,2013 г.
5. Колбина О.Н. Информационная система прогноза рисков наводнений в Санкт-Петербурге / Истомин Е.П., Колбина О.Н., Слесарева Л.С., Петров Я.А. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб.науч.тр./Вып. 10-СПб: ООО «Андреевский издательский дом» - 2013г.,156 с.
6. Колбина О.Н. Разработка методов анализа и синтеза системы управления производственным Предприятием / Соколов А.Г., Истомин Е.П., Кунин В.А., Сквородников А.П., Колбина О.Н. // Информационные технологии и системы: управление, экономика, транспорт, право: Сб.науч.тр./ Вып. 11-СПб ООО Андреевский издательский дом, СПб, 2013. – С. 146.
7. Колбина О.Н. База данных как центральный элемент маркетинга отношений/ Соколов А.Г., Колбина О.Н. //Информационные технологии и системы: управление, экономика, транспорт, право:Сб. науч. тр./Вып. 1 (12) – СПб: ООО «Андреевский издательский дом» - 2014 г., 116 с.
8. Колбина О.Н. Функциональное моделирование геоинформационной системы с применением распределенных гетерогенных баз данных //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.
9. 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.
 10. Колбина О.Н. ГИС. Обзор программного обеспечения ГИС / 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.
 11. Колбина О.Н. Анализ технологий создания виртуальных хранилищ данных/ Колбина О.Н., Улитина Т.С. //Информационные технологии и системы: управление, экономика, транспорт, право:Сб. науч. тр./Вып. 13 / Под ред. д.т.н., проф. Истомина Е.П. – СПб: ООО «Андреевский издательский дом» - 2014 г., 79 с.
 12. Пат. №2014621110 Российская Федерация, База данных Ice_Concentration /Истомин Е.П., Колбина О.Н., Степанов С.Ю., Миранков В.А.; заявитель и патентообладатель РГГМУ. – Заявл. № 2014620482, опубл. 07.08.14.
 13. Свидетельство о регистрации №19308 Российская Федерация, Программа расчета георисков и вероятности появления нагонных наводнений /Истомин Е.П., Колбина О.Н., Слесарева Л.С.; заявитель и патентообладатель РГГМУ. – Опубл. 24.06.13.