ФЕДЕРАЛЬНОЕ АГЕНТСТВО ЖЕЛЕЗНОДОРОЖНОГО ТРАНСПОРТА

Федеральное бюджетное образовательное учреждение
высшего образования
«Иркутский государственный университет путей сообщения»
(ФГБОУ ВО ИрГУПС)

УТВЕРЖДАЮ:
И.о. ректора
доктор технических наук,
профессор
Хоменко Андрей Павлович
«18» 02 2022 г.

ОТЗЫВ

ведущей организации на диссертацию Воробьева Андрея Владимировича «Методологические основы обработки пространственной информации для поддержки принятия решений на основе агрегированных цифровых двойников (на примере высокоширотных геомагнитных данных)», представленную на соискание ученой степени доктора технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации

Актуальность темы исследования

Проведенные в рамках диссертационной работы научные исследования направлены на разработку методологических основ обработки распределенной пространственной информации в отраслевых системах поддержки принятия решений для повышения их эффективности в условиях низкой ситуационной осведомленности на основе агрегированных цифровых двойников (на примере высокоширотных геомагнитных данных).

Актуальность проблемы модернизации методологии обработки распределенной пространственной информации в отраслевых системах поддержки принятия решений в условиях низкой ситуационной осведомленности не вызывает сомнений и подтверждается п. 5 «Технологии информационных, управляющих, навигационных систем» и п. 19 «Технологии мониторинга и прогнозирования состояния окружающей среды, предотвращения и ликвидации ее загрязнения» Перечня критических технологий Российской Федерации, а также п. 20, а «Переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам, новым материалам и способам конструирования, создание систем обработки больших объемов данных, машинного обучения и искусственного интеллекта» Стратегии научно-технологического развития Российской Федерации.

Важно отметить, что решение выделенной проблемы, применительно к рассматриваемым в работе отраслевым системам поддержки принятия решений, в значительной степени обеспечивает технологический базис для реализации актуальных стратегий государства (например, таких, как «Стратегия развития Арктической зоны Российской Федерации и обеспечения национальной безопасности») и во многом определяет возможность

УГАТУ
Вх. № 360-13
от 01.03.2022

эффективной реализации Государственных программ (например, Государственная программа «Социально-экономическое развитие Арктической зоны Российской Федерации»).

Оценка структуры и содержания работы

Диссертационная работа состоит из введения, 6 глав, заключения, списка литературы и приложений. Общий объем диссертации – 345 страниц основного текста, приложений – 11 страниц. Общее число рисунков – 62, таблиц – 38. Список литературы включает в себя 213 наименований.

Во введении на примере высокоширотных геомагнитных данных обосновывается актуальность разработки методологических основ обработки распределенной пространственной информации в системах поддержки принятия решений для повышения их эффективности в условиях низкой ситуационной осведомленности на основе агрегированных цифровых двойников. Определяются объект и предмет исследования, формулируются основные научные задачи и выбираются методы исследований.

В **первой главе** на примере высокоширотных геомагнитных данных проводится идентификация и анализ проблемы обработки распределенной пространственной информации в системах поддержки принятия решений, обозначаются пути ее решения и разрабатывается комплекс показателей для оценки их эффективности и надежности.

Так, на примере высокоширотных магнитных станций сети IMAGE было показано, что в 50 % случаев математическое ожидание размера пропущенного фрагмента превышает ~58,5 мин. Усредненный (по всем IMAGE-станциям) размер пропущенного фрагмента составляет ~1066 мин. Математическое ожидание числа отказов с восстановлением превышает 45 в год. При этом 50 % IMAGE-станций испытывают за год более 17 отказов, а вероятность безотказной работы всех магнитометров не превышает 0,47. В крайних случаях объем отсутствующих фрагментов одной станции может превышать 11,2 % от наработки, при этом среднее время восстановления достигает 10 суток и более.

Повысить эффективность и надежность систем сбора геомагнитных данных предлагается путем комплексного применения методов геоинформатики, машинного обучения и цифровых двойников.

Во **второй главе** автором разрабатывается и формализуется в теоретико-множественном базисе концепция обработки распределенной пространственной информации в условиях низкой ситуационной осведомленности (на примере геомагнитных данных) в системах поддержки принятия решений на основе технологии агрегированных цифровых двойников сложных технически объектов.

На примере DTI-двойника магнитных станций, функционально сопряженном с вычислительным модулем, показано, что предлагаемая концепция способна обеспечить информационную избыточность, и, вследствие создаваемого эффекта резервирования повысить надежность и отказоустойчивость соответствующей технической системы.

Здесь же автором предлагается подход к кластеризации источников информации, основанный на статистическом анализе временных рядов геомагнитных данных, получаемых с выхода пространственно-распределенных магнитных станций.

В **третьей главе** путем ансамблирования методов машинного обучения, геостатистики и статистического анализа автором разрабатывается комплекс математических моделей, методов и алгоритмов, способных в соответствии с предложенной выше концепцией реализовать обработку, анализ и ретроспективный прогноз пространственной информации, получаемой с выхода разнородных цифровых и физических двойников.

Разрабатывается, приводится и описывается логическая модель интеграции цифрового двойника в процесс сбора и обработки геомагнитных данных, а также алгоритм сбора геомагнитных данных с использованием технологии цифрового двойника.

На примере магнитной станции KIL показано, что внедрение цифрового двойника в процессы сбора и регистрации геомагнитной информации позволяет восстановить до 99,6 % данных, при этом для 86,7 % восстановленных значений MSE ≤ 12 нТл^2.

В **четвертой главе** основываясь на предложенном подходе к пространственной кластеризации опорных геодезических пунктов разрабатывается комплекс функциональных подсистем обработки (в частности, реконструирования) и динамической мультимасштабной визуализации пространственной информации, получаемой с выхода агрегированных цифровых двойников и их физических прототипов (на примере спутниковых данных, магнитных обсерваторий и вариационных станций)

В качестве основы разрабатываемой подсистемы соискателем предлагается использовать клиент-серверную архитектуру. При этом серверная составляющая предложенной автором архитектуры объединяет уровни данных и бизнес-логики: первый из них отвечает за обращение к провайдерам пространственной информации, а второй – за предварительную обработку поступивших от них откликов. Клиентская составляющая, в свою очередь, интегрирует функции визуализации пространственных данных для конечного пользователя, обеспечивая таким образом соответствующую поддержку интерактивного взаимодействия с приложением посредством набора элементов пользовательского интерфейса.

В результате предложенная архитектура обеспечивает динамическую визуализацию пространственного распределения возмущенной составляющей геомагнитного поля как в виде изодинам (линий уровня), так и в виде векторного поля.

В пятой главе на основе полученных результатов соискателем разрабатывается и анализируется экспериментальный прототип проблемно-ориентированной системы поддержки принятия решений, функционально связанный с системой агрегированных цифровых двойников, обеспечивающий повышение эффективности поддержки решений при эксплуатации технических систем высокоширотной инфраструктуры в условиях низкой ситуационной осведомленности.

На основе анализа значений коэффициента детерминации R^2 автор показывает, что в рамках решаемой проблемы подходы, основанные на методе множественной линейной регрессии и искусственной нейронной сети с функцией активации ReLu, являются наилучшими (R^2 = 0,816 и R^2 = 0,814 соответственно). Таким образом, например, на базе регрессионной модели уровень ГИТ в северной части Кольского полуострова в районе железнодорожной станции «Выходной» (VKH) можно оценить со среднеквадратической ошибкой, не превышающей 0,12 А^2.

Последующее сопряжение полученных результатов с подсистемой визуализации пространственного распределения геомагнитных возмущений позволило соискателю выделить географические границы регионов и принадлежащие им пространственно-распределенные технические системы, наиболее уязвимые в периоды ГМА.

В **шестой главе** осуществляется обработка и анализ данных, полученных в результате внедрения результатов работы в организации, предприятия и учреждения, представляющие различные отрасли народного хозяйства (наука, промышленность, образование), а также верифицируется и оценивается эффективность предложенных решений.

Так, например, значительная часть полученных результатов используется в Геофизическом центре РАН в рамках выполнения проекта «Системный анализ динамики геофизических процессов в российской Арктике и их воздействие на развитие и функционирование инфраструктуры железнодорожного транспорта»

Важно отметить, что созданная автором СППР для прогноза, визуализации и экспресс-анализа параметров в авроральной зоне «Aurora-Forecast» (http://aurora-forecast.ru/) используется в гелиогеофизическом центре ФГБУ Института прикладной геофизики им. академика Е.К. Федорова в цикле работ направленных на создание адаптируемых моделей авроральных овалов и каспов Земли по данным наземных и космических средств мониторинга солнечной активности и околоземного пространства с информационной независимостью от зарубежных источников информации и совместимостью с программным обеспечением технологии «Сюжет-МВ».

В **заключении** представлены основные результаты и выводы.

Новизна полученных результатов:

1. Адаптированный комплекс параметров для оценки показателей надежности систем, моделей, методов и алгоритмов сбора и обработки распределенной пространственной информации в отраслевых системах поддержки принятия решений (на примере высокоширотных геомагнитных данных) отличается тем, что включает процедуру отбора, ранжирования (по значимости) и анализа связанных с данными показателями свойств анализируемых систем, моделей, методов и алгоритмов путем экспертной оценки корреляций параметров статистических моделей и факторов, влияющих на получение первичной информации.

2. Концепция обработки распределенной пространственной информации (на примере геомагнитных данных) в отраслевых системах поддержки принятия решений в условиях низкой ситуационной осведомленности, отличается тем, что предлагается создание и применение информационной системы агрегированных цифровых двойников технически сложных объектов, пространственная кластеризация которых основана на аппроксимации и анализе хвостов статистик распределений значений целевого параметра.

3. Комплекс математических моделей, методов и алгоритмов, реализующих обработку, анализ и ретроспективный прогноз распределенной пространственной информации отличается тем, что

– предложено ансамблирование методов машинного обучения, геостатистики и статистического анализа, что позволило сократить вектор регрессоров модели цифрового двойника технического объекта и повысить надежность ее срабатывания;

– разработана логическая модель интеграции цифровых двойников в процесс сбора и обработки распределенной пространственной информации, обеспечивающая автоматизированный контроль качества восстановления данных, путем сравнительного анализа информации, получаемой с выхода цифрового двойника и его физического прототипа.

4. Комплекс функциональных подсистем для пространственного реконструирования и динамической мультимасштабной визуализации распределенной пространственной информации, получаемой с выхода разнородных агрегированных цифровых двойников и их физических прототипов, отличающийся тем, что в MVC-архитектуру программного средства представления данных интегрированы модели трансформации, реализующие возможность сопряжения распределенной пространственной информации, синтезируемой на различных уровнях кластерной иерархии.

5. Экспериментальный прототип проблемно-ориентированной отраслевой системы поддержки принятия решений, отличающийся тем, что благодаря кластерному подходу при интеграции разнородных физических и цифровых DTI-/DTA-двойников в единую систему, сопряженную с глобальной информационной средой пространственных данных и моделей, обеспечивается повышение эффективности принятия управленческих решений в условиях низкой ситуационной осведомленности.

Степень достоверности результатов

исследования подтверждается результатами сравнительного анализа эмпирической информации и данных, полученных в результате математического и численного моделирования, непротиворечивостью полученных результатов основополагающим законам физики, а также экспертной оценкой и повторяемостью полученных результатов.

Теоретическая и практическая значимость результатов, полученных автором диссертации

Значение результатов для теории сбора, обработки и анализа слабоструктурированной распределенной пространственной информации для задач поддержки принятия решений при эксплуатации технических систем высокоширотной инфраструктуры заключается в том, что предложен подход к оценке надежности источников пространственной информации, разработана концепция системы агрегированных цифровых двойников, а также предложен, обоснован и апробирован (на примере источников геомагнитных данных) метод пространственной кластеризации опорных геодезических пунктов, модернизировано математическое обеспечение для реконструирования пространственной информации, получаемой с выхода физических и цифровых двойников DTI и DTA типа, а также разработан комплекс математических моделей, методов и алгоритмов, обеспечивающих (на основе информации, получаемой с выхода системы агрегированных цифровых двойников) поддержку принятия решений в условиях низкой ситуационной осведомленности.

Значение полученных результатов для практики заключается в том, что внедрение цифровых двойников пространственно-распределенных технических объектов в процессы сбора и регистрации распределенной пространственной информации за счет информационной избыточности повышает отказоустойчивость их физических прототипов (так, например, среднее время безотказной работы источника геомагнитных данных – магнитной станции KIL повысилось на 12,4 %, а число необходимых для этого опорных геодезических пунктов сократилось в ~4,7 раза); разработка проблемно-ориентированных систем поддержки принятия решений на базе предложенных в работе концепции, моделей, методов и алгоритмов, обеспечивает (в том числе, за счет увеличения пространственного разрешения в ~7,4 раза) повышение качества процессов аналитического контроля распределенной пространственной информации, а также эффективности и своевременности (за счет осуществления краткосрочного 15–30 мин прогноза контролируемых параметров) принятия управленческих решений на ее основе.

Рекомендации по использованию результатов и выводов диссертации

Полученные в ходе выполнения диссертационного исследования результаты можно рекомендовать к внедрению и использованию в организациях, деятельность которых связана с принятием решений на основе пространственно-распределенных геомагнитных данных, например, в геофизике, навигации, в области техносферной безопасности и пр.

Отдельные научные и практические результаты могут быть использованы в учебном процессе высших учебных заведений при подготовке студентов по направлениям, связанным с информационными технологиями.

Замечания по диссертационной работе

1. В диссертационной работе не уделено достаточно внимания тому, что понимается под ситуационной неосведомленностью.
2. В работе явным образом не рассматривается возможность и ограничения применения предлагаемых решений в случае работы с большими данными.
3. В шестой главе диссертационной работы недостаточно явно при анализе эффективности полученных результатов прослеживается сопоставление их с мировым научным опытом.
4. В списке лиц, которые принимали участие в разработке данной тематике отсутствуют академик РАН А.А. Красовский, профессора Н.Н. Белоглазов, В.Г. Чикин и их ученики, которые внесли огромный вклад в разработке корреляционно-экстремальных систем навигации по геофизическим полям Земли.
5. Понятие логической модели в диссертации отличается от общепринятого, под которым обычно подразумевают использование того или иного логического исчисления и логический вывод на его основе.
6. К сожалению, диссертант далеко не в полной мере воспользовался отечественными публикациями последних лет в области регрессионного анализа. Это касается вопросов итерпретируемости разрабатываемых моделей, исследования мультиколлинеарности, комплексной оценки адекватности, снижения размерности, проведения «конкурса» моделей. В частности, переход от модели (3.3) к модели (3.6) (со снижением числа регрессоров с 14 до 3) в работе «рекомендуется выполнять вручную, например, с помощью попарного сравнения статистик доступных предикатов», с.123. Видимо, все-таки предикторов (независимых переменных)? В действительности эта задача легко может быть решена точно с привлечением алгоритмов, основанных на использовании аппарата линейно-булевого программирования. При этом полученное решение с большой вероятностью не совпадет с моделью (3.6).
7. Досадно много опечаток в нумерации рисунков и таблиц. В частности, дублируются номера рисунков 2.12, 3.1, 4.4., 5.3, 6.7, 6.8.

Указанные замечания не снижают общую ценность диссертационного исследования и не влияют на её положительную оценку.

Заключение

1. Диссертация Воробьева Андрея Владимировича на соискание ученой степени доктора технических наук является научно-квалификационной работой, в которой на основании выполненных автором исследований решена научная проблема модернизации методологических основ обработки пространственной информации в отраслевых системах поддержки принятия решений в условиях низкой ситуационной осведомленности при эксплуатации технических систем высокоширотной инфраструктуры, что соответствует требованиям п.9 «Положения о присуждении ученых степеней», а ее автор заслуживает присуждения искомой ученой степени.
2. Диссертация обладает внутренним единством, содержит новые научные результаты и положения.

3. Автореферат диссертации соответствует содержанию диссертации по основным квалификационным признакам: цели, задачам, новизне, актуальности, достоверности, научной и практической значимости.

4. Основные научные результаты диссертации изложены в 54 публикациях, в том числе 12 статьях в ведущих рецензируемых научных журналах, входящих в перечень изданий, рекомендованных ВАК, 17 публикациях в отечественных и зарубежных изданиях, индексируемых международными системами Scopus / Web of Science, из них 1 – Q1, 2 – Q2); 4 монографиях, изданных в России и за рубежом, 4 патентах на изобретение; 3 патентах на полезную модель; 4 свидетельствах о государственной регистрации программы для ЭВМ, трудах конференций и других работах.

5. Тема и содержание диссертации соответствуют специальности 2.3.1. Системный анализ, управление и обработка информации

Диссертация соответствует требованиям п.9 «Положения о присуждении ученых степеней», а ее автор, Воробьев Андрей Владимирович, заслуживает присвоения ученой степени доктора технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации.

Отзыв обсужден на заседании научного семинара кафедры «Информационные системы и защита информации». Дата 27.01.2022, протокол № 2.

Председатель семинара
доктор технических наук, профессор
Данеев Алексей Васильевич,
профессор, кафедры «Информационные системы и защита информации» ФГБОУ ВО «Иркутский государственный университет путей сообщения»

Докторская диссертация защищена
по специальности 05.13.18. Математическое моделирование, численные методы и комплексы программ

Адрес организации: 664074, Сибирский федеральный округ, Иркутская область, г. Иркутск, ул. Чернышевского, д. 15.
Рабочий телефон: +7(3952)638-399.
Адрес эл. почты: mail@irgups.ru

Подпись Данеева А.В.
ЗАВЕРЯЮ:
Начальник ... его отдела ИрГУПС
Подпись
«____» _____ 20___ г.