



**САМАРСКИЙ УНИВЕРСИТЕТ**  
SAMARA UNIVERSITY

Федеральное государственное автономное  
образовательное учреждение высшего образования  
«Самарский национальный исследовательский университет  
имени академика С.П. Королева»

ул. Московское шоссе, д. 34, г. Самара, 443086  
Тел.: +7 (846) 335-18-26, факс: +7 (846) 335-18-36  
Сайт: www.ssau.ru, e-mail: ssau@ssau.ru  
ОКПО 02068410, ОГРН 1026301168310,  
ИНН 6316000632, КПП 631601001

01 ОКТ 2020 № 104-И225

На № \_\_\_\_\_ от \_\_\_\_\_

УТВЕРЖДАЮ

Ректор Самарского университета

доктор экономических наук, профессор



В.Д. Богатырев

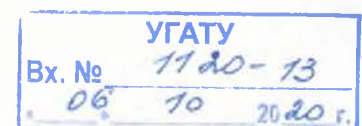
2020 г.

**ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ  
на диссертацию Воробьевой Гульнары Равилевны  
«Методологические основы обработки неоднородной пространственно-  
временной информации в системах поддержки принятия решений на  
основе технологий больших данных  
(на примере геомагнитных данных)»,  
представленную на соискание ученой степени доктора технических наук  
по специальности 05.13.01 - Системный анализ, управление и обработка  
информации (информационные и технические системы)**

Рассмотрев и обсудив диссертационную работу Воробьевой Г.Р. «Методологические основы обработки неоднородной пространственно-временной информации в системах поддержки принятия решений на основе технологий больших данных (на примере геомагнитных данных)», в соответствии с п.24 «Положения о присуждении ученых степеней», отмечаем следующее.

**Актуальность темы исследования**

Проведенные в рамках диссертационной работы научные исследования направлены на повышение эффективности процессов обработки информации в системах поддержки принятия решений посредством разработки единых методологических основ обработки, анализа и визуализации больших объе-



мов пространственных данных, полученных из территориально распределенных гетерогенных источников. Их актуальность обусловлена необходимостью решения задачи информационной перегруженности лиц, принимающих решения, которая выражается в необходимости подбора источников данных, фильтрации, унификации и обработке, предоставляемой ими информации, последующем анализе больших объемов данных доступными в системе поддержки принятия решений инструментально-программными средствами, которые, в свою очередь, не всегда эффективно справляются с возложенными на них задачами в условиях пространственной неоднородности соответствующих данных.

### **Оценка структуры и содержания работы**

Структура и содержание диссертации находятся в логическом единстве и соответствуют поставленным целям и задачам исследования. Диссертационная работа состоит из введения, шести глав, заключения, списка терминов, списка литературы и приложения с актами внедрения. Диссертация изложена на 318 страницах машинописного текста, содержит 43 рисунка и 4 таблицы.

Во **введении** обоснована актуальность, показана практическая и теоретическая ценности диссертационной работы. Сформулированы цель и задачи работы; представлены положения, выносимые на защиту; изложены краткая характеристика и сведения об апробации работы.

**Глава 1** посвящена анализу проблемы обработки информации из распределенных гетерогенных источников в процессе принятия решений и обсуждению предложенных методологических основ подхода к анализу и обработке пространственного распределения данных в информационных системах поддержки принятия решений.

Автором предложены и сформулированы методологические основы обработки неоднородной пространственно-временной информации, которые отличаются тем, что, с целью повышения оперативности доступа к информации, необходимой для принятия решений, выделяются критерии пространственной зависимости и пространственной гетерогенности для групп источ-

ников данных и на этой основе подстраиваются процессы сбора, анализа и визуализации информации.

Во **второй главе** обсуждаются разработанные модели и методы обработки данных при интеграции гетерогенных источников, целью реализации которых является повышение оперативности получения данных, необходимых для принятия решений, в частности, в области техносферной безопасности.

Предложенный автором комплекс моделей и методов обработки информации при интеграции гетерогенных источников данных в гибридные хранилище систем поддержки принятия решений в общем виде описан в составе единого информационного пространства, которое представляет собой совокупность гетерогенных источников данных, а также инфокоммуникационных технологий их интеграции, обработки, анализа и визуализации, функционирующих на основе единых принципов и обеспечивающих информационное взаимодействие поставщиков и потребителей данных, равно как и удовлетворение их информационных потребностей при решении прикладных и научно-исследовательских задач.

**Третья глава** посвящена обсуждению предложенных автором методов восстановления временных рядов данных (на примере геомагнитной информации) в информационных системах на основе принципов машинного обучения и информационного резервирования источников информации, обеспечивающих возможность импутации пропусков с ошибкой в пределах допустимого нормативами отклонения.

Основным элементом таких решений является ранжирование и агрегирование источников данных по результатам анализа пространственной зависимости теоретико-множественных, теоретико-информационных и статистических характеристик. Выявленные признаки совместной встречаемости временных рядов данных и степень связности описываемых ими процессов служат основанием для формирования доверительного списка для каждого задействованного в едином информационном пространстве источника дан-



ных. Попавшие в один доверительный список источники данных считаются резервирующими друг друга.

Другой метод восстановления данных – индуктивный. Его идея заключается в предположении, что если пара непоследовательных фрагментов ряда, разделенных отсутствующим фрагментом, оказываются близкими к паре фрагментов, разделенных известным фрагментом, то промежуточные между ними значения в соответствии с теоремой Такенса будут отличаться статистически незначительно. При этом признаковым описанием фрагмента временного ряда выступает пара предшествующего и следующего за ним фрагментов того же ряда, образующих обучающую выборку для поиска недостающего фрагмента по набору его признаков с последующим линейным масштабированием для восстановления исходного тренда информационного сигнала.

Восстановленные одним из предложенных способов данные размещаются в едином информационном пространстве и могут быть использованы для обработки, анализа и визуализации в процессе принятия решений.

В **четвертой главе** описаны особенности предложенной автором модели хранения данных в системах поддержки принятия решений, обеспечивающей сокращение вычислительных затрат на их физическое размещение и повышение скорости обработки запросов к данным.

Модель хранения данных представлена совокупностью трех компонент и использует правила ссылочной целостности для объединения реляционной, иерархической и колончатой моделей данных, применяемых для описания метаданных и непосредственно данных, а также реализует комбинацию текстового и бинарного форматов представления информации с целью повышения реактивности программных средств обработки данных, с одной стороны, и сокращения затрат требуемого объема физической памяти, с другой.

Предложенная модель рассмотрена на примере геомагнитной информации в виде гибридной структуры, в котором сочетаются особенности ре-

ляционного и нереляционного представления информации в составе систем поддержки принятия решений.

В **пятой главе** на примере геомагнитных данных рассматривается алгоритм визуализации пространственно-временного распределения данных на основе компьютерных методов обработки информации.

Автором предложен алгоритм визуализации пространственно-временного распределения геомагнитных данных в двух- и трехмерном режиме в веб-среде, позволяющий на основании данных из единого информационного пространства оценить распределение параметров геомагнитного поля и выделить, к примеру, потенциально опасные для наведения геоиндуцированных токов регионы.

Кроме того, автором предлагается усовершенствовать известные подходы к визуальному представлению геопространственных данных применительно к рендерингу геомагнитных данных посредством разработки адаптивного алгоритма, формирующего изображение на основании пространственной привязки данных.

**Шестая глава** посвящена обсуждению результатов оценки эффективности предложенных методологических основ, методов, моделей и алгоритмов на основе разработанного прототипа веб-ориентированного инструментально-программного средства ([www.geomagnet.ru](http://www.geomagnet.ru)) и выделенных показателей эффективности.

Преимущества решений показаны на примере ряда задач поддержки принятия решений на основе геомагнитной информации, предоставляемой распределенными гетерогенными источниками геомагнитных данных.

В **заключении** представлены основные результаты и выводы, а также определены перспективы дальнейшего развития.

### **Новизна полученных результатов**

1. Методологические основы обработки неоднородной пространственно-временной информации отличаются тем, что, с целью повышения оперативности доступа к информации, необходимой для принятия решений, выде-

ляются критерии пространственной зависимости и пространственной гетерогенности для групп источников данных и на этой основе подстраиваются процессы сбора, анализа и визуализации информации.

2. Комплекс моделей и методов обработки информации при интеграции гетерогенных источников данных в гибридные хранилище систем поддержки принятия решений отличается тем, что, с целью повышения вычислительной скорости сбора и обработки данных, преобразование оперативной информационной составляющей в постоянную определяется адаптированной моделью старения информации Бартона–Кеблера с исключенным динамическим компонентом.

3. Методы восстановления временных рядов данных, включающие:

– индуктивный метод, отличающийся тем, что, с целью повышения точности и скорости восстановления данных, наиболее вероятные значения определяются на базе статического сходства между массивом, образованным предшествующими и последующими за пропущенным фрагментом значениями, и массивами, построенными аналогично из известных значений;

– метод информационного резервирования источников данных, отличающийся тем, что, с целью обеспечения полноты временных рядов, наиболее вероятные значения определяются посредством формирования доверительного списка на основании оценки пространственной гетерогенности и зависимости синхронно регистрируемых данных, а также сравнительной оценки фрагментов рядов, зарегистрированных в предшествующий момент.

4. Модель хранения данных отличается тем, что, с целью повышения реактивности программных средств и сокращения затрат физической памяти, реляционная, иерархическая и колончатая модели данных объединены на базе правил ссылочной целостности, а также комбинирования текстового и бинарного форматов описания как собственно данных, так и их метаданных.

5. Алгоритм визуализации пространственно-временного распределения геомагнитных данных отличается тем, что для клиентского веб-рендеринга больших пространственных данных учитывается их пространственная анизо-



тропия посредством комбинирования подходов, демонстрирующих наилучшие показатели реактивности в соответствующих пространственных областях.

### **Степень достоверности результатов исследования**

Достоверность полученных научных результатов обеспечивается строгостью применяемого математического аппарата, результатами математического и компьютерного моделирования, подтверждается на примере обработки и анализа оригинальных геомагнитных данных, предоставляемых территориально распределенными гетерогенными источниками на примере магнитных станций и обсерваторий международной сети INTERMAGNET, а также результатами обработки и визуализации геомагнитных данных в рамках разработанного прототипа веб-ориентированного инструментально-программного средства Geomagnet (<https://www.geomagnet.ru>).

### **Теоретическая и практическая значимость результатов, полученных автором диссертации**

Разработанные в ходе диссертационного исследования методологические основы обработки, анализа и визуализации больших объемов пространственных данных позволяют формулировать подходы к поддержке принятия решений на основе временных рядов геопространственных данных, которые принимают в расчет не только характер самих данных, но и специфику пространственно-временной анизотропии описываемых ими объектов, процессов или явлений различной природы происхождения.

Практическая значимость полученных автором результатов в целом подтверждается представленными в диссертационной работе актами их внедрения на предприятия и организации различного профиля.

## **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 геопространственных данных, например, в геофизике, 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 достаточного внимания тому, что понимается под неоднородными пространственно-временными данными.

2. В работе явным образом не показаны ограничения по объему данных, к которым применимы предложенные методологические основы обработки, анализа и визуализации.

3. Из представленного в диссертационной работе описания полученных результатов не совсем понятно, каким образом формулируются требования и ограничения к той предметной области, в рамках которой предложенные положения могут быть использованы для поддержки принятия решений.

4. В диссертации при анализе эффективности полученных результатов недостаточно явно прослеживается сопоставление их с мировым научным опытом.

5. В работе не рассмотрены вопросы пространственной интерполяции исследуемого класса данных, в то время как такой вопрос представляет большую значимость во многих задачах поддержки принятия решений.

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



## **Заключение**

1. Диссертация Воробьевой Гульнары Равилевны на соискание ученой степени доктора технических наук является научно-квалификационной работой, в которой на основании выполненных автором исследований решена научная проблема повышения эффективности процессов обработки информации в системах поддержки принятия решений, имеющая важное хозяйственное значение, 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. Все основные научные результаты диссертации отражены в опубликованных работах, в том числе 13 статьях в рецензируемых журналах из списка ВАК; 22 статьях в изданиях, индексируемых международными системами Scopus / Web of Science; 3 монографиях, изданных в России и за рубежом; 2 патентах на изобретение; 5 свидетельствах о государственной регистрации программы для ЭВМ, трудах конференций и др.

5. Тема и содержание диссертации соответствуют паспорту специальности 05.13.01 – Системный анализ, управление и обработка информации (информационные и технические системы).

Диссертация соответствует требованиям п.9 «Положения о присуждении ученых степеней», а ее автор, Воробьева Гульнара Равилевна, заслуживает присвоения ученой степени доктора технических наук по специальности 05.13.01 – Системный анализ, управление и обработка информации (информационные и технические системы).

Диссертационная работа Воробьевой Гульнары Равиленовны «Методологические основы обработки неоднородной пространственно-временной информации в системах поддержки принятия решений на основе технологий больших данных (на примере геомагнитных данных)» обсуждена, отзыв заслушан и одобрен на заседании кафедры организации и управления перевозками на транспорте федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королёва», протокол № 2 от 30 сентября 2020 года.

Отзыв составлен:

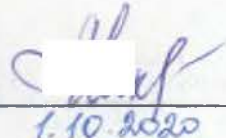
Руководитель научного направления

«Интеллектуальные транспортные геоинформационные системы»,  
профессор кафедры организации и управления перевозками на транспорте  
федерального государственного автономного образовательного учреждения  
высшего образования

«Самарский национальный исследовательский университет  
имени академика С.П. Королева»,

Заслуженный деятель науки и образования,

д-р техн. наук, профессор \_\_\_\_\_ Т.И. Михеева

  
1.10.2020

Докторская диссертация защищена

по специальности 05.13.01 – Системный анализ, управление и обработка информации (информационные и технические системы)

Адрес организации: 443086, Приволжский федеральный округ, Самарская область, г. Самара, Московское шоссе, д. 34.

Рабочий телефон: + 7 (846) 335-18-26, сотовый +79272033248

Адрес эл. почты: [ssau@ssau.ru](mailto:ssau@ssau.ru), [Mikheevati@gmail.com](mailto:Mikheevati@gmail.com)