

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.42-042.4:004.422.6

Судаков
Богдан Дмитриевич

МОДЕЛИ И МЕТОДЫ ПРОГНОЗИРОВАНИЯ В *ERP*-СИСТЕМАХ НА
ОСНОВЕ НЕПОЛНЫХ ДАННЫХ

АВТОРЕФЕРАТ

на соискание степени магистра технических наук
по специальности 1 - 40 80 02 Системный анализ, управление и обработка
информации

Научный руководитель

О.В. Герман,
кандидат технических наук,
доцент

Минск 2024

Нормоконтроль

О.В. Герман

ВВЕДЕНИЕ

Неполные данные – это повсеместная проблема в различных областях, начиная от здравоохранения и финансов и заканчивая социальными науками и техникой. Эффективная обработка недостающих данных имеет решающее значение для обеспечения целостности и надежности анализов и процессов принятия решений на основе неполных наборов данных.

Диссертация состоит из введения, трех глав, включающих описание архитектуры различных *ERP*-систем, обзор литературы на данную тему, сравнительный анализ существующих методов восстановления неполных данных, а также методологическое описание разработанного алгоритма, заключение, список использованных источников и приложение, содержащее схему разработанного алгоритма.

В первой главе работы описывается архитектура *ERP*-системы, программные модули и их функционал. Также приводятся примеры интеграции новых подходов разработки и анализа данных в системы, такие как машинное обучение, нейронные сети и анализ больших данных.

Во второй главе рассматриваются различные методы восстановления отсутствующих данных, включая традиционные методы вменения, подходы, основанные на машинном обучении, и стратегии вероятностного моделирования.

В третьей главе приводятся теоретические материалы, описывающие улучшенный метод восстановления неполных данных при помощи ядерных функций и машинного обучения, а также приводятся результаты эксперимента и тестирования метода на реальных данных.

Результаты, полученные в ходе исследования, могут быть использованы при интеграции программного модуля восстановления неполных данных в различные *ERP*-системы, тем самым улучшая дальнейшее качество аналитики на основе восстановленных данных.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследования

Диссертационное исследование выполнено в рамках направления «Цифровые информационно-коммуникационные и междисциплинарные технологии, производства на их основе» по реализации и внедрению программно-аппаратных решений с использованием искусственного интеллекта, машинного обучения, облачных технологий, интеллектуальных электронных терминалов в соответствии с Государственной программой инновационного развития Республики Беларусь на 2021-2025 годы, утвержденной Указом Президента Республики Беларусь от 15 сентября 2021 года № 348.

Цель и задачи исследования

Целью диссертационной работы является исследование восстановления неполных данных и разработка эффективного метода для обработки пропущенных значений с целью повышения качества анализа и принятия решений на основе этих данных. Поставленная цель работы определяет следующие основные задачи:

1. Проанализировать существующие методы восстановления неполных данных и их применимость к различным типам данных и контекстам исследования.
2. Провести обзор литературы и сравнительный анализ существующих методов восстановления неполных данных с целью выявления их преимуществ, недостатков и областей применения.
3. Разработать метод восстановления неполных данных, основанный на современных подходах в области машинного обучения, статистики и анализа данных.
4. Выполнить обоснование разработанного метода восстановления неполных для оценки его эффективности.
5. Подготовить обобщенные выводы и рекомендации для практического применения разработанных методов восстановления неполных данных в различных областях.

Область исследования

Содержание диссертации соответствует образовательному стандарту высшего образования второй ступени (магистратуры) ОСВО 1-40 80 02-2022

специальности 1-40 80 02 «Системный анализ, управление и обработка информации».

Теоретическая и методологическая основа исследования

В основу диссертации легли работы зарубежных исследователей в области восстановления неполных данных. А также технические нормативные правовые акты по тематике диссертационной работы.

Информационная база исследования сформирована на основе технической литературы, открытой информации, технических нормативно-правовых актов, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

Публикации автора

По результатам исследований, представленных в диссертации, опубликовано две авторские работы, в том числе: две статьи и тезисы докладов в материалах конференций.

Личный вклад соискателя

Соискателем выполнены все изложенные в работе разработки и исследования. Постановка задач и обсуждение результатов проводились совместно с научным руководителем. Обработка, интерпретация данных, а также выводы сделаны автором самостоятельно.

Плагиат

Проведена проверка диссертации «Модели и методы прогнозирования в ERP-системах на основе неполных данных» Судакова Богдана Дмитриевича на корректность использования заимствованных материалов с использованием сетевого ресурса «Антиплагиат» (адрес доступа: <https://antiplagiat.ru>) в режиме *on-line* 07.04.2024. В результате проверки установлена корректность использования заимствованных материалов (оригинальность диссертации составляет 82,5%).

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении дается обзор проблем, с которыми сталкиваются *ERP*-системы в случае потери данных, и подчеркивается важность разработки модифицированного алгоритма восстановления. Обозначены цель и задачи исследования, а также личный вклад соискателя в развитие знаний и научной степени. Также обсуждаются результаты диссертации с точки зрения их апробации и утверждения.

Общее описание работы демонстрирует соответствие исследования приоритетным направлениям научных исследований. В задачи исследования входит изучение существующих методов, разработка модифицированного метода восстановления отсутствующих данных, внедрение необходимых модулей, обоснование алгоритма и анализ результатов.

Личный вклад соискателя ученой степени включает разработку и обоснование алгоритма, формулировку выводов. Результаты исследования докладывались и обсуждались в Белорусском государственном университете информатики и радиоэлектроники.

Результаты диссертации также были распространены через публикации, включая статьи в научных журналах, рекомендованных 59-й научной конференцией студентов, магистрантов и аспирантов БГУИР и международной конференцией Информационные технологии и системы *ITS* 2023. Данные публикации вносят вклад в широкое научное сообщество и демонстрируют работу автора в данной области.

В первой главе работы описывается архитектура *ERP*-системы, программные модули и их функционал. Приводятся примеры интеграции новых подходов разработки и анализа данных в системы, такие как машинное обучение и нейронные сети. Наконец, в данной главе приводится пояснение, каким образом можно интегрировать инновационный подход восстановления неполных данных.

Во второй главе рассматриваются различные методы восстановления отсутствующих данных, включая традиционные методы и подходы, основанные на машинном обучении, и стратегии вероятностного моделирования. Указываются сильные стороны и ограничения каждого метода, подчеркивается их применимость в различных контекстах и наборах данных. Кроме того, рассматриваются последние достижения в области восстановления пропущенных данных, такие как модели вменения (*imputation*) на основе глубокого обучения.

Третья глава посвящена теоретическому описанию и программной реализации алгоритма. В ней подробно описывается реализация шагов

предложенного алгоритма, используемые средства программирования и описание кода. В главе также обсуждаются возможные пути дальнейшего развития и совершенствования алгоритма.

В целом, в работе решены проблемы, связанные с восстановлением неполных данных, и предложен модифицированный алгоритм. Результаты данной работы способствуют развитию технологии восстановления неполных данных в *ERP*-системах и ее применению в различных областях.

ЗАКЛЮЧЕНИЕ

В ходе проведенного исследования изучено актуальное направление развития информационных технологий в сфере управления предприятием (*ERP*), а именно проблема восстановления неполных данных в *ERP*-системах. Целью данного исследования являлись разработка и применение нового метода восстановления утраченных или поврежденных данных в *ERP*-системах, что способствует повышению надежности и качества информационных систем предприятий, а также улучшает качество дальнейшего анализа на основе восстановленных данных.

В ходе исследования проведен анализ существующих методов и подходов к восстановлению данных в *ERP*-системах. Выявлены основные проблемы и ограничения существующих подходов, такие как ограниченная точность восстановления, высокая вычислительная сложность и невозможность учета специфики отдельных предприятий.

На основе проведенного анализа разработан алгоритм восстановления неполных данных, позволяющий добиться высокой точности восстановления при приемлемом времени выполнения.

Приведено теоретическое обоснование алгоритма, позволяющее судить о его достаточно высокой скорости сходимости. Реализация предложенного метода позволит улучшить качество данных в *ERP*-системах, что способствует повышению эффективности управления предприятием и принятия обоснованных стратегических решений.

Исследование проводилось с целью разработки и применения эффективных методов восстановления утраченных или поврежденных данных в *ERP*-системах. Это направлено на повышение надежности и качества информационных систем предприятий, а также улучшение качества дальнейшего анализа на основе восстановленных данных. Анализ существующих методов и подходов к восстановлению данных в *ERP*-системах выявил основные проблемы и ограничения, такие как ограниченная точность восстановления, высокая вычислительная сложность и невозможность учета специфики отдельных предприятий.

Результаты данного исследования имеют практическое значение для предприятий, работающих с *ERP*-системами, а также вносят вклад в развитие теории и практики в области управления информационными технологиями.

Дальнейшее развитие исследований в этой области может включать улучшение алгоритмов восстановления данных, адаптацию методов к новым типам *ERP*-систем и расширение области применения данных подходов.

СПИСОК ПУБЛИКАЦИЙ АВТОРА

[1] Судаков, Б. Д. Архитектура *ERP*-систем / Б. Д. Судаков // Информационные технологии и управление: материалы 59-ой научной конференции аспирантов, магистрантов и студентов, Минск, 17–21 апреля 2023 года / Белорусский государственный университет информатики и радиоэлектроники; редкол.: Л. Ю. Шилин [и др.]. – Минск, 2023. – 44 с.

[2] Судаков, Б. Д. Методы восстановления неполных данных для повышения точности анализа / Б. Д. Судаков, О. В. Герман // Информационные технологии и системы 2023 (ИТС 2023) *Information Technologies and Systems 2023 (ITS 2023)*: материалы Международной научной конференции, Минск, 22 ноября 2023 / Белорусский государственный университет информатики и радиоэлектроники; редкол.: Л. Ю. Шилин [и др.]. – Минск, 2023. – 201–202 с.