

ОТЗЫВ

официального оппонента на диссертационную работу
Фёдорова Романа Константиновича

«Сервис-ориентированная информационно-аналитическая среда композиции сервисов обработки пространственных данных», представленную на соискание ученой степени доктора технических наук по специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки)»

Актуальность темы диссертации

Диссертационная работа посвящена автоматизации построения и применения композиций сервисов в гетерогенной информационно-вычислительной и информационно-аналитической среде. Современный этап развития методов и средств разработки программного обеспечения движется в сторону создания методов обработки данных в виде сервисов, обладающих свойствами независимости, заменимости и масштабируемости за счет оснащения стандартизированными интерфейсами для взаимодействия по стандартизированным протоколам. Эти свойства позволяют быстрее развертывать новые компоненты и упрощать создание программных систем в виде совокупностей взаимодействующих сервисов. Количество сервисов, разрабатываемых коллективами из разных организаций и предметных областей, в том числе в научной сфере, растет. Одним из ключевых моментов при их композиции является поиск программных сервисов, подходящих под решение задачи и обладающих необходимыми входными и выходными параметрами. Однако метаданные у сервисов часто отсутствуют, либо представлены слабо. Поэтому актуальна разработка новых методов создания композиций, использующих другую информацию об их применении, например, статистику.

В работе приведены и проанализированы существующие методы и технологии создания композиций сервисов. В отличие от существующих предлагаемый в работе подход ориентирован на анализ статистики применения сервисов с применением метаданных, онтологий и знаний (определяющих возможность применить сервис обработки и вычисления к совместимым с ним формату выдаваемым данным). Предлагаемая сервис-ориентированная информационно-аналитическая среда на основе цифровых платформ, позволяющая собирать статистику применения сервисов, сможет обеспечивать научных работников дополнительной информацией для создания полезных композиций сервисов. В статистике видны реальные примеры передачи данных между сервисами.

Выдвигаемые в работе положения обоснованы и доказаны. Для предложенных методов и информационно-аналитической среды приведены результаты апробации на задаче поддержки междисциплинарных научных исследований Байкальской территории. Созданы десятки сервисов обработки данных, сформированы композиции сервисов разного типа, развернуты геопорталы для коллективов из различных предметных областей. Результаты диссертации опубликованы, обсуждены на конференциях, получены свидетельства о регистрации программ для электронных вычислительных машин. Тем самым обосновывается успешное решение актуальной проблемы в диссертационной работе Фёдорова Романа Константиновича.

Оценка новизны исследований и полученных результатов

В диссертации Фёдорова Р.К. представлены результаты, обладающие научной новизной, а также имеющие практическую значимость.

1. Предложен принцип использования статистики применения сервисов и формирования оценок, построенных на основе многопользовательского опыта применения сервисов, для определения возможности создания композиций сервисов. Именно собираемая статистика больше всего отличает новые методы и средства.
2. Создана модель сервис-ориентированной информационно-аналитической среды для обработки пространственных данных.
3. Разработан метод создания композиций сервисов, учитывающий множество пользователей, множество сервисов и заданий, потребности в обмене ими и в организации междисциплинарных научных исследований.

Обоснованность, достоверность и значимость основных результатов и выводов

Обоснованность и достоверность полученных экспериментальных результатов, научных положений и выводов в диссертации не вызывает сомнений и подтверждается тем, что на основе выделенных при решении второй задачи элементов модели, отношений между ними и проанализированных свойств, был разработан метод создания композиций сервисов, учитывающий множество пользователей, сервисов и заданий; далее на основе модели и метода реализованы компоненты, помогающие собирать статистику выполнения сервисов, и реализованы инфраструктурные компоненты, включая каталоги, для выполнения сервисов. Разработан модуль WPSManager, который получает запросы в виде идентификатора сервиса, списка входных и выходных параметров, выходные параметры содержат пути сохранения результатов, параметры ссылки на файлы преобразуются в URL для и обеспечения доступа по HTTP. Каждый запрос сохраняется в таблицу статистики и отправляется на выполнение, после чего результаты сервиса добавляются в ту же строку таблицы.

Разработанные модели, метод и компоненты апробированы в различных задачах поддержки междисциплинарных научных исследований.

В работе применены методы информационного моделирования, теории графов, системного и объектно-ориентированного программирования, веб-технологий, построения распределенных комплексов, а также «специфичные» для проблематики диссертационного исследования методы формирования композиций сервисов на основе метаданных и онтологий. Важен задел к диссертации: в ИДСТУ СО РАН на протяжении нескольких десятков лет ведутся работы по развитию пакетов прикладных программ, параллельных и распределенных вычислительных систем, созданию геоинформационных систем и цифровых платформ, послужившие основой данной работы.

Личный вклад автора не вызывает сомнений и состоит в выборе методов исследования, моделировании, проектировании, разработке всех компонентов типовой среды.

Научные положения, выводы и рекомендации, сформулированные автором, подтверждаются широкой апробацией работы в открытой печати: результаты решения второй, третьей и последующих заявленных задач опубликованы в изданиях, входящих в базы цитирования Web of Science и Scopus. Предложенные методы реализованы в виде программного обеспечения (имеется 8 свидетельств о государственной регистрации программы для ЭВМ).

Практическая значимость работы определяется тем, что комплекс программных компонентов, реализующий модель сервис-ориентированной информационно-аналитической среды, активно используется на практике. В ходе выполнения различных проектов созданы более 200 *сервисов предоставления данных*, более 40 сервисов обработки данных и 250 сервисов публикации данных. Развернуты шесть геопорталов, ориентированных на различные предметные области и коллективы: Геопортал ИДСТУ СО РАН; Информационно-аналитическая система по фиторазнообразию Байкальской Сибири; Атлас ИГ СО РАН; Геопортал ИЗК СО РАН; Очаги распространения иксодовых клещей; Информационная система (ИС) «L.». В рамках этих геопорталов сформированы композиции сервисов, объединяющие сервисы данных, сервисы обработки и публикации, созданные разными коллективами. Практическая значимость результатов подтверждена полученными актами внедрения комплекса программных компонент ИГ СО РАН, СИФИБР СО РАН, ПАБСИ КНЦ РАН, ИППЭС КНЦ РАН, Самарский университет им. Королева, НЦ ПЗСРЧ. Автором в составе коллектива получено 8 свидетельств о регистрации программ для электронных вычислительных машин (ЭВМ).

Соответствие диссертации научной специальности

Представленные в диссертации результаты соответствуют паспорту специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей», а именно п.п.:

- модели, методы, архитектуры, алгоритмы, языки и программные инструменты организации взаимодействия программ и программных систем;
- модели, методы, алгоритмы, облачные технологии и программная инфраструктура организации глобально распределенной обработки данных.

Структура и содержание работы

Диссертационная работа состоит из введения, пяти глав, заключения, списка литературы, списка принятых сокращений и 12 приложений. Объем основного текста работы – 197 страниц, включая 5 таблиц и 81 рисунок. Список литературы содержит 201 источник.

Введение содержит обоснование актуальности темы исследования, цель и задачи, научную новизну.

В первой главе проведен обзор существующих направлений развития технологий распределенных вычислений, методов и подходов к созданию композиций сервисов и их выполнения. Доказана актуальность разработки сервис-ориентированной информационно-аналитической среды (СОИАС) обработки междисциплинарных пространственных данных в рамках научных исследований.

Во второй главе представлена новая вычислительная модель композиции сервисов, предложен метод создания композиций. Описаны его основные этапы. Показан пятишаговый процесс построения модели предметной области.

Приведен разработанный алгоритм формирования сети связанных сервисов, который может быть использован для упрощения разработки композиции сервисов. Построение сети связанных сервисов сводится к задаче определения передачи данных между параметрами сервисов. Сеть связанных сервисов определяет возможные связи сервисов по передаваемым данным. Для формирования сети связанных сервисов используют метаинформацию и онтологии.

Приводится алгоритм автоматического формирования композиций сервисов. Автоматизация процесса на основе статистических данных, знаний и онтологий формирует готовый инструмент анализа и обработки данных для других пользователей.

Показано получение общего графа, где все вызовы пользователей сервисов связаны по данным. Приводится постановка задачи поиска композиций сервисов, а также реализация поиска и эвристический метод ранжирования композиций сервисов.

Выполнено сравнение поддержки автоматизации формирования композиций сервисов в предлагаемой СОИАС с существующими системами управления научными рабочими процессами: BPEL Designer Project, GeoJModelBuilder, Everest, CLAVIRE, Pegasus и другими.

В третьей главе приводится алгоритмическая и программная реализация сервис-ориентированной информационно-аналитической среды. Описывается архитектура, реализующая вычислительную модель композиции сервисов, и обеспечивающая функциональную расширяемость и дающая возможность пользователю работать полностью удаленно.

Типовой геопортал реализует функции ввода, редактирования, отображения и анализа данных с помощью сервисов, он является точкой входа для поиска и использования сервисов данных и геообработки. Для упрощения работы пользователей для каждой предметной области создается отдельный геопортал, который предоставляет быстрый доступ к наиболее часто используемым сервисам.

Среди других компонентов – многопользовательский сервер для анализа данных, каталоги сервисов и данных, СУБД, фабрики сервисов. Создаваемые сервисы позволяют работать с пространственными атрибутами: вводить, отображать на карте, выполнять пространственные операции и т. д. Большинство сервисов обработки пространственных данных поддерживают стандарт WPS. Табличные данные часто используются в качестве входных данных для сервисов.

В четвертой главе рассматривается подсистема выполнения композиций сервисов обработки пространственно-временных данных. Сервис-ориентированная информационно-аналитическая среда композиции сервисов должна обеспечить эффективное распределение вычислительных ресурсов при выполнении сервисов. Распределение вычислительных ресурсов основывается на модификации спискового эвристического алгоритма составления расписания. Программный компонент учитывает изменения состояния вычислительной среды, например, сбой работы узлов, изменение времени выполнения задания, добавление новых заданий и т. д. С целью упрощения создания композиций сервисов предложено использование процедурного языка программирования для вызова сервисов и обработки промежуточных данных с помощью средств языка и его библиотек.

В пятой главе представлена реализация технологии на основе метода создания композиций, и апробация сервис-ориентированной информационно-аналитической среды обработки пространственно-временных данных. Охарактеризованы созданные геопорталы, ориентированных на разные предметные области, описаны методики работы в них, сформированные композиции сервисов, объединяющие сервисы данных, сервисы обработки и публикации, а также разработанная среда, которая обеспечивает создание композиций сервисов и их обмен между пользователями.

Замечания по диссертационной работе

1. Выводы из первой главы об актуальности информационно-аналитической среды обработки междисциплинарных пространственных данных для повышения эффективности научных исследований за счет

композиций сервисов и организации обмена ими, не получили продолжения в виде качественных или количественных оценок полученной эффективности по сравнению с прежним состоянием. Не прояснено, экономнее ли по затратам оформление WPS сервисов и регистрация каждого в нужном каталоге, по сравнению с формированием обычных библиотек повторно-используемых сервисов с их спецификациями, но без их статистики.

2. Для компонентов разработанной сервис-ориентированной информационно-аналитической среды не хватает диаграмм логической архитектуры (есть только пример одной модели наследования), чтобы можно было получить представление о сопровождаемости, стройности, повторной используемости или других аспектах архитектуры разработанного программного обеспечения вычислительных комплексов.

3. Несколько нетрадиционно применяются в работе термины онтологическая модель и экспертные знания, они не относятся к предметной области, в которой специалисты решают задачи, а относятся к самой технологии композиции.

4. Термин онтология употребляется по тексту в разных смыслах без уточняющего контекста: и как схема семантической разметки описания сервисов, и как прикладные онтологии (онтология наук о земле).

Заключение

В диссертационной работе изложены действительно новые методы, алгоритмы и технологическая среда для исследований, связанных с пространственно-временными данными. Полученные в работе результаты представляют существенное значение для развития междисциплинарных научных исследований.

Сделанные выше замечания не влияют на общую положительную оценку работы. Работа имеет комплексный характер и высокий уровень проработанности и будет интересна научному сообществу, что подтверждается публикационной активностью, участием в различных международных и всероссийских конференциях, успешным применением новых методов в поддержанных грантами проектах.

Оценивая диссертацию Фёдорова Р.К., можно сделать вывод, что она является самостоятельной, логически завершенной научно-исследовательской работой, содержащей научно обоснованные решения для создания композиций сервисов на геопорталах и распространения опыта их применения среди всех пользователей этих порталов. Работа актуальна, имеет научную новизну и практическую значимость. Содержание диссертации достаточно полно отражено в публикациях рецензируемых изданий. Автореферат диссертации соответствует ее основному содержанию. Результаты работы неоднократно докладывались и обсуждались на международных научных конференциях.

Диссертационная работа соответствует требованиям, которые предъявляются к диссертациям на соискание ученой степени доктора наук, а ее автор Фёдоров Роман Константинович заслуживает присуждения ученой степени доктора технических наук по специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей (технические науки)».

Официальный оппонент,
доктор технических наук,
доцент, ведущий научный сотрудник
ИАПУ ДВО РАН

Шалфеева Елена Арефьевна

«26» ноября 2024 г.

Федеральное государственное бюджетное
учреждение науки Институт автоматизации
и процессов управления
Дальневосточного отделения
Российской академии наук
690041, Россия, г. Владивосток,
улица Радио, д. 5
+7(423)-2-310-424
shalf@iacp.dvo.ru

«ЗАВЕРЯЮ»
РЕДАКТОР ИАПУ ДВО РАН

Д. А. ЦУКАНОВ