

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.

1. Актуальность темы исследований

В настоящее время в России и в мире на государственном уровне определяются цели экологического мониторинга обширных природных территорий и обработки данных наблюдений. При их реализации чрезвычайно актуальны задачи компьютерной обработки больших объемов разнородных пространственных данных. В связи с этим, задачи повышения эффективности и надежности процессов обработки и передачи данных в вычислительных системах, комплексах и компьютерных сетях путём внедрения решений на основе сервис-ориентированной архитектуры являются весьма актуальными. На основе сервис-ориентированной архитектуры реализуются соответствующие интеллектуальные технологии с использованием методов машинного обучения и системы обработки Big Data.

Объединение сервисов обработки больших объемов разнородных пространственных данных, т. е. создание их композиции, позволяет решать большое количество практических задач. Сервис-ориентированная архитектура значительно упрощает и ускоряет интеграцию программного обеспечения, созданного разработчиками из разных предметных областей, за счет упрощения и стандартизации интерфейсов. Композиции сервисов обеспечивают повышение уровня автоматизации решения задач, начиная от ввода данных и заканчивая публикацией результатов.

При этом создание композиций сервисов является весьма сложной задачей. Наличие большого количества сервисов, с одной стороны, увеличивает возможности исследователей, а с другой стороны, значительно усложняет поиск нужных сервисов для решения конкретной задачи. Составление найденных сервисов в композиции в некоторых случаях может оказаться комбинаторно сложной задачей, при этом возможна генерация достаточно большого числа альтернативных композиций сервисов. Все эти альтернативы необходимо будет оценить, выбрать наиболее релевантные для решаемой задачи, а также их проверить на возможность взаимодействия. Для практической реализуемости композиции сервисов необходимо учитывать различные структуры и форматы входных и выходных данных.

Композиции сервисов позволяют объединить результаты работы исследователей, но создание таких композиций в силу сложного процесса требует высокой квалификации. Поэтому актуальной научной проблемой является разработка новых моделей, алгоритмов, методов и технологии для создания сервис-ориентированной информационно-аналитической среды с целью повышения эффективности процессов подготовки и проведения научных экспериментов по решению задач в области геоинформатики за счет автоматизации построения и применения композиций сервисов.

2. Научная новизна исследований и полученных результатов

Автор диссертации получил следующие новые научные результаты:

- 1) создана модель сервис-ориентированной информационно-аналитической среды обработки пространственных данных междисциплинарных исследований, которая в сравнении с подобными моделями обеспечивает оценку композиций сервисов на основе многопользовательской статистики их применения;
- 2) разработан метод создания композиций сервисов, базирующийся на применении предложенной модели, который в отличие от существующих методов проводит комплексный анализ метаданных, онтологий, экспертных знаний и статистики применения сервисов, что позволяет находить композиции сервисов на основе комбинации данных;
- 3) разработан оригинальный программный инструмент создания сервисов ввода и публикации реляционных данных, обеспечивающий предоставление метаданных, пользовательский и программный интерфейс редактирования данных, поддержку передачи данных WPS сервисам. Создание сервисов данных впервые производится на основе иерархической модели данных с возможностью задания асинхронного вычисления значений атрибутов с помощью сервисов. Создаваемые сервисы можно сразу включать во множество композиций;
- 4) разработан оригинальный программный компонент выполнения композиций сервисов, заданных на процедурном языке, с обработкой промежуточных данных с помощью средств языка и его библиотек, для которого в отличие от других подходов обеспечивается формирование DAG с помощью процедурного языка и одновременно его планирование и выполнение с учетом добавляемых в процессе выполнения новых заданий в гетерогенной динамической вычислительной среде;
- 5) разработан комплекс программных компонентов, реализующий модель сервис-ориентированной информационно-аналитической среды, который обеспечил создание композиций сервисов и их обмен между пользователями.

Считаю, что все полученные автором результаты являются интересными, новыми, их научная новизна хорошо обоснована, апробирована в публикациях и выступлениях на конференциях, и признана научным сообществом.

3. Достоверность и обоснованность полученных результатов

В работе четко определены цели, объект и предмет исследования. Целью работы является повышение эффективности процессов подготовки и проведения научных экспериментов на основе сервис-ориентированной парадигмы за счет автоматизации построения и применения композиций сервисов, реализующих методы анализа и обработки пространственных данных; объектом - распределенная гетерогенная информационно-вычислительная среда, функционирующая на основе сервис-ориентированной архитектуры; предметом - модели, методы, алгоритмы и программное обеспечение автоматизации построения и применения композиций сервисов в распределенной гетерогенной информационно-вычислительной среде.

В работе корректно и творчески использовались методы информационного моделирования, теории графов, системного и объектно-ориентированного программирования, проектирования баз данных, построения распределенных комплексов проблемно-ориентированных программ, веб-технологий, планирования выполнения композиций сервисов в статических и динамических средах. Разработанные теоретические подходы профессионально реализованы, в том числе, в виде больших геопорталов.

Теоретическая значимость работы заключается в развитии методов и средств автоматизации построения и применения композиций сервисов на основе комплексного анализа метаданных, онтологий, экспертных знаний и статистики применения сервисов.

Основные результаты диссертационного исследования использованы при выполнении государственных заданий, гранта Министерства науки и высшего образования РФ на выполнение крупного научного проекта «Фундаментальные основы, методы и технологии цифрового мониторинга и прогнозирования экологической обстановки Байкальской природной территории», проектов Президиума РАН, проектов междисциплинарных интеграционных программ СО РАН, РФФИ и др.

Достоверность и обоснованность полученных в диссертации результатов подтверждается корректным применением классических методов исследования, анализом адекватности разработанных моделей и методов, решением прикладных задач, индексацией полученных результатов в базах РИНЦ, Web of Science, Scopus, активной эксплуатацией сервис-ориентированной информационно-аналитической среды большим числом пользователей.

Основные результаты диссертационного исследования докладывались на многих представительных конференциях, симпозиумах и семинарах.

4. Практическая ценность основных положений диссертации

Предложенные в рамках диссертационной работы методы, модели, алгоритмы и программное обеспечение позволяют снизить трудозатраты и сократить сроки разработки программного обеспечения за счет автоматизации построения композиций сервисов обработки междисциплинарных пространственных данных.

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.

Развернут ряд различных геопорталов, ориентированных на различные предметные области и коллективы. В рамках перечисленных геопорталов сформированы композиции сервисов, объединяющие сервисы данных, сервисы обработки и публикации, созданные разными коллективами.

Практическая значимость результатов подтверждена полученными актами внедрения комплекса программных компонент. Автором в составе коллектива получено восемь свидетельств о регистрации программ для ЭВМ.

5. Рекомендации по возможности использования результатов и выводов диссертации

В силу практической важности сервис-ориентированных технологий, разработанной автором, ему рекомендуется развить их, с учетом высказанных в настоящем отзыве замечаний, и применить их к другим предметным областям научных и прикладных исследований, выстроив сотрудничество с заинтересованными научно-образовательными организациями и коллективами.

6. Замечания и вопросы по диссертационной работе

1) Вычислительная модель композиции сервисов и ее реализация (с. 129, см. также главу 2) не содержит информации о вычислительной трудоемкости каждого из сервисов, а также о необходимых им ресурсах хранилищ данных. Таким образом, предлагаемая модель композиции, на наш взгляд, не позволяет оценить как вычислительную ее полную трудоемкость и, соответственно, запланировать необходимые вычислительные мощности и объемы систем хранения данных. Между тем, такого рода оценки исключительно важны для практической реализуемости композиции сервисов. На наш взгляд, это

недостаток предлагаемого подхода. Тем не менее, что автор предлагает для решения задачи оценки необходимых ресурсов?

2) При формировании сети связанных сервисов (глава 3) в случае разных предметных областей, ими обслуживаемых, каким образом автор предполагает объединять и связывать между собой онтологии, относящиеся к разным предметным областям?

3) При формировании сети связанных сервисов (глава 3) автор ничего не пишет о необходимости оценки объемов передаваемых между сервисами данных. Также нет информации об оценке необходимой скорости передачи данных в сети. Между тем, отсутствие такого рода оценок может привести к неработоспособности сети сервисов. Каким образом автор предлагает решать эти вопросы?

4) На наш взгляд, описание (п. 4.5, стр. 146) планировщика выполнения вычислительных заданий в гетерогенной сети вычислительных серверов или кластеров очень краткое. Но планировщик заданий является важной частью распределенной вычислительной системы, его функционал следует достаточно подробно определять. Может ли автор дать более подробное описание планировщика, его алгоритма работы?

5) Каким образом автор предлагает оценивать качество функционирования композиции сервисов или ее надежность? Какие метрики здесь можно использовать?

6) Автором не были сформулированы выводы о необходимости использования сервисов защиты данных при их хранении и передаче. Очевидно, что в современных условиях незащищенные сервисы подвержены атакам киберпреступников, данные могут быть украдены, фальсифицированы и т.п. Каким образом автор предполагает использовать сервисы защиты данных в своих разработках?

7) Каким образом автор предполагает оценивать уровень доверия к используемым в композициях сторонним сервисам и системам искусственного интеллекта?

8) При описании методики создания композиций сервисов для обработки данных дистанционного зондирования Земли (с. 182) описана система анализа снимков ДЗЗ с помощью искусственного интеллекта. На стр. 185 указано, что «всего размечены 134 космоснимка Байкальской природной территории». Достаточен ли указанный объем обучающей выборки для анализа данных ДЗЗ с требуемой на практике погрешностью?

9) На стр. 192 указано «Развернуты 7 геопорталов, ориентированных на различные предметные области и коллективы». То же количество геопорталов

указано на стр. 150. Но на стр. 195 написано «Развернуты 6 различных геопорталов, ориентированных на различные предметные области и коллективы». Такая же цифра приведена на стр. 7 автореферата. Так какие из этих цифр правильные?

В заключение отметим, что высказанные выше замечания не уменьшают научной значимости полученных автором результатов.

7. Заключение о работе

Представленная диссертация является завершенной научно-квалификационной работой, выполненной на высоком научном уровне, расширяющая существующий базис теории и практики сервис-ориентированных вычислений, в ней представлены теоретические положения и предложены новые научно обоснованные подходы к решению важных научных задач, имеющих большую практическую значимость.

Результаты исследований обладают научной новизной и достоверностью, все полученные выводы научно обоснованы.

Все выносимые на защиту научные положения получены соискателем лично. Из совместных исследований в диссертацию включены только те результаты, которые принадлежат непосредственно автору.

Диссертация состоит из введения, пяти глав, заключения, библиографии из 201 наименования, списка принятых сокращений и 12 приложений. Объем основного текста работы – 195 страниц, включая 5 таблиц и 81 рисунок. Общий объем диссертации 271 страница.

Результаты диссертационного исследования отражены в 45 научных работах. Основные публикации представлены в двух монографиях, девяти статьях в российских журналах, рекомендованных ВАК для опубликования научных результатов диссертации, а также в 32 статьях, проиндексированных в международных базах цитирования Web of Science и Scopus. Автором в составе коллектива получено восемь свидетельств о регистрации программ для ЭВМ.

Тема и основные результаты диссертации соответствуют следующим областям исследований паспорта специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»: 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.

Совокупность полученных Романом Константиновичем Фёдоровым результатов позволяет заключить, что он достоин присуждения учёной степени доктора технических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Официальный оппонент:

Марченко Михаил Александрович,

доктор физико-математических наук, профессор РАН,

Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук (ИВМиМГ СО РАН),

Телефон: +7 (383) 330-83-53

эл. почта: marchenko@sccc.ru

1.2.2. Математическое моделирование, численные методы и комплексы программ

Проспект академика Лаврентьева, 6, 630090, Новосибирск, Россия,

Телефон: +7 (383) 330-83-53

эл. почта: director@sccc.ru

сайт: <https://icmmg.nsc.ru/>

« 10 » декабря 2024 г.

Подпись Марченко М.А. заверяю,
ученый секретарь ИВМиМГ СО РАН,
к.ф.-м.н.



Л.В. Вшивкова