

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

УДК _____ 654 _____

**Серафат
Амин Хоссейн**

**РАЗРАБОТКА СИСТЕМЫ ОБНАРУЖЕНИЯ СЕМАНТИЧЕСКИХ WEB-
СЕРВИСОВ НА ОСНОВЕ АЛГОРИТМА СОПОСТАВЛЕНИЯ В
ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ**

АВТОРЕФЕРАТ

на соискание степени магистра технических наук
по специальности 1-45 80 02 Телекоммуникационные системы и компьютерные
сети

Научный руководитель

Лагутин Андрей Евгеньевич



Минск 2020

КРАТКОЕ ВВЕДЕНИЕ

Система URL-адресов и поисковые программы сформировали среду, в рамках которой информация подлежит не только передаче, но и интеллектуальной обработке. Интеллектуальная обработка означает, помимо технического преобразования информации, те процессы, схожие с процессами, происходящими в человеческом мышлении. Сюда можно включить автоматизированное группирование объектов на основании заданных критериев, установление отношений между данными множествами (распознавание образов) и т.п.

Изначально компьютерный интеллект не был способен загрузить произвольный документ, Web-страницу или файл, поняв его содержание; следовательно, требовался программист, который должен был понять семантику. Выходом из этого стала так называемая «семантическая сеть». Семантическая сеть стала дополнением WWW, которое оказалось понятным машинам с точки зрения смысла информации. Подобно тому, как семантическая сеть стала расширением обычной сети WWW, семантические Web-сервисы позволили расширить понятие Web-сервисов. Под Web-сервисом понимаются методы и данные, предоставляемые для использования внешними приложениями, работающими с сервисами через стандартные протоколы и форматы данных. Web-сервисы являются полностью независимыми от языка и платформы реализации. Соответственно, Web-сервис, обладающий каким-либо семантическим описанием, является семантическим Web-сервисом. Язык описания Web-сервисов предоставляет инструмент для описания того, каким образом взаимодействовать с Web-сервисом, тогда как семантическая разметка, в свою очередь, снабжает информацией о том, какие действия производятся данным сервисом. Сеть на сегодняшний день накопила колоссальное количество данных, поэтому актуальной и приоритетной становится задача их получения с извлечением нужного контента без привлечения избыточных ресурсов и промежуточного хранения. Описание семантики Web-сервисов, как и многих других информационных ресурсов распределенной гетерогенной среды Интернет, связывают с онтологическим подходом к представлению знаний (OWL-S). Совершенствование существующих и разработка новых подходов к сбору, хранению, обработке и распространению информации является

неотъемлемой частью процесса развития информационных технологий и информационных систем.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цели и задачи проводимых исследований. Целью диссертационного исследования является разработка новых алгоритмов обработки информации для обнаружения семантических Web-сервисов в глобальной телекоммуникационной сети.

Для выполнения указанной цели автором были поставлены следующие задачи:

1. Рассмотреть понятие семантического WEB и логики телекоммуникационных сетей;
2. Выявить главные спецификации для представления информационных ресурсов в глобальной сети;
3. Дать обзор инструментальному обеспечению процесса разработки для DAML+OIL;
4. Разработать следующие алгоритмы: алгоритм обнаружения, сопоставления входных параметров, алгоритм сопоставления выходных параметров, алгоритм сопоставления профиля, алгоритм сопоставления определенным пользователем;
5. Проанализировать результаты обнаружения Web-сервиса;
6. Выполнить разработку архитектуры системы поиска семантического Web-сервиса в распределенной информационной;
7. Представить систему контроля над версиями Web-сервисов;
8. Осуществить процедуру автоматического отображения OWL документов на языке программирования Java;
9. Реализовать экспериментальный анализ разработанной системы.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики. Работа проводилась на основании утвержденных принципов, представленных в Постановлении совета министров Республики Беларусь 12 марта 2015 г. № 190 «О приоритетных направлениях научных исследований Республики Беларусь на 2016–2020 годы». В соответствии с частью второй статьи 51 Закона Республики Беларусь от 21

октября 1996 года « О научной деятельности» Совет Министров Республики Беларусь определил в качестве приоритетных следующие направления в рамках разработки информационно-коммуникационных и авиакосмических технологий: разработка интегрированных систем автоматизации управления процессами и ресурсами организаций; информационные авиационно-космические технологии; средства технической и криптографической защиты информации; биоэлектроника, биоинформатика и информационные технологии в медицине; технологии и системы электронной идентификации; технологии развития информационного общества.

Личный вклад магистранта. Магистрантом самостоятельно выполнен информационный поиск и анализ научной литературы. Совместно с научным руководителем сформулированы цель и задачи исследования. Автором предложены и апробированы алгоритмы обнаружения семантического WEB-сервиса в сети, разработана архитектура и реализация системы обнаружения семантического WEB-сервиса в глобальной сети. Лично автором проведены все исследования, систематизация и обработка результатов, анализ полученных данных, подготовка публикаций. Автором самостоятельно написаны все разделы магистрантской работы. Совместно с научным руководителем проанализированы и обобщены результаты исследований, сформулированы научные выводы и практические рекомендации.

КРАТКОЕ СОДЕРЖАНИЕ

Диссертация состоит из введения, перечня терминов и сокращений, общей характеристики, 3 глав, заключения, списка литературных источников из 72 наименований и приложений

В подразделе «Обзор литературы» представлен спектр исследований, направленных на системы обнаружения семантических web-сервисов на основе алгоритма сопоставления в телекоммуникационных сетях. В научной среде достаточно неоднозначно интерпретируются подходы и сущность данных систем. Множество исследователей в этой связи предпринимают попытки определения исторических предпосылок появления семантических сервисов. Проведенный анализ научной литературы показал, что сегодня специалисты пришли к пониманию специфики разработки систем обнаружения семантических web-сервисов на основе алгоритма сопоставления, что на практике реализовалось во множестве научных публикаций, пособий, монографических работ.

Глава 1 «Специфика описания семантики в рамках глобальной телекоммуникационной сети» содержит подходы к определению понятия и ключевых принципов семантического WEB, содержит обзор ключевых спецификаций для представления информационных ресурсов в глобальной сети, а также дает представление об инструментальных средствах обеспечения процесса разработки для DAML+OIL.

Глава 2 «Алгоритм обнаружения семантического WEB-сервиса в сети» содержит разработанные автором алгоритмы обнаружения, сопоставления входных параметров, сопоставления выходных параметров, алгоритма сопоставления профиля, сопоставления определенным пользователем, а также результаты обнаружения Web-сервиса.

Глава 3 «Архитектура и реализация системы обнаружения семантического WEB-сервиса в глобальной сети» содержит описание архитектуры системы поиска семантического Web-сервиса в распределенной информационной системе, системы контроля над версиями Web-сервисов, системы автоматического отображения OWL документов на языке программирования Java и завершается экспериментальным анализом разработанной системы.

В заключении представлены итоги проделанной работы.

ЗАКЛЮЧЕНИЕ

В ходе исследования были разработаны архитектуры информационной системы поиска семантического Web-сервиса на основе алгоритма сопоставления. Архитектура «Выполнение на стороне сервера» включает алгоритм сопоставления под управлением сервера, который возвращает соответствующие адреса Web-сервисов клиенту по его запросу. Архитектура «Локальное выполнение алгоритма сопоставления» подразумевает локальное выполнение сочетающегося алгоритма. Следовательно, при поиске Web-сервиса клиент ответственен за реализацию алгоритма сопоставления. Преимуществом локального выполнения алгоритма сопоставления является то обстоятельство, что клиент выбирает необходимый ему алгоритм сопоставления, модифицируя функциональное назначение по своим требованиям.

Кроме того, был разработан механизм автоматического получения семантического описания Web-сервиса, т.е. DAML-S описания. Также разработана система контроля над версиями Web-сервиса, предоставлены шаблоны решения проблем, возникающих при выпуске новых версий Web-сервисов. Приведен пример реализации отображения OWL документов на языке программирования Java, который активно используется при разработке системы поиска семантического Web-сервиса. В конце практического исследования были проведены практические испытания разработанных алгоритмов, доказавшие, что средняя точность сопоставления составляет до 87 правильных поисков из 100 запросов.

Сущность полученных в работе результатов заключается в следующем:

1. Разработанные алгоритмы сопоставления увеличивают эффективность семантического поиска для обнаружения Web-сервисов.
2. Имеется принципиальная возможность использования разработанной модели архитектуры для автоматического обнаружения и композиции Web-сервисов.
3. Результаты исследования контроль над версиями Web-сервисов можно использовать в любом информационном реестре для упрощения задачи контроля над версиями Web-сервисов.
4. Разработанный реестр может быть использован в качестве замены других коммерческих информационных реестров семантических Web-сервисов, что может повысить эффективность поиска различных сервисов.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

Лагутин, А. Е. Потокбезопасные структуры данных с ослабленной семантикой выполнения операций / А. Е. Лагутин, С. А. Хоссейн // Современные средства связи: материалы XXIV Международной научно-технической конференции, Минск, 17-18 октября 2019 / редкол.: А. О. Зеневич [и др.]. – Минск: Белорусская государственная академия связи, 2019. – С. 136-138. <https://libeldoc.bsuir.by/handle/123456789/37176>.