



Труды XVI междунар. конф. (30 июня - 03 июля 2014 г., Самара, Россия). – Самара: СамНЦ РАН, 2014. - С. 228-244.

В.А. Серов, Н.И. Лиманова

ПОИСКОВЫЙ ИНФОРМАЦИОННО-БИБЛИОГРАФИЧЕСКИЙ WEB-СЕРВИС С ЭЛЕМЕНТАМИ НЕЧЁТКОЙ ЛОГИКИ

(Поволжский государственный университет
телекоммуникаций и информатики)

В настоящее время библиотеки активно используют в своей работе информационные технологии. Например, для автоматизации системы учёта фондов библиотек применяются автоматизированные информационные системы (АИБС). АИБС – это пакет программ с существенно разветвлённой структурой, созданный в СУБД. Ввиду отсутствия элементов нечёткой логики в стандартных WEB-сервисах поисковых информационно-библиографических систем выборка данных по запросу пользователя сопровождается определёнными трудностями. Например, при допущении орфографических ошибок в фамилии автора или неверно набранном названии книги стандартный WEB-поиск АИБС не сможет выполнить полноценную выдачу информации или вернёт нулевой результат.

Для уменьшения этого негативного фактора в разработанном WEB-сервисе был применён алгоритм, включающий в себя модель нечёткой логики. Для устранения девиаций, вызванных орфографическими ошибками в запросе, в предлагаемом алгоритме используются механизм вычисления редакционного расстояния и созвучия слов. Для многословных запросов, таких, как названия книг, используется разбиение запроса на элементарные составляющие — отдельные слова. Совокупность найденных результатов сверяется на идентичность, и на этом основании komponуется таблица упоминаний элементов. При превышении порога упоминаний результат считается найденным и отображается пользователю. Данный подход позволяет существенно увеличить процент удачных выборок при использовании нечётких запросов. Однако возрастает также и процент неверно найденных изданий. Для увеличения релевантности выборки данных из базы данных применён алгоритм, выполняющий поэтапный поиск документов.

Данный алгоритм выполняется в следующем порядке:

- 1) поиск точного соответствия запросу;
- 2) поиск соответствия при частичном совпадении правой части запроса;
- 3) поиск соответствия при частичном совпадении левой части запроса;
- 4) фильтрация результатов с помощью алгоритма нечёткого поиска на основе расстояния Левенштейна;
- 5) разбивка запроса на элементарные составляющие (слова);
- 6) поиск сочетаний пар слов;



- 7) составление таблицы соответствия;
- 8) получение данных из СУБД с использованием таблицы соответствия;
- 9) выдача результатов.

Такой подход не аннулирует полностью выдачу нерелевантных результатов, но существенно сокращает их и помогает выполнить сортировку результатов в порядке уменьшения соответствия. Алгоритм подразумевает множество циклических запросов и существенный объём баз данных, откуда производится выборка. Поэтому для уменьшения затрат процессорного времени на выполняемый запрос было решено не только перенести таблицы АИБС в СУБД MySQL, но и оптимизировать их перед использованием в разработанном WEB-сервисе. Оптимизация заключалась в сокращении количества таблиц по сравнению с оригинальными таблицами путём отказа от второстепенных и частичном объединении нескольких таблиц с учётом обеспечения соответствия базы данных первым трём законам нормализации формы. В результате таких манипуляций удалось получить работоспособный поисковый информационно-библиографический WEB-сервис, содержащий элементы нечёткой логики.

Главным недостатком стандартных WEB-сервисов, являющихся частями АИБС (таких, как Мега-ПРО, Ирбис, Марк-SQL) является требование использования выделенных серверов под управлением операционных систем семейства Windows и прав администраторов. Это существенно ограничивает использование таких сервисов библиотечными организациями. Поселенческие библиотеки, библиотеки малых городов и другие малобюджетные библиотеки не только не в состоянии оплатить или использовать выделенный сервер, но и просто не имеют в этом необходимости ввиду малой базы данных (в пределах 100 – 150 тыс. записей). Для решения этой проблемы разработанный WEB-сервис для программной реализации алгоритма использует язык PHP, а для хранения информации – систему управления базами данных MySQL. Это позволяет расширить ареал применения WEB-сервиса и использовать его под управлением более широкого спектра операционных систем. Кроме того, низкие требования к системным ресурсам позволяют разместить его на арендованной хостинг-площадке. Заинтересованность малобюджетных организаций в таком алгоритме достаточно велика, а реализация его в универсальном формате позволяет использовать данную разработку в самом широком спектре поисковых WEB-сервисов.

Установка WEB-сервиса заключается в обычном копировании набора файлов *.php на хостинг. Перенос информации из базы данных АИБС в СУБД MySQL возможен средствами доступными для пользователя. Например, с помощью условно-бесплатной программы dbForge Studio for MySQL. WEB-сервис с элементами нечёткой логики реализован и доступен по адресу <http://biblioteka-nsk.ru>. Он успешно выполняет поиск изданий по фонду библиотечной сети г. Новокуйбышевска.

Вышеизложенный подход к реализации поискового информационно-библиографического WEB-сервиса делает возможным применение такой



модели поиска в различных сферах, где есть потребность в on-line поиске по различным базам данных. Например, одновременно с внедрением разработанного поискового WEB-сервиса выполняется интеграция вышеизложенного алгоритма в WEB-сервис электронного каталога музейных экспонатов историко-краеведческого музея Волжского района. Заинтересованность малобюджетных организаций в таком алгоритме достаточно велика, а реализация его в универсальном формате позволяет использовать данную разработку в самом широком спектре поисковых WEB-сервисов.

Е.В. Симонова, Д.Е. Мишутин

РАЗРАБОТКА МЕТОДА СОЗДАНИЯ ПРЕДМЕТНО-ОРИЕНТИРОВАННОЙ ОНТОЛОГИИ ОБОРУДОВАНИЯ МАЛЫХ КОСМИЧЕСКИХ АППАРАТОВ ДЛЯ ОРГАНИЗАЦИИ ИНТЕЛЛЕКТУАЛЬНОГО ПРОЦЕССА ПРЕДОСТАВЛЕНИЯ УСЛУГ

(Самарский национальный исследовательский университет
имени академика С.П. Королёва)

Введение

В настоящее время расширяется сфера применения данных, получаемых из космоса. Одной из перспективных тенденций в космической отрасли является создание многоспутниковых орбитальных группировок, состоящих из малых космических аппаратов (МКА), способных значительно расширить нишу потребительских сервисов, по сравнению с реализуемыми в настоящее время. Для обеспечения целевого функционирования создаваемых орбитальных группировок и повышения качества предоставляемых услуг возникает необходимость использования современных схем взаимодействия с потребителями в сочетании с организацией эффективного планирования и управления ресурсами группировки МКА в режиме реального времени.

Один из новых подходов в сфере управления и планирования ресурсов связан с применением мультиагентных технологий и онтологий. Онтологии используются в качестве основы базы знаний о ключевых понятиях предметной области и связях между ними, этим обеспечивается единство терминологии и предоставляется возможность расширения знаний системы в случае появления новых концептов и связей, а также появляется возможность формально специфицировать ситуации в реальности.

Постановка задачи

Основная задача – разработка метода создания онтологии оборудования малых космических аппаратов для организации интеллектуального процесса предоставления услуг в реальном времени. Онтология должна позволять описывать знания о предметной области, строить концептуальные модели космической деятельности, формировать модели ситуаций (сцены), используемые для



ситуационного управления в интеллектуальных системах поддержки принятия решений на основе мультиагентных технологий, базирующихся на принципах самоорганизации и эволюции.

Язык представления онтологий OWL

В последнее время языки представления онтологий получают все более интенсивное развитие в связи с общим процессом «семантизации» веб. В рамках этого процесса каждый информационный ресурс Интернет должен получить свой семантический дескриптор на основе изначально заданной общей онтологии для того, чтобы позволить программам более содержательно использовать его информацию.

Язык описания онтологий должен удовлетворять следующим требованиям:

- быть достаточным для описания концептов предметной области;
- быть легко используемым на существующих платформах;
- давать возможность неспециалистам легко строить концептуальные модели.

Язык веб-онтологий OWL – это язык для определения и представления веб-онтологий [1]. Веб-онтология может включать описания классов, свойств и их примеры. Формальная семантика OWL описывает, как получить логические следствия, имея такую онтологию, т.е. получить факты, которые не представлены в онтологии буквально, но следуют из ее семантики. Эти следствия могут быть основаны на одном документе или множестве распределенных документов, которые комбинируются с использованием определенных механизмов OWL.

Классификация онтологий по назначению

На современном этапе развиваются следующие типы онтологий:

- 1) **Онтология представления.** Целью создания такой онтологии является описание области представления знаний, составление языка для спецификации других более детализированных онтологий. Например, описание понятий языка OWL средствами языков RDF/RDFS.
- 2) **Онтология предметной области.** Другое название – онтология домена. Область ее применения ограничена одной предметной областью, например, авиация, медицина, культура, и т.п. Онтология предметной области обобщает понятия, использующиеся в некоторых задачах домена. Во многих дисциплинах сейчас разрабатываются стандартные онтологии, которые могут использоваться экспертами для совместного аннотирования информации по своей тематике.
- 3) **Прикладная онтология.** Назначение такой онтологии состоит в создании концептуальной модели, используемой при решении конкретной задачи или при работе некоторого приложения. Прикладные онтологии описывают понятия, которые зависят как от онтологии задач, так и от онтологии предметной области. Для реализации поставленной задачи был выбран прикладной тип онтологии.