

УДК 004.738.52

А.С. Смелякова, Д.П. Галас, Т.С. Ткачёва

Харьковский национальный университет радиоэлектроники, Харьков

АВТОМАТИЗАЦИЯ ПОИСКА И АНАЛИЗ ЕГО ПРИМЕНЕНИЯ В ИНФОРМАЦИОННЫХ СИСТЕМАХ И ВЕБ-СЕРВИСАХ

В статье анализируются технологические аспекты использования алгоритмов поиска в различных классах социальных сетей и веб-сервисах. Приводится классификация существующих социальных сетей по различным признакам. Рассматриваются достоинства и недостатки существующих поисковых алгоритмов.

Ключевые слова: *информационная система, поиск, социальная сеть, веб-сервисы.*

Введение и постановка задачи

В настоящее время все большую популярность приобретают социальные сети. Социальная сеть – интерактивный многопользовательский веб-сайт, контент которого наполняется самими участниками сети. Сайт представляет собой автоматизированную социальную среду, позволяющую общаться группе пользователей, объединенных общим интересом. К ним относятся и тематические форумы, особенно отраслевые, которые активно развиваются в последнее время. Связь осуществляется посредством веб-сервиса, внутренней почты или мгновенного обмена сообщений.

Существует целый ряд социальных сетей для поиска не только людей по интересам, но и самих объектов этих интересов. Огромное количество пользователей, зарегистрированных в социальных сетях, создают и объединяются в группы по интересам. Однако, несмотря на то, что в социальных сетях существует большое количество групп по интересующей вас тематике, достаточно сложно найти пользователей по какому-либо интересу. Как правило, группы объединяют людей по узкой направленности в рамках нужной темы, и, к тому же, их количество очень велико. В этой ситуации найти действительно интересующую человека группу бывает затруднительно.

Большую популярность среди пользователей Интернета также приобретает увлечение экстремальными видами спорта, например такими видами велоспорта как: стрит, дерт, триал, даунхилл и фрирайд. На текущий момент еще не существует единой системы, которая бы объединяла людей с этим увлечением, и давала им соответствующие специфические функции. Поэтому актуальным является разработка сервисной социальной сети в которой пользователи, увлеченные подобными видами спорта, смогли бы найти не только возможность общаться с людьми с таким же интересом, но и обучаться новым трюкам и техникам езды, подбирать новые детали для своего велосипеда, рассказать другим, какие именно детали уже установлены на велосипеде, и возможность выложить объявление о продаже своих старых деталей. Следует отметить, что одной из главных функций разрабатываемой системы является функция поиска. Так как эта опция дает возможность пользователю в реальном масштабе времени перемещаться по сайту и получать информацию в соответствии с запросом пользователя.

Основная часть

Для решения этих задач была разработана информационная система, которая посвящена людям увлекающимся экстремальными видами велоспорта. В данном проекте на одном сайте объединены возможности, которые ранее пользователи могли получить только зарегистрировавшись на разных сайтах.

Информационная система «Экстремальный велоспорт» сочетает в себе следующие возможности:

- создан каталог деталей с различных магазинов Украины. На каждую деталь в каталоге показана ее стоимость в разных магазинах, и даны ссылки на магазины, где можно купить интересующую деталь;
- возможность комментирования и выставления рейтинга на детали в каталоге;
- создание личной странички пользователя, где он сможет разместить информацию о себе;
- после создания личной странички, пользователь может показывать на ней комплектацию своего велосипеда, а другие пользователи могут оставлять свои комментарии к комплектации;
- на странице пользователей есть стена, где другие участники данного веб-сервиса могут оставлять свои сообщения;
- создан раздел с описанием по выполнению различных трюков, а также с возможностью вставки видео из таких популярных сервисов как YouTube, Vimeo, vkontakte.ru и другие;
- раздел для размещения информации о прошедших или грядущих соревнованиях и других спортивных событиях;
- в каждом из разделов имеется возможность комментирования содержимого.
- поиск по разделам веб-сервиса.

Одной из актуальных задач при использовании различного рода социальных сетей и веб-сервисов является функция поиска. Эта опция дает возмож-

ность пользователю в реальном масштабе времени перемещаться по сайту и получать информацию в соответствии с запросом пользователя.

Существует большое количество алгоритмов поиска в Интернет, которые базируются на таких элементах, как внешние ссылки на сайт, уникальность контента, концентрация ключевого слова в тексте и другие. Разработка систем поиска по сайту схожа с разработкой поисковой системы Интернета. Основное отличие – это объемы информации. Попытки разработки системы поиска информации на сайте "на скорую руку" всегда приводят к снижению эффективности сайта, т.е. к снижению дохода и прибыли от сайта. Низкое качество результатов, выдаваемых поисковыми модулями, созданными без учета особенностей конкретного веб-сервиса, могут выдать некорректные результаты посетителям сайта.

Для избежания вышеперечисленных проблем, в данном веб-сервисе используется поиск, который ищет вхождения слова или словосочетания из поискового запроса, в соответствии с разделом сайта на котором находится пользователь (рис. 1), т.е. когда пользователь находится на страницах каталога, то вхождения поискового запроса ищется в рамках текущего каталога, если же пользователь находится, к примеру, на странице списка пользователей веб-сервиса, то поиск будет производиться по именам и фамилиям зарегистрированных пользователей.



Рис. 1. Диаграмма состояний алгоритма поиска

Функция поиска была реализована с помощью возможностей, которые дает маршрутизация фреймворка Symfony (рис. 2).

Для общего поиска используется единые контроллер и форма, но в зависимости от роутинга, выбирается нужная модель для поиска. Происходит это таким образом: пользователь находится на странице с URI:/catalog, при обращении к коллекции sfRouting, мы получаем название роутинга – catalog, который в данном веб-сервисе соответствует данному URI. Затем название роутинга подставляется с помощью возможностей, которые предоставляет PHP5.3, в назва-

ние соответствующей модели, а методы которые выполняют поиск в каждой модели имеют одинаковое название (например: `CatalogTable::getSearch($query)`). Таким образом описываемая возможность реализуется двумя строками кода:

```
$routename = $this->getRouting()->getCurrentRoutename();
$routename.Table::getSearch($query);
```



Рис. 2. Алгоритм выполнения выборки

Когда нужная модель получила запрос по которому будет осуществлен поиск происходит разбор запроса на слова. Это выполняется для нахождения большего количества соответствий с запросом. Когда запрос был разделен на слова, то сначала происходит поиск в базе на соответствие с запросом в том виде, в котором он был написан пользователем. И данные результаты поиска выводятся в первую очередь. Затем, если запрос был задан не одним словом, а словосочетанием, то поиск в базе производится на соответствие с данным словосочетанием, но в словосочетании меняется порядок слов. Эта выборка производится до тех пор, пока каждое слово в словосочетании не побывает на каждой из возможных позиций. Результаты этих выборок выводятся после

результатов выборки по полному словосочетанию запроса. В конец списка результатов также выводятся все найденные соответствия отдельно по каждому слову из запроса, начиная с первого слова и дальше, по порядку слов в словосочетании.

Описанный алгоритм позволяет сократить время поиска запроса и вывести наибольшее количество релевантных результатов.

Выводы

Таким образом, данная информационная система объединит пользователей, которые занимаются различными экстремальными видами велоспорта на одном веб-сайте. Даст им возможности сравнивать цены на детали в различных велосипедных магазинах Украины; узнавать отзывы пользователей на детали; находить нужную информацию по выполнению трюков; пользоваться собственной страницей, где размещены фотографии, которыми пользователь хочет поделиться с другими; размещать объявления о продаже своих старых деталей, а также дает возможность посвящать других пользователей в различные спортивные события которые уже прошли или будут. Эта система предоставляет все необходимые функции, которые могут понадобиться для пользователей, увлекающихся экстремальным велоспортом, и которые еще не были объединены в единой системе.

Список литературы

1. Кнут Д. Искусство программирования. Сортировка – 2-е изд. – М.: «Вильямс», 2007. – 824 с.
2. Алгоритмы: построение и анализ / Томас Х. Кормен, Чарльз И. Лейзерсон, Рональд Л. Ривест, Клиффорд Штайн. – 2-е изд. – М.: «Вильямс», 2006. – 1296 с.
3. Седжвик Р. Фундаментальные алгоритмы на Си. Анализ/Структуры данных/Сортировка/ Поиск / Р. Седжвик. – СПб.: ДиаСофтЮП, 2003. – 672 с.
4. Гуляев А. Поиск в Интернете / А. Гуляев. – Питер, 2006. – 355 с.
5. Курейчик В.М. Поисковая адаптация: теория и практика / В.М. Курейчик, Б.К. Лебедев, О.К. Лебедев – М.: Физматлит, 2006. — 272 с.

Поступила в редколлегию 23.03.2011

Рецензент: д-р техн. наук, проф. Е.П. Пуятин, Харьковский национальный университет радиоэлектроники, Харьков.

АВТОМАТИЗАЦІЯ ПОШУКУ І АНАЛІЗ ЙОГО ЗАСТОСУВАННЯ В ІНФОРМАЦІЙНИХ СИСТЕМАХ ТА ВЕБ-СЕРВІСАХ

А.С. Смелякова, Д.П. Галас, Т.С. Ткачова

У статті аналізуються технологічні аспекти використання алгоритмів пошуку в різних класах соціальних мереж і веб-сервісах. Наводиться класифікація існуючих соціальних мереж за різними ознаками. Розглядаються достоїнства та недоліки існуючих пошукових алгоритмів.

Ключові слова: інформаційна система, пошук, соціальні мережі, веб-сервіси.

AUTOMATIZATION OF SEARCHING AND ANALYSIS OF IT'S USING IN INFORMATION SYSTEMS AND WEB SERVICES

A.S. Smelyakova, D.P. Galas, T.S. Tkachova

In the article there are analysis of technological aspects of searching algorithms in different classes of social networks and Web services. A classification of existing social networks by different groups are given. The advantages and disadvantages of existing search algorithms are considered.

Keywords: information system, search, social networks, Web services.