

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

УДК 004.8912 00.4775

Рудый Александр Владимирович

ЭКСПЕРТНАЯ СИСТЕМА РЕКОМЕНДАЦИИ ВЫДАЧИ КОНТЕНТА

АВТОРЕФЕРАТ

на соискание академической степени
магистра технических наук

По специальности 1-40 80 04 – Математическое моделирование, численные
методы и комплексы программ

Научный руководитель
кандидат физико-математических наук, доцент
Сиротко С.И.

Минск 2015

ВВЕДЕНИЕ

В современном обществе четко прослеживается тенденция увеличения роста информационного потока. Так называемый информационный взрыв - постоянное увеличение скорости и объема информации в масштабах планеты, носит лавинообразный характер. Данный процесс является не случайным, а носит закономерный характер. Информационный взрыв таит в себе большую опасность и может вылиться в информационный кризис: человечество не будет успевать перерабатывать всю информацию, которую будет производить.

В условиях постоянного роста объема информации все острее встает вопрос о поиске информации. Обычные поисковые системы уже не могут справиться с таким объемом информации, так как выступают в роли обычных фильтров, следовательно нужно изменить сам процесс поиска. В последнее время большую популярность стали приобретать рекомендательные системы, которые пытаются предугадать интересы пользователя и удовлетворить его запросы посредством рекомендаций. Целью данной работы является построение такой экспертной системы.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Целью исследования является построение системы, способной осуществлять поиск в постоянно растущем объеме информации, без явного знания структуры искомого объекта.

Для достижение поставленной цели были выделены следующие задачи:

- обзор существующих типов поисковых систем и анализ их преимуществ и недостатков;
- разработка собственной системы, на основе полученных результатов;
- обзор путей повышения качества системы и внедрение собственных решений;
- построение прототипа системы на реальных данных и тем самым обоснование практической применимости авторских решений.

Магистерская диссертация относится к перспективному, активно развивающемуся и востребованному направлению поисковых, рекомендательных и экспертных систем, вырабатывающих решения на основе обработки больших объемов информации. Исследование затрагивает математические, алгоритмические и технологические основы рекомендательной системы с использованием явных количественных оценок объектов, адаптивных профилей пользователей и элементов самообучения.

Были разработаны собственные методы повышающие разнообразие выдаваемых рекомендаций, а также предложен подход к построению композиции рекомендаций.

В конце исследований была разработана экспериментальная версия системы для демонстрации и тестирования примененных решений на примере базы данных рейтингов кино- и видеопродукции MovieLens.

В ходе выполнения диссертаций автором был сделан ряд публикаций, список которых приведен в соответствующем разделе.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Диссертация состоит из трех глав, а также введения, заключения и приложения.

Во введении отражена постановка проблемы, а также возможное решение данной задачи путем построения экспертной системы.

В первой главе производится обзор современных поисковых систем. Среди них выделяются три класса: поисковые системы, вопросно-ответные системы и рекомендательные системы. В соответствующих разделах рассматриваются общие принципы их работы, историческое развитие, а также основные достоинства и недостатки. Также в конце раздела в качестве наилучшего решения для построения экспертной системы рекомендации контента выбирается рекомендательная система и дается обоснование этого выбора.

Во второй главе производится построение рекомендательной системы. Для этого в первом разделе осуществляется формализации задачи выдачи рекомендаций. В последующих разделах рассматриваются возможные методы решения поставленной задачи и соответственно методы построения рекомендательных систем с кратким описанием их достоинств и недостатков.

По результатам сделанного анализа, выбирается рекомендательная система, построенная на основе SVD-разложения, как имеющая наиболее лучшие показатели производительности, а в качестве тестируемых данных выбирается база данных рейтингов фильмов MovieLens, содержащая данные о примерно 1700 фильмах и 1000 пользователях.

Также автором производится обзор путей повышения качества рекомендаций и разрабатываются собственные решения, суть которых

описывается в соответствующих разделах. В конце главы делается обзор дальнейших путей развития повышения качества рекомендаций.

В третьей главе производится обоснование выбора Octave/Matlab в качестве технологии, на основе которой строится прототип рекомендательной системы. Листинг основных компонентов данной системы приводится в приложении А.

По результатам проделанных исследований делается общий вывод.

Также к работе прилагается библиографический список, содержащий, в том числе и публикации автора работы.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Был сделан обзор существующих подходов к решению задачи поиска, а также существующих на их основе систем. Сформулированы и перечислены их основные достоинства и недостатки, а также выбрана рекомендательная система в качестве экспертной системы рекомендации выдачи контента, как единственная, способная удовлетворить требованию поиска в постоянно растущем информационном потоке.

2. Перечислены основные способы построения рекомендательных систем, а также их достоинства и недостатки, а также построена рекомендательная система с использованием SVD-разложения.

3. Были сформулированы пути повышения качества построения списка рекомендаций, а также формализована оценка разнообразия посредством разработанного функционала.

4. Используя Octave/Matlab в качестве платформы была на практике реализована рекомендательная система, а также впоследствии внедрено решение, повышающее разнообразие выдаваемых рекомендаций.

Рекомендации по практическому использованию результатов

1. Полученные результаты формируют теоретическую и практическую базу для разработки алгоритмов повышения качества рекомендаций. Они могут быть использованы для модернизации и дальнейшего развития существующих систем.

2. Разработанная рекомендательная система может найти применение в различных сферах, т.к. не опирается на специфические особенности какой-либо конкретной области.

3. Результаты работы могут использоваться для дальнейшего исследования улучшения качества выдаваемых рекомендаций.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Рудый А.В. Защита рекомендательных систем от целенаправленного зашумления рейтингов.

// Технические средства защиты информации: материалы XII Белорусско-российская научно-технической конференции.

Минск, 28–29 мая 2014 г / БГУИР, Редакционная коллегия: М. Лыньков, А.М.

Прудник, В.Ф. Голиков, В. Давыдов,

О.Р. Сушко, В.К. Конопелько

2-А. Статья на электронном ресурсе habrahabr.ru "Функциональный Javascript. Пишем свои линзы"