

Данный метод балансировки нагрузки в РАС значительно повышает производительность системы, так как количество ДЦУ ограничено только вычислительной мощностью разработанной инфраструктуры системы.

Список литературы

1. Радченко Г.И. Распределенные вычислительные системы: учебное пособие / Г.И. Радченко. – Челябинск: Фотохудожник, 2012. – 184 с.
2. Родин А.В., Бурцев В.Л. Параллельные или распределенные вычислительные системы? // URL: http://gridclub.ru/library/publication.2006-02-07.1818516730/publ_file/ (дата обращения: 11.01.2016).
3. Балансировка нагрузки в распределенных системах // Интуит. URL <http://www.intuit.ru/studies/courses/1146/238/lecture/6153?page=2> (дата обращения: 09.12.2015).
4. Балансировка нагрузки NGINX // 8host. URL: <http://www.8host.com/blog/balansirovka-nagruzki-nginx/> (дата обращения 06. 01. 2016)
5. Исследование стратегий балансировки нагрузки в системах распределенной обработки данных // КиберЛенинка. URL: <http://cyberleninka.ru/article/n/issledovanie-strategiy-balansirovki-nagruzki-v-sistemah-raspredeleynoy-obrabotki-dannyh> (дата обращения 23. 11. 2015)

УДК 004

ВЫБОР ОПТИМАЛЬНОГО АЛГОРИТМА РЕКОМЕНДАЦИИ ОНЛАЙН-СЕРВИСА

Никитина К.С., Дорофеев В.А.

Научный руководитель: Дорофеев В.А., старший преподаватель кафедры ИПС ИК ТПУ

Национальный Исследовательский Томский политехнический университет,

634050, Россия, г. Томск, пр. Ленина, 30

E-mail: kseniya.niky@gmail.com

This article describes recommendation systems methods, algorithms and disadvantages of these methods. Also it explains choice of the optimal algorithm recommendation for the online service.

Key words: *system recommendations, collaborative filtering, user-based collaborative filtering, item-based collaborative filtering.*

Ключевые слова: *система рекомендаций, методы коллаборативной фильтрации, методы, анализирующие содержимое объектов, методы, основанные на знаниях.*

У каждого человека возникает желание посмотреть, послушать, почитать или купить что-то новое. Выбор из изобилия фильмов, музыки, книг или товаров – это не простая задача. Основываясь только на столь малом количестве информации о новом продукте, человек не может утверждать, что продукт ему понравится. Именно для решения таких задач используются рекомендательные системы.

Системы рекомендаций позволяют пользователю выбрать среди всех доступных объектов именно те, которые будут ему интересны. Данные системы обрабатывают информацию, как о самом объекте, так и о действиях пользователей с этими объектами. Наиболее популярны рекомендательные системы в интернет-магазинах, например, OZON.ru, но есть серви-

сы подбирающие своим пользователям видео (YouTube.com), фильмы (КиноПоиск), книги (Readly) или других пользователей (Facebook, ВКонтакте).

Основная задача системы рекомендации – на основе имеющихся данных об объектах и пользователях получить список объектов, которые будут наиболее интересны для конкретного пользователя. Исходя из данных, используемых для расчета рекомендации, системы делятся на три класса:

- методы коллаборативной фильтрации;
- методы, анализирующие содержимое объектов;
- методы, основанные на знаниях [1].

Методы коллаборативной фильтрации основываются на оценках, которые пользователи выставляют объектам, т. е. рассматривают реакцию пользователя на предложенный ему объект. Сбор оценок происходит явным способом, при котором пользователь ставит оценку, например, от 1 до 5, или неявным, когда система отслеживает поведение пользователя, например, количество просмотров определенного ролика [2]. Задача системы, в таких методах, сводится к предсказанию оценки, которую может поставить пользователь p объекту i . Алгоритм состоит из трёх шагов:

Сопоставим интересы каждого пользователя m с интересами пользователя n (вычислим близость пользователей).

Выборка множества пользователей наиболее близких к пользователю n .

Предсказываем оценку пользователя n на основе оценок объекта i «соседями» из предыдущего шага.

Описанный алгоритм основан на сравнении между собой пользователей (user-based collaborative filtering). Существует другой подход, основанный на сравнении объектов (item-based collaborative filtering). В этом случае алгоритм также состоит из трёх шагов, но схожесть проверяем не пользователей, а объектов, и в соответствии с наиболее близкими объектами предсказываем оценку рекомендуемого объекта.

Существенными недостатками данных методов является проблема «холодного старта»: система не знает, что рекомендовать новому пользователю и кому рекомендовать новый товар и необходимо большое количество данных о разных пользователях [3].

Методы, основанные на анализе содержания объекта, являются наиболее подходящими при рекомендации книг, статей и другой текстовой информации. В данном случае каждый пользователь имеет свой профиль, в котором указывает свои интересы. Затем на основе ключевых слов подбираются похожие объекты и предлагаются пользователю. Основным минусом является узконаправленность методов. Также данные система получает от пользователя с помощью обратной связи, но не каждый пользователь желает предоставлять информацию о своих интересах.

Предыдущие методы хорошо работают при высокой активности пользователя, например, пользователь довольно часто покупает книги в интернет-магазине. Методы, основанные на знаниях, решают проблему редких покупок, т. е. когда о действиях пользователя мало информации или её вовсе нет. Эти методы делятся на два вида: использование жестких ограничений и выбор близких объектов. Идея обеих следующая: пользователи формулируют свои требования к объекту, а система пытается найти нужный объект. Первый метод особенно требователен к подбору объекта, так как он должен соответствовать всем требованиям пользователя. Во втором допускается объекты с достаточно близкими характеристиками.

Выберем оптимальный алгоритм из выше предложенных для онлайн-сервиса рекомендаций художественной литературы. Для начала проанализируем, какие методы рекомендаций используются похожими сервисами. Интеллектуальный сервис «Узнай, что почитать» подбирает книги пользователю на основе фильтра, в котором пользователь явно указывает

цель чтения, что пользователь хочет получить в результате прочтения и основной критерий любой книги – жанр. По сути, фильтр является выбором ключевых слов, которые необходимы для подбора похожих книг.

Рассмотрим другой рекомендательный сервис. «Имхонет» является проводником не только в мир книги, но и в галактику фильмов и игр. Данный сервис собирает как можно больше информации от пользователя путем оценки некоторого количества книг, фильмов. Затем на основании поставленных оценок ищет единомышленников. Именно эти люди будут выступать в роли рекомендателей. Алгоритм, использующийся в сервисе, является коллаборативной фильтрацией, основанной на сравнении пользователей. Проблема «холодного старта» решена тем, что пользователь, желающий получить рекомендацию, будет вынужден оценить некоторое количество объектов.

На основе проведенного анализа, можно сделать вывод, что для онлайн-сервиса рекомендации художественной литературы наилучшим вариантом является метод анализа содержимого объекта. Из текстовой информации проще извлекать ключевые слова, чем отслеживать оценки или просмотры той или иной книги.

Список литературы

1. Гомзин А.Г., Коршунов А.В. Системы рекомендаций: обзор современных подходов // Труды Института системного программирования РАН. – 2012. – Т. 22. – С. 401–417.
2. Филиппова Е. Системы рекомендаций: задачи, подходы, алгоритмы. URL: <http://datareview.info/article/sistemyi-rekomendatsiy-zadachi-podhodyi-algoritmyi/> (дата обращения: 03.03.2016).
3. Джонс Т. Рекомендательные системы: Часть 1. Введение в подходы и алгоритмы. URL: <http://www.ibm.com/developerworks/ru/library/os-recommender1/index.html> (дата обращения 03.03.2016).

УДК 004

ПОСТРОЕНИЕ АРХИТЕКТУРЫ ЛОКАЛЬНОЙ СЕТИ МНОГОПРОЦЕССОРНОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

Петрухина И.А., Погребной А.В.

Научный руководитель: Погребной А.В., доцент кафедры ИПС

Национальный Исследовательский Томский политехнический университет,

634050, Россия, г. Томск, пр. Ленина, 30

E-mail: irk-2009@yandex.ru

The following article represents method of finding the best solution for connecting stations to the core network. The main criterion of choosing the variation is a minimum data transfer time, which is achieved by parallelizing dataflow in network. This method is used in modeling real-time system.

Key words: *real-time system, cutting dataflow graph, matrix for conflicts, software load, maximum empty subgraph, core network.*

Ключевые слова: *система реального времени, разрезание графа передачи данных, матрица наличия конфликтов, программная нагрузка, максимально пустой подграф, базовая сеть.*