

УДК 02 : 002.5
ББК 78.38

РЕКОМЕНДАТЕЛЬНЫЕ СИСТЕМЫ В ПУБЛИЧНЫХ БИБЛИОТЕКАХ

© А. С. Карауш, 2009

*Муниципальная информационная библиотечная система г. Томска
634034, г. Томск, ул. Красноармейская, 119*

Описывается технология, применимая в рамках автоматизированных библиотечных систем и основанная на коллаборативной фильтрации, предназначенная для формирования рекомендательных списков.

Ключевые слова: библиотека, автоматизированные системы, рекомендательные сервисы.

The article describes the technology, applicable in automated library systems and based on collaborative filtration, for forming recommendation lists.

Key words: library, automated systems, recommendation services.

В последнее время отмечается существенное ослабление интереса у населения к досуговому чтению и проведению свободного времени в виде чтения современной художественной или научно-популярной литературы. Многие читатели объективно полагают, что современные возможности библиотек не позволят им найти документы или данные, с помощью которых можно с пользой провести свое личное время. Намного проще использовать «подготовленный» для восприятия видеоконтент или звуковой контент. Для таких решений потенциальных читателей имеются следующие предпосылки:

- взрывной рост объемов доступных обществу данных (увеличение числа книг, фильмов, новостей, рекламных сообщений и пр.);
- увеличение объема он-лайн данных;
- реальный объем информации, окружающей человека, значительно выше того, что он может реально пропустить через себя, чтобы обнаружить необходимую и достаточную, а также понравившуюся ему.

Решения о фильтрации поступающих данных тесно связаны с технологиями.

1. Нужна технология, которая помогла бы читателю найти то, в чем он нуждается, и избежать того, на что не желает тратить свое время и внимание.

2. Для решения проблемы «информационного вала» часто используется метод фильтрации контента, который необходимо совершенствовать.

3. Для пользовательского потребления система должна рекомендовать объекты и данные на основе корреляции содержания с пользовательскими предпочтениями. Например, система может попытаться установить корреляцию между наличием

ключевых слов в статье и пользовательскими предпочтениями.

Сегодня библиотеки разных ведомств и уровней пытаются разобраться с показателями собственной деятельности и определить границы применимости этих показателей для новых сред распространения данных. Показатели (или индикаторы) работы библиотеки должны быть прежде всего объективные; легко вычисляемые; «конвертируемые» в «небиблиотечные» индикаторы и показатели (финансовые, кадровые, ресурсные и пр.); относительно «стабильные» во времени.

Для публичных и массовых библиотек особенно важно внимание населения к досуговому и развивающему чтению, поскольку это позволит библиотекам формировать и развивать пользовательские группы (целевые аудитории для потребления услуг предоставления информации).

Автор придерживается мнения, что основным критерием работы современной публичной библиотеки является качественное личностное время пользователя – время, которое человек субъективно оценивает как хорошо и полезно проведенное.

Основной показатель (индикатор) работы публичной библиотеки – количество качественного личностного времени, которое читатели приобрели, используя ресурсы конкретной библиотеки (книги, комфорт, тепло, цифровые данные, аудио- и видеодокументы и т. д.), присутствуя там непосредственно или дистанционно (удаленно).

В этой связи стоит рассмотреть наработки в библиотечном деле по направлению развития сотрудников до уровня «сетового библиотекаря» или «библиотекаря 2.0». В настоящее время для увеличения качественного личностного времени читате-

ля библиотеками активно используются следующие технологии и сервисы:

- интернет-справка и пр. (вопрос-ответ, форумы, блоги и пр.);
- создание профилей пользователей (скрытых и явных);
- фильтрация данных поиска (Интернет, пиринговые сети и пр.):
- автоматические системы рекомендаций;
- индивидуальные рейтинги (в том числе без возможности общения);
- развитие картотек пользователей библиотеки в Автоматизированной библиотечно-информационной системе (АБИС) до уровня CRM (система управления пользователями).

Коллаборативная фильтрация [1, 4] может и должна использоваться в АБИС как основная для фильтрации передаваемых данных между АБИС и пользователем.

Коллаборативная фильтрация, совместная фильтрация (collaborative filtration) – метод, дающий автоматические прогнозы (фильтрацию) относительно интересов пользователя по предварительно собранному данным о вкусах множества пользователей, сотрудничающих между собой, где определяется, какие из пользователей имеют тот же вкус на основании стандартных формул вычисления статистических корреляций. При этом основное допущение: те, кто соглашался в прошлом, склонны согласиться и в будущем.

Коллаборативная фильтрация отличается от рейтинга, дающего усредненную оценку для каждого объекта интереса, базирующуюся на количестве поданных за него голосов.

Сегодня уже можно описать круг задач и областей применения для автоматизированных библиотечных систем, основанных на использовании рекомендательных сервисов, в частности коллаборативной фильтрации [2, 5]:

- а) Аннотация в контексте. Фильтрация сообщений на структурированных форумах и блогах в целях принятия решения, какие из них стоит читать.
- б) Найти некоторые хорошие объекты. Пользователю предлагаются конкретные объекты, представленные в виде ранжированного перечня рекомендуемых объектов вместе с прогнозом, насколько сильно они ему понравятся.
- в) Найти ВСЕ хорошие объекты. Позволяет выделить из массива объектов не несколько хороших, а ВСЕ.
- г) Последовательность рекомендаций для пользователя из нескольких объектов. При чтении (изучении) каждого из них поочередно пользователь получит качественное личностное время, переходя от одного объекта (например, менее сложного в понимании) к другому (более сложному) и развиваясь как личность.

д) Только просматривание. Для кого-то это развлечение, для кого-то – процесс познания группы единомышленников.

е) Поиск пользователем надежной системы коллаборативной фильтрации. Пользователи «забавляются» с системой какое-то время, чтобы посмотреть, удовлетворяют ли рекомендуемые объекты их вкусам с учетом изменения ими поставленных ранее оценок для объектов.

Типовой пример алгоритма рекомендательной системы на основе коллаборативной фильтрации [1, 3]:

1. Каждый пользователь в некоторой группе оценивает определенное количество объектов потребления, например книг.

2. Каждой книге ставится оценка: «очень понравилось», «понравилось», «так себе», «скучновато» и «плохо».

3. На основании данных о предпочтениях для каждого пользователя формируется круг единомышленников, вкус которых по отношению к оцененным им книгам совпадает с его личными пристрастиями.

4. Выбираются произведения, прочитанные и высоко оцененные единомышленниками, но не прочитанные данным человеком, – эти произведения и составляют выданные ему рекомендации.

В результате лучше предсказывается оценка человеком неизвестных ему объектов, чем следование рейтингам и общесоциологическим данным: пол, возраст, образование, социальный статус и т. д.

При использовании вышеприведенного примера алгоритма можно уже сегодня получить новые сервисы для пользователей библиотек, проведя доработки в АБИС:

- рекомендательные списки к чтению для всех видов носителей и документов;
- индивидуальные рейтинги и списки (последовательность рекомендуемых объектов);
- навигация на основе предоставленных рекомендаций;
- индивидуальное ранжирование результатов поиска в АБИС.

Кроме очевидных достоинств, стоит упомянуть и про имеющиеся недостатки рекомендательных систем, построенных с использованием модели коллаборативной фильтрации:

1. Проблема накопления «личного отрицательного опыта» связана с чтением не всегда «приятных» для пользователя объектов и анализом отрицательной информации. Система коллаборативной фильтрации рекомендует только понравившиеся объекты для чтения.

2. Рекомендуются документы, уже кем-то прочитанные и оцененные. Нет возможности рекомендовать пользователю книгу, на которую нет отзыва в системе.

Новые задачи определяют проблемы и недоработки АБИС, которые необходимо решить для использования новых возможностей обслуживания пользователей:

1. Развитие АБИС по законам и направлениям развития CRM.

2. Использование весовых коэффициентов для каждого возвращенного объекта при поиске.

3. Обмен данными с другими современными CRM: единый формат обмена; формат профиля пользователя; особенности работы с «личной информацией».

4. Автоматически просчитываемые индивидуальные рейтинги для чтения.

5. Автоматические анализаторы спроса на литературу для определения объемов предполагаемого комплектования библиотеки.

6. Технологии передачи рекомендаций на объекты из других систем: электронные библиотеки; интернет-магазины, образовательные стандарты, списки победителей конкурсов и пр.

7. Использование сетевой модели баз данных для хранения и пересчета множества взаимных коэффициентов для связей пользователей между собой, а также для объединения их в группы «по интересам».

Выводы

Развитие современных технологий получения данных заставляет библиотеки искать новые под-

ходы для определения значимости в глазах пользователя. Попытки следовать только лишь в одном направлении – мемориальном – в настоящее время показали свою несостоятельность по причине уменьшения стоимости носителей информации и технологий хранения. Почти забытая «библиотечная технология» – рекомендовать непросвещенным пользователям новые книги (ссылки, файлы и пр.) с использованием новых информационных технологий позволяет создать новые возможности для пользователей библиотек и потенциальных читателей.

Список литературы

1. Долгин А. Экономика символического обмена. – М. : Инфра-М, 2006. – 632 с.
2. Evaluating collaborative filtering recommender systems / J. L. Herlocker [et al] // ACM Transactions on Inform/ Systems. – 2004. – Vol. 22, N 1. – P. 5–53.
3. Herlocker J. L., Konstan J. A. Content-Independent Task-Focused Recommendation // IEEE Internet Computing. – 2001. – Vol. 5, N 6. – P. 40–47.
4. Collaborative recommendation: A robustness analysis / M. O'Mahony [et al] // ACM Transactions on Internet Technology. – 2004. – Vol. 4, N 4. – P. 344–377.
5. Adomavicius G., Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions // IEEE Transactions on Knowledge and Data Engineering. – 2005. – Vol. 17, N 6.

Материал поступил в редакцию 20.11.2008 г.

Сведения об авторе: *Карауш Александр Сергеевич* – кандидат технических наук, директор, тел.: (382-2) 56-26-82, e-mail: ask@library.tomsk.ru