

Таблица 1 – Результаты проверки гипотез о сходимости процесса

№	Диапазон изменения параметров	Максимальное значение $n$ , при котором процесс сходится
1	(0; 0,1]	Сходится всегда
2	[0,1; 0,2]	25
3	[0,2; 0,3]	15
4	[0,3; 0,4]	9
5	[0,4; 0,5]	7
6	[0,5; 0,6]	5
7	[0,6; 0,7]	4
8	[0,7; 0,8]	3
9	[0,8; 0,9]	3
10	[0,9; 1]	2

Перспективными представляются исследования сходимости возвратных последовательностей в случае, когда агентами значения параметров выбираются из разных диапазонов.

### Библиографический список

1. Алгазин Г.И., Алгазина Д.Г. Коллективное поведение в модели Штакельберга в условиях неполной информации // Автоматика и телемеханика. – 2017. – № 9. – С. 91–105.
2. Маркушевич А.И. Возвратные последовательности. Популярные лекции по математике. – М.: Наука, 1950. – 52 с.
3. Глухов М.М., Елизаров В.П., Нечаев А.А. Алгебра. Т.2. – М.: Гелиос АРВ, 2003. – 416 с.

## УДК 004.93

### Основные принципы создания рекомендательных систем

*П.Е. Анафия<sup>1,2</sup>, Г.И. Алгазин<sup>1</sup>, А.С. Тлебалдинова<sup>2</sup>*

*<sup>1</sup>АлтГУ, г. Барнаул; <sup>2</sup>ВКГУ им. С.Аманжолова,*

*г. Усть-Каменогорск*

В этой статье описываются рекомендательные системы, реализованные в них алгоритмы и проблемы при разработке.

Рекомендательная система – это система, которая стремится прогнозировать или фильтровать предпочтения в соответствии с выбором пользователя. Рекомендательные системы используются в различных областях, включая фильмы, музыку, новости, книги, научные статьи, поисковые запросы, социальные теги и продукты в целом. Рекомендательные системы изменили способы взаимодействия неодушевленных вебсайтов со своими пользователями. Вместо предоставления статической информации, когда пользователи ищут и, возможно, покупают

продукты, рекомендательные системы увеличивают степень интерактивности для расширения предоставляемых пользователю возможностей. Рекомендательные системы формируют рекомендации независимо для каждого конкретного пользователя на основе его прошлых покупок и поисков, а также на основе поведения других пользователей.

**Базовые подходы.** В большинстве рекомендательных систем применяется один из двух базовых подходов: коллаборативная фильтрация (collaborative filtering) и контентная фильтрация (content-based filtering). Существуют также и другие подходы (в том числе гибридные) [1].

*Коллаборативная фильтрация* вырабатывает рекомендации, основанные на модели предшествующего поведения пользователя. Эта модель может быть построена исключительно на основе поведения данного пользователя или, что более эффективно, с учетом поведения других пользователей со сходными характеристиками. В тех случаях, когда коллаборативная фильтрация принимает во внимание поведение других пользователей, она использует знание о группе (group knowledge) для выработки рекомендаций на основе подобия пользователей. По существу рекомендации базируются на автоматическом сотрудничестве множества пользователей и на выделении (методом фильтрации) тех пользователей, которые демонстрируют схожие предпочтения или шаблоны поведения [1].

*Контентная фильтрация* формирует рекомендацию на основе поведения пользователя. Например, этот подход может использовать ретроспективную информацию о просмотрах (какие блоги читает пользователь и характеристики этих блогов). Если какой-либо пользователь обычно читает статьи о Linux или регулярно оставляет комментарии в блогах по проектированию программного обеспечения, то контентная фильтрация может использовать эту ретроспективную информацию для выявления подобного контента и предложения такого контента в качестве рекомендованного для этого пользователя (статьи в блогах по Linux или в других блогах по проектированию программного обеспечения). Этот контент может быть определен в ручном режиме или извлечен автоматически на основе других методов подобия [1,3].

В обоих случаях этот механизм рекомендаций имеет два шага:

1. Узнайте, сколько пользователей / элементов в базе данных похожи на данного пользователя / элемента.
2. Оцените других пользователей / элементов, чтобы предсказать, какую оценку вы бы дали пользователю этого продукта, учитывая общий вес пользователей/элементов, которые больше похожи на этот.

**Алгоритмы, используемые рекомендательными системами.** Известный подход, примененный в конкурсе [Netflix prize](#), наглядно демонстрирует, что в рекомендательных механизмах могут быть использованы самые различные алгоритмы. Получаемые результаты могут различаться в зависимости от проблемы, для решения которой спроектирован конкретный алгоритм, и от отношений, которые присутствуют в данных. Многие из этих алгоритмов пришли из области машинного обучения (которая, в свою очередь, представляет собой подобласть искусственного интеллекта), которая занимается алгоритмами для обучения, прогнозирования и принятия решений [1, 2].

*Корреляция Пирсона.* Сходство между двумя пользователями (и их атрибутами, такими как статьи, прочитанные в коллекции блогов) может быть точно вычислено с помощью т. н. *корреляции Пирсона*. Этот алгоритм измеряет линейную зависимость между двумя переменными (или пользователями) как функцию их атрибутов. Однако он не вычисляет эту меру по всей совокупности пользователей. Эту совокупность необходимо предварительно отфильтровать до *близких элементов* на основе высокоуровневых показателей сходства, таких как чтение схожих блогов. Корреляция Пирсона широко применяемая в исследовательской деятельности, является весьма популярным алгоритмом в сфере коллаборативной фильтрации [3].

*Алгоритмы кластеризации* – это разновидность т. н. «спонтанного обучения» (unsupervised learning), позволяющая выявить структуру в рядах на первый взгляд случайных (или немаркированных) данных. В общем случае такой алгоритм базируется на выявлении сходства между элементами (например, между читателями блога) посредством вычисления их расстояния от других элементов в пространстве признаков (feature space) (признаком в пространстве признаков может, например, быть количество прочитанных статей в наборе блогов). Количество независимых признаков определяет размерность пространства признаков. Если элементы «близки» друг к другу, то их можно объединить в один кластер [3].

Существует множество алгоритмов кластеризации. Самым простым из них является алгоритм *k-средних* (*k-means*), который разделяет элементы на *k* кластеров. Первоначально элементы распределяются по этим кластерам в произвольном порядке. Затем для каждого кластера вычисляется *центр масс* (или просто *центр*) как функция его членов. После этого проверяется расстояние каждого члена кластера от центра этого кластера. Если по результатам этой проверки член оказывается ближе к другому кластеру, то он перемещается в этот кластер. После

проверки всех расстояний для всех членов центры кластеров вычисляются заново. При достижении стабильного состояния (в процессе очередной итерации члены не перемещались) набор считается кластеризованным надлежащим образом, и алгоритм останавливается [3].

Вычисление расстояния между двумя объектами может быть трудным для визуализации. Один из распространенных методов решения этой задачи состоит в том, чтобы рассматривать каждый член кластера как многомерный вектор и вычислять для него т.н. евклидово расстояние.

К другим разновидностям кластеризации, относятся теория адаптивного резонанса (Adaptive Resonance Theory), нечеткая кластеризация методом С-средних (Fuzzy C-means), вероятностная кластеризация с помощью EM-алгоритма (Expectation-Maximization) и т. д.

**Проблемы рекомендательных систем.** Возможности сбора данных, которые предоставляет Интернет, существенно упростили использование «мудрости толпы» с помощью коллаборативной фильтрации. С другой стороны, огромное количество доступных данных усложняет реализацию этой возможности. К примеру, поведение некоторых пользователей вполне поддается моделированию, однако другие пользователи не демонстрируют типичного поведения. Наличие таких пользователей может приводить к смещению результатов рекомендательной системы и к снижению ее эффективности. Кроме того, пользователи могут задействовать рекомендательную систему для повышения предпочтительности одного продукта относительно другого продукта – например, посредством отправки позитивных отзывов об одном продукте и негативных отзывов о его конкурентах. Хорошая рекомендательная система обязана справляться с этими проблемами.

Еще одна проблема, свойственная крупным рекомендательным системам, связана с масштабируемостью. Традиционные алгоритмы хорошо работают со сравнительно небольшими объемами данных, однако с ростом этих наборов получение результатов на прежнем уровне качества при помощи традиционных алгоритмов может стать проблематичным. В случае оффлайн-обработки это может не составлять большой проблемы, однако для сценариев реального времени необходимы более специализированные подходы.

И, наконец, определенные проблемы порождает необходимость соблюдения конфиденциальности. Рекомендательные алгоритмы способны распознавать такие закономерности, о существовании которых люди могут даже не подозревать. Недавно пример такой ситуации имел место в крупной компании, которая смогла вычислять индекс прогнозирования беременности на основе покупательских предпочтений. После получения целевых рекламных объявлений отец дочери-подростка с

удивлением узнал о ее беременности. Прогнозирующая система этой компании оказалась настолько точной, что смогла предсказать ожидаемую дату родов у будущей мамы на основе приобретаемых ею товаров [4].

Сегодня рекомендательные механизмы активно применяются на большинстве популярных веб-сайтов социальной и коммерческой направленности. Они приносят огромную пользу владельцам этих сайтов и их пользователям, однако их использование имеет и определенные оборотные стороны. В этой статье были разьяснены некоторые концепции, на которых основаны рекомендательные системы, и описаны алгоритмы, которые реализуют эти концепции.

### **Библиографический список**

1. Луис Педро Коэльо, Вилли Ричарт. Построение систем машинного обучения на языке Python издание / пер. с англ. Слинкин А.А. –М.: ДМК Пресс, 2016. – 302 с.

2. Лутц М. Изучаем Python, 4-е издание / пер. с англ. – СПб.: Символ-Плюс, 2011. – 1280 с.

3. Федоровский А. Н., Логачева В. К. Архитектура рекомендательной системы, работающей на основе неявных пользовательских оценок // Электронные библиотеки: перспективные методы и технологии, электронные коллекции // Труды XIII Всероссийской научной конференции RCDL'2011. – Воронеж: Воронежский госуниверситет, 2011 – С. 76–82.

4. Rokach L., Shapira B. [et al.]. Recommender Systems Handbook. – New-York: Springer F. Ricci Science+Bussiness Media, 2011. – 842 p.

5. Jannach D., Zanker M., Felfernig A. [et al.]. Recommender systems: an introduction. – New-York: Cambridge University Press, 2011. – 352 p.

**УДК 339.138**

## **Информационные технологии в рекомендательном маркетинге**

***Р.В. Гончаров, А.Ю. Юдинцев***

*АлтГУ, г. Барнаул*

Появление компьютерных технологий и интернет-сетей приводит к переменам, проникающим во все сферы общества, меняются условия рыночной среды, происходит ускорение экономических процессов. Это предъявляет к управлению компанией новые требования, заставляет пересматривать стратегические ориентиры компаний. [1, с.3]

На сегодняшний день владельцы бизнеса, особенно малого бизнеса, имеют ряд ключевых проблем: