

А.В. ПОНОМАРЕВ
**ОБЗОР МЕТОДОВ УЧЕТА КОНТЕКСТА
В СИСТЕМАХ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ**

Пономарев А.В. Обзор методов учета контекста в системах коллаборативной фильтрации.

Аннотация. По мнению многих исследователей, одним из наиболее действенных способов повышения качества рекомендующих систем является использование этими системами информации о текущем контексте. В статье произведен обзор основных методов использования информации о контексте в системах коллаборативной фильтрации. Особое внимание уделено разновидностям метода предварительной контекстной фильтрации и метода разложения матрицы предпочтений в связи с их перспективностью и широким распространением.

Ключевые слова: рекомендующие системы, коллаборативная фильтрация, контекст, персонализация.

Ponomarev A. V. Incorporating context into collaborative filtering systems: a survey.

Abstract. It is recognized that incorporating context information into recommender systems is one of the most effective ways to increase their quality and predictive abilities. The paper surveys primary methods of enhancing collaborative filtering systems by taking actual context information into account. The focus is mostly on different flavours of contextual pre-filtering and matrix factorization approaches which are the most popular and promising.

Keywords: recommender systems, collaborative filtering, context, personalization.

1. Введение. Рекомендующие системы в настоящее время получают все более широкое распространение. С одной стороны, это связано с тем, что они позволяют в какой-то степени снизить сложность принятия решений в условиях взрывного роста объемов информации, с которой сталкивается современный человек. С другой — их развитие подкрепляется возникновением новых приложений, основанных на использовании Интернет, мобильных вычислений и других информационных технологий, обеспечивающих быстрое увеличение той фактографической базы, на которой строится работа рекомендующих систем.

Рекомендующие системы обычно подразделяются на три вида [5]: а) *контентные* (content-based), основанные на схожести существенных характеристик объектов; б) *коллаборативной фильтрации* (collaborative filtering), основанные на схожести предпочтений пользователей; в) *гибридные* (hybrid), комбинирующие различные черты контентных систем и систем коллаборативной фильтрации.

Некоторые исследователи выделяют, кроме того, рекомендуемые системы, основанные на знаниях (knowledge-based) [10]. Такие системы построены на совершенно иных принципах, по сравнению с традиционными рекомендуемыми системами, а их разработка связана с решением совершенно специфических проблем. И хотя использование контекста в системах, основанных на знаниях, происходит зачастую совершенно естественно, эти вопросы находятся за рамками настоящей статьи, а заинтересованный читатель может обратиться к работам [13, 14].

Несмотря на то, что традиционные системы (в особенности, системы коллаборативной фильтрации) получили широкое распространение и в настоящее время разработаны эффективные алгоритмы, позволяющие получать достаточно точные рекомендации, некоторые исследователи [26] констатируют наличие кризиса в области создания таких систем. Этот кризис связывается с исчерпанием возможностей подхода, в соответствии с которым рекомендации формируются исключительно на основе зафиксированных ранее оценок и предпочтений пользователей. Одним из путей выхода из кризиса видится расширение перечня используемой при формировании рекомендаций информации, в частности, использование контекста — как того, в котором происходит фиксация предпочтений пользователя, так и того, в котором происходит формирование рекомендаций [1, 3, 5, 26].

Для многих приложений важным видом информации о контексте являются время и географическое положение. Повсеместное распространение мобильных устройств с модулями позиционирования в пространстве (GPS, ГЛОНАСС) делает этот вид информации о контексте легко доступным и открывает новые перспективы для создания интеллектуальных персонализируемых проактивных окружений, важным компонентом которых являются рекомендуемые системы [2, 16].

Использование информации о контексте в рекомендуемых системах связано с проработкой множества аспектов, среди которых: выбор тех составляющих контекста, которые могут оказывать влияние на качество рекомендаций, методы сбора значений, составляющих контекст задачи, алгоритмы формирования перечня рекомендаций с учетом контекста, архитектуры построения контекстно-зависимых систем. В статье рассматривается только один из перечисленных аспектов, а именно — алгоритмы форми-

рования перечня рекомендаций.

Следует заметить, что для ряда терминов в русскоязычной научной литературе не сложилось канонического или хотя бы общепринятого перевода. В этих случаях, во избежание неточности, наряду с русскоязычным вариантом термина, будет приводиться и его принятое англоязычное написание.

2. Системы коллаборативной фильтрации. Традиционная рекомендующая система имеет дело с двумя видами сущностей: пользователь и объект. Здесь пользователь — это получатель рекомендаций и источник данных о предпочтениях, а объектом, в зависимости от области применения системы, может быть товар, информационный ресурс, место; то есть, объект это то, что предлагается пользователю в качестве рекомендации. Системы коллаборативной фильтрации основываются на следующем наблюдении: пользователи, демонстрирующие схожие предпочтения в прошлом, скорее всего, будут демонстрировать схожие предпочтения и в будущем. Такие системы собирают сведения о предпочтениях пользователей в матрицу Пользователь–Объект, количество строк которой равно количеству пользователей системы, количество столбцов равно количеству объектов, а в ячейки помещаются данные об оценке соответствующего объекта пользователем. Далее, основываясь на схожести строк, система пытается предсказать отсутствующие оценки пользователя и формирует перечень рекомендаций с учетом предсказанных оценок.

Чаще всего в матрицу попадают оценки, предоставленные пользователем явно. Наиболее распространенными системами оценок являются: унарная («нравится»), бинарная («нравится»/«не нравится»), пяти- и десятибалльные. Впрочем, матрица Пользователь–Объект не обязательно должна заполняться явно. При ее заполнении, помимо оценки, может учитываться время взаимодействия пользователя с объектом, дополнительные действия (покупка, создание комментария, добавление в «Избранное») и т.п. В конечном итоге (возможно, после дополнительной обработки) содержимое заполненных ячеек этой матрицы должно соотноситься с тем, насколько положительно (или отрицательно) пользователь оценивает объект в рамках выбранной авторами рекомендующей системы модели учета предпочтений.

В дальнейшем, матрицу Пользователь–Объект с оценками будем называть *матрицей предпочтений*. В некоторых случаях ока-

зывается удобным говорить не о матрице, а о множестве троек — тернарном отношении — (Пользователь, Объект, Оценка), где каждой заполненной ячейке матрицы предпочтений соответствует ровно одна тройка.

3. Контекст в рекомендующих системах. Во многих приложениях условия, в которых пользователь оценивает объект, существенным образом влияют на оценку. Подобные условия в исследованиях по рекомендующим системам получили название *контекста*. Этот термин является очень нагруженным и в различных областях знаний обладает своими специфическими смысловыми оттенками. обстоятельный анализ его использования в области рекомендующих систем и связанных с ней областях знаний приводится, например, в [6].

В [26] сформулированы основные виды контекста, актуальные для рекомендующих систем (в первую очередь, мобильных):

- Физический контекст (время, положение, вид деятельности пользователя, погода, освещенность и т.п.).
- Социальный контекст (наличие и роль окружающих людей).
- Контекст устройства (вид и характеристики устройства, с которого осуществляется доступ к информации).
- Модальный контекст (настроение пользователя, цель, опыт, когнитивные способности).

В большей части работ, вошедших в обзор, под контекстом понимаются атрибуты, так или иначе описывающие *ситуацию*, в которой пользователь оценил объект или получает рекомендации. То есть, будучи смоделированными с помощью UML или другого языка объектного моделирования, эти атрибуты стали бы атрибутами ассоциации между классами Пользователь и Объект или, если воспользоваться многомерной моделью данных, схожей с OLAP, то отдельным измерением гиперкуба. Однако есть и работы [23, 27], где под контекстом понимаются атрибуты самих объектов, пользователей или ассоциаций Объект-Объект, Пользователь-Пользователь, то есть под контекстом *в широком смысле* может пониматься любая информация, выходящая за рамки традиционных для систем коллаборативной фильтрации троек (Пользователь, Объект, Оценка).

Достаточно необычное значение термину «контекст» придается в работе [29]. Здесь контекст связывается только с фактом оценивания объекта и под этим контекстом понимается набор тех объектов, которым данный объект был предпочтен.

Если в случае традиционных систем коллаборативной фильтрации исходными данными были либо тернарное отношение (Пользователь, Объект, Оценка), либо двумерная матрица Пользователь–Объект, содержащая оценки, то исходные данные контекстно-зависимых систем удобно представлять в виде либо n -арного отношения (Пользователь, Объект, K_1, \dots, K_k , Оценка), либо в виде многомерного гиперкуба или тензора. Здесь K_i — контекстные факторы. Часто применительно к контекстно-зависимой рекомендации применяют термин многомерная рекомендация, противопоставляя ее традиционной двумерной рекомендации.

В области использования контекста в рекомендующих системах наметилось три подхода [3]: а) *предварительная контекстная фильтрация* (contextual pre-filtering); б) *контекстная постфильтрация* (contextual post-filtering); в) *моделирование контекста* (contextual modeling). В оставшейся части статьи рассмотрены основные приемы, применяемые в каждом из подходов.

4. Контекстная фильтрация. Общей особенностью методов контекстной фильтрации (то есть, предварительной контекстной фильтрации и контекстной постфильтрации) является то, что сам алгоритм обработки матрицы предпочтений может быть таким же, как и в традиционных рекомендующих системах, а учет контекста осуществляется путем трансформации исходных или выходных данных этого алгоритма. В случае предварительной контекстной фильтрации данные, не относящиеся к текущему контексту, отбрасываются перед применением алгоритма формирования рекомендаций, а в случае контекстной постфильтрации итоговый перечень рекомендаций упорядочивается или фильтруется с учетом текущего контекста.

Существенным преимуществом методов контекстной фильтрации является возможность непосредственного использования многочисленных алгоритмов и программных решений, разработанных для двумерной рекомендации [3].

Предварительная контекстная фильтрация При предварительной контекстной фильтрации перед применением алгоритма формирования рекомендаций из исходной матрицы (или ба-

зы) предпочтений пользователей отбрасываются оценки, которые не имеют отношения к текущему контексту. Например, если в качестве контекста в системе рекомендации развлечений используются погодные условия, то при получении рекомендации в пасмурный день все оценки пользователей, полученные в солнечные дни, будут отброшены. Формально, пусть U – множество пользователей, I – множество объектов, C – множество возможных контекстов, R – множество значений шкалы предпочтений (оценок), $H \subset U \times I \times C \times R$ – все известные оценки пользователей с учетом того контекста, в котором они были получены, $c^* \in C$ – текущий контекст, а $Fr_{U,I,R}^D(u, i)$ – некоторая функция предсказания неизвестной оценки на основе известных оценок, задаваемых множеством $D \subset U \times I \times R$. Если рассматривать множества H и D как отношения реляционной алгебры, то функцию предсказания неизвестной оценки с учетом контекста можно определить следующим образом [3]:

$$\forall(u, i, c) \in U \times I \times C, Fr_{U,I,C,R}^H(u, i, c) = Fr_{U,I,R}^{H[c^*]}(u, i). \quad (1)$$

Здесь $[c^*]$ обозначает контекстный фильтр, а именно:

$$H[c^*] = \pi_{U,I,R}(\sigma_{C=c^*}(H)),$$

где π – реляционная операция проекции, а σ – реляционная операция выборки.

Одной из важнейших проблем традиционных систем коллаборативной фильтрации является проблема разреженности таблицы предпочтений, и, как следствие, сложности с оценкой степени сходства между предпочтениями пользователей и ухудшение качества генерируемых рекомендаций. Предварительная контекстная фильтрация еще более обостряет эту проблему, поскольку после ее проведения исходная матрица предпочтений теряет часть значений. Чем точнее описан контекст, тем строже условие в контекстном фильтре, тем меньше остается после контекстной фильтрации данных и тем менее надежными являются результаты применяемого к ним алгоритма коллаборативной фильтрации. В конечном счете, основная задача, решаемая методами предварительной контекстной фильтрации, сводится к тому, чтобы учесть текущий контекст задачи, но при этом не допустить излишнего прореживания матрицы предпочтений.

Метод обобщения контекста (context generalization) предложен в работе [4]. Прежде чем приступить к его рассмотрению, заметим, что во многих случаях контекстные атрибуты допускают построение одной или нескольких иерархий обобщения. Например, атрибут Время может иметь точное значение, однозначно идентифицирующее некоторый момент времени, а может быть обобщен до времени суток, дня недели, месяца и т.п. Для обозначения того, что c' является обобщением c , будем использовать запись $c \rightarrow c'$. Можно ввести определение обобщения и для составного контекста (контекста, описываемого несколькими атрибутами). А именно, контекст $c'(c'_1, \dots, c'_k)$ является обобщением контекста $c(c_1, \dots, c_k)$, тогда и только тогда, когда $c_i \rightarrow c'_i$ для каждого $i \in 1, \dots, k$.

В соответствии с методом обобщения контекста при осуществлении предварительной фильтрации предлагается использовать не точное значение текущего контекста, а одно из его обобщений. В [4] такие обобщения названы контекстными сегментами (contextual segments). Если c^* - текущий контекст, а $c^{*'}$ - текущий контекстный сегмент, то выражение для функции предсказания неизвестной оценки с учетом контекста из (1) трансформируется в $Fr_{U,I,R}^{H[c^{*'}]}(u, i)$, где $H[c^{*'}] = \gamma_{U,I,AGGR(R)}(\sigma_{C \rightarrow c^{*'}}(H))$. Запись $AGGR(R)$ означает, что все оценки, относящиеся к актуальному контекстному сегменту, должны быть каким-то образом агрегированы (обычно, усреднены).

Как правило, для заданного контекста существует несколько возможных обобщений. Для выбора обобщения в [4] предлагается два способа: ручной и автоматизированный. При ручном способе в рекомендующую систему экспертами закладываются предпочтительные обобщения для каждого контекстного атрибута. Автоматизированный же способ предполагает проведение эмпирической оценки предиктивных характеристик каждого из обобщений с последующим выбором наилучшего. Полный перебор возможных обобщений при этом может оказаться слишком дорогим в вычислительном плане, поэтому применяются различные эвристики.

В работе [12] предлагается другой способ формирования фильтра. А именно, вместо обобщения текущего контекста по контекстной иерархии авторы вводят понятие семантически-близких контекстов и подают на вход алгоритму двумерной коллаборативной фильтрации не только данные, относящиеся к текущему контексту, но и относящиеся ко всем семантически-близким контекстам. Под

семантически-близкими они понимают такие контексты, которые оказывают схожий эффект на предпочтения пользователей. Таким образом, выражение для функции предсказания неизвестной оценки с учетом контекста из (1) трансформируется в $Fr_{U,I,R}^H[c^*](u, i)$, где $H[c^*] = \gamma_{U,I,AGGR(R)}(\sigma_{sim(C,c^*) \geq \epsilon_{tr}}(H))$, sim – мера семантической близости контекстов, а ϵ_{tr} – необходимый порог семантической близости для включения в выборку.

Для оценки семантической близости контекстов авторы предлагают подход, напоминающий латентный семантический анализ (Latent Semantic Analysis). Подход включает три этапа:

1. Создается матрица «контекст-объект», в ячейки которой записываются числа, выражающие степень влияния определенного сочетания контекстных параметров на оценку объекта. При формировании этой матрицы используется базовая модель предсказания оценки без учета контекста, и в матрицу записываются значения, характеризующие расхождение между известной оценкой с учетом контекста и оценкой, вычисленной с помощью базовой модели.
2. Размерность матрицы «контекст-объект» понижается с помощью сингулярного разложения.
3. Непосредственно для оценки близости используется косинусная мера между семантическими векторами в пространстве латентных факторов.

В работах [8, 9, 11] для осуществления предварительной фильтрации предлагается метод разделения объектов (item-splitting). Этот метод в значительной мере отличается от уже описанных методов предварительной фильтрации тем, что как таковой фильтрации в нем не производится, однако авторы сами классифицируют его как разновидность подхода предварительной фильтрации, поскольку суть этого метода сводится к трансформации исходного набора данных и последующему применению двумерного алгоритма формирования рекомендаций.

Предлагаемая авторами трансформация исходного набора данных заключается в том, что те объекты, оценки которых демонстрируют существенную зависимость от контекста, разделяются на фиктивные объекты, связанные с соответствующими значениями контекстных атрибутов. Например, если в системе рекоменда-

ций вариантов отдыха оказывается, что оценки некоторого курорта S зависят от сезона, то вместо одного столбца, соответствующего этому курорту, в таблице предпочтений появится два: S зимой и S летом. При разделении используется подход, очень напоминающий построение деревьев решений (например, [20]). А именно, для различных способов разделения множества оценок объекта на два множества в зависимости от значения какого-то контекстного атрибута, вычисляется значение некоторого критерия, характеризующего неравномерность оценок в двух результирующих подмножествах. Выбирается разделение, дающее наибольшую неравномерность. Авторы экспериментируют с различными критериями оценки неравномерности — t -тестом Стьюдента, z -тестом, мерой Кульбака–Лейблера.

Контекстная постфильтрация В методах контекстной постфильтрации сначала используется какой-либо традиционный (двумерный) алгоритм формирования рекомендаций, а затем полученный перечень рекомендаций трансформируется в зависимости от текущего контекста. Трансформация может быть двоякой [3]:

- удаление рекомендаций, не согласующихся с текущим контекстом;
- переупорядочение рекомендаций в зависимости от текущего контекста.

Примером применения контекстной постфильтрации является работа [22]. В ней предлагается анализировать данные пользователя, полученные в определенном контексте, и вычислять вероятность $P_k(u, i)$ того, что пользователь u заинтересуется объектом i в контексте k . Вычисляется эта вероятность как доля пользователей, обладающих схожими с u предпочтениями, которые заинтересовались объектом i в контексте k , по отношению к общему количеству таких пользователей. После применения двумерного алгоритма рекомендации, результатом работы которого является некоторое предсказание оценки пользователя, эти предсказания «контекстуализируются» с учетом вычисленных вероятностей. А именно, предлагается два метода такой контекстуализации: удаление тех рекомендаций, для которых вероятность оказывается ниже определенного порога, или домножение предсказываемого рейтин-

га на $P_k(u, i)$ и упорядочение рекомендаций по убыванию этого произведения.

5. Моделирование контекста. Отличительной особенностью подхода, получившего в [3] название моделирование контекста (contextual modeling), является то, что дополнительные параметры, описывающие контекст, используются непосредственно в модели, на которой базируется работа рекомендующей системы. В этом случае при построении контекстно-зависимой рекомендующей системы уже невозможно непосредственно использовать алгоритмы двумерной рекомендации, а необходимо строить специализированную многомерную предиктивную модель для учета требуемых контекстных параметров.

В настоящее время наибольшее распространение [17, 19, 24, 25, 27, 28] получили различные модификации метода разложения матрицы предпочтений, учитывающие контекстные параметры.

Метод разложения матрицы предпочтений Основная идея, лежащая в основе всех разновидностей метода разложения матрицы предпочтений (matrix factorization), заключается в том, что и пользователи, и объекты отображаются в некоторое единое пространство латентных факторов размерности f таким образом, что взаимодействие между пользователями и объектами представляется как скалярное произведение в этом пространстве [18, 19]. Каждому объекту ставится в соответствие вектор $q_i \in R^f$, а каждому пользователю — вектор $p_u \in R^f$. Для заданного объекта i , элементы вектора q_i выражают степень, в которой объект обладает соответствующими факторами. Эти значения могут быть положительными или отрицательными. Аналогично для пользователя u элементы вектора p_u выражают степень заинтересованности пользователя в объектах, для которых ярко выражен соответствующий фактор. Скалярное произведение $q_i^T p_u$ выражает заинтересованность пользователя u в объекте i . Вспомним, что известные оценки пользователей занесены в матрицу предпочтений. Обозначим эту матрицу R , а ее элементы r_{ui} . Тогда значение скалярного произведения $q_i^T p_u$ должно совпадать с r_{ui} для тех элементов матрицы, значение которых известно. Для неизвестных элементов скалярное произведение будет давать некоторую оценку, на основании которой будет приниматься решение о рекомендации объекта пользователю. Таким образом, задача вычисления отсутствующих элементов матрицы предпочтений сводится к поиску таких векторов

q_i и p_u , чтобы их скалярные произведения совпадали с известными элементами матрицы предпочтений. Эта задача тесно связана с задачей о сингулярном разложении матрицы R , но непосредственное применение сингулярного разложения осложняется наличием в матрице неизвестных значений, поэтому в современных системах чаще применяются модели, в явном виде содержащие только известные оценки r_{ui} [19].

Вычисление векторов q_i и p_u сводится к решению следующей оптимизационной задачи:

$$\min_{q^*, p^*} \sum_{(u,i) \in K} (r_{ui} - q_i^T p_u)^2 + \lambda (\|q_i\|_F^2 + \|p_u\|_F^2). \quad (2)$$

Здесь K – множество тех пар (u, i) , для которых известны значения r_{ui} (обучающая выборка), а $\|\cdot\|_F$ – норма Фробениуса. Параметр λ контролирует степень соответствия модели обучающей выборке — чем он больше, тем «грубее» модель описывает обучающую выборку, при маленьких значениях высок риск «сверхподгонки» (overfitting). Конкретные значения параметра λ определяются в результате перекрестной валидации.

Для решения этой задачи оптимизации в рекомендуемых системах чаще всего используется либо стохастический метод градиентного спуска, либо метод чередующихся наименьших квадратов (alternating least squares).

Широкое распространение, которое метод разложения матрицы получил в последние годы, во многом связано с легкостью расширения формулы (2) для учета различных факторов, оказывающих влияние на предпочтения пользователей. Например, в [18] делается предположение, что оценка, которую получает объект от пользователя, помимо выражения того, насколько объект понравился пользователю, содержит 3 вида систематических ошибок: μ — общая ошибка рейтинга (оцениваемая как среднее значение всех оценок), b_u — ошибка, связанная с пользователем (пользователи по-разному оценивают объекты — одни склонны давать более высокие оценки, другие более низкие), b_i — ошибка, связанная с объектом. Тогда оценка рейтинга будет описываться уже формулой $\hat{r}_{ui} = \mu + b_i + b_u + q_i^T p_u$, а оптимизационная задача будет записываться следующим образом:

$$\min_{b^*, q^*, p^*} \sum_{(u,i) \in K} (r_{ui} - \mu - b_i - b_u - q_i^T p_u)^2 + \lambda (\|q_i\|_F^2 + \|p_u\|_F^2 + b_u^2 + b_i^2). \quad (3)$$

Схожий прием используется и для внесения в модель информации о контексте.

В работе [7] предлагается три модификации описанной выше модели для использования контекста, задаваемого значениями номинальных атрибутов. Весь подход назван авторами *контекстно-зависимое разложение матрицы предпочтений* (context-aware matrix factorization, CAMF), а три частные модели получили названия CAMF-C, CAMF-CI и CAMF-CC. В первой, CAMF-C, предполагается, что условия контекста оказывают одинаковое влияние на все оценки вне зависимости от объекта. Соответственно, в этой модели вводится по одному параметру для каждого условия контекста. Эти параметры выражают степень отклонения (систематическую ошибку) оценки пользователей в результате воздействия контекста. В другой модели, CAMF-CI, предполагается, что условия контекста оказывают различное влияние на все объекты. В третьей модели, CAMF-CC, авторы предлагают экспертно группировать объекты и учитывать влияние контекста на группы объектов.

Авторами [7] предлагается следующая модель контекстно-зависимой оценки привлекательности объекта:

$$\hat{r}_{uic_1\dots c_k} = q_i^T p_u + \mu + b_u + \sum_{j=1}^k B_{ijc_j}.$$

Здесь B_{ijc_j} выражают систематическую ошибку оценки привлекательности объекта i , возникающую в силу действия контекстного фактора j , имеющего значение c_j . В модели CAMF-CI все значения B_{ijc_j} могут быть различны, в CAMF-CC для пары объектов i и f , принадлежащих к одной группе, $B_{ijc_j} = B_{fjc_j}$, а в CAMF-C, для любой пары объектов i и f выполняется $B_{ijc_j} = B_{fjc_j}$. Задача минимизации записывается следующим образом:

$$\min_{q^*, p^*, b^*, B^*} \sum_{r \in R} \left[\left(r_{uic_1\dots c_k} - q_i^T p_u - \mu - b_u - \sum_{j=1}^k B_{ijc_j} \right)^2 + \lambda \left(b_u^2 + \|q_i\|_F^2 + \|p_u\|_F^2 + \sum_{j=1}^k \sum_{c_j=1}^{z_j} B_{ijc_j}^2 \right) \right].$$

Здесь R — это множество всех известных оценок с учетом контекста, то есть множество кортежей $r_{uic_1\dots c_k}$, а z_j — количество возможных значений j -того контекстного фактора. Решается эта задача минимизации методом стохастического градиентного спуска.

В работе [27] используется не совсем обычное понятие контекста — под контекстом в ней понимаются отношения схожести (по ключевым словам, настроению) между объектами (фильмами). Предлагается метод, названный авторами *совместное разложение матрицы* (joint matrix factorization). Название отражает основную идею метода, а именно, при решении задачи минимизации используются три матрицы: предпочтений пользователей и схожести объектов по двум параметрам. Предлагается потребовать, чтобы схожие объекты отображались в близкие вектора в пространстве латентных факторов. Для этого вводится функция потерь, которая добавляется в минимизируемое выражение, аналогичное (2). Функция потерь определяется следующим образом:

$$L(q) = \sum_{j=1}^n \sum_{z=1}^n (S_{jz} - q_i^T q_z)^2.$$

Здесь \mathbf{S} — матрица схожести объектов, S_{ij} — элемент матрицы схожести.

В [24, 25] предлагается обобщение методов разложения матрицы, названное автором factorization machines.

В работе [15] предлагается полностью отойти от двумерной матрицы предпочтений и моделировать данные в виде многомерного тензора, два измерения которого соответствуют пользователям и объектам, а остальные — контекстным факторам. Тензор, содержащий предпочтения пользователей, предлагается раскладывать на множители по аналогии с матрицей. Предсказание неизвестных оценок осуществляется через нахождение элементов произведения множителей. Метод разложения тензора, предложенный в работе, опирается на описанный выше метод разложения матрицы: вводится функция отклонения между предсказываемым значением оценки и известным, а затем методом стохастического градиентного спуска решается задача по минимизации этой функции с учетом регуляризации (во избежание «сверхподгонки»).

В [16] также используется факторизация 6-мерного тензора в

системе рекомендации мобильных приложений. В системе принимаются во внимание следующие факторы контекста: статус движения, местоположение (дом, офис, другое), время суток, день недели.

Многомерные модели машинного обучения Некоторые классические модели и алгоритмы машинного обучения, применяющиеся при разработке рекомендующих систем, допускают естественное расширение от двумерного случая к многомерному. В [23], например, предлагаются модификации классических техник — метода kNN и индуктивного логического программирования — для построения контекстно-зависимой системы рекомендации фильмов. Под контекстом в этой работе понимается информация об отношениях между пользователями и дополнительная информация о фильмах (режиссер, актеры и т.п.). kNN (k-Nearest Neighbors) — популярная эвристика, заключающаяся в том, что предсказание оценки, которую получит объект i от пользователя u , производится на основе комплексирования оценок объекта i , полученных от пользователей, ближайших — в смысле какой-то меры — к пользователю u . Ключевыми вопросами при разработке методов, относящихся к этой категории, является проблемно-ориентированное понятие близости пользователей и способ комплексирования оценок.

В [23] предлагается несколько мер близости, учитывающих не только то, насколько схожим образом пользователи оценивали объекты, но и схожесть самих объектов по жанру, составу актеров, а также факт знакомства пользователей (отношение «дружбы» в терминах социальной сети). То есть учет контекста здесь производится именно в мере близости между объектами.

В [21] предлагается контекстно-зависимый метод опорных векторов (C-SVM). Известные данные о предпочтениях рассматриваются как точки в многомерном пространстве (контекстные измерения моделируются наряду традиционной парой Пользователь-Объект) и в этом пространстве с помощью метода опорных векторов строится классификатор, различающий объекты (ситуации), вызвавшие положительную реакцию пользователя и отрицательную. Практическое сравнение, проведенное авторами работы, показало, что подобный классификатор, учитывающий информацию о контексте, по своим предиктивным характеристикам значительно превзошел классификатор, не учитывающий контекст.

6. Заключение. В статье рассмотрены основные методы учета контекста в системах коллаборативной фильтрации. В соответствии с принятой классификацией, эти методы разделены на две группы: методы контекстной фильтрации, в которых контекстные фильтры накладываются на исходные данные или результат работы двумерного алгоритма рекомендаций, и методы моделирования контекста, в которых контекстные факторы непосредственно учитываются алгоритмом. Наибольшее распространение в настоящее время получили методы предварительной контекстной фильтрации и различные модификации метода разложения матрицы предпочтений, учитывающие контекстные факторы. Преимущества методов предварительной фильтрации связаны с тем, что они позволяют в явном виде моделировать и обрабатывать данные о контексте, используя, в том числе знания экспертов, применять существующие методы и алгоритмы двумерной рекомендации, и получать более интерпретируемые результаты. Разновидности метода разложения матрицы предпочтений, с другой стороны, зачастую позволяют добиться более высокой точности рекомендаций.

Литература

1. *Шолов Н.Г.* Групповые рекомендуемые системы для конфигурирования гибких сетевых организаций // Информационно-управляющие системы. 2012. №5. С. 69–74.
2. *Ackerman E., Guizzo E.* 5 Technologies That Will Shape the Web // IEEE Spectrum. June 2011. P. 33–37.
3. *Adomavicius G., Mobasher B., Ricci F., Tuzhilin A.* Context-aware recommender systems // AI Magazine. 2011. 32(3), P. 67–80.
4. *Adomavicius G., Sankaranarayanan R., Sen S., Tuzhilin A.* Incorporating contextual information in recommender systems using a multidimensional approach // ACM Trans. Inf. Syst. January 2005. vol. 23, no. 1. P. 103–145.
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. June 2005. vol. 17, no. 6, P. 734–749.

6. *Adomavicius G., Tuzhilin A.* Context-Aware Recommender Systems // In *Recommender Systems Handbook*, Ricci F., Rokach L., Shapira B., Kantor P.B. (eds.), Springer, 2011. P. 217–256.
7. *Baltrunas L., Ludwig B., Ricci F.* Matrix factorization techniques for context aware recommendation // In *Proceedings of the fifth ACM conference on Recommender systems (RecSys '11)*. P. 301–304.
8. *Baltrunas L., Ricci F.* Context-based splitting of item ratings in collaborative filtering // In *Proceedings of the third ACM conference on Recommender systems (RecSys '09)*, ACM, 2009. P. 245–248.
9. *Baltrunas L., Ricci F.* Context-Dependent Recommendations with Items Splitting // *Proceedings of the 1st Italian Information Retrieval Workshop (IIR'10)*, January 27–28, 2010, Padua, Italy. P. 71–75.
10. *Burke R.* Knowledge-based recommender systems // *Encyclopedia of Library and Information Science*, 2000. 69(32). P. 180–200.
11. *Campos P.G., Cantador I., Díez F.* Exploiting time contexts in collaborative filtering: an item splitting approach // In *Proceedings of the 3rd Workshop on Context-awareness in Retrieval and Recommendation (CaRR '13)*, New York, NY, USA, 2013. P. 3–6.
12. *Codina V., Ricci F., Ceccaroni L.* Semantically-enhanced pre-filtering for context-aware recommender systems // In *Proceedings of the 3rd Workshop on Context-awareness in Retrieval and Recommendation (CaRR '13)*, New York, NY, USA, 2013. P. 15–18.
13. *Felfernig A., Burke R.* Constraint-based recommender systems: technologies and research issues // In: *Proceedings of the 10th International Conference on Electronic Commerce, ICEC'08*, ACM, New York, NY, USA, 2008. P. 1–10.
14. *Jannach D., Kreutler G.* Rapid development of knowledge-based conversational recommender applications with advisor suite // *Journal of Web Engineering*, 2007. vol. 2, no. 6. , P. 165–192.
15. *Karatzoglou A., Amatriain X., Baltrunas L., Oliver N.* Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering // In *Proceedings of the 4th ACM Conference on Recommender Systems (RecSys'10)*, Barcelona, Spain, 2010. P. 79–86.
16. *Karatzoglou A., Baltrunas L., Church K., Böhmer M.* Climbing the app wall: enabling mobile app discovery through context-aware

- recommendations // In Proceedings of the 21st ACM international conference on Information and knowledge management (CIKM '12). P. 2527–2530.
17. *Koren Y.* Factorization meets the neighborhood: a multifaceted collaborative filtering model // In Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2008. P. 426–434.
 18. *Koren Y., Bell R.* Advances in Collaborative Filtering // In Recommender Systems Handbook. Ricci F., Rokach L., Shapira B., Kantor P.B. (eds.) Springer, 2011. P. 145–186.
 19. *Koren Y., Bell R., Volinsky C.* Matrix factorization techniques for recommender systems // IEEE Computer, Aug 2009. vol. 42, no. 8. P. 30–37.
 20. *Murthy S.* Automatic construction of decision trees from data: A Multi-disciplinary survey // Data Mining and Knowledge Discovery, 1997. no. 2, P. 345–389.
 21. *Oku K., Nakajima S., Miyazaki J., Uemura S.* Context-aware svm for context-dependent information recommendation // In Proceedings of the 7th international Conference on Mobile Data Management, 2006. P. 109.
 22. *Panniello U., Tuzhilin A., Gorgoglione M. et al.* Experimental comparison of pre- vs. post-filtering approaches in context-aware recommender systems // In Proceedings of the third ACM conference on Recommender systems (RecSys '09). ACM, New York, NY, USA, P. 265–268.
 23. *Rahmani H., Piccart B., Fierens D., Blockeel H.* Three complementary approaches to context aware movie recommendation // In Proceedings of the Workshop on Context-Aware Movie Recommendation (CAMRa '10). P. 57–60.
 24. *Rendle S.* Factorization Machines with libFM // ACM Trans. on Intell. Syst. Technol., May 2012. vol. 3, no. 3. P. 57:1–57:22.
 25. *Rendle S., Gantner Z., Freudenthaler C., Schmidt-Thieme L.* Fast context-aware recommendations with factorization machines // In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval (SIGIR '11). P. 635–644.

26. *Ricci F.* Contextualizing Useful Recommendations // Personalization — UMAP 2012, Montreal, July 16-20, 2012 [Электронный ресурс] URL: <http://www.inf.unibz.it/ricci/Slides/Context-UMAP-2012-Ricci.pdf> (Дата обращения: 31.05.2013).
27. *Shi Y., Larson M., Hanjalic A.* Mining contextual movie similarity with matrix factorization for context-aware recommendation // ACM Trans. Intell. Syst. Technol., February 2013. vol. 4, no. 1. P. 16:1–16:19.
28. *Wermser H., Rettinger A., Tresp V.* Modeling and Learning Context-Aware Recommendation Scenarios using Tensor Decomposition // International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2011, P. 137–144.
29. *Yang S.-H., Long B., Smola A.J., Zha H., Zheng Z.* Collaborative competitive filtering: learning recommender using context of user choice // SIGIR 2011: P. 295–304.

Пономарев Андрей Васильевич — к.т.н.; старший научный сотрудник лаборатории интегрированных систем автоматизации СПИИРАН. Область научных интересов: рекомендующие системы, системы персонализации, технологии баз данных, дискретная оптимизация. Число научных публикаций — 15. ponomarev@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-8071, факс +7(812)328-4450.

Ponomarev Andrei — Ph.D.; senior researcher, Laboratory of Computer Aided Integrated Systems, SPIIRAS. Research interests: recommender systems, personalization, database technologies, discrete optimization. The number of publications — 15. ponomarev@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-8071, fax +7(812)328-4450.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проекты №№ 13-07-00271-а и 13-07-00039-а), Президиума РАН (проект № 213), Отделения нанотехнологий и информационных технологий РАН (проект № 2.2).

Рекомендовано лабораторией интегрированных систем автоматизации, ведущий лабораторией Смирнов А.В., д.т.н., проф.

Статья поступила в редакцию 19.06.2013.

РЕФЕРАТ

Пономарев А.В. Обзор методов учета контекста в системах коллаборативной фильтрации

Системы коллаборативной фильтрации — это разновидность рекомендующих систем, работа которых основана на следующем наблюдении: пользователи, демонстрирующие схожие предпочтения в прошлом, скорее всего, будут демонстрировать схожие предпочтения и в будущем.

В области использования контекста в рекомендующих системах наметилось три подхода: а) предварительная контекстная фильтрация; б) контекстная постфильтрация; в) моделирование контекста.

Общей особенностью методов контекстной фильтрации (то есть, предварительной контекстной фильтрации и контекстной постфильтрации) является то, что сам алгоритм обработки матрицы предпочтений может быть таким же, как и в традиционных рекомендующих системах, а учет контекста осуществляется путем трансформации исходных или выходных данных этого алгоритма. В случае предварительной контекстной фильтрации данные, не относящиеся к текущему контексту, отбрасываются перед применением алгоритма формирования рекомендаций, а в случае контекстной постфильтрации итоговый перечень рекомендаций упорядочивается или фильтруется с учетом текущего контекста.

В статье рассмотрены следующие методы контекстной фильтрации: метод обобщения контекста, метод семантического расширения предварительного фильтра, метод разделения объектов, а также один из базовых методов контекстной постфильтрации.

Отличительной особенностью подхода, получившего название моделирование контекста, является то, что дополнительные параметры, описывающие контекст, используются непосредственно в модели, на которой базируется работа рекомендующей системы.

В настоящее время наибольшее распространение получили различные модификации метода разложения матрицы предпочтений, учитывающие контекстные параметры. В статье рассмотрено несколько таких модификаций.

Преимущества методов контекстной фильтрации (в особенности, предварительной) связаны с тем, что они позволяют в явном виде моделировать и обрабатывать данные о контексте, используя, в том числе знания экспертов, применять существующие методы и алгоритмы двумерной рекомендации, и получать более интерпретируемые результаты. Разновидности метода разложения матрицы предпочтений, с другой стороны, зачастую позволяют добиться более высокой точности рекомендаций.

SUMMARY

Ponomarev A.V. Incorporating context into collaborative filtering systems: a survey.

Collaborative filtering system is the kind of recommender system, which is based on the observation that users showing similar preferences in the past, most likely, will exhibit similar preferences in the future.

There are three approaches to incorporating context information into recommender systems: a) contextual pre-filtering, b) contextual post-filtering, c) contextual modeling.

A common feature of the contextual filtering methods (i.e., contextual pre-filtering and contextual post-filtering) is that they employ one of the preference matrix processing algorithms that are developed for traditional recommender systems and account for context by transforming either input or output of that algorithm. In the contextual pre-filtering methods data that is not relevant to the current context are filtered out before applying the basic traditional recommender algorithm, while in the case of contextual post-filtering resulting list of recommendations is ordered or filtered based on the current context.

The article describes the following methods of contextual filtering: context generalization method, semantically-enhanced pre-filtering, item-splitting, and the basic contextual post-filtering method.

A distinctive feature of the contextual modeling approach is that the additional parameters describing the context are used directly in the recommender model. Hence systems following this approach can not employ traditional (or two-dimensional) recommender algorithms and new multidimensional algorithms must be developed.

Currently, the most widely used methods taking this approach are different modifications of the matrix factorization with respect to some context variables. The idea of matrix factorization is to project users and items into some latent factor space where relationship between them is represented by cross-product. Projection matrixes are found as the solution of some minimization task which can be extended to contain contextual variables. The paper describes several contextual modifications of matrix factorization.

The advantages of context filtering techniques (in particular, pre-filtering) due to the fact that they allow the explicit modeling and processing context information, using for example expert knowledge, apply the existing techniques and algorithms of the two-dimensional recommendation, and get more interpretable results. Variations of matrix factorization method, on the other hand often allow to achieve higher precision.