

Групповые контекстно-управляемые рекомендуемые системы¹

Аннотация. В статье предложена архитектура и основные модели контекстно-управляемых рекомендуемых систем на основе коллаборативной фильтрации. Выделены основные проблемы, возникающие при создании таких систем, предложены пути их решения. Приведено обоснование преимуществ использования методов предварительной фильтрации для учета контекста. Описаны основные процессы формирования рекомендаций. Применение предлагаемой архитектуры продемонстрировано на примере рекомендуемой системы для мобильного туристического приложения.

Ключевые слова: рекомендуемые системы, коллаборативная фильтрация, контекст, онтология, управление профилями.

Введение

Групповые рекомендуемые (рекомендательные) системы широко используются в Интернете для подбора изделий и услуг индивидуальным пользователям, учитывая их предпочтения и вкусы [1], а также в различных бизнес-приложениях [2, 3]. Определение рекомендаций для групп пользователей усложняется необходимостью учитывать не только индивидуальные интересы, но и искать компромисс между интересами группы пользователей и их индивидуальными интересами [4, 5].

Большинство существующих рекомендуемых систем оперируют в двумерном пространстве «Пользователь-Товар» [6]. Это значит, что они выдают рекомендации, основываясь исключительно на информации о пользователе или о товаре (объекте), и обходят стороной контекстуальную информацию, которая может оказаться первостепенно важной в некоторых приложениях. Например, во многих случаях полезность товара или услуги может зависеть от того, когда происходит потребление (время года, день недели, время суток). Полезность может также зави-

сеть от того, с кем, в какой компании, при каких обстоятельствах потребляется продукт. В таких случаях простая рекомендация продукта клиенту недостаточна, в результате при выработке рекомендации система должна обратиться к дополнительной контекстной информации о времени и обстоятельствах предполагаемого потребления.

Если пользователь (турист) предпочитает посещать выставки живописи эпохи ренессанса, то при выборе им музеев система должна давать предпочтения тем, в которых есть соответствующие выставки.

В предыдущих работах авторов [4, 7-9] были сформулированы требования к групповым рекомендуемым системам. В данной статье предлагается общая архитектура подобных систем (с учетом сформулированных требований) и обсуждаются способы реализации их отдельных составных частей.

1. Групповые рекомендуемые системы

В [10] предложена архитектура групповой рекомендуемой системы (Рис. 1), основанная

¹ Исследование частично поддержано проектами РФФИ № 13-07-00271-а, 13-07-00336-а, 11-07-00045-а, 12-07-00298-а, 13-07-12095-офи, проектом 2.13 программы фундаментальных исследований президиума РАН «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация» и проектом 2.2 программы «Интеллектуальные информационные технологии, системный анализ и автоматизация» ОНИТ РАН.

на трех компонентах: поиск характерных элементов в профилях индивидуальных пользователей, группировка (кластеризация) пользователей на основе их предпочтений [11], разработка окончательных рекомендаций на основе созданных групп пользователей. Разработка алгоритмов кластеризации, способных непрерывно улучшать структуру групп на основе постоянно поступающей информации, может сделать возможным самоорганизацию пользователей в группы [12].

Принято выделять два основных типа рекомендующих систем:

– контентные или «content-based» (рекомендующие системы, основанные на ранее принятых решениях пользователя);

– коллаборативной фильтрации или «collaborative filtering» (рекомендующие системы, основанные на ранее принятых решениях пользователей, чьи интересы и характеристики сходны с рассматриваемым пользователем).

Оба типа рекомендующих систем имеют свои достоинства и недостатки. Ниже приведено краткое описание типов систем.

В контентных рекомендующих системах вывод о полезности объекта для пользователя делается, исходя из полезности, присвоенной

пользователем объектам, сходным с данным объектом. Например, в системе, рекомендующей произведения искусства, для того, чтобы рекомендовать объект пользователю, контентная рекомендующая система пытается найти сходство между объектами, выбранными и высоко оцененными пользователем ранее (например, тот же автор, то же направление или школа). И только объекты, обладающие высокой степенью общности с предпочтениями пользователя, будут рекомендованы.

Контентный подход к рекомендующим системам основан на работах, изучающих методы поиска информации [13, 14] и информационной фильтрации [15]. Однако если система рекомендует объекты, характеристики которых совпадают с содержащимися в профиле пользователя, то пользователь получит рекомендации только о таких объектах, которые сходны с объектами, ранее уже получившими его оценку. Например, в системе, рекомендующей произведения искусства, человек, не имевший дела с творчеством прерафаэлитов, никогда не получит совета ознакомиться с творчеством Д.Г. Россетти. Для решения этой проблемы часто применяется фактор случайности.



Рис. 1. Типовая архитектура групповой рекомендующей системы

В отличие от контентных рекомендуемых систем коллаборативные рекомендуемые системы (или системы коллаборативной фильтрации) пытаются предсказать полезность объекта для отдельного пользователя, исходя из оценок, полученных этим объектом от других пользователей. Более формально полезность объекта для пользователя оценивается исходя из полезностей, приписанных объекту, теми пользователями, которые «похожи» на данного пользователя. Например, в системах, рекомендуемых посещения достопримечательностей, для того, чтобы предлагать объекты пользователю, система пытается найти пользователей, схожих в своем выборе с данным пользователем, и тогда рекомендовать будут только объекты, наиболее часто выбираемые и наиболее высоко оцененные «похожими» пользователями.

Однако, поскольку коллаборативные системы при выработке рекомендаций руководствуются только предпочтениями пользователей, они не могут рекомендовать объект, пока он не получит достаточное количество оценок. Данная проблема решается с использованием гибридных методов.

2. Основные проблемы и пути их решения

Ниже приведены основные проблемы, возникающие при создании и использовании рекомендуемых систем.

Использование контекстной информации при выработке рекомендаций. Рекомендуемые системы должны быть контекстно-ориентированными. Для этого предлагается применение технологии управления контекстом вместе с онтологическим описанием проблемной области. Применение методов управления контекстом позволит выбирать только ту информацию и знания, которые являются наиболее актуальными для конкретной задачи, а также позволит существенно снизить размерность решаемых задач, и, следовательно, упростить их решение.

При использовании контекстной информации условия выработки рекомендаций постоянно меняются, что требует разработки *механизмов самоорганизации групп* [12]. Разработка алгоритмов кластеризации, способных непрерывно совершенствовать составы групп на основе поступающей новой информации, позволит реализовать вышеуказанные механизмы.

Поскольку условия выработки рекомендаций постоянно меняются, необходимы *механизмы, позволяющие наиболее полно и многократно использовать имеющуюся информацию.* Двухуровневое моделирование проблемной области, в отличие от одноуровневого, позволит разделить структурное и контекстное представление проблемной области. Под контекстным представлением проблемной области в данном случае понимается модель последней с актуальными данными, соответствующими конкретной ситуации.

Представление комплексной проблемной области. Описание проблемной области на основе комплекса мультидисциплинарных интеллектуальных моделей, интегрируемых посредством общей онтологии проблемной области, предоставит возможность разбиения комплексной проблемной области на более простые подзадачи, а также возможность использования существующих моделей и онтологий, описывающих различные аспекты рассматриваемой проблемной области.

Объединение мультидисциплинарных моделей. Онтологическое описание предметной области с помощью единого формализма позволит избежать многочисленных трансформаций информации и знаний из одного формализма в другой, а также повторной формализации проблемы при структурном синтезе их облика.

Представление знаний об области применения рекомендуемой системы на основе формализма объектно-ориентированных сетей ограничений [9] позволит интегрировать объектно-ориентированное представление информации и знаний с технологией удовлетворения ограничений.

Интеграция информации, получаемой из различных источников. Использование технологий интеллектуальных пространств вместе со стандартами Интернет-сервисов позволит обеспечить стандартизированный и легко реализуемый обмен информацией и знаниями между различными элементами системы.

3. Методы учета контекста в системах коллаборативной фильтрации

Коллаборативная фильтрация является достаточно универсальным подходом к получению эмпирических данных о предпочтениях пользователей и использованию этих эмпирических данных для формирования рекомендаций. Такая универсальность связана с тем, что для ме-

тодов коллаборативной фильтрации не требуется моделирования самих объектов и их свойств, достаточно лишь данных об отношении пользователей к объектам. При этом объекты могут быть самой различной природы: книги, музыка, новостные заметки, достопримечательности. Конечно, эту универсальность нельзя и преувеличивать, например, коллаборативная фильтрация хорошо показывает себя только применительно к объектам массового спроса и слабо применима к рекомендации редких или конфигурируемых объектов.

В основе системы коллаборативной фильтрации находится матрица «Пользователь-Объект», количество строк которой равно количеству пользователей системы, количество столбцов равно количеству объектов, а в ячейки помещаются данные об оценке соответствующего объекта пользователем. В дальнейшем матрицу «Пользователь-Объект» будем называть *матрицей предпочтений*.

В области использования контекста в системах коллаборативной фильтрации наметилось три подхода [16]: предварительная контекстная фильтрация (contextual pre-filtering); контекстная постфильтрация (contextual post-filtering); моделирование контекста (contextual modeling).

Общей особенностью методов контекстной фильтрации (т.е. 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.

При предварительной контекстной фильтрации перед применением алгоритма формирования рекомендаций из исходной матрицы 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. Однако это усугубляет и без того важную проблему, присущую системам коллаборативной фильтрации – проблему разреженности матрицы предпочтений. 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.

В литературе описано несколько методов, которые могут быть отнесены к методам предварительной фильтрации. В [17] предложен метод обобщения контекста (context generalization), в соответствии с которым при формировании рекомендаций рассматривается не только точное значение контекста, но и возможные его обобщения. В [18] вводится понятие семантически-близких контекстов, и при формировании рекомендаций используются данные, полученные во всех контекстах, близких к текущему. В работах [19, 20] для осуществления предварительной фильтрации предлагается метод разделения объектов (item-splitting).

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 [16]:

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

Существенным преимуществом методов контекстной фильтрации является возможность непосредственного использования многочисленных алгоритмов и программных решений, разработанных для двумерной рекомендации [16].

Отличительной особенностью подхода, получившего в [16] 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

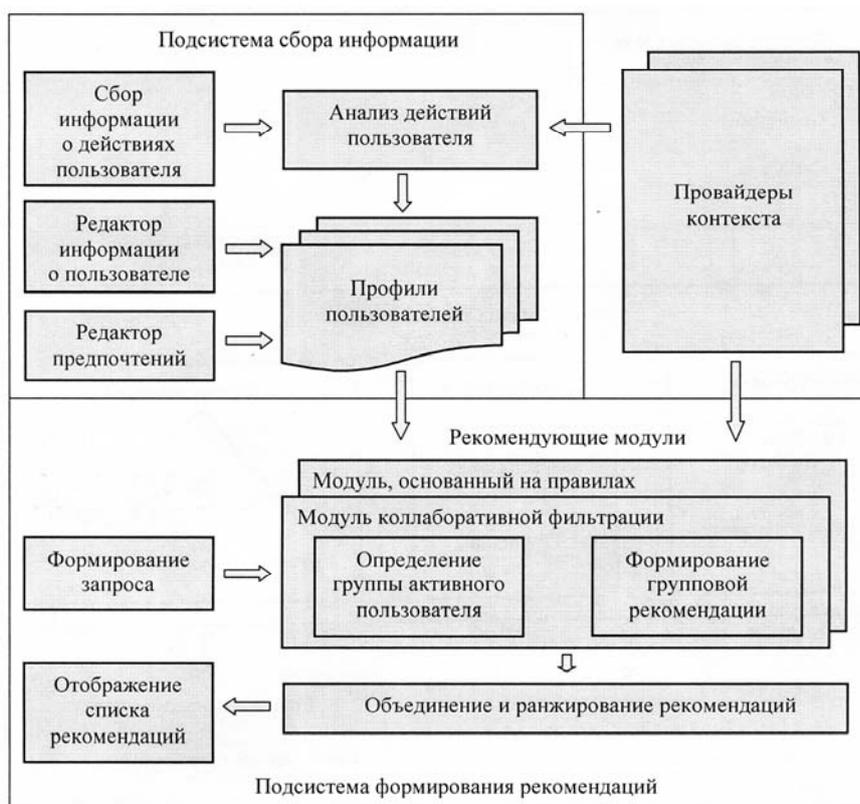


Рис. 2. Уточненная архитектура групповой рекомендующей системы

для учета требуемых контекстных параметров. Среди методов моделирования контекста [21-23] наибольшее распространение в настоящее время получили различные модификации метода разложения матрицы предпочтений, учитывающие контекстные параметры.

Несмотря на то, что методы разложения матрицы предпочтений зачастую позволяют добиться более высокой точности рекомендаций, методы предварительной фильтрации позволяют в явном виде моделировать и обрабатывать данные о контексте, используя знания экспертов, применять существующие методы и алгоритмы двумерной рекомендации и получать более интерпретируемые результаты. Следовательно, методы предварительной фильтрации в целом лучше подходят для реализации функций коллаборативной рекомендации в интеллектуальных системах.

4. Архитектура

Уточненная архитектура групповой рекомендующей системы приведена на Рис. 2. В системе выделяется две подсистемы: сбора информации и

формирования рекомендаций. Задачей подсистемы сбора информации является заполнение профилей пользователя с учетом всех используемых источников информации: явных и неявных предпочтений, текущего контекста и действий пользователя в системе. Следует отметить, что рекомендующая система чаще всего создается не обособленно, а тесно связана с какой-либо другой системой, как правило, системой электронной коммерции (будем называть такую систему «базовой»). В этом случае задача «Сбор информации о действиях пользователя» решается именно в рамках базовой системы, которая должна быть модифицирована с целью предоставления рекомендующей системе необходимой информации о действиях пользователя.

Задачей подсистемы формирования рекомендаций является построение ранжированного перечня рекомендаций на основе известных профилей пользователей, текущего контекста и запроса.

Обмен информацией между компонентами системы осуществляется на основе общей онтологии, т.е. текущий контекст и предпочтения пользователей описываются в терминах единой концептуальной модели предметной области.

Далее будут более подробно рассмотрены основные аспекты и элементы предложенной схемы.

5. Представление контекста

Контекст задачи предлагается описывать контекстами двух типов: абстрактным и оперативным. Абстрактный контекст представляет собой модель задачи, построенную на основании интеграции знаний проблемной области, релевантных для данной задачи. Оперативный контекст является экземпляром абстрактного контекста с присвоенными текущими значениями переменным, входящим в модель задачи [24]. В качестве средства представления знаний используется онтологическая модель, описывающая знания через множества классов, объединяющих типовые объекты реального мира, свойств, характеризующих данные объекты, и отношений, существующих между ними.

В качестве основного метода учета контекста при формировании рекомендаций используется предварительная фильтрация, а именно метод обобщения контекста [17]. Для использования этого метода модель представления контекста (в данном случае, онтология) должна поддерживать возможность обобщения, т.е. все или большая часть контекстных атрибутов, описывающих ситуацию, должны определяться на множествах, для элементов которых заданы иерархии обобщения.

6. Профиль пользователя

Информационная модель профиля пользователя рекомендующей системы представлена на Рис. 3. В профиле предлагается выделить три составляющие: информацию о пользователе, историю взаимодействия пользователя с системой и информацию о предпочтениях. Информация о пользователе задается явным образом и в зависимости от области приложения рекомендующей системы может включать в себя контактные данные, антропометрическую информацию и т.п.

В истории взаимодействия фиксируются все факты оценки объектов пользователем. Важно, что фиксируются не только оценки и контекст, в рамках которого эти оценки были сделаны, но и предпочтения пользователя, действовавшие на момент оценивания. В результате в профиле пользователя накапливаются исторические данные, характеризующие применимость объектов в определенных условиях с учетом определенных целей использования. Именно эти данные обрабатываются алгоритмами коллаборативной фильтрации.

Наконец, предпочтения пользователя, задаваемые в виде ограничений, позволяют фильтровать и упорядочивать перечень формируемых рекомендаций в соответствии с потребностями пользователя.



Рис. 3. Информационная модель профиля пользователя

7. Процессы, связанные с формированием рекомендаций

Анализ действий пользователя представляет собой процесс, в ходе которого информация о первичных, примитивных действиях пользователя системы (в том числе, базовой системы) преобразуется в оценки. В основе этого преобразования могут стоять различные модели. В самом простом случае этот процесс сводится просто к фиксации оценки, явно присвоенной пользователем объекту. В менее тривиальных случаях может учитываться время ознакомления пользователя с описанием объекта, факт покупки (в случае, если подобные возможности поддерживаются напрямую системой), количество возвратов к описанию объекта и т.п. Само построение модели, связывающей действия пользователя относительно некоторого объекта и оценку этого объекта пользователем, является одним из шагов построения рекомендующей системы и зависит как от области применения рекомендующей системы, так и от функциональных возможностей программного комплекса, в составе которого рекомендующая система функционирует.

Фиксация контекста производится с помощью набора провайдеров контекста. Эти провайдеры представляют собой программные (в некоторых случаях программно-аппаратные) модули, формирующие на основе текущего состояния физического или виртуального окружения его описание в виде фрагмента онтологии, т.е. с использованием определенного, единого для всей системы, словаря. Фрагменты контекста, полученные от провайдеров, согласовываются между собой и далее либо сохраняются в профиле пользователя при сборе истории взаимодействий пользователя с системой, либо используются в запросе для формирования рекомендаций. Множество провайдеров контекста и их конкретный состав могут варьироваться в зависимости от назначения рекомендующей системы.

Процесс формирования рекомендаций для пользователя состоит из двух шагов: выяснение группы, к которой относится пользователь, и формирование рекомендаций на основе предпочтений пользователей, входящих в эту группу. Для локализации группы можно воспользоваться одним из двух подходов: либо проведение кластеризации пользователей в зависимости от

предпочтений, либо использование какого-либо метода поиска ближайших соседей (например, метода локально-чувствительного хеширования). Однако подход, связанный с проведением кластеризации, осложняется тем, что для обеспечения учета контекста необходимо одновременно поддерживать несколько разбиений множества пользователей на кластеры (группы) в зависимости от установленного уровня детализации контекста.

Существенным фактором, ограничивающим применение методов коллаборативной фильтрации, является то, что для их работы необходимо накопление информации о предпочтениях пользователей. Как правило, накопление необходимого объема такой информации занимает определенное время и требует значительного количества пользователей с разнообразными предпочтениями. Даже для систем, в которых рекомендации применяются к объектам массового спроса (книги, музыка, фильмы), существует так называемая проблема «холодного старта», когда накопленных сведений о предпочтениях еще не достаточно для формирования сколько-нибудь обоснованных рекомендаций. В узкоспециализированных рекомендующих системах эта проблема проявляется еще острее. Традиционно одним из путей решения этой проблемы является гибридизация.

В разработанной архитектуре рекомендующей системы используются несколько модулей, результаты работы которых объединяются перед выдачей пользователю. Основу составляют два модуля: модуль контекстно-зависимой коллаборативной фильтрации и рекомендующий модуль, основанный на знаниях. Первый модуль предназначен для учета фактически демонстрируемых предпочтений пользователей, а второй – для учета мнений экспертов. Объединение рекомендаций производится в отдельном модуле на основе задаваемых при создании рекомендующей системы правил. С помощью этих правил, в частности, для каждого рекомендующего модуля определяется вероятность того, что в результирующую выборку попадет рекомендация, сформированная этим модулем. Простейший набор правил устанавливает вероятности таким образом, чтобы изначально предпочтение отдавалось рекомендующему модулю, основанному на знаниях, а по мере накопления фактической информации увеличивает значимость модуля коллаборативной фильтрации.

8. Реализация рекомендующей системы для мобильного приложения поддержки туриста

Представленные в предыдущих главах принципы построения и архитектура контекстно-управляемых рекомендующих систем легли в основу прототипа приложения поддержки туриста, которое позволяет туристу на основе текущей ситуации и его профиля получать рекомендации по посещению достопримечательностей и ресторанов, выбору гостиниц и транспорта с использованием персонального мобильного устройства.

На Рис. 4 представлена концептуальная модель приложения поддержки туриста. Основными компонентами приложения являются: информационный сервис, профиль туриста, сервис управления контекстом, сервис планирования маршрута, сервис планирования посещения достопримечательностей и сервис планирования использования транспорта. Взаимодействие перечисленных компонент осуществляется посредством интеллектуального пространства, что позволяет использовать для этих целей доступные вычислительные устройства, вместо проведения всех ресурсоемких расчетов на мобильном устройстве, и обеспечить инфраструктуру для взаимодействия различным мобильным устройствам туристов. Сервисы информации о транспорте предоставляют расписание движения общественного транспорта. В прототипе представленного приложения используется сервис «Яндекс.Расписания» (<http://rasp.yandex.ru/>), который

включает в себя наилучшую на сегодняшний день базу расписаний общественного транспорта по России. Сервисы информации о достопримечательностях представляют собой ресурсы Интернет, которые содержат какую-либо информацию об искомой достопримечательности. Например, для поиска текстовой информации может использоваться Wikipedia ресурс - WikiVoyage, в то время как для поиска изображений помимо перечисленных источников может также использоваться Panoramio, Flickr и др.

На Рис. 5 представлен макросценарий взаимодействия туриста с сервисом, который разделяется на три основных этапа: перед поездкой, во время поездки и после поездки. Перед поездкой турист получает предварительный план экскурсии с рекомендациями к посещению тех или иных достопримечательностей в зависимости от его предпочтений, а также выдержку из законов, которые непосредственным образом могут его коснуться во время путешествия, и культурных норм, которых необходимо придерживаться во время путешествия по региону.

Во время поездки туристу также предлагается ознакомиться с различной важной информацией о регионе. На основе текущей ситуации, профиля туриста, информации из сервиса поиска попутчика (Ridesharing) и информации о достопримечательностях из различных источников строится актуальный экскурсионный план. Сервис поиска попутчика позволяет [25] осуществить совместное использование автотранспорта туристами и водителями региона, что позволяет существенно повысить комфорт



Рис. 4. Концептуальная модель туристического сервиса



Рис. 5. Сценарная модель мобильного туристического приложения

и снизить стоимость перемещения по региону для туристов. Данный сервис оказывает существенное влияние на построение актуального экскурсионного плана, так как зачастую туристу безразлична последовательность просмотра достопримечательностей, а экономия финансов и комфорт поездки играют большую роль.

После поездки мобильный туристический сервис предлагает туристу оценить его работу для дальнейшего использования этой информации в модуле коллаборативной фильтрации.

Оценки достопримечательностей, полученные от пользователей, дополняются значениями параметров контекста и сохраняются в базе данных. Среди используемых параметров контекста - компания, в которой пользователь посетил достопримечательность (один, вдвоем, с семьей), время (обобщаемое до разных уровней), погода, способ путешествия и другое. Значения параметров предоставляются провайдерами контекста (Рис. 2), часть из них предполагает взаимодействие с пользователем (например, указание компании), а остальные опираются на собственные сервисы устройства (например, часы или определение местоположения) или Интернет-сервисы (предоставляющие, например, данные о погоде).

Пользователь оценивает объекты по пятибалльной шкале (1 – худшая оценка, 5 – лучшая). Полученная от пользователя оценка

нормализуется для снижения влияния индивидуальных особенностей оценивания, проявляемых различными пользователями (простановка завышенных или заниженных оценок). Нормализованная оценка достопримечательности j пользователем u (далее \tilde{r}_{uj}) вычисляется по формуле:

$$\tilde{r}_{uj} = r_{uj} - \frac{1}{|K_u| + 1} \left(3 + \sum_{k \in K_u} r_{uk} \right),$$

здесь r_{uj} – исходная оценка объекта j , полученная от пользователя u , а K_u – множество объектов, оцененных пользователем u .

Оценки нормализуются относительно среднего значения, а дополнительное слагаемое позволяет учитывать среднее значение шкалы. Это особенно важно, когда пользователь поставил одну или две оценки, и оценивать их среднее не имеет смысла.

При формировании рекомендаций в соответствии с используемым методом предварительной контекстной фильтрации рассматривается несколько обобщений текущего контекста. Для каждого обобщения отбрасываются оценки, полученные в несовместимом с этим обобщением контексте, а оставшиеся обрабатываются рекомендующими модулями.

Определение группы пользователя происходит с помощью метода «к ближайших соседей» (kNN). В качестве оценки схожести интересов

пользователей (s_{uv}) при определении соседей используется мера косинуса угла между векторами нормализованных оценок пользователей [26]:

$$s_{uv} = \frac{\sum_{o \in O} \tilde{r}_{uo} \tilde{r}_{vo}}{\sqrt{\sum_{o \in O} \tilde{r}_{uo}^2} \sqrt{\sum_{o \in O} \tilde{r}_{vo}^2}}.$$

Здесь O – множество тех объектов, которые оценены и пользователем u , и пользователем v .

Предсказание оценки достопримечательности j для пользователя u (далее r_{ij}^*)

Предсказание оценки достопримечательности для активного пользователя (r_{ij}^*) производится на основе оценок этой достопримечательности, оставленных другими пользователями группы. Для этого находится средневзвешенная нормализованная оценка достопримечательности с учетом степени схожести интересов пользователей:

$$r_{ij}^* = \frac{\sum_{v \in G} \tilde{r}_{vj} s_{uv}}{\sum_{v \in G} |s_{uv}|},$$

здесь G – группа, к которой отнесен активный пользователь. В перечень рекомендаций включаются объекты, обладающие наибольшими значениями предсказанной оценки.

Как уже отмечалось, метод коллаборативной фильтрации требует накопления значительной фактической базы оценок, поэтому для формирования рекомендаций на ранних этапах функционирования туристического сервиса используется, по большей части, рекомендующий модуль, основанный на знаниях. На основе данных профиля пользователь с помощью правил относится к одной из predeterminedных групп, для каждой из которых заданы рекомендуемые объекты с учетом условий контекста.

На Рис. 6 представлены экранные формы разработанного мобильного туристического сервиса. Сервис разработан на базе операционной системы Android, которая на сегодняшний день занимает лидирующую позицию как в России, так и в мире. На первом экране сервис автоматически определяет регион местонахождения пользователя и предлагает ему посмотреть достопримечательности, расположенные рядом. Также допускается возможность ручного выбора страны, региона и города или поиска достопримечательности по названию. Во втором окне представлен результат работы модуля генерации рекомендаций, который на основе предпочтений туриста и текущей ситуации выбирает и предлагает ему соответствующие достопримечательности. Информация о достопримечательности

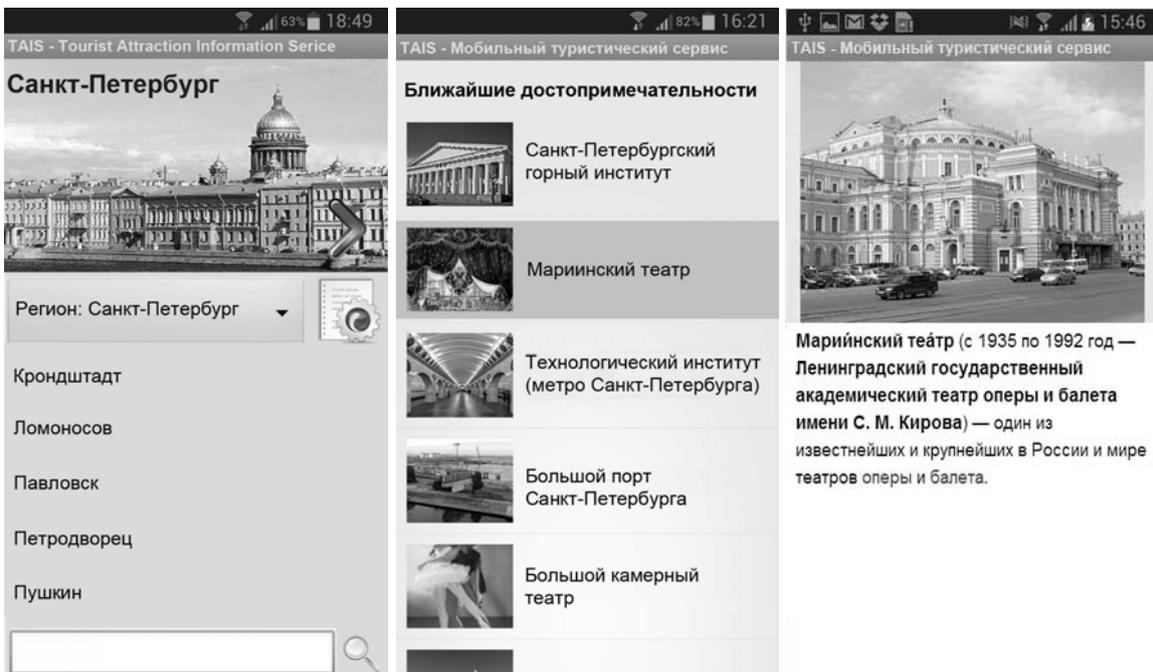


Рис. 6. Экранные снимки прототипа мобильного туристического приложения

мечательностях (окно три, Рис. 6) извлекается «на лету» из общедоступных и постоянно пополняемых источников (на данный момент реализован доступ к ресурсу Wikipedia), что позволяет говорить об актуальности предоставляемой информации.

На Рис. 7 представлен пример рекомендации туристу плана посещения достопримечательностей в Санкт-Петербурге. С учетом предпочтений туриста и текущей ситуации в регионе приложение рекомендует туристу начать посещение достопримечательностей с Эрмитажа, затем проследовать в Кунсткамеру, потом в музей истории СПИИРАН и школы К. Мая, после это в Исаакиевский собор, и, наконец, в музей Достоевского. В зависимости от текущей ситуации данный план посещения может меняться, когда турист закончит посещение того или иного музея. Причем, текущая ситуация в этом случае включает в себя как дорожную обстановку в регионе и возможность следования с попутчиком до следующего музея, так и статистику очередей на вход в музеи в данный день и данное время, например на вход в Эрмитаж турист зачастую может потратить до трех-четырёх часов ожидания в очереди.

Заключение

В работе рассматриваются групповые контекстно-зависимые системы коллаборативной фильтрации. В отличие от контентных рекомендующих систем, коллаборативные рекомендующие системы (или системы коллаборативной фильтрации) в большей степени ориентируются на оценки многих пользователей, а не одного, что позволяет учесть опыт и знания групп экспертов при решении той или иной проблемы.

Показано, что групповые рекомендующие системы должны быть контекстно-ориентированными, поскольку применение методов управления контекстом позволяет выбирать наиболее актуальные для конкретной задачи информацию и знания. Для учета контекста предложено использование методов предварительной фильтрации, поскольку они позволяют в явном виде моделировать и обрабатывать данные о контексте.

Предложена архитектура групповой контекстно-зависимой системы коллаборативной фильтрации, основанная на выделении двух подсистем: подсистемы сбора информации и подсистемы формирования рекомендаций. Кроме

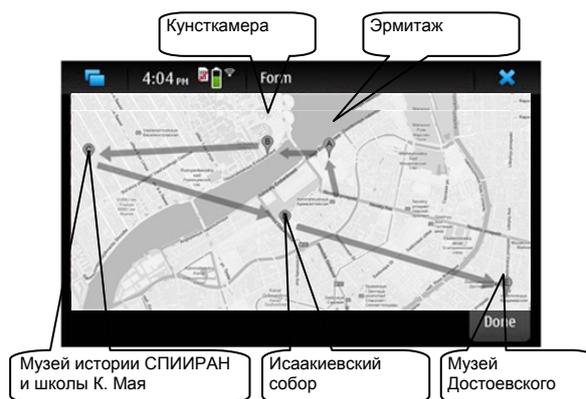


Рис. 7. Пример рекомендации туристу плана посещения достопримечательностей

того, предложена структура профиля пользователя и выделены основные процессы, выполняемые при функционировании групповой рекомендующей системы.

Применение предлагаемой архитектуры частично продемонстрировано на примере рекомендующей системы для мобильного приложения поддержки туриста.

Литература

1. Garcia I., Sebastia L., Onaindia E., Guzman C. A. Group Recommender System for Tourist Activities // EC-Web 2009: Proceedings of E-Commerce and Web Technologies, The 10th International Conference (2009). LNCS 5692. Springer, 2009. Pp. 26–37.
2. Moon, S. K., Simpson T. W., Kumara S. R. T. An agent-based recommender system for developing customized families of products // Journal of Intelligent Manufacturing, Springer, 2009. Vol. 20, No. 6. Pp. 649–659.
3. Chen Y.-J., Chen Y.-M., Wu M.-S. An expert recommendation system for product empirical knowledge consultation // ICCSIT2010: The 3rd IEEE International Conference on Computer Science and Information Technology, IEEE, 2010. Pp. 23–27.
4. Шилов Н. Групповые рекомендующие системы для конфигурирования гибких сетевых организаций, Информационно-управляющие системы. №5, СПб.: «Издательство «Политехника», 2012, ISSN 1684-8853, С. 69–74.
5. McCarthy K., Salamo M., Coyole L., McGinty L., Smyth B., Nixon P. Group Recommender Systems: A Critiquing Based Approach // IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces, 2006. Pp. 267–269.
6. 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), IEEE Educational Activities Department, 2005. Vol. 17, No. 6. Pp. 734–749.
7. Шилов Н. Проблемы поддержки принятия решений при конфигурировании гибких сетевых организаций. Труды СПИИРАН, №22, Санкт-Петербург, 2012, С. 224–233.
8. Smirnov A., Shilov N. Collaborative Recommendation Systems for PLM: Approach and Technological Frame-

- work // Proc. of the 8th International Product Lifecycle Management Conference - PLM11 (July 11–13, 2011, Eindhoven, the Netherlands)
9. Смирнов А.В., Шилов Н.Г. Групповая рекомендательная система для управления жизненным циклом изделий: подход и технологии // Известия ЮФУ. Технические науки. – Таганрог: Технологический институт Южного Федерального Университета, 2011. – № 5 (118). – Тематический вып. «Компьютерные и инженерные технологии в науке, инженерии и управлении». – С. 203–206.
 10. Baatarjav E.-A., Phithakkitnukoon S., Dantu R. Group Recommendation System for Facebook / OTM 2008 // Proceedings of On the Move to Meaningful Internet Systems Workshop (2008), LNCS 5333, Springer, 2009. Pp. 211–219.
 11. Romesburg H. C. Cluster Analysis for Researchers, Lulu Press, California, 2004. 340 p.
 12. Flake G. W., Lawrence S., Giles C. L., Coetzee F. Self-Organization and identification of Web Communities // IEEE Computer, IEEE, 2002. Vol. 35, No. 3. Pp. 66–71.
 13. Baeza-Yates R., Ribeiro-Neto B. Modern Information Retrieval, Addison-Wesley, 1999. 513 p.
 14. Salton G. Automatic Text Processing: The Transformation Analysis and Retrieval of Information by Computer, Addison-Wesley, 1989. 543 p.
 15. Belkin N., Croft B. Information Filtering and Information Retrieval // Communications of the ACM, Special issue on information filtering, ACM New York, 1992. Vol. 35, No. 12. Pp. 29–37.
 16. Adomavicius G., Mobasher B., Ricci F., Tuzhilin A. Context-aware recommender systems // AI Magazine. 2011. 32(3), Pp. 67–80.
 17. 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. Pp. 103–145.
 18. 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. Pp. 15–18.
 19. 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. Pp. 245–248.
 20. 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. Pp. 71–75.
 21. Koren Y., Bell R., Volinsky C. Matrix factorization techniques for recommender systems // IEEE Computer, Aug 2009. vol. 42, no. 8. Pp. 30–37.
 22. Rendle S. Factorization Machines with libFM // ACM Trans. on Intell. Syst. Technol., May 2012. vol. 3, no. 3. Pp. 57:1–57:22.
 23. 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. Pp. 16:1–16:19.
 24. Смирнов А.В., Пашкин М.П., Шилов Н.Г., Левашова Т.В. Подход к построению распределенной системы интеллектуальной поддержки принятия решения в открытой информационной среде // Труды СПИИРАН. Вып. 4. т. 1 – СПб.: Наука, 2007. С. 36–49.
 25. Кашевник А.М., Тесля Н.Н., Автоматизированная система совместного использования автотранспорта. Материалы конференции «Информационные технологии в управлении» (ИТУ-2012). - СПб.: ОАО «Концерн «ЦНИИ «Электронприбор», 2012, стр. 471–479.
 26. Boratto, L., Carta, S., Chessa, A., Agelli, M. Group Recommendation with Automatic Identification of Users Communities. In: Proceedings of IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies, 2009 (WI-IAT '09). Vol. 3, pp. 547–550, 2009.

Смирнов Александр Викторович. Заведующий лабораторией СПИИРАН. Окончил Ленинградский государственный политехнический университет в 1979 году. Доктор технических наук, профессор, заслуженный деятель науки РФ. Автор более 300 печатных работ. Область научных интересов: управление знаниями, веб-сервисы, системы групповой поддержки принятия решений, виртуальные предприятия, управление цепями поставок. E-mail: smir@iias.spb.su

Шилов Николай Германович. Старший научный сотрудник СПИИРАН. Окончил Санкт-Петербургский государственный политехнический университет в 1998 году. Кандидат технических наук. Автор более 150 печатных работ. Область научных интересов: управление знаниями, виртуальные предприятия, самоорганизующиеся системы. E-mail: nick@iias.spb.su

Пономарев Андрей Васильевич. Старший научный сотрудник СПИИРАН. Окончил Тюменский государственный нефтегазовый университет в 2003 году. Кандидат технических наук. Автор 15 печатных работ. Область научных интересов: рекомендующие системы, профилирование, методы дискретной оптимизации, СУБД. E-mail: ponomarev@iias.spb.su

Кашевник Алексей Михайлович. Старший научный сотрудник СПИИРАН. Окончил Санкт-Петербургский государственный политехнический университет в 2005 году. Кандидат технических наук. Автор 121 печатной работы. Область научных интересов: управление знаниями, профилирование, онтологии, интеллектуальные пространства, логистические системы. E-mail: alexey@iias.spb.su

Парфенов Владимир Глебович. Декан Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики, заведующий кафедрой информационных систем. Окончил Ленинградский институт точной механики и оптики в 1972 году и Ленинградский государственный университет в 1973 году. Доктор технических наук, профессор, лауреат Премий Президента и Правительства Российской Федерации в области образования. Автор более 150 научных работ. Область научных интересов – математическое моделирование оптоэлектронных систем, технологии программирования, управление знаниями, технологии дистанционного обучения. E-mail: parfenov@mail.ifmo.ru