

**658.310.7 РЕШЕНИЕ ЗАДАЧИ КОЛЛАБОРАТИВНОЙ ФИЛЬТРАЦИИ ПРИ
ФОРМИРОВАНИИ МАЛЫХ ПРОЕКТНЫХ ГРУПП ИТ-РАЗРАБОТЧИКОВ
SOLVING THE PROBLEM OF COLLABORATIVE FILTERING IN THE
FORMATION OF SMALL PROJECT TEAMS OF IT DEVELOPERS**

А.Ю. Харитонов*, *старший преподаватель*

В.С. Андрианов

М.А. Саковский

ФГАОУ ВО «Сибирский федеральный университет»

(Россия, г. Красноярск)

Россия, 660074, г. Красноярск, ул. Академика Киренского 26, корпус 1

Андрей Юрьевич Харитонов – pharret31@gmail.com

В наше время в интернете очень активно используется метод коллаборативной фильтрации, но многие ничего не знают об этом методе. Мы взяли данную тему, для того чтобы понятным языком рассказать принцип действия данного метода. В связи с возможностью широкого использования данного метода, его изучение представляет практический и научный интерес. На примере, рассмотренном в данной статье, мы доказали работоспособность данного метода, его принципы и возможность применения его на практике. Решение задачи метода коллаборативной фильтрации при формировании малых проектных групп ит-разработчиков мы рассмотрели на примере формирования группы для выполнения командного курсового проекта по направлению "Программная инженерия". Также мы рассказали о важных критериях для успешного формирования группы. Для понимания сути данного метода мы с помощью объединения множеств показали принцип работы метода коллаборативной фильтрации. В конечном итоге мы сделали вывод, что такая адаптивная система обучения поможет студентам без проблем подобрать команду для выполнения программных проектов. Благодаря такой системе студент будет более углубленно изучать то, что ему нравится.

Ключевые слова: коллаборативная фильтрация, адаптивная система, малые проектные группы, проектный подход, программный проект, командный курсовой проект.

В наше время рекомендательные системы активно используются в жизнедеятельности человека – мы сталкиваемся с ними повсеместно. С их помощью нам предлагают товары и услуги, которые могут нас заинтересовать.

Коллаборативная фильтрация – это метод, позволяющий предсказать неизвестные предпочтения пользователя на основе известных оценок и/или поведения похожих пользователей. Суть коллаборативной фильтрации заключается в том, что пользователи, одинаково оценившие, какие-либо элементы системы склонны одинаково оценивать и другие элементы системы [1].

Алгоритм включает в себя 3 шага:

1. Для каждого пользователя вычислить насколько его интересы совпадают с интересами пользователя текущего пользователя;
2. Выбрать множество пользователей, наиболее близких к текущему;
3. Предсказать оценку i -го объекта на основе оценок объекта соседей.

Для определения соседства пользователей существует множество различных алгоритмов, например, косинусное, манхэттенское и евклидово расстояния.

Косинусное расстояние определяется как:

$$d(u_1, u_2) = \frac{\vec{u}_1 \cdot \vec{u}_2}{\|\vec{u}_1\|_2 \times \|\vec{u}_2\|_2} = \frac{\sum_i u_{1i} u_{2i}}{\sqrt{\sum_i u_{1i}^2} \sqrt{\sum_i u_{2i}^2}},$$

где u_{1i} и u_{2i} – оценки пользователей.

Манхэттенское расстояние имеет следующий вид:

$$d(u_1, u_2) = \sum_i |u_{1,i} - u_{2,i}|.$$

Метрика в Евклидовом пространстве вычисляется при помощи следующей формулы:

$$d(u_1, u_2) = \sqrt{\sum_i (u_{1,i} - u_{2,i})^2}.$$

Для выбора пользователей, наиболее похожих на текущего пользователя, задается константа k . Затем все пользователи сортируются по убыванию меры близости и из этого списка отбираются первые k , наиболее близких к текущему.

Предсказание оценки для i -го объекта на основе оценок объекта соседей осуществляется по следующей формуле:

$$P(u_{j,i}) = \frac{\sum_{k \in K} u_{k,i} d(u_j, u_k)}{\sum_{k \in K} d(u_j, u_k)},$$

где K – множество наиболее похожих на u_j пользователей.

Вторым крупным видом коллаборативной фильтрации является фильтрация, основанная на анализе модели данных. Для данного подхода характерно сначала построение описательной модели предпочтений по совокупности оценок, а затем формирование рекомендации на основании полученной модели. Примером этого вида является кластерная модель.

Одним из самых известных алгоритмов в кластерном анализе является метод k -means или k -средних. Он основан на разделении объектов или пользователей на группы – кластеры, которые создаются по некоторым общим признакам, а количество которых задается заранее. Суть алгоритма состоит в случайном выборе k центров кластера и уменьшении суммарного квадратичного отклонения пользователей или объектов от центра кластера. Формально это вычисляется с помощью следующей формулы:

$$d = \sum_i \sum_j (x_j - u_i)^2.$$

Цель метода заключается в том, чтобы как можно более точно предсказать оценку, которую поставит текущий пользователь системы ранее не оцененным им объектам (рис.1).

Принцип коллаборативной фильтрации



Рисунок 1.

На сегодняшний день коренные изменения в экономической сфере российского общества требуют научного подхода к переосмыслению устоявшихся социальных норм. Важнейшей областью в пространстве социальных отношений является взаимодействие внутри трудового коллектива. От социально-психологической атмосферы в коллективе, специфики организационной культуры во многом зависит не только эффективность совместной деятельности, выражающейся в экономических показателях, но также отношение сотрудников к своей работе, эмоциональный настрой, внешняя мотивация и, в конечном итоге, удовлетворенность работой. В связи с этим, актуальной становится задача разработки новых подходов к формированию рабочих коллективов в малых проектных группах.

По результатам множества опросов менеджеров проектов в России и за рубежом, до 80% успеха при реализации проектов обусловлены слаженной работой проектной команды, которая, в свою очередь, обеспечивается верным распределением ролей среди участников[2]. При распределении ролей необходимо соблюдать два принципа: принцип компетентности и принцип предпочтения[3]. Принцип компетентности – навыков сотрудника должно быть достаточно для исполнения своей роли. Члены команды будут отдавать большее предпочтение тем целевым ролям, которые больше соответствуют их умениям и индивидуальным потребностям, т.е. принижать умения человека тоже не стоит. Принцип предпочтения – как можно большее число членов команды должны получить те роли, которые они предпочитают. Существуют такие ситуации, в которых сотрудник не согласен с

данной ему ролью, в связи с этим снижается продуктивность этого сотрудника и нарушается взаимодействие между членами команды.

Как было сказано выше, коллаборативная фильтрация — подход к предсказанию предпочтения с использованием исключительно информации о связях пользователей и объектов рекомендации. В нашем случае объектом будет являться роль, которой интересовался пользователь. Сформулируем задачу более строго. Пусть U — множество пользователей (users), I — множество объектов (items), информация об известных предпочтениях представлена в виде набора троек: $D = \{(u, i, r_{ui})\} | (u, i) \in R$, где $r_{ui} \in \mathbb{R}$ — вещественная степень предпочтения объекта $i \in I$ пользователем $u \in U$; $R \subseteq U \times I$ — множество пар (пользователь, объект), про которые известна степень предпочтения. Для дальнейшего удобства, введем также обозначения: $R(u) = \{i : (u, i) \in R\}$ — множество объектов, смежных с пользователем u , аналогично: $R(i) = \{u : (u, i) \in R\}$. По известной информации D требуется уметь строить предсказание предпочтения $\hat{r}_{ui} \approx r_{ui}$ для новых пар $(u, i) \notin R$. Будем называть матрицей оценок матрицу $R \in (\mathbb{R} \cup \emptyset)^{|U| \times |I|}$, строки которой соответствуют пользователям, столбцы — объектам (ролям), а элементы принимают значение r_{ui} , если $(u, i) \in R$, иначе — пропуск \emptyset . На задачу коллаборативной фильтрации можно смотреть как на задачу объединения пользователей с разными предпочтениями к ролям в одну малую группу it разработчиков. Помимо предсказания значения предпочтения, на практике могут быть интересны следующие задачи:

- построение списка рекомендаций из объектов, на которые не известна степень предпочтения (новые для пользователя).
- определение степени схожести объектов и построение списков наиболее похожих.
- обоснование списка рекомендаций: некоторое человеко-понятное пояснение, почему пользователю u был порекомендован объект i .

В малых проектных группах невозможно распределить роли по всем участникам проекта - участников проекта меньше, чем ролей. Но в тоже время важно, чтобы были учтены все первостепенные роли – такие, как руководитель проекта, программист, консультант и т.д. В этом случае роли участников проекта совмещаются, но не стоит забывать, что не все роли можно совмещать друг с другом. Не следует допускать совмещения ролей, которые имеют конфликтные или противоречивые интересы. Разработка – роль которая нуждается в большом количестве времени, так что сотрудников, у которых основной ролью является разработка не стоит нагружать второстепенными ролями, это увеличит время на реализацию проекта.

В нашем случае, необходимо распределить роли между участниками проекта, у каждого из которых может быть более чем одна роль. Для этого нужно сравнить текущую команду с другими подобными командами. К примеру, если в большинстве подобных команд разработчик является одновременно и специалистом по интерфейсу, то, соответственно, в этой команде нужно назначить разработчика специалистом по интерфейсу.

В наше время в высших учебных заведениях активно используются электронные курсы, как замена привычным для нас учебникам и задачникам. У электронных курсов есть огромное преимущество перед бумажными носителями в том, что обновление учебных-методических материалов происходит быстрее, но так же как и учебники они рассчитаны на стандартного пользователя и не учитывают индивидуальности каждого студента. Здесь на помощь и приходит адаптивное обучение, благодаря которому учебный курс «подстраивается» под учащегося. Без сомнений, такой вид обучения можно отнести к перспективному направлению в развитии образования [4]. Для демонстрации эффективности данного метода в качестве примера можно рассмотреть его применение для реализации профессионального обучения, например, по направлению «Программная инженерия».

В процессе обучения студентам предстоит выполнить командный курсовой проект. Для его выполнения учащиеся должны разбиться на группы, в которых каждый студент выполняет отведенную ему роль.

Например, студент интересуется тестированием программных продуктов, найдя подходящий материал, студент сам того не сознавая внесет эту информацию в систему [5]. В дальнейшем от результата работы коллаборативной фильтрации система будет подбирать командам или отдельным разработчикам людей на недостающие роли (рис 2).

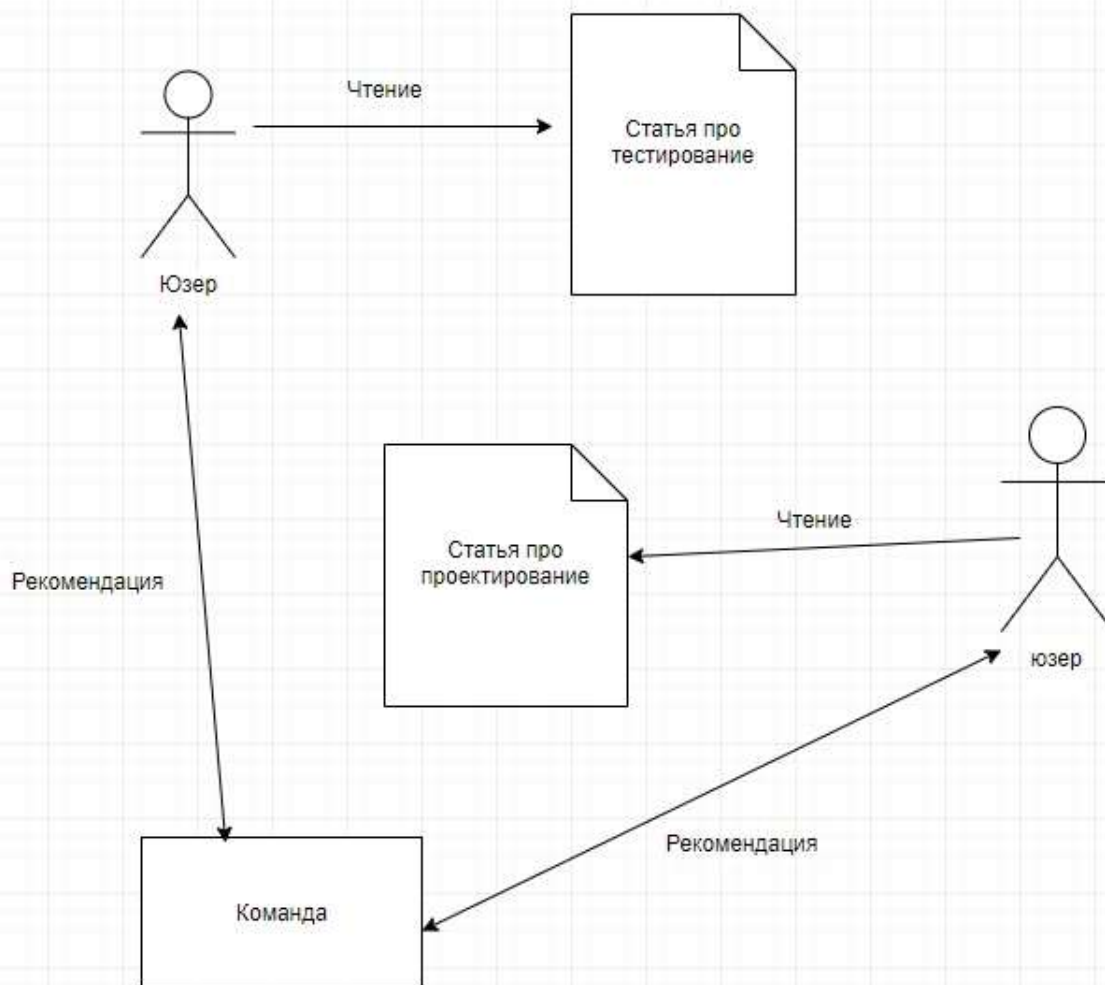


Рисунок 2.

Такая адаптивная система обучения поможет студентам без проблем подобрать команду для выполнения программных проектов. Благодаря такой системе студент будет более углубленно изучать то, что ему нравится.

ЛИТЕРАТУРА

1. M. D. Ekstrand, J. T. Riedl and J. A. Konstan. Collaborative Filtering Recommender Systems. Minneapolis: University of Minnesota, 2011.
2. Определение ролей участников проектной команды. URL: https://iteam.ru/publications/project/section_37/arti.. (дата обращения: 20.07.2018).
3. Распределение ролей в команде. URL: http://www.bikr.ru/informatsiya/kak_upravlyat_persona.. (дата обращения: 20.07.2018)

4. Евдокимов И.В. ТРИЗ-инструментарий программной инженерии / И.В. Евдокимов, А.С. Михалев, В.В. Ковалев // Теоретические, методологические и прикладные вопросы науки и образования: материалы международной научно-практической конференции. – Самара, 2018. – С. 8–16.

5. Евдокимов И.В. Аспекты внедрения информационных технологий на предприятиях г. Братска // Труды Братского государственного университета. Серия: Экономика и управление. – 2006. – Т. 1. – С. 144–148.

Solving the problem of collaborative filtering in the formation of small project teams of it developers

A.Y. Kharitonov*, senior lecturer

V.S. Andrianov

M.A. Sakovskiy

Siberian Federal University

(Russia, Krasnoyarsk)

Kirenskogo st. 26, building 1, Krasnoyarsk, 660074, Russia

Kharitonov Andrew – pharret31@gmail.com

Nowadays, the Internet is very actively used method of collaborative filtering, but many do not know anything about this method. We took this topic in order to explain in plain language the principle of operation of this method. In connection with the possibility of widespread use of this method, its study is of practical and scientific interest. On the example considered in this article, we have proved the efficiency of this method, its principles and the possibility of its application in practice. The solution of the problem of the method of collaborative filtering in the formation of small project groups of it-developers, we have considered the example of the formation of a group to perform a team course project in the direction of "Software engineering". We also talked about the important criteria for the successful formation of the group. To understand the essence of this method, we have shown the principle of the method of collaborative filtering by combining sets. In the end, we concluded that such an adaptive learning system will help students easily find a team to implement software projects. Thanks to this system, the student will be more in-depth study of what he likes.

Keywords: collaborative filtering, adaptive system, small project groups, project approach, program project, team course project.