
ВЫРАБОТКА РЕКОМЕНДОВАННОГО НАБОРА ФИЛЬМОВ ДЛЯ КЛИЕНТОВ ОНЛАЙН-КИНОТЕАТРА

Казаков Олег Леонидович¹, д-р техн. наук, проф.

Казаков Леонид Олегович², рук. отдела анализа

¹ Московский политехнический университет, ул. Большая Семеновская, 38, Москва, Россия, 107023; e-mail: olekazakov@gmail.com

² Общество с ограниченной ответственностью «Нейрософт», Дербеневская набережная, 11, Москва, Россия, 115114; e-mail: leonid.kazakov@gmail.com

Цель: статья посвящена совершенствованию методики формирования перечня упорядоченного по интересам любого конкретного клиента онлайн-кинотеатра новых для него фильмов на основе ранее полученных проявлений интересов всех клиентов к имеющимся фильмам. *Обсуждение:* составление набора фильмов по интересам требует применения математических и инструментальных методов, учитывающих особенности метрик для определения схожести строк в матрицах неструктурированных предпочтений большой размерности. Поэтому актуальным является поиск в области Big Data новых способов решения и их сочетаний. *Результаты:* предлагается использование свойств множеств Парето для выделения близких по интересам клиентов и их ранжирование по близости к интересам каждого клиента с помощью оценочных векторов. Разработанный в статье способ решения неформальной задачи большой размерности отличается оригинальностью. Он позволит автоматизировать составление рекомендаций и повысить их объективность и точность. Кроме того, использование технологии Big Data обеспечит самоадаптацию предлагаемого средства к изменяющимся интересам клиентов и возможность объяснения предоставляемых советов.

Ключевые слова: набор (перечень) фильмов, интерес клиентов онлайн-кинотеатра, размещение как упорядоченная последовательность, матрица как множество векторов, ранжирование по близости интересов, множество Парето, оценочный вектор.

DOI: 10.17308/meps.2021.7/2630

Введение

Клиентам онлайн-кинотеатра предоставляется перечень имеющихся фильмов, из которого они могут выбрать любой для просмотра. Иногда клиент выбирает, даже неоднократно, понравившийся ему фильм из уже просмотренных им ранее.

Выбор фильма из обширного их перечня может быть неудобным для клиента. Поэтому предлагается предоставить клиенту возможность выбора фильма из упорядоченного перечня, в котором на первом месте будут предположительно наиболее интересные для него фильмы, кроме уже просмотренных им [5]. Рассмотрим создание такого упорядоченного «интересного» перечня на основе ранее полученных проявлений интересов всех клиентов к имеющимся фильмам.

Такая задача относится к т.н. коллаборативной фильтрации [12] и обычно решается с использованием матрицы предпочтений, в которой строки соответствуют клиентам, а столбцы – фильмам. Для определения схожести строк в таких матрицах применяют методы классификации, кластеризации, декомпозиции и агрегирования, а также вычисления средне-квадратичных ошибок [2, 3]. В качестве метрик используют расстояние Жаккара, косинусное расстояние, округление данных, нормировку оценок. Но все это не позволяет окончательно избавиться от погрешностей в решении задачи [9]. Поэтому актуальным является поиск новых способов решения и их сочетаний.

Выработка рекомендованного набора фильмов повышает удобство клиентов и популярность онлайн-кинотеатра [1, 4].

Постановка задачи

Имеется n клиентов компании и m фильмов, которые компания может предложить клиентам для просмотра. Выберем одно из размещений m из m , т.е. упорядоченных последовательностей фильмов, которыми будем постоянно пользоваться для представления данных о просмотрах клиентами фильмов за определенный период времени в виде векторов. Тогда все известные данные о просмотрах клиентами фильмов за определенный период времени можно представить матрицей, или множеством векторов:

$$F = \{(f_1, a_1), (f_2, a_2), \dots, (f_m, a_m)\}_{i = \overline{1, n}},$$

где каждый компонент вектора включает f_j – название j -го фильма, a_j – количество просмотров j -го фильма ($a_j \in \{0, 1, 2, \dots, \infty\}$, $j = \overline{1, m}$).

Если сохранить порядок представления имеющихся фильмов, закрепить за названием фильма его индекс в этом векторе, то можно не использовать их названий и множество векторов F можно заменить множеством векторов A :

$$A = \{(a_1, a_2, \dots, a_m)\}_{i = \overline{1, n}} = \{a_{(m)i}\}_{i = \overline{1, n}}.$$

Отсутствие просмотра i -м клиентом j -го фильма $a_{ij} = 0$ еще не означает, что этому клиенту не нравится такой фильм. Также единичный просмотр фильма $a_{ij} = 1$ не может свидетельствовать о возникшем интересе к нему. Поэтому интерес клиента к фильму определяется числом его просмотра, большим единицы. Кроме того, считается, что нет ни одного клиента, у которого не было бы просмотра ни одного из заданных m фильмов:

$$\sum_{j=1}^m a_{ij} \neq 0, i = \overline{1, n}.$$

Будем считать, что количество просмотров фильма численно выражает интерес клиента к этому фильму, и далее использовать эти термины как синонимы.

Можно ранжировать [6, 7, 8] фильмы по интересам для каждого i -го клиента, т.е. составлять упорядоченное множество, или вектор, или перечень фильмов в такой последовательности, чтобы интерес клиента к каждому последующему фильму был бы не больше, чем к предыдущему фильму в этой последовательности:

$$P_i = \{a_{pq} | p_q \in \{1, 2, \dots, m\}, q = \overline{1, m}\}_i = (a_{p_1}, a_{p_2}, \dots, a_{p_m})_i = a_{p_{qm}i},$$

где a_{pq} – количество просмотров pq -го фильма, численно выражающее интерес клиента к этому фильму; $a_{pq} \geq a_{p_{q+1}}$, $i = \overline{1, n}$.

По существу, это другое размещение m из m по сравнению с принятым нами выше эталонным размещением в векторах множества A .

Таким образом, можно осуществить переход от множества векторов $A = \{a_{mi} | i = \overline{1, n}\}$ к множеству векторов $P = \{a_{p_{mi}i} | i = \overline{1, n}\}$, в которых векторы различаются только размещением одних и тех же компонентов: в векторе a_{mi} размещение компонентов эталонное, а в векторе $a_{p_{mi}i}$ – упорядоченное по убыванию величин тех же компонентов.

Предоставлять каждому i -му клиенту для выбора последовательность фильмов $a_{p_{mi}i}$ с ненулевым интересом нецелесообразно, т.к. он и так о них знает. А вот предоставление ему ранжированной последовательности фильмов из $a_{p_{mi}i}$ с нулевыми интересами может вызвать интерес.

Для составления такой ранжированной последовательности фильмов из $a_{p_{mi}i}$, ранее не просмотренных каждым i -м клиентом, будем использовать интересы тех клиентов к фильмам, которые наиболее близки к интересам этого клиента к фильмам. Предполагается, что использование данных о выборе фильмов клиентами с похожими интересами позволит актуализировать упорядоченный перечень $a_{p_{mi}i}$ приоритетными для каждого i -го клиента фильмами, которые он еще не просматривал.

С другой стороны, если предположить, что существуют два абсолютно совпадающих вектора $a_{p_{mi}i}$ и $a_{p_{mj}j}$, то такое совпадение совершенно не позволяет предложить какие-нибудь рекомендации по актуализации этих векторов. То же самое можно, без сомнения, заявить для совершенно разных векторов $a_{p_{mi}i}$ и $a_{p_{mj}j}$.

Поэтому предлагается сначала для каждого i -го клиента определить близких по интересам к нему клиентов, затем ранжировать этих клиентов по близости к его интересам и, наконец, использовать прежде всего интересы тех клиентов, которые наиболее близки к интересам этого клиента, для составления упорядоченной по его интересам последовательности ранее не просмотренных им фильмов. Подобная технология Big Data обеспечивает

самоадаптацию к изменяющимся условиям и возможность объяснить получаемые результаты [10].

Обозначим искомые актуализированные упорядоченные по интересам перечни фильмов, в том числе новых, для каждого i -го клиента вектором a_{pmi}^A . Составление таких векторов и будет решением поставленной задачи выработки рекомендованного набора фильмов клиентам онлайн-кинотеатра.

Способ решения поставленной задачи

Выделение клиентов, близких по интересам к каждому i -му клиенту

Используемые векторы с одинаковым количеством компонентов могут быть сравнимыми и несравнимыми между собой. Результатами сравнения между собой пары векторов могут быть пять оценок: больше, меньше, равно, не больше, не меньше. Для несравнимых между собой векторов не подходит ни одна из этих оценок.

В любом рассматриваемом множестве векторов можно выделить называемое множеством Парето [11] подмножество всех несравнимых или совершенно одинаковых между собой векторов, каждый из которых или меньше, или несравним с оставшимися в исходном множестве векторами. В данном случае такое множество Парето содержит самые наименьшие векторы среди всех остальных сравнимых с ними векторов в исходном множестве.

Введем следующие обозначения:

PA_1 – множество Парето наименьших векторов из множества векторов A , A_1 – множество векторов A без множества Парето PA_1 ;

PA_2 – множество Парето наименьших векторов из множества векторов A_1 , A_2 – множество векторов A_1 без множества Парето PA_2 ;

...

PA_e – пустое множество Парето наименьших векторов из множества векторов A_{e-1} .

Каждому вектору $a_{pmi} = (a_1, a_2, \dots, a_m)_i$ из множества векторов A соответствует каждый i -й клиент. Этот вектор может присутствовать только в одном множестве Парето наименьших векторов, пусть в PA_r . Доказательством этого служит то, что он может вообще присутствовать в множестве Парето наименьших векторов только на какой-нибудь итерации изымания из множества векторов текущего множества Парето наименьших векторов и после изымания вместе с множеством Парето наименьших векторов не участвует в образовании следующего множества Парето наименьших векторов.

Тогда наиболее близкие к каждому вектору a_{pmi} векторы находятся в множестве векторов A_r , полученном извлечением множества Парето наименьших векторов, включающего этот вектор a_{pmi} , из общего множества векторов. Обосновывается это тем, что близость здесь понимается как сравнимость векторов. Наиболее близкие к каждому вектору a_{pmi} будут сравнимые с ним векторы из A_r , множество которых обозначим A_{ri} . Векторы из A_{ri}

не будут меньше вектора a_{mi} , но не всегда будут сравнимы между собой.

Обратим внимание на то, что размерность матрицы A_{ri} не больше размерности матрицы A , т.е. речь идет о снижении размерности.

Составим словесный алгоритм выделения, например, множества векторов A_{ri} из всего множества векторов A .

Организуем цикл перебора всех n векторов A . Для каждого рассматриваемого вектора организуем цикл перебора его m компонентов. Тогда можно сравнить каждого рассматриваемого компонента с соответствующим компонентом вектора a_{mi} и если встретится рассматриваемый компонент, меньший соответствующего компонента вектора a_{mi} , то этот рассматриваемый вектор исключается из множества векторов A . По завершении всех циклов множество векторов A будет преобразовано в множество векторов:

$$A_{ri} = \{a_{mj} | j = \overline{1, n_{ri}}\}.$$

Таким образом, имеется способ выделения клиентов, близких по интересам к любому конкретному i -му клиенту. Теперь нужно ранжировать этих клиентов по близости к его интересам, обеспечивающей актуализацию интересов одного клиента по данным других клиентов.

Ранжирование клиентов по близости к интересам каждого i -го клиента

От найденного множества векторов A_{ri} перейдем к множеству векторов P_{ri} по аналогии с тем, как это делали выше при переходе от множества векторов A к множеству векторов P .

Естественно считать, что чем больше будет совпадений между векторами a_{pmi} и a_{pmj} из множеств P_{ri} и P_{rj} , тем ближе по интересам будут i -й и j -й клиенты. Причем совпадения могут устанавливаться по двум условиям: по составу больших единицы компонентов и по их месту в упорядоченном перечне.

Поэтому перед поиском совпадений уберем из множества P_{ri} векторы, у которых все компоненты не больше единицы. Если при этом окажется, что множество векторов P_{ri} стало пустым, то для выбора i -му клиенту предоставляются те фильмы, у которых в векторе a_{pmi} компоненты равны нулю, затем равны единице и т.д. по возрастанию. Если же этого не произойдет, то для ранжирования предлагается определить векторный способ численного оценивания совпадений.

Рассмотрим такой способ на примере ранее определенных вектора a_{pmi} и множества векторов $P_{ri} = \{a_{pmk} | k = \overline{1, n_{ri}}\}$, в котором n_{ri} – количество векторов после вышеописанного удаления из множества P_{ri} векторов с компонентами не больше единицы.

Компоненты каждого вектора, характеризующего определенного клиента, можно разбить по убывающей величине этих компонентов на последовательность классов, которых будет не больше m . Каждый класс будет включать компоненты с одинаковой величиной, соответствующие индексам фильмов с одинаковым количеством просмотров этим клиентом. По-

следним классом будет совокупность единичных и нулевых компонентов.

Для каждого вектора $a_{p\sigma i k}$ из P_i введем оценочный вектор $o_{p\sigma i k}$ с числом компонентов, равным числу l_{pi} классов вектора $a_{p\sigma i}$. Каждый компонент оценочного вектора приравняем количеству индексов фильмов, совпадающих с индексами фильмов соответствующего класса вектора $a_{p\sigma i}$. Таким образом, каждый оценочный вектор $o_{p\sigma i k}$ ($i = \overline{1, n}$, $k = \overline{1, n_{ri}}$, $i \neq k$) будет векторным способом численно оценивать совпадения интересов i -го клиента с каждым соответствующим k -м клиентом.

Можно утверждать, что для $l_{pi} = l_{pk}$ выполняется $o_{p\sigma i k} = o_{p\sigma i k i}$. Это может учитываться для снижения числа переборов в используемых алгоритмах, т.е. для увеличения их быстродействия.

Для того чтобы ранжировать клиентов по близости к интересам i -го клиента, нужно сначала составить множество оценочных векторов:

$$O_i = \{o_{p\sigma i k} | k = \overline{1, n_{ri}}\}.$$

Затем среди этого множества оценочных векторов найти векторы, у которых первые компоненты не меньше, чем у остальных.

Если будет найден только один такой оценочный вектор, то на этом поиски заканчиваются и этот вектор будет использоваться для создания вектора $a_{p\sigma i}^A$, т.е. упорядоченного перечня фильмов, в котором на первом месте будут предположительно наиболее интересные фильмы для i -го клиента, кроме уже просмотренных им фильмов.

Если таких оценочных векторов будет несколько, то среди них нужно искать оценочные векторы, у которых вторые компоненты не меньше, чем у остальных. Если будет найден только один такой оценочный вектор, то, как и в первом случае, переходим к созданию вектора $a_{p\sigma i}^A$. Если же таких оценочных векторов будет несколько, то процедура с ними повторяется относительно следующего компонента оценочных векторов.

Такая процедура будет повторяться до тех пор пока не будет найден только один оценочный вектор или будут перебраны все компоненты оценочных векторов и будут найдены несколько из оценочных векторов.

Таким образом, осуществляется ранжирование клиентов по близости к интересам каждого i -го клиента. Полученный окончательно один оценочный вектор указывает на того клиента, который наиболее близок к интересам i -го клиента к фильмам. Если таких результирующих оценочных векторов будет несколько, то они образуют множество Парето оценочных векторов, которые указывают на наиболее близких клиентов к интересам i -го клиента к фильмам. В последнем случае для создания вектора $a_{p\sigma i}^A$ может быть использован любой из оценочных векторов множества Парето.

Формирования перечня рекомендованных фильмов для выбора каждым i -м клиентом

По выбранному оценочному вектору определяется некоторый k -й клиент, наиболее близкий к интересам каждого i -го клиента к фильмам. Сравни-

ваются векторы $a_{p\sigma\tau i}$ и $a_{p\sigma\tau k}$ и находятся те индексы фильмов, которые присутствуют в векторе $a_{p\sigma\tau k}$ и отсутствуют в векторе $a_{p\sigma\tau i}$. Сохраняя порядок следования этих присутствующих индексов фильмов в векторе $a_{p\sigma\tau k}$, их вставляют в начало формируемого вектора $a_{p\sigma\tau i}^A$, а затем за ними в этом векторе могут следовать индексы фильмов из вектора $a_{p\sigma\tau i}$ в том же порядке, в котором они были в этом векторе. На этом создание каждого вектора $a_{p\sigma\tau i}^A$ заканчивается.

Возможен вариант, когда найденных при сравнении векторов $a_{p\sigma\tau i}$ и $a_{p\sigma\tau k}$ индексов фильмов будет чрезвычайно мало или вообще не будет. Тогда можно использовать другие оценочные векторы из результирующего множества Парето или, если такого множества нет, то полученные на предыдущем этапе процедуры нахождения оценочных векторов. Так или иначе, всегда будет сформирован приоритетный перечень фильмов для выбора каждым i -м клиентом.

Таким образом, задача выработки рекомендованного набора фильмов клиентам онлайн-кинотеатра решена.

Заключение

Предоставление клиентам онлайн-кинотеатра возможности выбора фильма из упорядоченного перечня, в котором на первом месте будут предположительно наиболее интересные фильмы для каждого из них, кроме уже просмотренных ими, обеспечит несомненное удобство и не помешает выбору полюбившегося фильма из уже ранее просмотренных. Причем предоставляемый для выбора упорядоченный по интересам клиента перечень фильмов может иметь любую необходимую длину.

Предлагаемый способ решения задачи составления рекомендованного набора фильмов для клиентов онлайн-кинотеатра отличается оригинальностью по сравнению с другими известными методами оценивания интересов клиентов, нахождения их схожести и выработки рекомендаций. Его реализация позволит повысить объективность полученного результата, автоматизировать этот процесс, увеличить точность предложений, адаптировать их к изменяющимся условиям, обеспечить гибкость по отношению к интересам клиентов, а также объяснимость предоставляемых советов. Причем быстрое действие такого процесса, конечно, будет зависеть от числа клиентов и фильмов, но не будет критичным. Тем более, что предлагаются методы снижения размерности используемых матриц (множеств).

Список источников

1. Березина А.Д. Развитие рынка онлайн-видеосервисов в экосистеме цифровой экономики // *Экономика и бизнес: теория и практика*, 2019, no. 10-1 (56), с. 50-53.
2. Бова В.В., Кулиев Э.В., Щеглов С.Н. Оценка эффективности метода поиска ассоциативных правил для задач обработки больших данных // *Известия ЮФУ. Технические науки*, 2020, no. 2 (212), с. 66-78.
3. Бочкарев П.В., Киреев В.С. Разработка ансамбля алгоритмов кластеризации на основе изменяющихся метрик расстояний // *В сборнике: Аналитика и управление данными в областях с ин-*

тенсивным использованием данных. XVIII Международная конференция, 2016, с. 69-73.

4. Гергиев И.Э., Данильянц Г.Г. Рынок онлайн-видеосервисов как фактор цифровизации и глобализации российской экономики // *Инновационная наука*, 2019, no. 11, с. 53-55.

5. Загиров А.Р., Пыжов С.И. Разработка модуля рекомендательной системы по выбору кинофильмов в онлайн-кинотеатре // *В книге: Новые направления научной мысли. Материалы Международной научно-практической конференции*, 2018, с. 167-169.

6. Казаков О.Л. Метод учета зависимости инвестиционных проектов при их ранжировании // *Научно-аналитический журнал: Актуальные проблемы социально-экономического развития России*, 2012, no. 1.

7. Казаков О.Л., Грушин О.И. Проведение ранжирования инвестиционных проектов с учетом их взаимного влияния друг на друга // *Управление экономиче-*

скими системами: электронный научный журнал, 2012, no. 7 (43).

8. Казаков О.Л., Петренко А.А. Выбор показателей для ранжирования инвестиционных проектов // *Управление экономическими системами: электронный научный журнал*, 2014, no. 1 (61).

9. Курейчик В.М., Сафроненкова И.Б. Метод формирования кластеров задач компоновки // *Известия ЮФУ. Технические науки*, 2018, no. 4 (198), с. 66-74.

10. Ларионцев М.М. BIG DATA в сфере культуры: тренды и проблемы // *Культурологический журнал*, 2020, no. 2 (40), с. 2.

11. Рабинович Я.И. Универсальная процедура построения множества Парето // *Журнал вычислительной математики и математической физики*, 2017, т. 57, no. 1, с. 30-48.

12. Юре Лесковец, Ананд Раджараман, Джеффри Д. Ульман. *Анализ больших наборов данных* / Пер. с англ. Слинкин А.А. Москва, ДМК Пресс, 2016.

DEVELOPING A RECOMMENDED SET OF MOVIES FOR ONLINE MOVIE THEATER CUSTOMERS

Kazakov Oleg Leonidovich¹, Dr. Sc. (Eng.), Prof.

Kazakov Leonid Olegovich², Head of the Analysis Department

¹ Moscow Polytechnic University, Bolshaya Semyonovskaya str., 38, Moscow, Russia, 107023; e-mail: olekazakov@gmail.com

² Neurosoft Limited Liability Company, Derbenevskaya embankment, 11, Moscow, Russia, 115114; e-mail: leonid.kazakov@gmail.com

Purpose: the article is devoted to the improvement of the methodology for forming a list of new films ordered by the interests of any particular client of an online cinema based on previously received expressions of interests of all clients for existing films. *Discussion:* compiling a set of films based on interests requires the use of mathematical and instrumental methods that take into account the features of metrics to determine the similarity of rows in matrices of unstructured preferences of large dimension. Therefore, it is relevant to search in the field of Big Data for new solutions and their combinations. *Results:* it is proposed to use the properties of Pareto sets to select clients who are close in interests and to rank them by proximity to the interests of each client using evaluation vectors. The method of solving an informal problem of large dimension developed in the article is distinguished by its originality. It will automate the preparation of recommendations and increase their objectivity and accuracy. In addition, the use of Big Data technology will ensure the self-adaptation of the proposed tool to the changing interests of customers and the ability to explain the advice provided.

Keywords: set (list) of films, interest of customers of the online cinema, placement as an ordered sequence, matrix as a set of vectors, ranking by proximity of interests, Pareto set, evaluative vector.

References

1. Berezina A.D. Razvitie rynka onlajn-videoservisov v ekosisteme cifrovoj ekonomiki [Development of the market of online video services in the ecosystem of the digital economy]. *Ekonomika i biznes: teoriya i praktika*, 2019, no. 10-1 (56), pp. 50-53. (In Russ.)
2. Bova V.V., Kuliev E.V., SHCHeglov S.N. Ocenka effektivnosti metoda poiska asociativnyh pravil dlya zadach obrabotki bol'shikh dannyh [Evaluation of the effectiveness of the method of searching for associative rules for big data processing tasks]. *Izvestiya YUFU. Tekhnicheskie nauki*, 2020, no. 2 (212), pp. 66-78. (In Russ.)
3. Bochkarev P.V., Kireev V.S. Razrabotka ansamblya algoritmov klasterizatsii na osnove izmenyayushchihsvyay metrik rasstoyanij [Development of an ensemble of clustering algorithms based on changing distance metrics]. *V sbornike: Analitika i upravlenie dannymi v oblastiakh s*

intensivnym ispol'zovaniem dannyh. XVIII mezhdunarodnaya konferenciya, 2016, pp. 69-73. (In Russ.)

4. Gergiev I.E., Danil'yanc G.G. Rynok onlajn-videoservisov kak faktor cifrovizacii i globalizacii rossijskoj ekonomiki [Market of online video services as a factor of digitalization and globalization of the Russian economy]. *Innovacionnaya nauka*, 2019, no. 11, pp. 53-55. (In Russ.)

5. Zagirov A.R., Pyzhov S.I. Razrabotka vyboru kinofil'mov v onlajn-kinoteatre [Development of a module of the recommendation system for the selection of films in an online cinema]. *V knige: Novye napravleniya nauchnoj mysli. Materialy Mezhdunarodnoj nauchno-prakticheskoy konferencii*, 2018, pp. 167-169. (In Russ.)

6. Kazakov O.L. Metod ucheta zavisimosti investicionnyh proektov pri ih ranzhirovanii [Method of accounting for the dependence of investment projects when ranking them]. *Nauchno-analiticheskij zhurnal: Aktual'nye problemy social'no-ekonomicheskogo razvitiya Rossii*, no. 1, 2012. (In Russ.)

7. Kazakov O.L., Grushin O.I. Provedenie ranzhirovaniya investicionnyh proektov s uchetom ih vzaimnogo vliyaniya drug na druga [Ranking of investment projects taking into account their mutual influence on each other]. *Upravlenie ekonomicheskimi*

sistemami: elektronnyj nauchnyj zhurnal, 2012, no. 7 (43). (In Russ.)

8. Kazakov O.L., Petrenko A.A. Vybory pokazatelej dlya ranzhirovaniya investicionnyh proektov [Selection of indicators for ranking investment projects]. *Upravlenie ekonomicheskimi sistemami: elektronnyj nauchnyj zhurnal*, 2014, no. 1 (61). (In Russ.)

9. Kurejchik V.M., Safronenkova I.B. Metod formirovaniya klasterov zadach komponovki [Method of forming clusters of layout tasks]. *Izvestiya YUFU. Tekhnicheskie nauki*, 2018, no. 4 (198), pp. 66-74. (In Russ.)

10. Larioncev M.M. BIG DATA v sfere kul'tury: trendy i problem [Big Data in the field of culture: trends and problems]. *Kul'turologicheskij zhurnal*, 2020, no. 2 (40), p. 2. (In Russ.)

11. Rabinovich YA.I. Universal'naya procedura postroeniya mnozhestva Pareto [Universal procedure for constructing a Pareto set]. *Zhurnal vychislitel'noj matematiki i matematicheskoy fiziki*, 2017, T. 57, no. 1, pp. 30-48. (In Russ.)

12. Yure Leskovec, Anand Radzharaman, Dzheffri D. Ul'man. *Analiz bol'shih naborov dannyh* [Analyze large data sets] / Per. s angl. Slinkin A.A. Moscow, DMK Press, 2016. (In Russ.)