

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 378

Халецкий Андрей Михайлович

«Система и моделирование организации по рекомендации курсов учащихся»

### **АВТОРЕФЕРАТ**

на соискание степени магистра технических наук  
по специальности 1-40 80 06 Искусственный интеллект

Научный руководитель  
Голенков Владимир Васильевич  
Доктор технических наук

---

Минск 2023

## КРАТКОЕ ВВЕДЕНИЕ

Данная работа посвящена улучшению и модифицированию существующих систем рекомендации. Учащиеся часто выбирают курсы, соответствующие их интересам и способностям. Это решение может сильно повлиять на их мотивацию и их будущие достижения, как правило, без какой-либо поддержки, которая могла бы помочь им в этом.

Среди следующих – неправильный выбор учебной программы, в результате которого учащийся не может в полной мере воспользоваться возможностью учиться или получать пользу. Необходимо помочь учащимся сделать оптимальный выбор по рекомендации курсов. Система рекомендаций может представлять собой набор инструментов и методов, используемых для предложения различных наборов знаний и курсов, когда количество доступных вариантов не позволяет получить дополнительную информацию.

Рекомендация курсов — непростая задача, поскольку рекомендация предмета должна быть не только интересна учащемуся, но и соответствовать его навыкам и учитывать любые ограничения (например, предварительные условия обучения). Современные рекомендательные системы обычно рекомендуют курсы на основе их содержания, независимо от того, как курс влияет на общую успеваемость учащихся. Кроме того, некоторые решения предполагают использование чрезмерно субъективных критериев при выработке рекомендаций, в то время как другие требуют слишком активного участия учащихся.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

### Цель и задачи исследования

*Целью* диссертации является исследование и описание эффективного алгоритма в рекомендательной системе в сфере образования для решения задачи прогнозирования предпочтений пользователей на основе их предыдущего поведения, позволяющего формировать рекомендации с положительным уровнем релевантности в условиях большого числа курсов и учащихся при неполной информации об их оценках и предпочтениях.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Изучить существующую литературу, источники и концепции, связанные с проблемой системы рекомендаций.
2. Изучить применение интеллектуального анализа данных в образовательном контексте.
3. Оценка существующих рекомендательных систем в образовательном контексте.
4. Описать применение сингулярного разложения матрицы в дополнение к пользовательской коллаборативной фильтрации.
5. Описать практический пример, основанный на реальных данных учащихся.
6. Оценить предложенный подход рекомендаций и соответствующие результаты.

*Объектом* исследования являются рекомендательные системы и сервисы в образовательной сфере.

*Предметом* исследования является интеллектуальное, математическое и программное обеспечение компьютерных систем для решения интеллектуальных задач системы рекомендации, алгоритмы, методы и программные платформы построения классификации данных, оценка быстродействия, эффективности и точности существующих и разрабатываемых алгоритмов.

Основной *гипотезой*, положенной в основу диссертации, является возможность повышения точности работы алгоритмов в процессе практического использования системы. Этот подход, основанный на уже существующих алгоритмах коллаборативной фильтрации, позволяет нам проанализировать основные элементы и характеристики, обобщить результаты и сравнить применимость различных моделей классификации, которые доступны на сегодняшний день.

### Новизна полученных результатов

Разработана рекомендательная система на основе совместной фильтрации пользователей, позволяющая формировать персонализированные выборки интересных образовательных мероприятий с учетом предыдущего образовательного опыта пользователя.

Данная разработанная система направлена на улучшения существующих рекомендательных систем, а также образовательных услуг. Также следует отметить, что примеры, представленные в качестве примеров, в предлагаемом подходе не являются искусственно созданными данными, что повышает эффективность информационной поддержки при отработке аспектов использования открытой семантической технологии на практике.

Актуальность темы обусловлена тем, что современные рекомендательные системы базируются на методах, которые строят рекомендации на основе содержания курсов, при этом не учитывают интересы и общую успеваемость учащихся.

### **Структура и объём диссертации**

Диссертация состоит из:

1. Введения
2. Четырёх глав
3. Заключения
4. списка использованных источников
5. списка публикаций автора

В первой главе исследуются существующие ключевые понятия и литература, связанные с проблематикой систем рекомендаций. Во второй главе описывается применение интеллектуального анализа данных в образовательном контексте. Также определяется метод оценки, на основании которого можно классифицировать системы рекомендаций в области сбора данных в сфере образования. Затем представляется текущее состояние рекомендаций в образовательной среде и доступные методы моделирования навыков и достижений учащихся. Данная глава завершается критическим взглядом на существующие работы, приближающиеся к рекомендациям в области образования. В третьей главе рассказывается о сингулярном разложении матриц и описывается его применение в дополнение к подходу, основанному на рекомендациях пользователей. Также будет показано, как использовать модель ASAP-классификаторов без данных об учащемся, в качестве подготовительного шага для рекомендации. В четвертой главе представлено тематическое исследование, содержащее данные учащегося, на которых будет тестироваться система, а также объяснение того, как оценивать данную рекомендацию и соответствующие результаты. Общий объём работы составляет 87 страниц, из которых основного текста 71 страниц, 32 рисунка, 7 таблиц, список использованных источников из 74 наименований на 6 страницах.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** подтверждается актуальность темы диссертационной работы, даётся краткая характеристика исследуемых вопросов и проблем, указаны актуальные задачи, решению которых посвящена диссертационная работа.

В **первой** главе описаны основные понятия систем рекомендаций, такие как основные проблемы рекомендаций, влияние социальных проблем, проблема обоснования оценки, оценку релевантности созданных рекомендаций, рассмотрение типов систем фильтрации рекомендаций.

Систему рекомендаций можно описать как набор инструментов и методов, используемых для предоставления предложений по различным объектам лицам, которые не имеют достаточной квалификации для оценки огромного числа доступных альтернатив.

Основная проблема рекомендательных систем заключается в том, что рекомендация, как правило, определяется не на всем оценочном пространстве, а лишь на некоторых его подмножествах. Это означает, что утилите, измеряющей полезность элементов, должна быть экстраполирована от известных оценок к неизвестным, либо с помощью эвристики, либо оценочных функций.

Системы рекомендаций предлагают три класса метрик точности рекомендательной системы:

1. *Метрики прогностической точности* измеряют, насколько предсказанные оценки системы рекомендаций близки к истинным оценкам пользователей;
2. *Метрики точности классификации* измеряют, сколько раз система имеет правильное суждение о том, является ли элемент хорошим для пользователя или нет;
3. *Метрики ранговой точности* измеряют способность алгоритма составлять упорядоченный список элементов, который соответствует потребностям пользователя.

Типы систем фильтрации рекомендаций:

1. Контентная фильтрация;
2. Коллаборативная фильтрация;
3. Фильтрация, основанная на знаниях;
4. Фильтрация на основании сокращения размерности;
5. Гибридные фильтрации.

Во **второй** главе описаны сбор данных об образовании, система оценок для классификации поведения рекомендательной системы, описание существующих систем и алгоритмов.

Можно использовать систему оценки для правильной классификации поведения рекомендательной системы при сборе образовательных данных в соответствии с фиксированным набором четко определенных параметров анализа:

*Рекомендуемые пункты.* Рекомендуемые пункты можно разделить на:

1. Учебные материалы;
2. *Обратная связь;*
3. *Группы;*
4. *Обязательные курсы.*

*Пользователи.* С учетом образовательного контекста можно разделить образовательную систему на три различных типа заинтересованных сторон:

1. *Учащиеся;*
2. *Обучающие;*
3. *Разработчики курсов.*

*Метод рекомендаций:* Система рекомендаций может быть разбита по категориям в соответствии со своей методикой следующим образом:

1. Коллаборативная фильтрация ориентированная на пользователей;
2. Предметная коллаборативная фильтрация;
3. Контентная фильтрация;
4. Фильтрация основанная на знаниях;
5. Фильтрация основанная на сокращении размерности.

Можно классифицировать каждую систему рекомендаций в соответствии с подходами по добыче данных, которые они используют для извлечения информации:

1. Неконтролируемые техники;
2. *Контролируемые техники.*

Выработка рекомендаций в образовательном контексте:

1. Рекомендационные системы учебных материалов;
2. Рекомендационные системы основанные на обратной связи;
3. Рекомендационные системы основанные на группах;
4. Рекомендационные системы, основанные на курсах;
5. Моделирование учащихся и прогнозирование успеваемости.

Представлены рисунки основных сервисов демонстрирующие базовые возможности работы системы.

В **третьей** главе рассматривается техника CRM и то, как она используется в рекомендации. Представлен подход по рекомендации курсов, которые максимально повышает вероятность хорошей академической успеваемости учащихся с учетом их предыдущих оценок, работа с неполными профилями оценок учащихся до выработки какой-либо рекомендации.

В рамках этой главы дается обоснование использования CPM, применение CPM к рекомендациям, градиентный подход. Описана методология рекомендаций применения коллаборативной фильтрации. Описана подготовка рекомендаций учащимся с неполными профилями оценок с использованием ASAP-классификаторов.

В **четвертой** главе представлено и рассмотрено применение разработанного алгоритма коллаборативной фильтрации рекомендаций с использованием ASAP-классификаторов. В качестве примера использованы данные пользователей. В результате обработки которых создаются потенциально релевантные выборки курсов. Также рассматривается использование ASAP-классификаторов, сравниваются различные вариации использования параметров, с которыми создаются рекомендации.

Для анализа прогнозов оценок курсов необходимо было использовать одну из общепринятых метрик точности статистического прогнозирования. Эти метрики предназначены для сравнения значения прогноза с фактическими реальными значениями.

Было принято использовать САО (среднеарифметическая абсолютная ошибка), так как она наиболее часто используется и проще всего интерпретируется.

Использование ASAP-классификаторов позволяет нам оценить данные о учащемся до процесса рекомендации, если профиль оценок учащегося неполный. Полученные результаты лучше, некоторых базовых показателей. Также диссертационная работа дает вариант повышения точности работы алгоритмов в процессе практического использования системы. Подход основан уже на существующих алгоритмах коллаборативной фильтрации и позволяет проанализировать основные элементы и характеристики, обобщить результаты применимости различных моделей классификации, которые доступны на сегодняшний день.

## ЗАКЛЮЧЕНИЕ

Выбор курсов для обучения – один из самых ответственных шагов студента – сложность в выборе оптимального курса, отсутствие поддержки и консультации для принятия решения. В диссертации описывается использование известных методов рекомендаций курсов обучающимся. Описана возможность совмещения успешных рекомендательных методик, такие как сингулярное разложение матрицы и пользовательская коллаборативная фильтрация, чтобы исследовать оценки учащихся, объяснять и предсказывать оценки учащихся по рекомендуемым курсам.

В результате рассмотрения существующей литературы и основных понятий были обозначены наиболее актуальные проблемы существующих подходов, а также выявлены их преимущества и недостатки.

Система оценивания была определена, на основе которой были классифицированы системы рекомендаций в области сбора данных в сфере образования, оценено текущее состояние рекомендательных систем в образовательной среде и существующие методы моделирования навыков и успеваемости учащихся.

Были описаны преимущества и недостатки применения сингулярного разложения матрицы в дополнение к пользовательской коллаборативной фильтрации и модели ASAP-классификаторов.

На основе тематического исследования, основанного на реальных данных учащихся, был оценен предложенный подход рекомендаций и показаны соответствующие результаты, основанные на различных исходных данных. Данный подход основан на сочетании и адаптации двух наиболее успешных методов рекомендаций в образовательном контексте □ использование разложения по единичному значению и пользовательская коллаборативная фильтрация. Обнаруженные скрытые функции используются для прогнозирования оценок курсов для каждого учащегося. Полученные результаты достигают низкого уровня средней арифметической ошибки в представленном примере. Такой подход позволяет определить, какие курсы являются актуальными для учащегося, поскольку рекомендации - это курсы, по которым учащиеся ставят схожие оценки. Предлагаемый подход предоставляет персональные рекомендации для разных учащихся, а не курсы с лучшими результатами.

Использование ASAP-классификаторов позволяет нам оценить данные об учащемся до процесса рекомендации, если профиль оценок учащегося неполный. Полученные результаты лучше, некоторых базовых показателей.

Может оказаться так, что только подгруппа курсов будет доступна в различных доменах или некоторые курсы могут потребовать сначала прийти другие. Следует изучить получение данных из разных источников информации для создания рекомендаций, специфичных для целевой области, принимая во внимание ряд академических факторов, которые помогают объяснить связи или ассоциации в отношении учащихся, поставивших оценки. Следует дополнительно изучать различные исследовательские возможности, когда данные о целевом учащемся ограничены.

Для уточнения рекомендаций необходимо не только рекомендовать курсы, соответствующие навыкам студентов, но и изучить содержательную составляющую для удовлетворения их потребностей.

Успеваемость учащегося связана с мотивацией, интересами, навыками, условиями работы, методами оценки, обучающими и любыми другими ресурсами, которые помогают ему во время обучения. Таким образом, существует множество дальнейших различных путей для изучения, будущие исследования могут быть объединены в одну систему рекомендаций курсов. Это лишь первый шаг в исследовании данной научной области и создании более точной системы рекомендаций.

Представленные в статье результаты экспериментальных исследований представляют интерес для проектирования новых и модернизации существующих систем рекомендаций, однако необходимо продолжать изучать разработки различных систем рекомендаций различных целевых областей.

## Список опубликованных работ

1. Халецкий А.М. Система оценки знаний учащихся с прогнозированием успеваемости : Информационные технологии и системы 2021 (ИТС 2021) = И74 Information Tehnologies and Systems 2021 (ITS 2021) : материалы международной научной конференции, Минск, Беларусь, 24 ноября 2021 г., С 103-104. / [Электронный ресурс]. — Режим доступа: URL: [https://its.bsuir.by/m/12\\_130111\\_1\\_157684.pdf](https://its.bsuir.by/m/12_130111_1_157684.pdf)
2. Халецкий А.М. Система рекомендаций в области образования и курсов : Информационные технологии и системы 2022 (ИТС 2022) = И74 Information Tehnologies and Systems 2022 (ITS 2022) : материалы международной научной конференции, – Минск : БГУИР, 2022., С 59-60. / [Электронный ресурс]. — Режим доступа: URL: [https://its.bsuir.by/m/12\\_130111\\_1\\_168455.pdf](https://its.bsuir.by/m/12_130111_1_168455.pdf)