

DOI: 10.15514/ISPRAS-2024-36(3)-15



Разработка алгоритма формирования команд ИТ-проектов на основе данных цифрового следа студентов

А.В. Мельникова, ORCID: 0009-0006-1011-5225 <a.v.melnikova@utmn.ru>

М.С. Воробьева, ORCID: 0000-0002-1508-4089 <m.s.vorobeva@utmn.ru>

Е.В. Егорова, ORCID: 0009-0005-2106-506X <egorovaelizaveta.study@gmail.com>

Е. Д. Чеканова, ORCID: 0009-0005-6731-8026 <lisachekanova@yandex.ru>

*Тюменский государственный университет,
Россия, 625003, г. Тюмень, ул. Володарского, д. 6.*

Аннотация. В статье рассматривается разработка алгоритма для формирования команд ИТ-проектов. Материалами послужили данные цифрового следа студентов ИТ-направления. Цифровым следом студента является постоянно пополняемый набор данных, включающий отчетные документы проектных дисциплин, промежуточные результаты по дисциплинам, практическим подготовкам. В работе приводится пример решения задачи формирования команд с применением графов. Был предложен алгоритм, в основе которого лежит графовая модель, позволяющая построить граф, отражающий взаимодействие студентов в прошлых проектах. На основе построенного графа формируются команды. Внутри графовой модели предложены два подхода формирования команд: на основе кластеризации вершин и с помощью обхода графа. Для определения лучшей команды строится граф связи студентов с текстовыми тегами, представляющими технологии, языки программирования, фреймворки и пр. Алгоритм был апробирован на данных студентов ИТ-направления Математическое обеспечение и администрирование информационных систем Школы компьютерных наук и требованиях к реальному проекту и протестирован на спонтанном распределении студентов по проектам в рамках дисциплины. С помощью алгоритма можно оценить на сколько удачно было разбиение. Создание эффективных команд играет ключевую роль в успешной реализации проектов, поэтому предложенный алгоритм может быть полезен для преподавателей и руководителей проектов в ИТ-сфере. Разработанный алгоритм планируется интегрировать в веб-сервис поиска исполнителей ИТ-проектов.

Ключевые слова: формирование команд, цифровой след, текстовые теги, студенты, проект, исполнители, команда, графы, спектральная кластеризация, поиск в глубину.

Для цитирования: Мельникова А.В., Воробьева М.С., Егорова Е.В., Чеканова Е. Д. Разработка алгоритма формирования команд ИТ-проектов на основе данных цифрового следа студентов. Труды ИСП РАН, том 36, вып. 3, 2024 г., стр. 213–224. DOI: 10.15514/ISPRAS-2024-36(3)-15.

Development of an Algorithm of the Formation of IT Project Teams Based on Data from the Digital Footprint of Students

A.V. Melnikova, ORCID: ORCID: 0009-0006-1011-5225 <a.v.melnikova@utmn.ru>

M.S. Vorobeve, ORCID: 0000-0002-1508-4089 <m.s.vorobeve@utmn.ru>

E.V. Egorova, ORCID: 0009-0005-2106-506X <egorovaelizaveta.study@gmail.com>

E.D. Chekanova, ORCID: 0009-0005-6731-8026 <lisachekanova@yandex.ru>

University of Tyumen,

6, Volodarskogo st., Tyumen, 625003, Russia.

Abstract. The article discusses the development of an algorithm for the formation of IT project teams. The materials were data from the digital footprint of IT students. The student's digital footprint is a constantly updated set of data, including accounting documents of project disciplines, intermediate results in disciplines, and practical training. The paper provides an example of solving the problem of forming teams using graphs. An algorithm based on a graph model has been proposed, which allows you to build a graph reflecting the interaction of students in past projects. Commands are formed based on the constructed graph. Two approaches to command formation are proposed inside the graph model: based on vertex clustering and using graph traversal. To determine the best team, a graph of student communication is built with text tags representing technologies, programming languages, frameworks, etc. The algorithm was tested on data from students of the IT department of Mathematical Support and Administration of Information Systems of the School of Computer Science and requirements for a real project and tested on the spontaneous distribution of students on projects within the discipline. Using the algorithm, you can estimate how successful the split was. The creation of effective teams plays a key role in the successful implementation of projects, therefore, the proposed algorithm can be useful for teachers and project managers in the IT field. The developed algorithm is planned to be integrated into the IT project executors search web service.

Keywords: team formation, digital footprint, skills, students, project, performers, team, graphs, spectral clustering, depth-first search.

For citation: Melnikova A.V., Vorobeve M.S., Egorova E.V., Chekanova E.D. Development of an algorithm of the formation of IT project teams based on data from the digital footprint of students. *Trudy ISP RAN/Proc. ISP RAS*, vol. 36, issue 3, 2024. pp. 213-224 (in Russian). DOI: 10.15514/ISPRAS-2024-36(3)-15.

1. Введение

В настоящее время все больше организаций ориентируются на реализацию проектов, требующих специализированных навыков и знаний. Однако процесс подбора команд для выполнения проектов остается одной из центральных проблем, с которой можно столкнуться. Студенты стремятся получить опыт работы с реальными проектами еще на этапе обучения, но многие из них встречаются с трудностями при поиске соответствующих проектов и образовании команд, которые успешно справятся с поставленными задачами.

Асанов А. З. и Мышкина И. Ю. в статье [1] полагают, что успешность нового проекта напрямую зависит от компетентности выбранных исполнителей. Существуют разные подходы к формированию команд, в которых учитывают различные характеристики: успешность проекта, образовательные результаты участников, индивидуальные достижения студентов и др. Но важной составляющей, которую необходимо учитывать при работе в ИТ-проекте, являются навыки участников и предыдущий опыт командной работы.

В современном мире эффективное формирование команд играет ключевую роль в успешной реализации проектов, особенно в области информационных технологий. В связи с проведенным анализом была выявлена потребность в разработке алгоритма для формирования команд на основе данных цифрового следа студентов. Цифровой след студента – постоянно пополняемый набор данных, включающий отчетные документы проектных дисциплин, промежуточные результаты по дисциплинам, практическим подготовкам и другие [2]. Современные методы анализа данных позволяют извлечь

объективную информацию из цифрового следа студентов для оценки их профессиональной компетентности, на основе которой возможна разработка алгоритма процесса формирования команд [3].

Цель работы – проектирование и разработка алгоритма для формирования команд ИТ-проектов. Материалами для исследования послужили данные цифрового следа студентов ИТ-направления Математическое обеспечение и администрирование информационных систем (МОАИС) Школы компьютерных наук (ШКН). В рамках выполнения работы исследуется возможность применения графов в проектном управлении для решения задачи, которая заключается в подборе такой команды студентов, чтобы было достигнуто максимальное соответствие требованиям проекта. Таким образом, предложенный алгоритм не только актуален для современной образовательной среды ИТ-области, но и представляет собой инновационный шаг в области автоматизации формирования команд. Данный алгоритм может быть полезен для преподавателей и руководителей проектов в сфере ИТ.

2. Обзор существующих решений

Для решения поставленной проблемы могут применяться разнообразные подходы: графовые модели, критерий справедливости, когнитивные карты и генетические алгоритмы, алгоритмы кластеризации для создания команд, обнаружение сообществ в социальных графах, обход графа в глубину и т. д., рассмотрим некоторые из них.

Сервис «Команда как сервис» [4] представляет собой инструмент, разработанный для эффективного подбора команды исполнителей. Данный сервис позволяет заказчику найти уже готовую команду для выполнения проекта, но не предусматривает автоматизированного процесса подбора команды, что может затруднять работу заинтересованных сторон. В свою очередь, зарубежная платформа OpenNTF [5] предоставляет возможность автоматического формирования команды на основе заданных параметров. Однако, использование данного сервиса требует регистрации на платформе и активного ее использования для наполнения профиля, что для нашей задачи является невозможным, так как наполнение датасета цифрового следа студентов происходит периодически (раз в семестр) и может привести к недостаточной информации для эффективного подбора команды. Профессиональная платформа для исследователей ResearchGate использует анализ публикаций и рекомендует пользователям релевантные проекты и экспертов в заданной области. Но применение этого ресурса нецелесообразно для нашей работы, потому что студентам не характерны высокая исследовательская активность, четко выраженные научные интересы и соответствующий уровень специфики данных, на которых базируется рассматриваемый сервис.

Основная цель процесса формирования команды – собрать группу участников с определенными навыками. В работе [6] представлен обзор библиотеки PyTFL в Python с открытым исходным кодом для формирования команд. Для создания проектной группы применяются алгоритмы на основе нейронных сетей. Инструментарий может решать различные подзадачи, включая генерацию графов совместной работы, обучение модели, оценку и прогнозирование. Однако данная библиотека имеет недостаточно подробную документацию, что усложняет ее применение.

В статье [7] Воробьева М. С., Мельникова А. В., Захаров С. В. представили подход формирования команды на основе критерия справедливости, в котором все студенты выбранного ИТ-направления разбиваются на группы для выполнения проектных работ. В работе используются данные для выделения навыков студентов и необходимых требований для проектов. Применение данного подхода в первоначальном виде невозможно для решения нашей цели, потому что при использовании данного подхода главной задачей является формирование команд с одинаковым уровнем знаний.

Асанов А. З., Мышкина И. Ю., Грудцына Л. Ю. в статье [8] исследуют возможность применения графовых моделей в проектном управлении для задачи формирования команды

и распределения работ проекта между потенциальными исполнителями. Авторы реализовали онтологию проекта и исследовали возможность построения векторных представлений элементов графовых моделей для распределения задач между исполнителями проекта.

В работе [9] для решения задачи формирования команд предлагается вариационная байесовская архитектура нейронной сети для подбора команд таких участников, которые сотрудничали друг с другом в прошлом. Преимущество предлагаемого авторами подхода заключается в том, что учитывается история сотрудничества ее участников, и улавливаются отношения между исполнителями и их набором навыков. Недостатком является ограничение, накладываемое на участников, которые сотрудничали друг с другом в прошлом, это может привести к созданию команд узкой специализации и исключению новых перспективных возможностей сотрудничества.

Команда будет работать слаженно, если участники активно взаимодействуют друг с другом и участвуют в совместных мероприятиях. В работе [10] предложен алгоритм обнаружения сообществ в социальных сетях, объединяющий узлы с низкой степенью связей с ближайшими узлами более высокой степени и основанный на методе сжатия графа.

В статье [11] Н. В. Гриневой приводятся результаты сравнительного анализа спектральных методов выделения сообществ на основе реальной сети, представляющей сотрудничество ученых в некотором институте, и рекомендации по применению методов для обнаружения структуры сложных сетей в зависимости от их свойств и целей разбиения на сообщества. Для задачи формирования студенческих команд необходимо найти подмножества графа, в котором наблюдаются тесные взаимосвязи между предполагаемыми сообществами, эффективнее использовать классический метод спектральной кластеризации, основанный на K-средних.

Авторы в исследовании [12] описывают важность формирования эффективной команды для успешного завершения новых проектов в компании. Основным методом, предложенным в статье, является алгоритм генетического программирования с подкреплением (RL-GP), который использует стратегии ансамбля популяций и суррогатные модели для ускорения обучения, а также эвристические правила, которые могут быть использованы при формировании проектных групп.

В работе [13] авторы используют библиометрические методы и сетевой анализ для изучения результатов исследований и сотрудничества в поддержку Целей устойчивого развития ООН (ЦУР). Обнаружено, что, хотя количество взаимодействий увеличивается со временем, большая часть сотрудничества в исследованиях, связанных с ЦУР, происходит только один раз. Авторы также установили, что повторная совместная работа происходит чаще на уровне институтов, чем на уровне авторов, но после более длительного периода времени.

В ходе исследования [14] было выявлено, что в небольших командах (от 2 до 4 участников) частое изменение состава приводит к снижению эффективности работы. В то же время команды среднего (5-8 участников) и крупного размера (более 8 участников) могут получить больше выгоды от обновления состава команды до того момента, пока доля новых участников не достигнет критической точки. Обнаружена обратная зависимость между уровнем новых участников в коллективе и уровнем цитируемости научных статей в различных областях знания и в разные периоды.

3. Постановка задачи

Дана предметная область X , представленная множеством текстовых тегов, множество наборов документов $D = \{d_1, \dots, d_b\}$, где $b \in \mathbb{N}, \forall d \cap X = \emptyset$, и множество кандидатов $A = \{a_1, \dots, a_m\}$, в котором каждый обладает набором документов $\hat{d}, \hat{d} \subseteq D, \forall a (\exists d \in D)$.

Из неструктурированных текстов документов для каждого кандидата a_j извлекается набор текстовых тегов $S_j = \{s_1, \dots, s_n\}$, где $s \in X, n \in \mathbb{N}, j \in [1; m]$.

Каждый набор текстовых тегов имеет вектор весов $W_j = \{w_1, \dots, w_n\}$, в котором значение веса w_i вычисляется на основе среднего значения нормализованных оценок за проекты с использованием соответствующего текстового тега s_i , где $w_i \in [0; 1], n \in N, i \in [1; n]$.

Пусть $P = \{p_1, \dots, p_e\}$ – множество проектов, в котором каждый проект p_u описывается набором необходимых тегов $R_u = \{r_1, \dots, r_t\}$, где $r \in X, e \in N, u \in [1; e]$. Каждый набор требуемых тегов имеет вектор значений $Z_u = \{z_1, \dots, z_t\}$, в котором каждое значение z_k обозначает уровень значимости тега r_k для выполнения проекта p_u , где $z_k \in [0; 1], t \in N, k \in [1; t]$. Для каждого проекта нужна команда с заданным количеством участников $q, q \in N, q \in [1; m]$.

Найти команду T^* для заданного проекта p_u , состоящую из q участников, для которой выполняется условие (1) и требуемые для проекта теги совпадают с набором тегов кандидатов.

$$T^* = \operatorname{argmax} \sum_{j=1}^m \sum_{k=1}^t \sum_{r_k=s_i}^n (z_k * w_{ji}) \quad (1)$$

Для решения задачи формирования команды T^* используются подходы, описанные в работах [7-8], в которых рассматривается применение графовых моделей и критерия справедливости. К существующим решениям добавляются весовые коэффициенты для тегов и требуемых тегов для проектов.

Для алгоритма формирования студенческих команд ИТ-проектов существуют ограничения: предметная область X – ИТ-сфера, количество участников в командах q определяется заказчиком, а веса Z для требуемых проекту тегов R задаются экспертом (куратором проекта), а также студенты A должны быть действующими на момент формирования команды, наборы документов D проверены экспертами (преподавателями) и оформлены по стандартному шаблону.

4. Разработка алгоритма процесса формирования команд

Формирование команд является ключевым процессом, определяющим успешность проекта. Команды, состоящие из квалифицированных участников, способны достигать значимых результатов и решать сложные задачи. Для процесса формирования команд был спроектирован алгоритм на основе данных о студентах, проектах и дисциплинах. Рассмотрим этапы работы алгоритма, представленные на схеме (рис. 1).

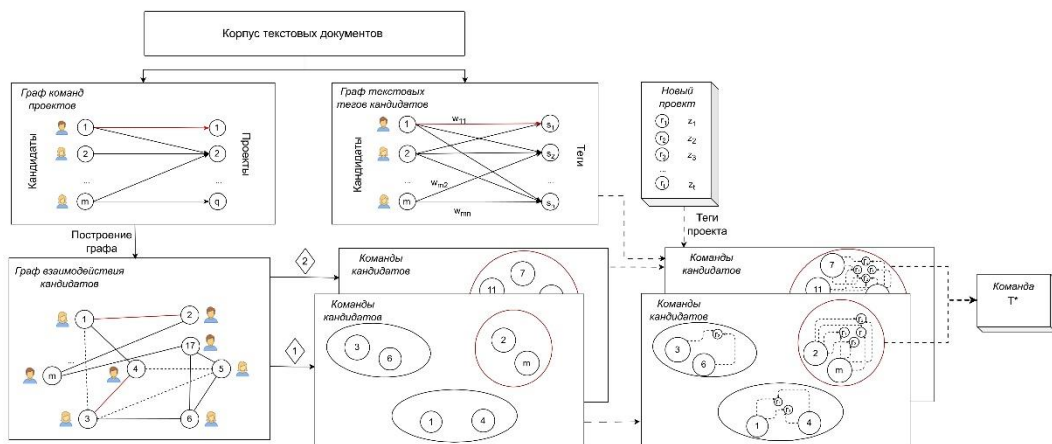


Рис. 1. Этапы работы алгоритма формирования команды студентов
Fig. 1. Stages of the algorithm for forming a team of students

Этап 1. На основе информации о составах команд из корпуса текстов происходит построение графа кандидатов, связанных участием в общих проектах, в котором вершинами являются элементы множества кандидатов, а ребрами - пары вершин с весами, где вес определяется, как количество общих документов для вершин. Таким образом, происходит представление взаимодействий студентов в общих проектных работах в виде графа.

Этап 2. На основе графа взаимодействия кандидатов, построенного на этапе 1, возможны два случая формирования команд с фиксированным или с заданным диапазоном числа участников: 1) применение обхода графа в глубину (Deep First Search) для формирования команд с участниками, которые работали вместе ранее; 2) использование спектральной кластеризации вершин графа. В результате получается множество команд.

Этап 3. На основе корпуса текстов выполняется построение графа текстовых тегов кандидатов, в котором вершины – идентификаторы кандидатов и тегов, а ребра – наличие тегов у кандидатов. Для каждого ребра между кандидатом и тегом определяется вес на основе данных об оценках тех документов, где кандидатом использовался этот тег.

Этап 4. Для каждой команды, полученной на этапе 2, из графа текстовых тегов кандидатов (этап 3) выделяются теги и их веса, с которыми связаны участники этих команд. Если у нескольких участников в команде есть один и тот же тег, то вес учитывается как среднее между весами для каждого из участников.

Этап 5. В результате происходит поиск команды для заданного проекта с определенными требованиями. Для каждой команды вычисляется значение согласно формуле (1), и выбирается команда с наибольшим значением как наиболее подходящая.

Данный алгоритм формирования команды применяется повторно для других проектов, но уже без учета студентов, назначенных на другие проекты.

5. Обработка исходных данных

Исходными данными выступают неструктурированные данные цифрового следа студентов, представленные в табл. 1.

Для создания корпуса текстов использовалось 114 наборов отчетных документов, полученных в ходе прохождения проектных дисциплин и практической подготовки (Управление проектами, Разработка систем обработки данных в предметных областях, Проектно-технологическая практика), которые были реализованы 96 студентами 2019–2020 года поступления. Большинство проектов были реализованы в командах от 2 до 5 человек, и для получения личных тегов студентов были использованы индивидуальные дневники и характеристики.

Для извлечения тегов студентов выполнены следующие шаги: корпус текстов был переведен на английский язык с использованием библиотеки Translate для Python, затем произведена очистка от знаков пунктуации и стоп-слов, а также лемматизация. Предобработка текстов отчетных документов происходила на языке программирования Python с применением библиотек Pandas, Numpy, nltk и re. После этого к текстам была применена открытая модель распознавания сущностей SkillNER, предварительно обученная на базе данных навыков в различных областях, в том числе и ИТ области. Для распознавания профессиональных навыков был использован соответствующий класс модели SkillNER.

Для оценки точности распознавания тегов проведен сравнительный анализ результатов работы методов извлечения тегов на основе регулярных выражений и SkillNER. Оценка осуществлялась путем расчета среднего значения коэффициента Жаккара на 12 размеченных текстах ВКР студентов, поступивших в 2019 году. Коэффициент Жаккара используется для измерения степени сходства между двумя множествами, в данном случае между размеченным набором текстовых тегов и результатами работы метода извлечения тегов.

Метод, использующий SkillNER, показал средний коэффициент Жаккара 0,752, а метод на основе регулярных выражений — 0,426. Исходя из этих результатов, для дальнейшей работы

был выбран метод извлечения текстовых тегов с применением SkillNER, так как он обеспечивает более точное распознавание сущностей.

В отчетных документах каждого студента было выделено в среднем по 22 текстовых тега, обозначающие технологии, языки программирования, фреймворки и другие аспекты.

Табл. 1. Перечень используемых данных

Table 1. List of data used

Тип данных	Источник	Количество документов	Формат
Описания проектов Большой Математической Мастерской (БММ) за 2022 и 2023 год	Официальный сайт БММ	83	XML
Рабочие программы дисциплин (РПД) образовательных программ направления МОиАИС	Официальный сайт ТюмГУ	18	PDF
Выпускные квалификационные работы (ВКР) студентов направления МОиАИС	Локальное хранилище кафедры Программного обеспечения ШКН ТюмГУ	27	PDF
Отчеты студентов МОиАИС по проектным дисциплинам	Локальное хранилище кафедры Программного обеспечения ШКН ТюмГУ	114	DOCX
Результаты учебного модуля студентов МОиАИС	Modeus (Сервис для управления учебным процессом и организации академической деятельности)	30	CSV
Промежуточные результаты по дисциплинам и практическим подготовкам	Электронный журнал преподавателей кафедры ШКН ТюмГУ	63	XLSX

6. Результаты

Рассмотрим пример процесса формирования команды для ИТ-проекта с использованием разработанного алгоритма. Заказчику для проекта «Технология NLP: сценарное проектирование и городская среда», реализуемого в рамках Большой Математической Мастерской, необходимо подобрать команду, состоящую из 2-3 студентов 2019 года поступления, причем для проекта определены текстовые теги и уровни значимости (табл. 2).

Табл. 2. Требования для проекта «Технология NLP: сценарное проектирование и городская среда»

Table 2. Requirements for the NLP Technology: Scenario Design and Urban Environment project

Тег	Значимость	Тег	Значимость
python	1.0	scipy	0.5
numpy	1.0	keras	0.4
pandas	1.0	nlTK	0.4
matplotlib	0.9	beautifulsoup	0.4
seaborn	0.8	opencv	0.3
tensorflow	0.6	flask	0.1

Примененный метод кластеризации выделил 29 кластеров, характеристики которых представлены в табл. 3. Каждый кластер состоит из определенного числа участников (от 1 до 3), которые ранее сотрудничали друг с другом в учебных проектах. Для каждой команды подсчитано общее количество тегов, которые выделены из корпуса документов участников. Участники 2 и 17 из кластера 1 использовали в своих отчетных документах 168 текстовых тегов, соответствующих ИТ-области (языки и технологии программирования, технологии баз данных, библиотеки, инструменты для анализа данных, системы управления версиями и др.).

Табл. 3. Характеристики кластеров, полученных после кластеризации

Table 3. Characteristics of clusters obtained after clustering

Кластер	ID участников	Количество тегов
1	[2, 17]	168
2	[40, 47]	58
3	[37, 49]	52
4	[7, 41]	22
5	[3, 14]	80
6	[16, 18, 35]	91
7	[19, 34]	26
8	[21, 29, 48]	49
...	...	...
29	[24, 50]	46

Этап 3. На основе текстовых тегов, извлеченных из наборов отчетных документов, строится неориентированный граф с 399 вершинами (59 вершин – студенты, 340 вершин – теги) и 1302 ребрами. Каждое ребро между студентом и тегом имеет вес, который определяется как среднее значение нормализованных оценок студента по проектам и дисциплинам, в которых использовался данный тег. Из работ каждого студента выделено в среднем 22 тега, причем максимальное количество тегов у одного студента – 102.

Этап 4. Для каждой команды, сформированной на этапе 2, из графа текстовых тегов были извлечены только те теги, которые соответствуют заданным требованиям проекта БММ «Технология NLP: сценарное проектирование и городская среда», описанным в табл. 2. Максимально для команд было выделено 8 тегов для данного проекта.

Этап 5. В результате для проекта «Технология NLP: сценарное проектирование и городская среда» была выбрана команда с номером кластера 1, в котором оказались студенты с идентификаторами 2 и 17. В табл. 4 представлены кластеры, в которых соответствие команды проекту больше 50%. Выбранная команда кластера 1 соответствует требованиям проекта на 69%.

Табл. 4. Итоговая таблица с оценками соответствия команд для заданного проекта

Table 4. Final table with team compliance scores for a given project

Кластер	T	Команда	Соответствие проекту
1	5.1	[2, 17]	69%
17	4.9	[12, 22]	67%
26	4.1	[46, 58]	56%
8	4.0	[21, 29, 48]	55%
5	3.9	[3, 14]	53%
25	3.8	[44, 52]	52%
24	3.7	[4, 15]	50%

После применения алгоритма для формирования команд к заданному проекту на данных студентов 2019 года поступления было проведено тестирование на спонтанном распределении 44 студентов 2020 года поступления по проектам в рамках дисциплины «Интеллектуальные системы поддержки принятия решений». В результате применения алгоритма для формирования команд проектов дисциплины у 65% команд средний балл увеличился.

7. Заключение

В результате на основе данных цифрового следа студентов направления МОАИС был предложен и разработан алгоритм для формирования команд ИТ-проектов на языке программирования Python с использованием библиотек для предобработки текстов отчетных документов (Pandas, Numpy, nltk, re), для выделения текстовых тегов, отражающих профессиональные навыки (Spacy, SkillNer), для кластеризации и расчета евклидова расстояния для выделения команд на графе взаимодействий студентов (scikit-learn, scipy), для получения данных о проектах БММ (BeautifulSoup), для визуализации графов (Gephi, NetworkX, Seaborn).

Алгоритм был протестирован на данных студентов МОАИС обучающихся в период с 2019 по 2024 годы. В результате были сделаны следующие выводы:

- Формирование студенческой команды осуществляется с учётом необходимых требований проекта, и определяется процент соответствия «команда-проект», который может изменяться в зависимости от исходных данных. Низкий уровень соответствия может возникнуть из-за недостаточной информированности о взаимодействии участников или несоответствия их навыков и компетенций требованиям проекта.
- Результаты работы могут быть полезны для преподавателей и руководителей образовательного процесса в ситуации, когда требуется распределить большую группу студентов с разным уровнем профессиональной подготовки по множеству проектов в рамках дисциплины, образовательного трека, практической подготовки, проектно-ориентированного мероприятия.

В дальнейшем планируется применение других методов выделения сообществ в социальных графах для формирования студенческих команд и, не нарушая общности предложенного подхода формирования команд, включить дополнительные параметры: экспертные оценки гибких студенческих навыков, динамику развития студента (личные достижения, вклад в групповой проект, оценка сложности реализованных задач), что позволит учитывать индивидуальные предпочтения каждого студента для формирования компетентной и слаженной команды.

Список литературы / References

- [1]. Асанов А. З., Мышкина И. Ю. Процедура формирования команды исполнителей проекта на основе когнитивных карт и генетических алгоритмов /А. З. Асанов, И. Ю. Мышкина // Проблемы управления и моделирования в сложных системах: Труды XXI Международной конференции. В 2-х томах, Самара, 03-06 сентября 2019 года / Под редакцией С.А. Никитова, Д.Е. Быкова, С.Ю. Боровика, Ю.Э. Плешивцевой, Том II. – Самара: Общество с ограниченной ответственностью "Офорт", 2019. – С. 354-358. / Asanov A.Z., Myshkina I.Yu. The procedure for forming a team of project performers based on cognitive maps and genetic algorithms. Problems of control and modeling in complex systems: Proceedings of the XXI International Conference. Samara: Ofort, 2019. P. 354-358. (in Russian) Доступно по ссылке: <https://elibrary.ru/item.asp?id=41101033>, 01.10.2023.
- [2]. Захарова И.Г., Боганюк Ю.В., Воробьева М.С., Павлова Е.А. Диагностика профессиональной компетентности студентов ИТ-направлений на основе данных цифрового следа // Информатика и образование. 2020. № 4. – С. 4-11. / Zakharova I.G., Boganyuk Yu.V., Vorobyova M.S., Pavlova E.A.

- Diagnosis of professional competence of IT students based on digital footprint data. *Informatics and Education*, 2020, no. 4, p. 4-11. (in Russian) <https://doi.org/10.32517/0234-0453-2020-35-4-4-11>.
- [3]. Захарова И.Г. Методы машинного обучения для информационного обеспечения управления профессиональным развитием студентов. *Образование и наука*. 2018;20(9):91-114. / Zakharova I.G. Machine Learning Methods of Providing Informational Management Support for Students' Professional Development. *The Education and science journal*. 2018;20(9):91-114. (in Russian) <https://doi.org/10.17853/1994-5639-2018-9-91-114>.
- [4]. Решение TEAM AS A SERVICE – Команда как сервис – Текст: электронный // *Effective technologies*: [сайт]. – 2022. Доступно по ссылке: <https://www.effective-group.ru/solutions/TAAS.html>, 04.10.2023.
- [5]. The Open Source Community for Collaboration Solutions – Текст: электронный // *openntf*: [сайт]. – 2021. Доступно по ссылке: <https://www.openntf.org/main.nsf>, 04.10.2023.
- [6]. Radin Hamidi Rad, Aabid Mitha. PyTFL: A Python-based Neural Team Formation Toolkit // Radin Hamidi Rad, Aabid Mitha, Hossein Fani, Mehdi Kargar, Jaroslaw Szlichta, Ebrahim Bagheri / In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM '21)*. Association for Computing Machinery 4716–4720. – 2021. <https://doi.org/10.1145/3459637.3481992>.
- [7]. Воробьева М. С., Мельникова А. В., Захаров С. В. Формирование студенческих команд для IT-проектов на основе отчетных документов по практике / М. С. Воробьева, А. В. Мельникова, С. В. Захаров // *Информатизация образования и методика электронного обучения : цифровые технологии в образовании : материалы VII Междунар. науч. конф. Красноярск, 19-22 сентября 2023 г. / под общ. ред. М.В. Носкова*. – Красноярск: Красноярский государственный педагогический университет им. В.П. Афанасьева, 2023. – С. 1024-1028. / Vorobeva M.S., Melnikova A.V., Zakharov S.V. Formation of student team for IT-projects on the basis of reporting documents on practice. *Informization of education and e-learning methods: digital technologies in education: materials VII International Scientific Conference Krasnoyarsk, September 19-22, 2023*, ed. by M.V. Noskov, Krasnoyarsk: Krasnoyarsk State Pedagogical University named after V.P. Afanasyev, 2023, p. 1024-1028. (in Russian) Доступно по ссылке: <https://www.elibrary.ru/item.asp?id=54778516>, 14.10.2023.
- [8]. Асанов А.З., Мышкина И.Ю., Грудцына Л.Ю. Применение графовых моделей в проектном управлении // *Онтология проектирования*. 2023. Т.13, №2(48). С.232-242. / Asanov AZ, Myshkina IY, Grudtsyna LYu. Application of graph models in project management. *Ontology of designing*. 2023; 13(2): 232-242. (in Russian) DOI: 10.18287/2223-9537-2023-13-2-232-242.
- [9]. Radin Hamidi Rad, Hossein Fani, Ebrahim Bagheri, Mehdi Kargar, Divesh Srivastava, and Jaroslaw Szlichta. 2023. A Variational Neural Architecture for Skill-based Team Formation. *ACM Trans. Inf. Syst.* 42, 1, Article 7 (January 2024), 28 pages. <https://doi.org/10.1145/3589762>.
- [10]. Zhaoa X., Lianga J., Wangab J. A community detection algorithm based on graph compression for large-scale social networks / Xingwang Zhaoa, Jiye Lianga, Jie Wangab // *Information Sciences*, No 551. – 04.2021. – P. 358–372. <https://doi.org/10.1016/j.ins.2020.10.057>.
- [11]. Гринева, Н. В. Сравнительный анализ спектральных методов выделения сообществ / Н. В. Гринева // *Актуальные проблемы прикладной математики, информатики и механики: сборник трудов Международной научной конференции, Воронеж, 12–14 декабря 2022 года / Воронежский государственный университет*. – Воронеж: Научно-исследовательские публикации, 2023. – С. 389-396. / Grineva, N. V. Comparative analysis of spectral methods of community allocation / N. V. Grineva // *Actual problems of applied mathematics, computer science and mechanics: proceedings of the International Scientific Conference, Voronezh, December 12-14, 2022 / Voronezh State University*. Voronezh: Scientific Research Publications, 2023. p. 389-396. (in Russian) Доступно по ссылке: <https://www.elibrary.ru/item.asp?id=54251945>, 14.10.2023.
- [12]. Yangyang Guo, Hao Wang. A reinforcement learning-assisted genetic programming algorithm for team formation problem considering person-job matching / Yangyang Guo, Hao Wang, Lei He, Witold Pedrycz, P. N. Suganthan, Yanjie Song // *arXiv preprint arXiv:2304.04022*. – 2023.
- [13]. Payumo Jane, He Guangming. Mapping Collaborations and Partnerships in SDG Research / Payumo Jane, He Guangming, Manjunatha Anusha Chintamani, Higgins Devin, Calvert Scout // *Frontiers in Research Metrics and Analytics*. – 2021. <https://doi.org/10.3389/frma.2020.612442>.
- [14]. Liu M. et al. Team formation and team performance: The balance between team freshness and repeat collaboration / Meijun Liu, Ajay Jaiswal, Yi Bu, Chao Min, Sijie Yang, Zhibo Liu, Daniel Acuña, Ying Ding // *Journal of Informetrics Volume 16 Issue 4*, 2022. <https://doi.org/10.1016/j.joi.2022.101337>.

Информация об авторах / Information about authors

Антонина Владимировна МЕЛЬНИКОВА – ассистент и аспирант кафедры программного обеспечения Школы компьютерных наук Тюменского государственного университета. Сфера научных интересов: обработка естественного языка, анализ данных, языки программирования.

Antonina Vladimirovna MELNIKOVA is an assistant at the Software Department of University of Tyumen, a postgraduate student. Research interests: natural language processing, data analysis, programming languages.

Марина Сергеевна ВОРОБЬЕВА – кандидат технических наук, доцент, заведующая кафедрой программного обеспечения Школы компьютерных наук Тюменского государственного университета. Сфера научных интересов: построение математических и информационных моделей, исследование методов и технологий машинного обучения, анализ цифрового следа студента.

Marina Sergeevna VOROBIEVA – Cand. Sci. (Tech.), Associate Professor, Head of the Software Department of University of Tyumen. Research interests: construction of mathematical and information models, research methods and machine learning technologies, methods for students' digital footprint analysis.

Елизавета Владимировна ЕГОРОВА – студент-бакалавр направления Математического обеспечения и администрирования информационных систем Тюменского государственного университета. Сфера научных интересов: анализ данных, машинное обучение, обработка естественного языка, глубокое обучение, большие данные, генеративные модели.

Elizaveta Vladimirovna EGOROVA is a bachelor's student in the field of Mathematical Support and Administration of Information Systems at University of Tyumen. Research interests: data analysis, machine learning, natural language processing, deep learning, big data, generative models.

Елизавета Дмитриевна ЧЕКАНОВА – студент-бакалавр направления Математического обеспечения и администрирования информационных систем Тюменского государственного университета. Сфера научных интересов: языки программирования, технологии программирования, машинное обучение, анализ данных.

Elizaveta Dmitrievna CHEKANOVA is a bachelor's student in the field of Mathematical Support and Administration of Information Systems at University of Tyumen. Research interests: programming languages, programming technologies, machine learning, data analysis.