

УДК 004.4

Д.М. Жевакин, С.Н. Широбокова

Южно-Российский государственный политехнический университет (НПИ)
имени М.И. Платова, Новочеркасск

ИНСТРУМЕНТАРИЙ FINDPROFESSION ДЛЯ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО ВЫБОРУ ПРОФЕССИИ И МЕСТА ТРУДОУСТРОЙСТВА

Целью проектного решения, рассмотренного в данной статье, является разработка инструментария, который позволяет проводить анализ требований рынка труда из различных информационных источников. Рассмотрены аспекты расчета комплексного рейтинга вакансии, позволяющего подобрать вакансии, наилучшим образом соответствующие требованиям пользователя и при этом находящиеся в пределах области поиска, радиус которого он задал. Разрабатываемая информационная система призвана помочь пользователям определить подходящую вакансию или подобрать свою будущую профессию, исходя из их личных предпочтений.

Ключевые слова: расчет расстояния по координатам; рейтингование вакансий; инструментарий FindProfession; выбор профессии; оценка быстродействия.

D.M. Zhevakin, S.N. Shirobokova

Platov South-Russian State Polytechnic University (NPI), Novocherkassk

FINDPROFESSION TOOLKIT FOR INFORMATION SUPPORT OF DECISION-MAKING ON THE CHOICE OF A PROFESSION AND LOCATION OF EMPLOYMENT

The purpose of the project solution discussed in this article is to develop tools that allow you to analyze the requirements of the labor market from various information sources. The aspects of calculating a comprehensive job rating that allows you to select vacancies that best meet the user's requirements and at the same time are within the search area, the radius of which he set. The developed information system is designed to help users determine a suitable vacancy or choose their future profession based on their personal preferences.

Keywords: distance calculation by coordinates; vacancy rating; FindProfession toolkit; choice of profession; performance evaluation.

Инструментарий FindProfession предназначен для информационной поддержки принятия решений различными категориями пользователей по выбору места трудоустройства, а также выбору профессии [1]. Общая архитектура приложения представлена на рис. 1. Для реализации поддержки принятия решений инструментарий получает и обрабатывает большие объемы информации с различных источников. В частности, источниками данных выступают порталы, которые предоставляют информацию об опубликованных работодателями вакансиях. Инструментарий агрегирует эти данные в рамках одной системы. Пользователь, при необходимости, может получить нужную информацию несколькими способами, например, через Telegram бота,

мобильное, десктопное или web-приложение. Обмен между приложениями, входящими в состав инструментария, производится в формате JSON, который удобен для парсинга и предоставления информации пользователю в итоговом виде.

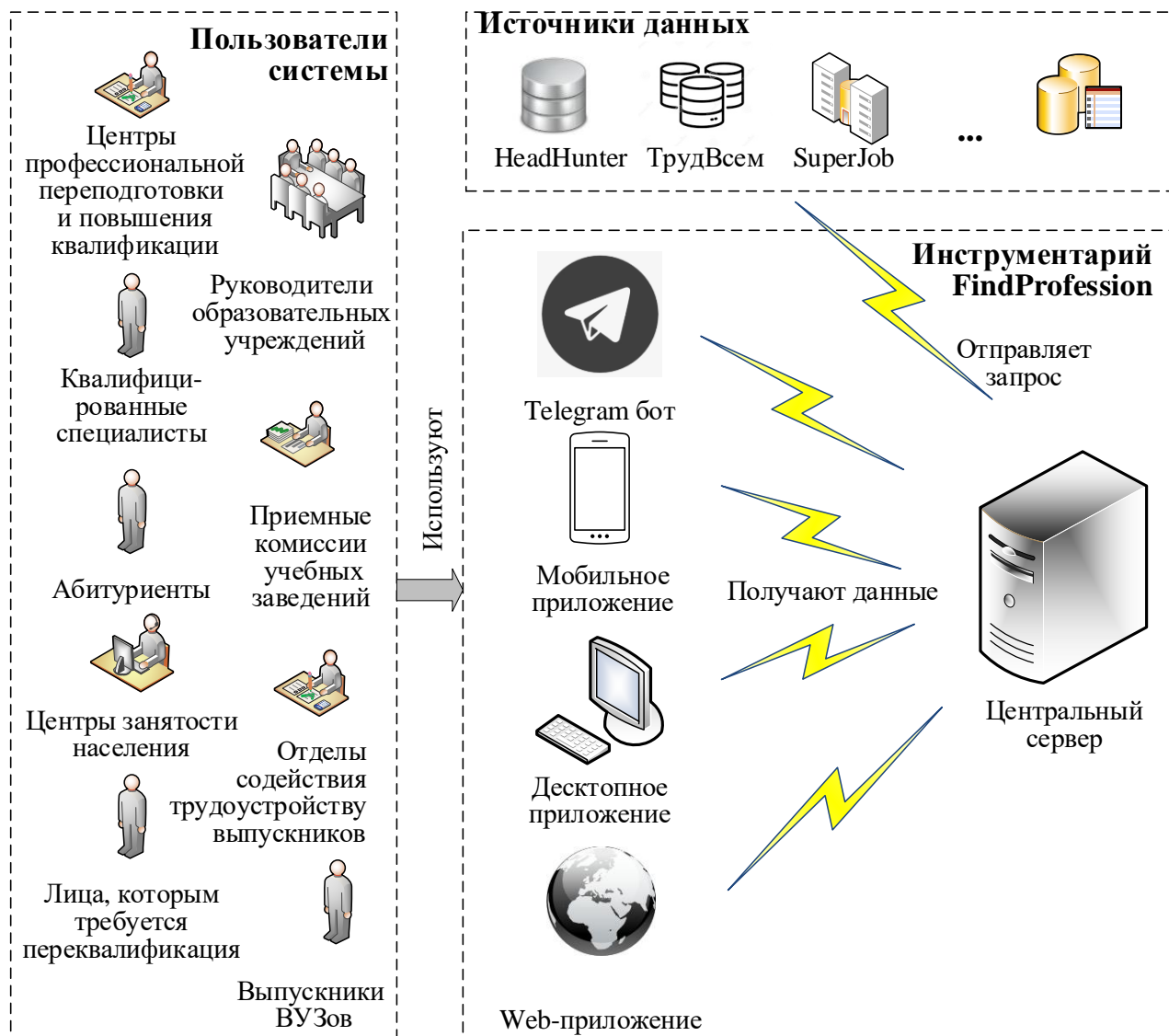


Рисунок 1 – Архитектура инструментария FindProfession

Одним из аспектов функционала инструментария является рейтингование вакансий. Подсчет рейтинга разделен на два этапа:

- 1) Расчет рейтинга при формировании SQL запроса.
- 2) Расчет дополнительного рейтинга на основе вхождения в область поиска (рис. 2).



Рисунок 2 – Поиск вакансий по области в указанном радиусе

В инструментарии реализован свой механизм для оценивания вакансий. Пользователь может задать предпочтительные настройки отбора по определенным полям (наименование профессии, регион, график работы, занятость, минимальная и максимальная заработные платы) для получения индивидуальной выборки из множества всех вакансий. Затем рассчитываются значения *Rating* для каждой вакансии. Для ограничения количества вакансий можно поставить условия на минимальный рейтинг. *ConstRating* – базовый рейтинг, заданный в инструментарии. Значение рейтинга для *j*-ой вакансии рассчитывается по следующей формуле:

$$Rating^j = Param^j + radius^j, j = \overline{1, n},$$

где *Rating^j* – общий рейтинг *j*-ой вакансии, *Param^j* – рейтинг, подсчитанный при формировании SQL запроса, *radius^j* – рейтинг, основанный на вхождении вакансии в область поиска, определенную пользователем. Значение *Param^j* в общем случае рассчитывается по следующей формуле:

$$Param^j = \sum_{i=1}^m \begin{cases} ConstRating * 1, \text{ если } param_i^j = UserParam_i^j \\ ConstRating * 0, \text{ если } param_i^j \neq UserParam_i^j \text{ или } UserParam_i^j = \emptyset \end{cases}$$

где *param_i^j* – *i*-ый параметр *j*-ой вакансии, *UserParam_i^j* – выбранное пользователем значение *i*-го параметра для *j*-ой вакансии.

Более подробно расчет рейтинга при формировании SQL запроса был описан в работе [2]. Дополнительный расчет по вхождению дополняет основной рейтинг и показывает, какие вакансии находятся непосредственно в том радиусе поиска, который указал пользователь. В дальнейшем на основе значения этого комплексного рейтинга производится сортировка вакансий.

Для того, чтобы определить какие вакансии, предлагаемые в каком-либо городе или в его окрестностях, попадают в область поиска по критерию

удаленности от местоположения пользователя, используется формула расчёта расстояния по координатам. Также в расчетах необходимо учитывать, что Земля не плоская, а имеет форму, приближенную к сфере. Исходя из этого, кратчайшим расстоянием между двумя точками является длина дуги круга между ними. Для расчета необходимо знать широту и долготу населенного пункта, обозначенного пользователем как его местоположение, широту и долготу, указанные в вакансии. Координаты населенных пунктов хранятся в справочнике «Населенные пункты». Тогда lat_1 и lng_1 – широта и долгота, указанные в вакансии, а lat_2 и lng_2 – координаты населенного пункта пользователя. Предварительно необходимо перевести координаты в радианы. Расстояние между двумя этими точками рассчитывается по формуле:

$$\Delta\sigma = \arctan\left(\frac{\sqrt{(\cos lat_2 \sin \Delta lng)^2 + (\cos lat_1 \sin lat_2 - \sin lat_1 \cos lat_2 \cos \Delta lng)^2}}{\sin lat_1 \sin lat_2 + \cos lat_1 \cos lat_2 \cos \Delta lng}\right),$$

где Δlng – абсолютная разность между долготой, указанной в вакансии, и долготой населенного пункта.

Итоговое расстояние равно:

$$S = r * \Delta\sigma,$$

где r – радиус Земли (используется среднее значение радиуса Земли, равное 6372.795 км, позволяющее получать результат расчета с точностью 99,5%).

Поскольку планируется хранить большое количество данных (суммарно с разных источников может быть более одного миллиона вакансий), то необходимо провести замеры времени на то, как инструментарий может работать с разным объемом информации при подсчете рейтинга вакансий.

При подсчете первоначального рейтинга во время формирования SQL запроса берутся данные из сущности «Вакансия», в которую ранее были загружены данные из информационных источников. Затем, когда необходимо рассчитать расстояние от выбранного пользователем населенного пункта до местоположения вакансии, необходимо произвести дополнительные вычисления, используя математические операции, которые не поддерживает SQL запрос. Отсюда появляется необходимость выделения дополнительного рейтинга, который вместе с первоначальным отображает те вакансии, которые не только подходят пользователю, но и находятся в непосредственной близости. Также нужно учитывать, что поскольку используется многоуровневая архитектура инструментария, то все данные отображаются на клиентской части, что может повлечь увеличение времени ожидания выполнения расчетов. Этим обуславливается необходимость провести замеры времени с различным количеством вакансий, чтобы определить наиболее подходящий и быстродействующий способ отображения данных для пользователя. Замеры производились со следующим количеством записей выборке: 1000, 5000, 10000 вакансий, находящихся в одинаковых условиях отбора по трем различным методам:

1) Когда не все данные передаются в SQL запросе и системе необходимо дополнительно обращаться к сущности «Вакансия» и получить данные по координатам. Постобработка данных, а именно формирование таблицы на форме

с результатами, формирование дополнительного рейтинга и сортировка производятся на стороне сервера, а клиенту возвращается результат.

2) Когда все данные, включая координаты, передаются в SQL запросе и системе не нужно обращаться к сущности «Вакансия» дополнительно. Постобработка данных производится на стороне сервера, а клиенту возвращается результат на форму.

3) Когда все данные, включая координаты, передаются в SQL запросе и системе не нужно обращаться к сущности «Вакансия» дополнительно. При этом частичная постобработка, а именно формирование дополнительного рейтинга, происходит на стороне сервера, а клиенту возвращается результат, где формируется таблица на форме и производится сортировка данных.

Так как в третьем способе идет частичная постобработка данных на клиентской части приложения, то целесообразно проанализировать, повлечет ли это увеличение времени обработки данных в целом. Результаты замеров представлены в табл. 1. Результаты, связанные с выполнением первой части расчета, а именно формирование рейтинга при запросе, представлены слева от разделителя, а расчет дополнительного рейтинга представлен справа от разделителя. Все значения отображены в секундах. Замеры проводились три раза над каждым набором записей, округление было до первой значащей цифры, итоги разделяются на две части – это время самого запроса и время постобработки данных. Для третьего способа данные постобработки представлены в двух частях, а именно выполнение замера на стороне сервера и на стороне клиента.

Таблица 1 – Замеры функционала подсчета рейтинга вакансий (сек.)

Способ	1000 записей	5000 записей	10000 записей
Первый	0,323 / 1,538	0,361 / 7,894	0,371 / 15,934
Второй	0,318 / 0,155	0,351 / 1,151	0,366 / 2,516
Третий	0,314 / 0,161+ 0,235	0,345 / 1,025 + 1,174	0,351 / 2,278 + 2,386

Проведенные эксперименты позволяют сделать вывод, что при формировании первой части комплексного рейтинга время выполнения запроса не существенно возрастает с ростом количества записей и, следовательно, не оказывает значительного влияния на выполнение общей обработки во всех трех способах. Поэтому время ожидания в большой мере зависит от постобработки данных. Формирование второй части комплексного рейтинга происходит на сервере, но при этом в первом способе происходит дополнительный подзапрос к сущности «Вакансия», чтобы получить координаты вакансии, а во втором способе эти координаты были получены во время запроса. Затем производится расчет второй части комплексного рейтинга, после этого данные отображаются на форме. В третьем случае после формирования второй части комплексного рейтинга данные были сформированы и отправлены клиенту, где в дальнейшем происходит их отображение и сортировка.

Эффективную отдачу от использования инструментария могут получить несколько категорий пользователей:

- абитуриенты ВУЗов и колледжей смогут оценить ситуацию с трудоустройством по выбираемой для поступления профессии;
- приемные комиссии образовательных учреждений могут использовать агрегированную информацию для обоснования востребованности профессий и соответствующих им направлений подготовки в профориентационных целях;
- люди в поисках новой работы или желающие получить переквалификацию могут получить необходимую информацию о вакансиях и обобщенные требования рынка труда по выбранной профессии;
- центрам занятости населения предоставляется возможность подбора вакансии для нетрудоустроенного лица с учетом его индивидуальных предпочтений, выделения с помощью семантического анализа набора обобщенных требований работодателей к специалисту данной профессии для отправки соискателя на соответствующие курсы переподготовки или повышения квалификации для получения недостающих компетенций;
- отделы содействия трудоустройству выпускников в ВУЗах получают полезный инструмент подбора рекомендаций выпускникам с учетом полученной ими квалификации не только в регионе их проживания (обучения), но и в других регионах, что позволит увеличить мобильность выпускников и удовлетворить потребность отраслей в молодых специалистах;
- центрам профессиональной переподготовки и повышения квалификации инструментарий позволит владеть актуальной информацией о потребностях рынка труда и своевременно актуализировать или открывать новые программы профессиональной переподготовки.

Список литературы

1. Жевакин Д.М., Широбокова С.Н., Диков М.Е. О проектировании информационной системы подбора подходящих вакансий // Моделирование и конструирование в образовательной среде: сб. матер. V Всеросс. (с междунар. участием) науч.-практ., методол. конф. для науч.-педагог. сообщества, 18 апреля 2020 года, г. Москва.– М.: Издательство ГБПОУ «Московский государственный образовательный комплекс», 2020.– С. 141-146.
2. Широбокова С.Н., Жевакин Д.М., Диков М.Е., Перекрестова Т.И. Формализованная модель формирования рейтинга вакансий и выделения требований рынка труда к ключевым профессиональным компетенциям // Перспективы науки.– 2020. – № 9. – С. 28–32.

Сведения об авторах

Жевакин Дмитрий Михайлович – магистрант 2 курса направления подготовки «Прикладная информатика», кафедры «Информационные и измерительные системы и технологии, Южно-Российский государственный политехнический университет (НПИ) имени М.И. Платова, Новочеркасск, email: dimas-zhevakin@yandex.ru.

Широбокова Светлана Николаевна – кандидат экономических наук, доцент кафедры «Информационные и измерительные системы и технологии, Южно-Российский государственный политехнический университет (НПИ) имени М.И. Платова, Новочеркасск, email: shirobokova_sn@mail.ru.

About the authors

Zhevakin Dmitry Mikhailovich – 2nd year undergraduate of the direction of training "Applied Informatics", department "Information and measuring systems and technologies, South-Russian State Polytechnic University (NPI) named after M.I. Platova, Novocherkassk, email: dimas-zhevakin@yandex.ru.

Shirobokova Svetlana Nikolaevna – Candidate of Economic Sciences, Associate Professor of the Department of Information and Measuring Systems and Technologies, South-Russian State Polytechnic University (NPI) named after M.I. Platova, Novocherkassk, email: shirobokova_sn@mail.ru.