

УДК 004.635

О РЕАЛИЗАЦИИ ПОДХОДА К ПОИСКУ ПОЛЬЗОВАТЕЛЕЙ ИЗ ОПРЕДЕЛЕННОГО ГОРОДА В СОЦИАЛЬНОЙ СЕТИ В УСЛОВИЯХ ОГРАНИЧЕНИЙ API МЕТОДОВ

М. Е. Диков, Д. М. Жевакин, С. Н. Широбокова

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

Ситуация с пандемией показала актуальность развития инструментов, позволяющих эффективно автоматизировать ряд этапов онлайн-работы с целевой аудиторией в социальных сетях. В статье описан алгоритм поиска страниц в определенном городе с учетом ограничений социальной сети на использование API метода `users.search`, а также представлена блок-схема алгоритма, реализованного в рамках развития разработанного ранее инструментария для автоматизированного поиска страниц пользователей в социальной сети ВКонтакте (`Vkuserssearch`).

Ключевые слова: мотивация школьников, IT-образование, профориентация, поиск страниц в сети, социальная сеть, API методы ВКонтакте, `users.search`.

ABOUT THE IMPLEMENTATION OF THE APPROACH TO FINDING USERS FROM A CERTAIN CITY IN A SOCIAL NETWORK UNDER THE RESTRICTIONS OF API METHODS

M. E. Dikov, D. M. Zhevakin, S. N. Shirobokova

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

The situation with the pandemic has shown the urgency of developing tools that can effectively automate a number of stages of online work with the target audience in social networks. The article describes the approach and algorithm for searching pages in a particular city, taking into account the restrictions of the social network on the use of the API of the `users.search` method, and also presents a block diagram of the algorithm implemented as part of the development of previously developed Tools for automated search of user pages in the social network Vkontakte (`VKUSERSSEARCH`).

Keywords: motivation of schoolchildren, IT education, career guidance, search for pages in the network, social network, VK API methods, `users.search`

Введение. Организационные условия проведения профориентационной работы с абитуриентами образовательных учреждений существенно изменились во время пандемии. Проведение традиционных очных мероприятий затруднено ограничительными мерами, и, как следствие, на первый план выходит применение инструментов дистанционной работы с целевой аудиторией, в частности, работа в социальных сетях. Важным этапом проведения профориентации в социальной сети является поиск целевого контингента (выпускников школ и колледжей). В работах [1–2] описан инструментарий, позволяющий автоматизировать ряд этапов работы с потенциальными абитуриентами.

Для поиска целевой аудитории требовалось охватить контингент пользователей заданной возрастной категории из определенного населенного пункта. В основе всех алгоритмов поиска в разработанном приложении лежит API метод `users.search`, который возвращает список пользователей в соответствии с заданным критерием [2].

Согласно документации, метод может вернуть только первые 1000 результатов поиска с заданным критерием поиска: «Обратите внимание, даже при использовании параметра `offset` для

получения информации доступны только первые 1000 результатов». В качестве основных параметров поиска выступают:

- возраст от;
- возраст до;
- город, в котором требуется искать страницы.

Данное ограничение не позволит получить все страницы пользователей по заданным критериям, т. к. существуют города, в которых большое количество страниц, удовлетворяющих заданным критериям поиска. Если целевая аудитория находится в диапазоне возраста от 16 до 19 лет, количество страниц может в десятки раз превышать возможности использования API метода поиска.

На основе выявленных ранее особенностей и ограничений на использование API методов сети можно сделать вывод, что разрабатываемый алгоритм должен содержать как можно больше запросов поиска, но не превышать частотные и количественные ограничения.

Алгоритм также должен использовать основные параметры либо каким-то образом их варьировать, использовать различные сочетания, другие параметры, которые не повлекут за собой получение некорректных данных (пользователей из другой целевой аудитории).

В качестве неизменяемого параметра требуется выбрать город, т. к. функция направлена на получение страниц потенциальных абитуриентов в определенном городе.

Для увеличения количества запросов, а, следовательно, увеличения количества полученных данных можно в качестве дополнительных параметров выбрать следующие значения:

- пол;
- день рождения.

Далее требуется определить все используемые параметры для метода поиска:

- city: город;
- age_from: возраст от;
- age_to: возраст до;
- sex: пол (1 — женский, 2 — мужской);
- birth_day: день рождения.

Рассмотрим пример запроса на получение списка страниц для города с идентификатором 101 метода users.search:

```
https://api.vk.com/method/users.search?access_token=587d345f57109862179683f22280485c0ec86fb9a4d335ff1ff9bca3fc3a2ffee&v=5.120&city=101&age_from=16&age_to=16&sex=1&birth_day=1&count=1000
```

Данный запрос вернет список страниц пользователей женского пола, возрастом 16 лет, которые родились в первый день месяца. Далее для получения большего количества данных требуется в данном запросе инкрементировать параметр birth_day. После того как получены все данные, изменяя параметр birth_day, требуется изменить параметр sex на 2 и произвести аналогичные действия с параметром birth_day. Затем требуется инкрементировать параметры age_from и age_to, вернуться на предыдущий этап и повторить действия.

Этапы алгоритма поиска следующие:

1. Установить начальные параметры для метода:
 - city = 101 (идентификатор города Новочеркаска);
 - age_from, age_to = 16;
 - sex = 1 (женский);
 - birth_day = 1.

2. Итеративно инкрементировать параметр birth_day на единицу по достижении значения 31 (максимальное количество дней в месяце).

3. При достижении значения 31 параметра `birth_day` установить параметр `sex = 2` и вернуться на этап 2.

4. После получения страниц пользователей женского и мужского пола, возрастом 16 лет, требуется инкрементировать значения `age_from`, `age_to` на единицу и вернуться на этап 2.

Теоретическое количество полученных данных (без учета конкретного города) можно рассчитать по следующей формуле:

$$N = C_m * C_s * C_a * 1000, \quad (1)$$

где C_m — количество дней в месяцах,

C_s — количество полов;

C_a — количество значений в диапазоне поиска по возрасту.

Подставим значения в формулу (1) для диапазона возраста 16–19 и получим максимально теоретическое значение полученных данных:

$$N = 31 * 2 * 4 * 1000 = 248000 \text{ страниц.}$$

Для того чтобы определить, достаточно ли такого количества данных для получения как можно большего количества страниц, осуществим запрос на получение списка страниц в городе Новочеркасске в диапазоне возраста от 16 до 19 с помощью программы Postman (рис. 1).

Результат запроса показывает, что количество страниц пользователей равно 8083, это означает, что в теории можно получить все страницы пользователей описанным выше алгоритмом.

Разработан алгоритм, с помощью которого возможно получить больше 1000 страниц пользователей с учетом использования текущих ограничений, касающихся количества возвращаемых страниц в результате поиска по основным параметрам запроса. Также стоит учесть, что в результатах поиска содержатся страницы, которые непригодны для дальнейшей обработки: неактивные, а также страницы, которые содержат сведения об университетах и высшем образовании. Их требуется удалить из конечного результата поиска. На рис. 2 представлена блок-схема алгоритма поиска.

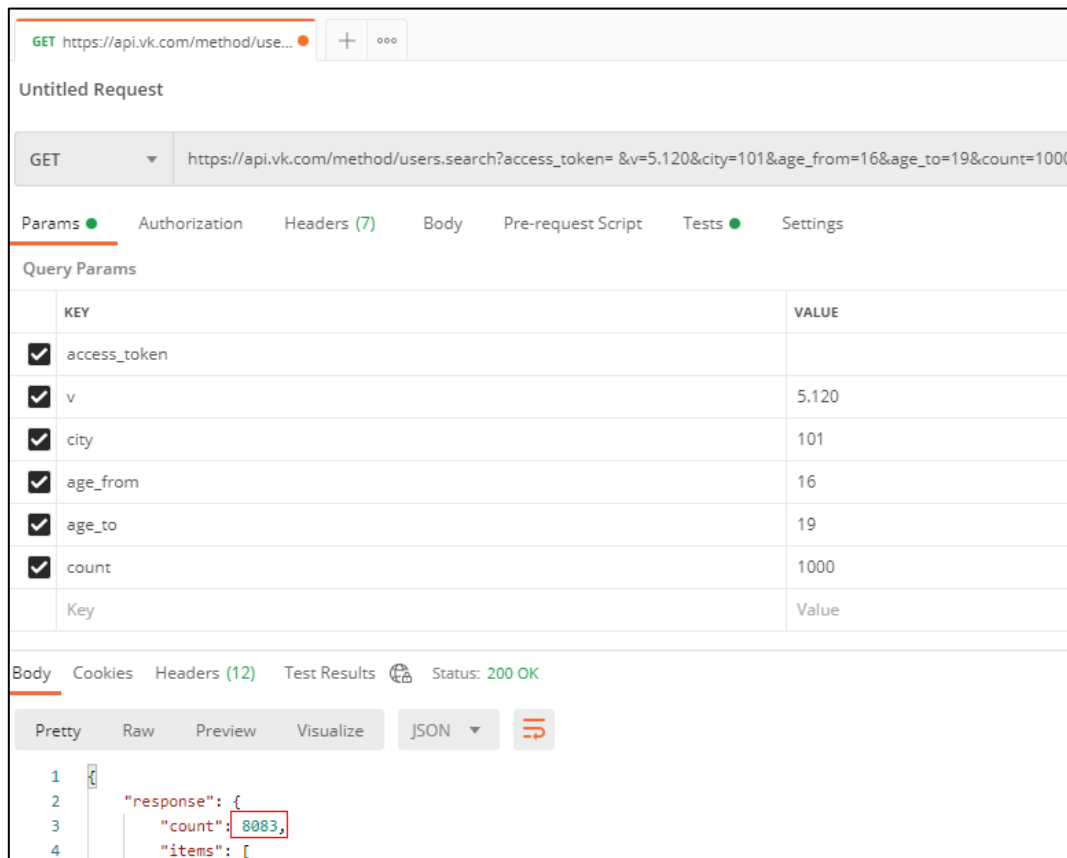


Рис. 1. Результат запроса метода `users.search` на получение списка пользователей в диапазоне возраста от 16 до 19 в городе Новочеркасске

Заключение. Таким образом, разработанный подход к поиску в определенном городе в диапазоне выбранного возраста позволяет получить более тысячи страниц активных пользователей необходимой целевой аудитории несмотря на количественные ограничения на результат api метода users.search.

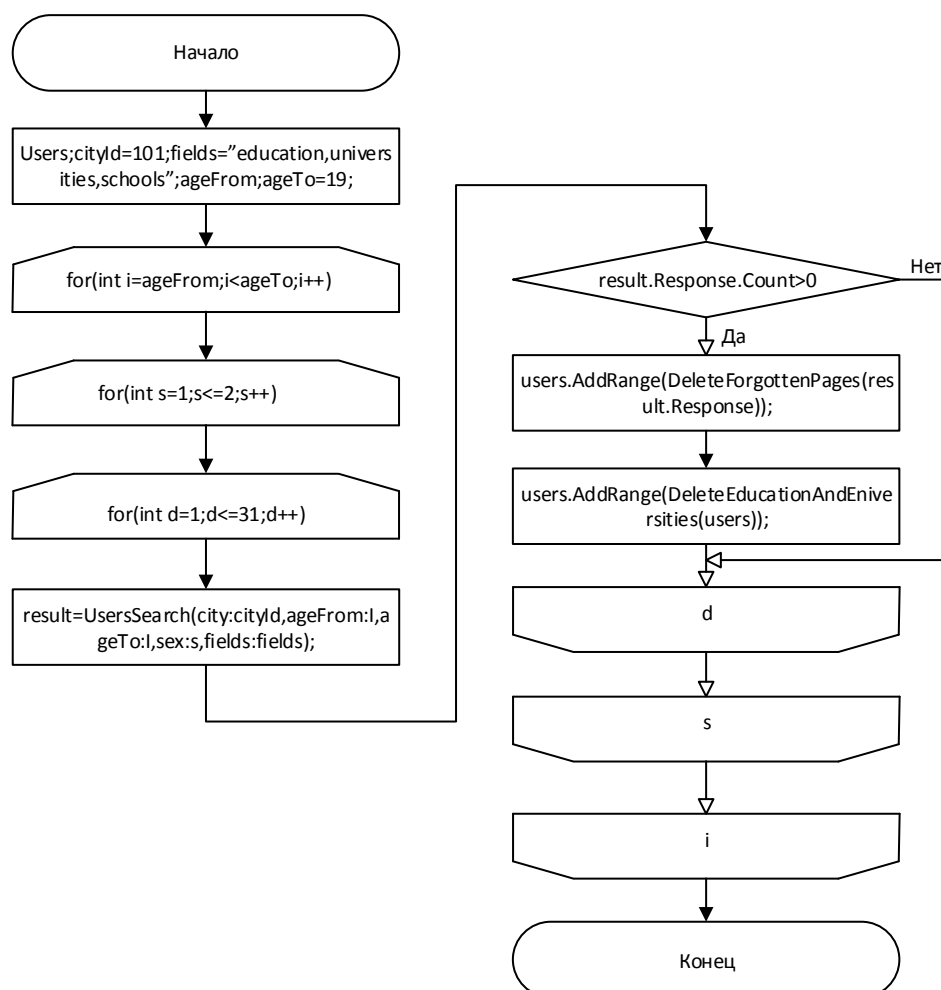


Рис. 2. Блок-схема алгоритма поиска пользователей определенного города в социальной сети

В условиях вынужденных временных ограничений на проведение массовых мероприятий в связи с пандемией применение подобных автоматизированных инструментов для работы в социальных сетях способствует переформатированию профориентационной работы образовательных учреждений и эффективному переходу ее преимущественно в формат онлайн.

Библиографический список

1. Свидетельство о госрегистрации программы для ЭВМ. Российская Федерация. Инструментарий для автоматизированного поиска страниц пользователей в социальной сети ВКонтакте (VKUSERSSEARCH) / С. Н. Широбокова, М. Е. Диков ; правообладатель Южно-Российский государственный политехнический университет (НПИ) имени М. И. Платова — № 2019617814 ; зарегистр. 04.06.2019 ; опублик. 20.06.2019. Бюл. № 6.

2. Диков, М. Е. Программное обеспечение для автоматизированного поиска страниц пользователей в социальной сети ВКонтакте / М. Е. Диков, О. Н. Сериков, С. Н. Широбокова // Мир компьютерных технологий: сборник статей Всеросс. науч.-техн. конф. студентов, аспирантов и молодых ученых. — Севастополь : СевГУ, 2019.— С. 235–239.



Об авторах:

Диков Максим Евгеньевич, магистрант кафедры «Информационные и измерительные системы и технологии» Южно-Российского государственного политехнического университета (НПИ) имени М. И. Платова (346428, РФ, г. Новочеркасск, ул. Просвещения, 132), maxjust@inbox.ru

Жевакин Дмитрий Михайлович, магистрант кафедры «Информационные и измерительные системы и технологии» Южно-Российского государственного политехнического университета (НПИ) имени М. И. Платова (346428, РФ, г. Новочеркасск, ул. Просвещения, 132), dimas-zhevakin@yandex.ru

Широбокова Светлана Николаевна, доцент кафедры «Информационные и измерительные системы и технологии» Южно-Российского государственного политехнического университета (НПИ) имени М. И. Платова (346428, РФ, г. Новочеркасск, ул. Просвещения, 132), кандидат экономических наук, доцент, shirobokova_sn@mail.ru