

Секция 2.
ОСНОВНЫЕ ТРЕНДЫ
ЭМПИРИЧЕСКИХ ИССЛЕДОВАНИЙ

УДК 316

**МЕТОДОЛОГИЯ ИДЕНТИФИКАЦИИ СЕЛЬСКИХ
СООБЩЕСТВ В СОЦИАЛЬНОЙ СЕТИ «ВКОНТАКТЕ»¹**

П. А. Басина, Д. О. Дунаева

*Центр прикладного анализа больших данных
Томского государственного университета
г. Томск, Россия*

E-mail: polya.basina@yandex.ru, ddo@data.tsu.ru

Цифровые следы пользователей в социальных сетях являются ценным источником для изучения сельских территорий, поскольку экономят временные и материальные ресурсы исследователя, а также открывают возможности для изучения сельских жителей как социальной группы. В статье представлена методология идентификации сообществ сельских поселений в социальной сети «ВКонтакте» и сообществ, связанных с тематикой сельской жизни.

Ключевые слова: цифровые следы, «ВКонтакте», сельские сообщества, сельская местность, социальные сети.

**METHODOLOGY OF IDENTIFICATION
OF RURAL COMMUNITIES
IN THE «VKONTAKTE» SOCIAL NETWORK**

P. A. Basina, D. O. Dunaeva

Digital traces of users in social networks are a valuable source for studying rural areas, since they save time and material resources of the researcher, as well as open up opportunities for studying rural residents as a social group. The article presents a methodology for identifying communities of rural settlements in the «VKontakte» social network and communities related to rural life.

¹ Исследование выполнено при поддержке РФФИ, проект № 20-011-00391.

Keywords: digital traces, «VKontakte», rural communities, rural area, social network.

В условиях цифровизации и информатизации общества одним из ключевых источников информации являются цифровые следы пользователей в Интернете, которые применяют для различных задач в коммерческом и научном секторе. Например, в работе [3; 7] данные цифровых следов использованы для оценки качества жизни населения с применением алгоритмов машинного обучения. При этом следует отметить, что цифровые следы как источник социологической информации обладают своими достоинствами и ограничениями в сравнении с традиционными методами [3].

Изучение жизнедеятельности в сельской местности является на сегодняшний день актуальной проблемой, поскольку в целом по стране наблюдается отрицательная динамика численности сельского населения [6], и государственные органы активно разрабатывают меры эффективной долгосрочной политики в отношении устойчивого развития сельских территорий в различных регионах России [1]. Фокус работы сосредоточен на разработке методологии идентификации в социальной сети «VKontakte» сообществ, отражающих информацию о поведенческих моделях сельской жизни и типичных особенностях и проблемах сельских территорий. В частности, объект исследования представляют целевые группы сельских жителей и потенциальных переселенцев в сельскую местность в зависимости от их интересов и ценностей. Представленная методология апробирована для 10 регионов, выбор которых обусловлен показателями миграции по данным Росстата [4]: в Кировской, Тамбовской и Саратовской областях отмечается значительный отток населения из сельских территорий; республику Башкортостан, Свердловскую область и Ставропольский край можно считать нейтральными регионами, где существенного оттока населения из сел не наблюдается; в Краснодарском крае, Самарской, Белгородской и Псковской областях отмечается значительная положительная миграция на сельские территории.

При разработке методологии мы придерживались гипотезы, что сельские сообщества в социальной сети «VKontakte» могут

быть представлены в явной и неявной формах. Явные сообщества – это сообщества сельских поселений конкретных территориальных единиц, а неявные – это тематические сообщества, транслирующие ценности и особенности сельского образа жизни.

Для идентификации неявных сельских сообществ был проведен анализ тематических Интернет-блогов и выдвинуты гипотезы о целевых группах сельских жителей и потенциальных переселенцах в сельскую местность в зависимости от их интересов и ценностей. Далее был составлен список маркерных слов по каждой из выдвинутых гипотез, на основании которого целевым способом отобраны сообщества, освещающие вопросы, связанные с переездом в сельскую местность или особенностями жизнедеятельности в ней. В результате был сформирован реестр сельских сообществ неявной формы, который включает в себя 200 сообществ. В реестр были включены только те сообщества, которые публиковали свежие посты в 2021 году, имели подписчиков из интересующих регионов, и тематика их контента соответствовала выдвинутым гипотезам.

Для идентификации явных сельских сообществ был разработан следующий алгоритм поиска. Во-первых, был сформирован реестр региональных сообществ социальной сети «ВКонтакте» [2, с. 90], затем из этого реестра были отобраны сообщества регионов исследования, и на основании списка маркеров жизни в сельской местности сформирована первая часть реестра явных сельских сообществ, в который вошли только те сообщества, в названии или описании которых встречались маркерные слова. Во-вторых, был составлен полный список всех населенных пунктов сельской местности исследуемых территорий, на основании которого были отобраны сообщества, в названии или описании которых встречались географические наименования сельских поселений регионов. Полученные реестры были объединены и очищены от дубликатов: сформированный реестр сообществ сельских поселений включает в себя 4855 сообществ.

Таким образом, с помощью разработанной методологии был составлен реестр региональных сообществ социальной сети «ВКонтакте», связанных с тематикой сельской жизни. Выделены 2 тематические группы сообществ: 200 неявных сообществ по ги-

потезам о пользователях, которые потенциально могли бы рассматривать переезд в сельскую местность, и 4855 явных сообществ сельских поселений регионов исследования. Полученная база данных включает в себя следующую информацию: название и описание сообщества, количество подписчиков, регион, ссылку на сообщество, его тематику, статус, тип, информацию об открытости, и может быть использована в качестве источника данных для исследований различных характеристик сельских жителей и особенностей жизни в сельской местности. База доступна по ссылке [5].

Список использованных источников

1. Проблемы и перспективы социально-экономического развития сельских территорий: региональный аспект / под ред. Н.М. Харитонов. М.: Издание Государственной Думы, 2021. 320 с.
2. Щекотин Е.В., Гойко В.Л., Басина П.А., Бакулин В.В. Использование машинного обучения для изучения качества жизни населения: методологические аспекты // Цифровая социология. 2022. Т. 5, № 1. С. 87–97
3. Щекотин Е. В., Коварж Г. Ю., Гойко В.Л., Петров Е. Ю., Бакулин В. В. Оценка качества жизни населения регионов РФ на основе цифровых данных: методологические аспекты // Векторы благополучия: экономика и социум. 2020. № 3(38). С. 138-156.
4. Миграция населения в России по регионам: коэффициент и численность/ [Электронный ресурс]. URL: <http://www.statdata.ru/migraciya-naseleniya-v-rossii-po-regionam> дата обращения: 28.03.2022)
5. Репозиторий GitHub Методология идентификации сельских сообществ в социальной сети ВКонтакте [Электронный ресурс]. URL: <https://github.com/datacentr/Methodology-of-identification-of-rural-communities-in-the-social-network-Vkontakte>
6. Численность и состав населения РФ / [Электронный ресурс]. URL: <https://rosstat.gov.ru/folder/12781#> (дата обращения: 28.03.2022)
7. Shchekotin E., Myagkov M., Goiko V., Kashpur V. Digital methods of analysis of subjective quality of life: case of Russian Regions // Administrative si Management Public. – 2021. – № 36. – P. 25-48.