

Выявление активных пользователей групп районных администраций Санкт-Петербурга в социальной сети ВКонтакте

Б.А. Низомутдинов, Л.А. Видясова

Университет ИТМО

boris-wels@yandex.ru, bershadszkaya.lyudmila@gmail.com

Аннотация

В статье представлены результаты выявления активных пользователей, оставляющих комментарии в 18 официальных группах районных администраций Петербурга ВКонтакте. В качестве основного метода исследования был использован парсинг социальных сетей и статистический анализ данных. По результатам анализа были выделены активные пользователи в городских сообществах, оставляющие большинство комментариев. На основе автоматизированного анализа были выделены 33 активных пользователя городских сообществ, оставляющие больше всего комментариев под постами районных администраций. Анализ социально-демографических характеристик группы пользователей позволил сделать некоторые предложения о формировании ядра активистов из людей среднего (40+) и старшего возраста (60+) с преобладанием женщин, а также активистов, занятых в сфере образования.

Ключевые слова: социальные сети, общественные активисты, активные пользователи, лидеры мнений, парсинг, анализ социальных сетей

Библиографическая ссылка: Низомутдинов Б.А., Видясова Л.А. Выявление активных пользователей групп районных администраций Санкт-Петербурга в социальной сети ВКонтакте // Государство и граждане в электронной среде. Выпуск 5 (Труды XXIV Международной объединенной научной конференции «Интернет и современное общество», IMS-2021, Санкт-Петербург, 24 – 26 июня 2021 г. Сборник научных статей). — СПб.: Университет ИТМО, 2021. С. 126 – 137. DOI: 10.17586/2541-979X-2021-5-126-137.

1. Введение

В последние годы активное вовлечение горожан в развитие городской среды является одним из ключевых приоритетов городского управления на всех уровнях. Создаются различные онлайн площадки для взаимодействия представителей органов власти и населения: порталы решения городских проблем, сервисы электронных обращений и петиций, официальные группы в социальных сетях.

За последние годы социальные сети стали мощным инструментом распространения информации, влияющей на мнения и поведение людей. Популярные аккаунты в социальных сетях охватывают широкую аудиторию, что приводит к появлению в сети профессиональных влиятельных лиц, способных влиять на поведение и выбор потребителей.

В современной медиа-среде лидеры мнений и влиятельные пользователи играют все более важную роль в регулировании доступа к онлайн-контенту [1]. При оценке достоверности информации люди склонны полагаться на свои социальные сети. Исследование поведения людей, читающих новости в Интернете, показывает,

что при определении достоверности новостной статьи чрезвычайно важно, кто ею делится [2]. Люди склонны полагаться на информацию, которую рекомендует человек, которому они доверяют, до такой степени, что одна и та же новость воспринимается как более заслуживающая доверия, если она рекомендована другом из социальной сети, чем при чтении ее на исходном новостном сайте [2].

Официальные площадки для обсуждения вопросов городского развития в сети публикуют новости и информацию по ключевым проблемам, а также собирают комментарии от жителей города. Для того, чтобы оценить реагирование на публикуемый контент, следует учитывать специфику таких реакций, а также возможности определения ключевых пользователей, активно выражающих гражданскую позицию и оказывать влияние на общественное мнение по поднятым вопросам.

В данной статье представлены результаты исследования, проведенного с целью апробации механизма выявления активных пользователей среди подписчиков официальных групп администраций районов Санкт-Петербурга в социальной сети ВКонтакте. Структура статьи состоит из пяти частей. Во введении обоснована актуальность проведения исследования. В секции 2 приводится анализ исследований в предметной области изучения активных пользователей. Секция 3 посвящена описанию методологии исследования. В секции 4 приведены основные результаты и анализ полученных данных. Секция 5 систематизирует выводы по результатам работы и описывает перспективы для дальнейших исследований.

2. Анализ развития предметной области

Впервые, термин «лидеры мнений» был использован Полом Лазарсфельдом в 1955 г. для обозначения людей, имеющих авторитет в какой-либо социальной группе и способных оказать на нее влияние в процессе обсуждения сообщений [3]. В современной научной литературе лидеры мнений определяются как заинтересованные и компетентные люди, мнения которых сторонники считают честными и заслуживающими доверия [2]. Лидеры мнений уделяют пристальное внимание рассматриваемой теме, часто обсуждают ее с аудиторией, убеждая других принять то или иное мнение [4]. В то же время, лидеры общественного мнения не всегда занимают влиятельную позицию. Чаще всего, они занимают такое же социальное положение, что и те, на кого они стремятся повлиять, но воспринимаются как хорошо осведомленные по обсуждаемой теме [3; 4]. Лидеры общественного мнения обладают несколькими атрибутами, связанными с личностью, опытом или сетями, включая доверие, знания, энтузиазм, взаимосвязь и центральность [5; 6]. Размер сети обычно является хорошим показателем популярности, но сам по себе не показывает, насколько влиятельным является пользователь: популярные учетные записи в социальных сетях не обязательно вызывают значительное взаимодействие с точки зрения вовлеченности аудитории [7].

Одной из основных сфер, в которой часто наблюдается влияние на поведение, является политическая коммуникация в Интернете. В социальных сетях люди все чаще сталкиваются с политической информацией и мнениями людей в их сетях, что создает потенциальное пространство для политического убеждения [8]. Социальный контекст и особенно сообщения, полученные от группы сверстников, играют ключевую роль в формировании политических взглядов [8].

Социальные сети имеют огромные преимущества для проведения политических кампаний, например, путем повышения осведомленности людей и обсуждения соответствующих тем [4]. В то же время, социальные сети также могут способствовать распространению дезинформации, что может иметь серьезные социальные последствия [9]. Например, распространение вредоносного контента с целью повлиять на общественное мнение и решения людей при голосовании в преддверии выборов [10].

В связи с этим, особый интерес приобретает изучение механизмов онлайн-влияния и распространения информации.

Идентификация лидеров мнений стала важным предметом исследований в научной среде, в связи с чем появились различные методы выявления таких лиц [11 – 13]. В последнее время исследователи начали использовать крупномасштабные наборы сетевых данных для определения влияющего поведения. Для выявления онлайн-влияния использовались несколько показателей, таких как распространение контента [5; 7], качество сети [11] и количество подписчиков. и упоминания, содержащие имя пользователя [7].

Некоторые исследования сделали прогнозы влияния на поведение на основе предыдущей деятельности пользователей [12]. Также были определены качественные индикаторы влияния, которые выявляют типичные поведенческие модели влиятельных лиц на основе естественного языка, который они используют в обсуждениях в социальных сетях [13; 14]. Однако, несмотря на большое разнообразие новых методов исследования, выявление влиятельных лиц - непростая задача. Исследования показали, что влиятельные лица встречаются редко: в большинстве разговоров есть только один или ноль влиятельных лиц [14], а объекты влияния на поведение, подверженные влиянию, встречаются еще реже [15].

Некоторые критические пробелы существуют в существующих исследованиях, касающихся обнаружения влияния. Используемые методы могут обнаруживать влиятельных лиц только после того, как они покажут очевидные признаки влияния на других [12]. Также исследователи расходятся во мнениях относительно традиционной модели поведенческого заражения, согласно которой влиятельные люди являются движущей силой влияния сверстников. Есть также данные, позволяющие предположить, что восприимчивость к влиянию является ключевой чертой социального заражения [15]. Кроме того, состав сетей влиятельных и восприимчивых людей играет роль в успешном влиянии [11].

Проведенный обзор научных исследований подчеркивает важность разработки методов выявления лидеров мнений и их апробации в среде городских сообществ.

3. Сбор информации

Исследование было проведено в апреле 2021 г. В качестве объекта исследования были использованы 18 официальных групп администраций районов Санкт-Петербурга в социальной сети ВКонтакте (Таблица 1). По результатам анализа было выявлено, что сообщества имеют разный стиль организации, оформления и подхода к ведению его деятельности.

В данной работе изучались именно группы в социальной сети администраций районов Санкт-Петербурга, Администрация района Санкт-Петербурга является исполнительным органом государственной власти Санкт-Петербурга, проводящим государственную политику Санкт-Петербурга и осуществляющим государственное управление на территории района Санкт-Петербурга. Такой подход решает 2 задачи, во-первых, состав подписчиков таких групп является схожим, во-вторых, если пользователь подписан на группу своего района, можно предположить, что он и проживает в данном районе.

В анализируемых сообществах подписчики не могут создавать свои записи, все записи публикует модератор, как правило, это посты о благоустройстве, мероприятиях района и другие новости из жизни администрации. У пользователей есть возможность оставлять комментарии под такими постами. Комментарии под записями модераторов - единственный доступный канал исследования коммуникации в официальных сообществах между гражданами и администрацией района. Используя эти данные, можно оценить, на какие темы пишут жители района в комментариях, что их беспокоит, и как используется данный канал коммуникации с органами власти.

Таблица 1. Общий список исследуемых сообществ

Район	URL
Адмиралтейский район Санкт-Петербурга	https://vk.com/officialadmiral
Выборгский район Санкт-Петербурга	https://vk.com/vyb_spb
Василеостровский район Санкт-Петербурга	https://vk.com/vonews
Калининский район Санкт-Петербурга	https://vk.com/kalin_spb
Кировский район Санкт-Петербурга	https://vk.com/club.kimews
Колпинский район Санкт-Петербурга	https://vk.com/kolpr
Красногвардейский район Санкт-Петербурга	https://vk.com/krqv_live
Красносельский район Санкт-Петербурга	https://vk.com/krsel
Курортный район Санкт-Петербурга	https://vk.com/vesti_kurort
Кронштадтский район Санкт-Петербурга	https://vk.com/kronshtadtcity
Московский район Санкт-Петербурга	https://vk.com/zhit_v_moskovskom
Невский район Санкт-Петербурга	https://vk.com/nevnews
Петроградский район Санкт-Петербурга	https://vk.com/petropressa
Петродворцовый район Санкт-Петербурга	https://vk.com/pet_spb
Приморский район Санкт-Петербурга	https://vk.com/primorsknews_ru
Пушкинский район Санкт-Петербурга	https://vk.com/pushkin_adm
Фрунзенский район Санкт-Петербурга	https://vk.com/frunzr
Центральный район Санкт-Петербурга	https://vk.com/newstucentr

В ходе исследования был проведен анализ комментариев, которые оставляют под постами в каждом сообществе пользователи. На основе анализа данных были выявлены активные пользователи, которые генерируют максимальное количество комментариев.

Сбор информации для исследования проводился с помощью набора инструментов, который включал в себя API Вконтакте, парсера контента и публичного сервиса. Были использованы доступные методы API VK для сбора записей и комментариев из сообществ, а также публичной информации профилей пользователей.

При проведении исследования социальных сетей возникает сложная проблема безопасности персональных данных при парсинге. Персональными данными, согласно ст. 3 федерального закона «О персональных данных», называется «любая информация, относящаяся к прямо или косвенно определенному, или определяемому физическому лицу (субъекту персональных данных)». Традиционно при скачивании данных пользователей социальных сетей образуются поля `first_name` (Имя) и `last_name` (Фамилия), их можно отнести к персональным данным, в таком случае появится объект персональных данных. Для того, чтобы избежать любых связей с реальным человеком, в момент сохранения информации был использован модуль деперсонализации, который позволяет не сохранять такие данные. Каждому профилю присваивается ID вида `User-n`, где `n` порядковый номер строки, поля `first_name` и `last_name` не сохраняются, и в ходе работы исследователи использовали обезличенные данные.

Процессы сбора и обработки были разделены на 3 этапа:

- сбор статистической информации о сообществах методами API VK, сохранение всех комментариев сообщества;
- анализ сохраненных комментариев, выделением самых активных пользователей, расчет средних показателей каждого сообщества;
- сбор информации о профилях самых активных пользователей.

На основе информации о самых активных участниках групп был проведен анализ социально-демографических и иных параметров, характеризующих активистов в городских сообществах в Петербурге. Оценка параметров проводилась на основе данных, скаченных из общедоступной информации о профиле пользователя в социальных сетях.

4. Результаты исследования

В результате автоматизированного сбора данных была получена информация о количестве участников, публикаций и комментариев в каждом сообществе, представленная в табл.2. По данным исследования, численность участников районных сообществ не превышает 3-4% от общей численности населения, проживающего на данной территории, за исключением Курортного (16%), Калининского (10%) и Выборгского районов (6%).

Таблица 2. Основные показатели районных сообществ

Район	Население района (2020)	Участ. в сообществе	% участн. от населения	Всего коммент. от всех	Всего постов	Сред. кол-во коммент. на пост
Адмиралтейский	159795	4331	2,7	4708	5590	0,8
Выборгский	207482	12498	6,0	21829	15412	1,4
Василеостровский	522482	9168	1,8	12738	11392	1,1
Калининский	529187	53440	10,1	96014	26630	3,6
Кировский	336157	3717	1,1	3533	4557	0,8
Колпинский	193839	9482	4,9	9229	14900	0,6
Красногвардейский	356628	17198	4,8	14764	5217	2,8
Красносельский	408026	8175	2,0	6704	7598	0,9
Курортный	44461	7284	16,4	7187	11014	0,7
Кронштадтский	79067	2845	3,6	2298	2424	0,9
Московский	352172	6282	1,8	9411	3464	2,7
Невский	536137	10060	1,9	9787	7585	1,3
Петроградский	128075	5205	4,1	8238	8700	0,9
Петродворцовый	143823	4797	3,3	8971	4975	1,8
Приморский	573024	6263	1,1	4586	2653	1,7
Пушкинский	226336	7142	3,2	5686	4943	1,2
Фрунзенский	386537	9062	2,3	14263	7532	1,9
Центральный	214572	6491	3,0	11954	7942	1,5

На диаграмме (рисунок 1) представлено распределение общего количества комментариев пользователей в каждом районном сообществе. Следует отметить, что среди сообществ с максимальной долей участников от всего населения Калининский и Выборгский район также лидируют по количеству публикаций.

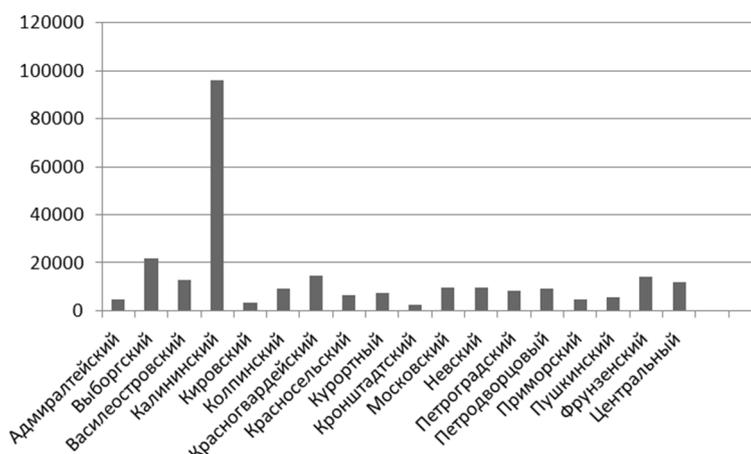


Рис. 1. Общее количество комментариев в каждом районе

Далее в ходе исследования был проведен анализ собранной информации, рассчитаны средние показатели активности пользователей в районных сообществах Санкт-Петербурга. С использованием разработанного скрипта были выделены самые активные подписчики в каждой из групп, подсчитано количество комментариев и лайков в расчете на каждый пост (табл. 3).

Таблица 3. Показатели активности пользователей в районных сообществах в Санкт-Петербурге

Район	Просм. на пост	Лайков на пост	% участников, поставивших лайк	Коммент. на пост	Репостов на пост	Всего постов
Адмиралтейский	503	6	0,1	1	1	5590
Выборгский	890	9	0,1	1	1	15412
Василеостровский	819	9	0,1	1	1	11392
Калининский	2971	26	0,0	4	4	26630
Кировский	550	5	0,1	1	1	4557
Колпинский	949	9	0,1	1	1	14900
Красногвардейский	1910	19	0,1	3	2	5217
Красносельский	968	5	0,1	1	2	7598
Курортный	520	6	0,1	1	1	11014
Кронштадтский	1006	15	0,5	1	2	2424
Московский	1284	14	0,2	3	3	3464
Невский	945	9	0,1	1	2	7585
Петроградский	600	8	0,2	1	1	8700
Петродворцовый	938	13	0,3	2	2	4975
Приморский	1363	12	0,2	2	2	2653
Пушкинский	794	8	0,1	1	2	4943
Фрунзенский	1489	10	0,1	2	2	7532
Центральный	696	7	0,1	2	1	7942

Полученные данные позволяют говорить о достаточно низкой активности пользователей в сообществах районных администраций в Санкт-Петербурге. Об этом свидетельствуют показатели среднего количества оставленных реакций в виде лайков (одобрения) и комментариев под постами от администрации групп.

Однако дальнейший анализ позволили сделать вывод о существовании в каждом сообществе определенных активных пользователей, которые оставляют в разы больше

реакций, чем остальные члены группы. В таблице 4 собраны самые активные подписчики и количество их комментариев в своем районном сообществе. Как было отмечено ранее, все данные при сборе и обработке были деперсонализированы. Также, при сборе комментариев, была использована настройка, которая позволила исключить ответы модераторов в комментариях, так как их записи могли исказить статистику.

Таблица 4. Выявленные в ходе анализа самые активные пользователи

Район подписчика	ID самого активного	Кол-во комментариев в своем районном сообществе для самого активного
Адмиралтейский	User_1_1	126
	User_1_2	110
Выборгский	User_2_1	375
	User_2_2	341
	User_2_3	340
	User_2_4	315
	User_2_5	310
Василеостровский	User_3_1	356
	User_3_2	249
Калининский	User_4_1	1170
	User_4_2	888
Кировский	User_5_1	98
	User_5_2	75
Колпинский	User_6_1	397
Красногвардейский	User_7_1	167
	User_7_2	151
Красносельский	User_8_1	195
Курортный	User_9_1	407
Кронштадтский	User_10_1	135
Московский	User_11_1	231
Невский	User_12_1	173
Петроградский	User_13_1	442
Петродворцовый	User_14_1	419
	User_14_2	389
	User_14_3	340
Приморский	User_15_1	115
	User_15_2	109
Пушкинский	User_16_1	229
Фрунзенский	User_17_1	389
Центральный	User_18_1	740
	User_18_2	562

В исследовании самые активные пользователи определялись по числу комментариев, оставленных под записями в группе. Эти пользователи являются активистами в анализируемых сообществах, они активно разворачивают дискуссии и выражают свою позицию в онлайн среде. Из расчетов были исключены лайки, так как в ходе анализа было установлено, что лайки часто ставят администраторы сообществ. Таким образом, реальная картина реакций в виде лайков имеет искажение.

В ходе анализа было выявлено 33 самых активных пользователя в районных сообществах. Как видно из таблицы 4, в некоторых районах это были единичные активные пользователи, а в других – по несколько человек. Вместе с тем, разброс активности в виде количества комментариев среди районов достигает 10 раз.

На следующем этапе был проведен анализ профилей 33 самых активных участников сообществ. Была собрана информация о профиле (в случае, если профиль открытого типа),

а также произведен подсчет общего количества комментариев от пользователя, которые он оставлял в других сообществах и страницах социальной сети ВКонтакте. Дополнительно, оценивалась дата регистрации профиля и наличия реального человека на фотографии, данная информация использовалась для исключения рекламных ботов. В случае, если аккаунту менее одного года, а на фото нет реального человека, можно предположить, что это фейковый аккаунт.

Таблица 5. Информация о пользователях

ID	Район подписчика	Общее кол-во комментариев во всем ВК	Пол	Лет	Текущая занятость
User_1_1	Адмиралтейский	1202	Ж	37	Университет
User_1_2	Адмиралтейский	173	Ж		
User_2_1	Выборгский	1326	Ж	60	НИИ
User_2_2	Выборгский	1192	Ж		
User_2_3	Выборгский	1528	М	43	
User_2_4	Выборгский	1508	Ж	59	Образование
User_2_5	Выборгский	844	Ж		
User_3_1	Василеостровский	829	Ж	55	
User_3_2	Василеостровский	1364	Ж	53	Экономика
User_4_1	Калининский	1218	Ж		
User_4_2	Калининский	613	Ж		
User_5_1	Кировский	1534	Ж	41	
User_5_2	Кировский	1361	Ж	67	
User_6_1	Колпинский	322	Ж	53	Университет
User_7_1	Красногвардейский	949	Ж		
User_7_2	Красногвардейский	1664	М	40	
User_8_1	Красносельский	1534	Ж		
User_9_1	Курортный	1505	М		Завод
User_10_1	Кронштадтский	1391	Ж	67	Поликлиника
User_11_1	Московский	1017	М	37	
User_12_1	Невский	1334	Ж	52	
User_13_1	Петроградский	192	М		
User_14_1	Петродворцовый	1625	М	67	Университет
User_14_2	Петродворцовый	272	Ж		
User_14_3	Петродворцовый	1988	М	45	
User_15_1	Приморский	596	М	39	
User_15_2	Приморский	1085	М	48	
User_16_1	Пушкинский	1698	М	31	
User_17_1	Фрунзенский	1941	М		Школа
User_18_1	Центральный	621	М		Университет
User_18_2	Центральный	1862	Ж		

Для сбора всех комментариев подписчика в социальной сети ВКонтакте был задействован внешний сервис, который также использует API VK. Важно отметить, что в чистом формате во ВКонтакте нельзя собрать все комментарии подписчика в открытых группах, из-за вопросов приватности. Но есть косвенный метод для оценки общего количества комментариев пользователя во всем ВКонтакте, и он основан на упоминаниях. Для упоминаний в сети ВКонтакте существует определённый тип поиска, что дает возможность найти исходные ответы-комментарии и перейти от них к комментариям заданного профиля. Сам поиск был создан для сообществ, но работает он и для подписчиков.

На данный момент этот метод работает весьма ограниченно (есть лимиты на сбор, упоминания могут удаляться, либо их наоборот может быть слишком много), однако, позволяет косвенно оценить активность во всей сети ВКонтакте. Так, для активных

профилей из Петроградского, Колпинского и Центрального района не удалось оценить общую активность во всей социальной сети, так как собранные показатели оказались неполноценными, по той причине, что собирая данные об общем количестве комментариев активных профилей из этих районов API VK отдавало для них ошибку при достижении определенного числа комментариев (для решения данного вопроса был отправлен запрос в техническую поддержку, но на момент написания материала, ответ не получен).

В таблице 5 собрана обобщенная информация о социально-демографических характеристиках группы активистов в группах районных администраций Санкт-Петербурга. По данным исследования, среди активистов 39% составляют мужчины и 61% - женщины. Определить возраст удалось только у 59% активных пользователей. Как показывает статистика, в основном это люди старше 30, причем 12% составляют активисты старше 60 лет.

Сложнее обстояло дело с определением основных сфер занятости среди группы активных пользователей. В графе "Текущая занятость" в таблице 5. Была произведена замена названий организаций на тип учреждения, чтобы исключить возможность обратной идентификации, тем самым, дополнительно обезличив данные. Почти у 70% в изучаемой группе было невозможно определить сферу занятости. Однако среди пользователей, чья сфера занятости была указана в открытом доступе основную долю составляют работники сферы образования и науки.

Данные исследования свидетельствуют о проявлении активности выявленных лидеров мнений в виде комментариев к постам и на других площадках (сообществах). Эти факты позволили сделать вывод о существовании определенного социально-психологического типа лидеров мнений в районных сообществах. У собранных профилей активность начинается от 192 комментариев во всей социальной сети ВКонтакте, и достигает значения 1625 комментариев, что можно считать высоким показателем активности. В расчете общего числа комментариев во всем ВКонтакте учитываются только записи из открытых сообществ. В закрытых сообществах количество комментариев рассчитать невозможно.

5. Выводы

В результате проведения исследования был отработан механизм выявления активных пользователей в районных сообществах Санкт-Петербурга в социальной сети ВКонтакте. Механизм был реализован на основе API VK и позволяет собирать деперсонализированные наборы данных. На основе применения данного метода было выявлено существование в каждом сообществе от 1 до 5 явно выраженных активных пользователей, которые оставляют больше всего реакций на посты, публикуемые в группах, а также развивают дискуссии.

На основе автоматизированного анализа были выделены 33 активных пользователя городских сообществ, оставляющие больше всего комментариев под постами районных администраций. Анализ социально-демографических характеристик группы активистов позволил сделать некоторые предложения о формировании ядра активистов из людей среднего (40+) и старшего возраста (60+) с преобладанием женщин, а также активистов, занятых в сфере образования.

Определенные ограничения исследования связаны с недоступностью части информации о социально-демографических характеристиках выявленных активных пользователей. Это затруднение напрямую связано с закрытостью части профилей, а также недостаточной наполненностью данных в открытой их части.

Возможным решением данного вопроса при проведении дальнейшего анализа может стать персонализированная рассылка активным пользователям приглашения для прохождения участия в более детальном опросе. Таким образом, можно будет получить

согласие на проведение дальнейшего исследования от целевой группы активных подписчиков районных сообществ.

Литература

- [1] Mutz D., Young L. Communication and public opinion: Plus a change? // *Public Opinion Quarterly*. 2011. Vol. 75 (5). P. 1,018–1,044. DOI: 10.1093/poq/nfr052.
- [2] Turcotte J., York C., Irving J., Scholl R.M., Pingree R.J. News recommendations from social media opinion leaders: Effects on media trust and information seeking // *Journal of Computer-Mediated Communication*. 2015. Vol. 20 (5). P. 520–535. DOI: 10.1111/jcc4.12127.
- [3] Katz E., Lazarsfeld P.F. *Personal influence: The part played by people in the flow of mass communication* // Glencoe, Ill.: Free Press. 1955. 400 p. DOI:10.4324/9781315126234.
- [4] Nisbet M.C., Kotcher J.E. A two-step flow of influence? Opinion-leader campaigns on climate change // *Science Communication*. 2009. Vol. 30 (3). P. 328–354. DOI: 10.1177/1075547008328797.
- [5] Bakshy E., Hofman J.M., Mason W.A., Watts D.J. Everyone’s an influencer: Quantifying influence on Twitter // *WSDM ’11: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. 2011. P. 65–74. DOI: 10.1145/1935826.1935845.
- [6] Katz E. The two-step flow of communication: An up-to-date report on an hypothesis // *Public Opinion Quarterly*. 1957. Vol. 21 (1). P. 61–78. DOI: 10.1086/266687.
- [7] Cha M., Haddadi H., Benevenuto F., Gummadi K.P. Measuring user influence in Twitter: The million follower fallacy // *ICWSM ’10: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*. 2010. P. 10–17.
- [8] Diehl T., Weeks B.E., de Zúñiga H.G. Political persuasion on social media: Tracing direct and indirect effects of news use and social interaction // *New Media & Society*. 2016. Vol. 18 (9). P. 1,875–1,895. DOI: 10.1177/1461444815616224.
- [9] Bennett W.L., Livingston S. The disinformation order: Disruptive communication and the decline of democratic institutions // *European Journal of Communication*. 2018. Vol. 33 (2). P. 122–139 DOI: 10.1177/0267323118760317.
- [10] Ferrara E. Manipulation and abuse on social media’ by Emilio Ferrara with Ching-man Au Yeung as coordinator // *ACM SIGWEB Newsletter*. 2015. Article 4. DOI: 10.1145/2749279.2749283.
- [11] Huang P.Y., Liu H.Y., Lin C.T., Cheng P.J. A diversity-dependent measure for discovering influencers in social networks // *Information retrieval technology. Lecture Notes in Computer Science*. 2013. Vol. 8281. P. 368–379. DOI: 10.1007/978-3-642-45068-6_32.
- [12] Mahmud J. Why do you write this? Prediction of influencers from word use // *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. 2014. P. 603–606.
- [13] Rosenthal S., Mckeown K. Detecting influencers in multiple online genres // *ACM Transactions on Internet Technology*. 2017. Vol. 17 (2). Article 12. DOI: 10.1145/3014164.
- [14] Rosenthal S. Detecting influencers in social media discussions // *XRDS: Crossroads*. 2014. Vol. 21 (1). P. 40–45. DOI: 10.1145/2659889.
- [15] Пабрики городской власти в социальных сетях. URL: <https://read.infometer.org/gorodsm> (дата обращения 05.03.2021).
- [16] Aral S., Walker D. Identifying influential and susceptible members of social networks // *Science*. 2012. Vol. 337 (6092). P. 337–341. DOI: 10.1126/science.1215842.

Identification of Opinion Leaders Among Active Users of Groups of District Administrations of St. Petersburg in the Social Network VKontakte

B.A. Nizomutdinov, L.A. Vidyasova

ITMO University

The paper presents the results of identifying opinion leaders among users who leave comments in 18 official groups of the district administrations of St. Petersburg on VKontakte. Parsing of social networks and statistical data analysis were used as the main research method. According to the results of the analysis, opinion leaders in urban communities were identified as the most active users of the groups, leaving the most comments. On the basis of an automated analysis, 33 active users of urban communities were identified, leaving the most comments under the posts of district administrations. Analysis of the socio-demographic characteristics of a group of opinion leaders made it possible to make some proposals on the formation of a core of activists from people of middle age (40+) and older people (60+) with a predominance of women, as well as activists employed in the field of education.

Keywords: social networks, social activists, opinion leaders, parsing, social network analysis

Reference for citation: Nizomutdinov B.A., Vidyasova L.A. Identification of Opinion Leaders Among Active Users of Groups of District Administrations of St. Petersburg in the Social Network VKontakte // *The State and Citizens in the Electronic Environment. Vol. 5 (Proceedings of the XXIV International Joint Scientific Conference «Internet and Modern Society», IMS-2021, St. Petersburg, June 24-26, 2021).* - St. Petersburg: ITMO University, 2021. P. 126 – 137. DOI: 10.17586/2541-979X-2021-5-126-137

Reference

- [17] Mutz D., Young L. Communication and public opinion: Plus a change? // *Public Opinion Quarterly*. 2011. Vol. 75 (5). P. 1,018–1,044. DOI: 10.1093/poq/nfr052.
- [18] Turcotte J., York C., Irving J., Scholl R.M., Pingree R.J. News recommendations from social media opinion leaders: Effects on media trust and information seeking // *Journal of Computer-Mediated Communication*. 2015. Vol. 20 (5). P. 520–535. DOI: 10.1111/jcc4.12127.
- [19] Katz E., Lazarsfeld P.F. *Personal influence: The part played by people in the flow of mass communication* // Glencoe, Ill.: Free Press. 1955. 400 p. DOI: 10.4324/9781315126234.
- [20] Nisbet M.C., Kotcher J.E. A two-step flow of influence? Opinion-leader campaigns on climate change // *Science Communication*. 2009. Vol. 30 (3). P. 328–354. DOI: 10.1177/1075547008328797.
- [21] Bakshy E., Hofman J.M., Mason W.A., Watts D.J. Everyone's an influencer: Quantifying influence on Twitter // *WSDM '11: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. 2011. P. 65–74. DOI: 10.1145/1935826.1935845.
- [22] Katz E. The two-step flow of communication: An up-to-date report on an hypothesis // *Public Opinion Quarterly*. 1957. Vol. 21 (1). P. 61–78. DOI: 10.1086/266687.
- [23] Cha M., Haddadi H., Benevenuto F., Gummadi K.P. Measuring user influence in Twitter: The million follower fallacy // *ICWSM '10: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*. 2010. P. 10–17.
- [24] Diehl T., Weeks B.E., de Zúñiga H.G. Political persuasion on social media: Tracing direct and indirect effects of news use and social interaction // *New Media & Society*. 2016. Vol. 18 (9). P. 1,875–1,895. DOI: 10.1177/1461444815616224.

- [25] Bennett W.L., Livingston S. The disinformation order: Disruptive communication and the decline of democratic institutions // *European Journal of Communication*. 2018. Vol. 33 (2). P. 122–139 DOI: 10.1177/0267323118760317.
- [26] Ferrara E. Manipulation and abuse on social media' by Emilio Ferrara with Ching-man Au Yeung as coordinator // *ACM SIGWEB Newsletter*. 2015. Article 4. DOI: 10.1145/2749279.2749283.
- [27] Huang P.Y., Liu H.Y., Lin C.T., Cheng P.J. A diversity-dependent measure for discovering influencers in social networks // *Information retrieval technology. Lecture Notes in Computer Science*. 2013. Vol. 8281. P. 368–379. DOI: 10.1007/978-3-642-45068-6_32.
- [28] Mahmud J. Why do you write this? Prediction of influencers from word use // *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. 2014. P. 603–606.
- [29] Rosenthal S., Mckeown K. Detecting influencers in multiple online genres // *ACM Transactions on Internet Technology*. 2017. Vol. 17 (2). Article 12. DOI: 10.1145/3014164.
- [30] Rosenthal S. Detecting influencers in social media discussions // *XRDS: Crossroads*. 2014. Vol. 21 (1). P. 40–45. DOI: 10.1145/2659889.
- [31] Pabliki gorodskoj vlasti v social'nyh setyah. URL: <https://read.infometer.org/gorodsm> (data obrashcheniya 05.03.2021).
- [32] Aral S., Walker D. Identifying influential and susceptible members of social networks // *Science*. 2012. Vol. 337 (6092). P. 337–341. DOI: 10.1126/science.1215842.