

Возможности идеографического исследования сетевых сообществ средствами интеллектуального анализа текстов*

Ю.М. Кузнецова, Н.В. Чудова, А.А. Чуганская

Федеральный исследовательский центр «Информатика и управление»
Российской академии наук, г. Москва, Россия

Аннотация. Представлено описание групп участников городских сетевых сообществ, построенное на результатах математической обработки данных о сетевой активности пользователей и на полученных с помощью инструмента TITANIS характеристиках их текстовой деятельности в ходе сетевых обсуждений. Результаты демонстрируют обоснованность проведения исследований для задач социального мониторинга, выполненных в смешанном – номотетическом и идеографическом – ключе, когда социально-психологический «портрет» большой группы строится с опорой на методы искусственного интеллекта и математической статистики.

Ключевые слова: автоматический анализ текста, психолингвистические показатели, сетевое сообщество, идеографический подход.

DOI: 10.14357/20790279250102 **EDN:** ZMUYAU

Введение

Различение в науке номотетической (генерализующей, количественной) методологии установления объективных законов на основе обобщения сведений о разных процессах, и идеографической (индивидуализирующей, качественной) методологии, задачей которой является описание уникальных явлений, рассматривается как проявление оппозиции общего и индивидуального в познании [1]. В зависимости от познавательной задачи познающий субъект может рассматривать наблюдаемый феномен «либо как что-то отличное от всех других явлений данного множества, либо как носитель свойств и функций, присущих множеству других вещей, повторяющихся в них. В первом случае предмет вычленяется из тотальности, противостоящей субъекту действительности, ... во втором случае предмет включается во множество, с точки зрения субъекта, взаимозаменяемых, подобных явлений» [2].

Исследователи отмечают, что суждения номотетического (в виде научных законов и высказываний о необходимости) и идеографического (в

виде дескриптивного знания) характера имеются в любой области науки [3], хотя традиционно идеографический метод атрибутировался т. н. наукам о духе, или социогуманитарным, требующим «понимания», по контрасту с науками о природе, ориентированными на «объяснение» [4]. К специфическим особенностям социогуманитарных наук основоположники данного различия Дильтей, Риккерт, Виндельбанд относили такие характеристики, которые определяют их непринадлежность парадигме классической научной рациональности, как ценностная компонента, историчность, уникальность исторического развития, невозможность использования применяемого в естествознании экспериментального метода; в плане гносеологических установок их отличает «относительность, диалектика существенного и несущественного в знании об объекте: существенное в одном отношении – несущественно в другом и наоборот» [4]. Если номологические науки, изучающие общие правила и законы, требуют повторных наблюдений, то идеографические имеют возможность оперировать данными однократного наблюдения [5]. По сравнению с номотетическим подходом, претендующим на распространение полученных выводов на все однородные объекты, цель идеографической методологии заключается в получе-

* Исследование выполнено в рамках научной программы Национального центра физики и математики, направление № 9 «Искусственный интеллект и большие данные в технических, промышленных, природных и социальных системах».

нии «локального знания внутри четко очерченного контекста» [6]. Сами предметы, составляющие исследовательское поле идеографического знания, также специфичны: они «не существуют сами по себе, но имеют место тогда, когда человек говорит о них и проблематизирует их. Субъекту не даны в опыте готовые вещи, он конструирует предмет в процессе познания. Однако этот предмет более нагляден, очевиден и конкретен, чем какой-нибудь физический или химический конструктор» [7].

В практике изучения сложных и сверхсложных социальных систем, таких как государство [8,9] или современная производственная организация [10] формируется представление о том, что предпочтение того или иного типа исследовательских средств приводит к искажению картины изучаемого явления, что выдвигает новую для отечественной науки задачу интеграции («смешивания») качественных и количественных методов и формирования т. н. стратегий “mixed methods research” [6], подразумевающих, среди прочего, и широкое применение методов математической обработки на разных стадиях сбора и интерпретации данных эмпирического исследования «понимаемых» предметов [9]. В качестве примеров реализации такой стратегии можно привести исследование механизмов социального взаимодействия в организации [10], где раскрыты возможности компенсации ограниченности классических номотетических методов путем использования идеографического подхода, чувствительного к индивидуальности, субъективным предпочтениям, мотивам и ценностям социальных субъектов. В работе [11] применение номотетического и идеографического подходов в их сочетании для исследования социально-экономической деятельности отдельного города, как государственно-территориального субъекта, позволило выявить и описать существенные характеристики и связи в городском социуме и их влияние на макроуровень в масштабах страны в привязке к определенному временному этапу развития данного субъекта.

Помимо процессов социального взаимодействия традиционными предметами социогуманитарных исследований являются: специфика сознания социальных групп, стереотипы общественного сознания, механизмы формирования политических предпочтений и социальных представлений и т. п. Сочетание номотетического и идеографического подходов к изучению феноменов данного порядка реализовано в исследовании [12] путем психосемантического анализа политического дискурса. В ряде работ [13,14] показаны возможности извлечения информации о структуре и содержании социальных представлений из контента социальных

сетей с помощью инструмента интеллектуального анализа текста TITANIS [15].

Целью настоящей работы явилось идеографическое описание групп участников сообществ социальной сети ВКонтакте жителей одного поволжского города на основе результатов математической обработки данных об их сетевой активности и о полученных с помощью инструмента TITANIS характеристиках их текстовой деятельности в ходе сетевых обсуждений.

1. Определение психологически значимых параметров активности комментирования

На основе обзора литературы, посвященной анализу сетевой активности, был составлен список показателей активности, релевантный возможностям процедуры оценки и комментирования постов, размещаемых в сети ВКонтакте в новостных городских сетевых сообществах. Таких показателей оказалось восемь: число дискуссий, отмеченных пользователем лайком или прокомментированных им; число комментариев, созданных пользователем; средняя длина комментария; число лайков, поставленных пользователем; число лайков, полученных ко всем комментариям, оставленным пользователем; число отправленных реплав; число полученных реплав; число начатых тредов (веток). Был собран материал дискуссий пяти наиболее популярных новостных городских сетевых сообществ (малый город в Поволжье, население – 96 тыс.чел.). Всего за всю историю существования этих сообществ с 2011 года по июль 2023 было собрано 289593 дискуссий, из них 170102 содержало хотя бы один текстовый комментарий. Методом кластеризации k-means были выделены 8 кластеров пользователей, различающихся характером своей активности в этих сообществах.

Рассмотрим полученные характеристики кластеров с точки зрения задач, связанных с определением иерархических отношений пользователей (в настоящем исследовании – комментаторов) в сетевых дискуссиях. Для этого исследуем представленные в табл.1 первичные показатели, использованные для кластеризации, а также несколько вычисленных на их основе производных показателей. Позиция комментаторов в иерархии сетевой коммуникации может быть охарактеризована по тем признакам, значение которых в кластере значительно отличается от средних величин, что свидетельствует об определенных специфических особенностях участия представителей кластера в сетевых дискуссиях. В табл. 1 указаны границы первого и третьего квартилей для каждого признака.

Было выделено пять первичных и три производных показателей кластеризации в качестве значимой для исследования специфики сетевого общения с применением методов текстового анализа.

1. Количество дискуссий, в которых комментатор принимал участие (строка 5 табл. 1). По этому признаку выделяются кластеры, объединяющие коммуникантов, постоянно (в период наблюдения) оставлявших комментарии в дискуссиях (Кл4 и Кл1), принявших участие менее чем в двух дискуссиях за весь период (Кл0, Кл7 и Кл3), а также имеющих средние значения по данному показателю. К категории «разовых коммуникантов» относятся комментаторы, которые крайне редко либо посещают анализируемые сетевые сообщества, либо принимают

деятельное участие в обсуждениях. Не исключено, что некоторая часть таких пользователей является одноразовыми ботами, созданными в целях формирования определенной реакции на конкретные сообщения.

2. Способ участия в сетевом общении (строка 6 табл.1). Различаются кластеры, представители которых оставляют в семь-девять раз больше лайков, чем текстовых комментариев (Кл0 и Кл6), кластеры с относительно сбалансированным соотношением лайков и комментариев (Кл4 и Кл1) и занимающие промежуточное положение по данному признаку.

3. Позиция в эмоциональном обмене (строка 6 табл.1) – характеристика, определяющая в настоящем исследовании соотношение между ак-

Табл. 1

Значения первичных и производных признаков для выделенных кластеров комментаторов

№	Признак	Кластеры								Квартили	
		0	1	2	3	4	5	6	7	I	III
1	comments_count_mean	1,5	2499,1	1103,7	9,4	4590,5	504,2	133,7	2,1	7,6	1452,6
2	likes_given_mean	13,9	4009,2	3891,6	42,4	6960,7	1190,3	938,4	8,8	35,2	3921,0
3	likes_received_comments_mean	2,1	4643,1	2098,9	22,7	9731,5	887,9	240,3	5,1	18,3	2734,9
4	average_comment_length_mean	1,2	14,6	14,0	30,3	15,2	12,1	11,3	255,0	12,6	22,7
5	discussions_mean	1,1	1455,3	645,3	5,5**	2500,2	328,5	90,0	1,4	45,7	1050,3
6	способ участия (likes_given_mean / comments_count_mean)*	9,32	1,60	3,53	4,50	1,52	2,36	7,02	4,25	2,17	5,13
7	позиция в эмоциональном обмене (likes_given_mean / likes_received_comments_mean)*	6,69	0,86	1,85	1,86	0,72	1,34	3,90	1,73	1,22	2,37
8	продуктивность текстовой деятельности (comments_count_mean X average_comment_length_mean X discussions_mean)*	1,77	36449	15426**	284	69698	6113	1509	540**	476	20682
9	size	216929	62	191	10737	15	730	3212	154		
10	% комментаторов в кластере от их общего количества	93,49	0,03	0,08	4,63	0,01	0,31	1,38	0,07		

Примечания: знаком * отмечены производные признаки, вычисленные на основе исходных; знаком ** выделены значения, близкие к границе квантиля для признака

Табл. 2

Описание представителей кластеров

Кластер	Описание
Кл0	Малопишущие разовые коммуникаторы, активно ставящие лайки и редко их получающие, редко оставляющие аномально короткие комментарии
Кл6	Средне вовлеченные коммуникаторы, активно ставящие лайки и редко их получающие, редко оставляющие короткие комментарии
Кл3	Малопишущие разовые коммуникаторы, оставляющие среднее количество развернутых комментариев
Кл7	Как представители Кл3, однако менее активные в целом и оставляющие очень длинные комментарии
Кл4	Многопишущие постоянные коммуникаторы, оставляющие комментарии почти так же часто, как лайки, и получающие лайков больше, чем оставляют сами
Кл1	Как представители Кл4, но имеют менее высокие показатели
Кл5	Имеют по всем показателям средние значения; пишущие короткие комментарии
Кл2	Как представители Кл5, но все показатели более высокие; условно многопишущие

тивностью комментатора как продуцента эмоциональной реакции, выражаемой в сетевых дискуссиях путем проставления лайков чужим сообщениям, и количеством лайков, которые получают собственные комментарии этого пользователя как реципиента эмоциональной реакции. Позиция продуцента в эмоциональном обмене означает отсутствие адекватного эмоционального ответа в виде поддержки и одобрения, то есть ситуацию эмоциональной депривации. Реципиент в эмоциональном обмене получает более высокую поддержку со стороны сетевого окружения по сравнению со своим собственным вкладом, что делает его позицию эмоционально «профицитной». Представители Кл0 и Кл6 в эмоциональном обмене являются продуцентами, а представители Кл4 и Кл1 – реципиентами.

- Средняя длина комментария в словах (строка 7 табл.1) непосредственно характеризует значимый для текстового анализа параметр сетевого поведения коммуникантов. Представители Кл7 и Кл3 склонны оставлять развернутые комментарии, при этом комментарии объединяемых в Кл7 пользователей можно считать аномально длинными. В то же время комментарии пользователей из Кл6 и Кл5 отличаются краткостью, которая у представителей Кл0 достигает крайней выраженности, когда на один комментарий приходится в среднем менее двух слов.
- Интегральный показатель текстовой продукции (строка 8 табл.1) позволяет определить средний объем текстов, которые могут быть получены от одного комментатора за все время наблюдения. К Кл4, Кл1 и Кл2 относятся «многопишущие» комментаторы, тексты которых в сумме составляют в среднем от 70 до 15 тыс. слов, в то время как Кл0, Кл3 и Кл7 объединяют комментаторов с очень низкой суммарной текстовой продукцией.

Описанные показатели в своей совокупности позволяют охарактеризовать те особенности положения комментаторов в иерархии сетевого общения, которые послужили основанием для их объединения в кластеры. Можно видеть, что характеристики в парах Кл3 и Кл7, Кл1 и Кл4 воспроизводятся полностью, а в паре Кл0 и Кл6 – в основных чертах. Различия между входящими в названные пары кластерами заключается в интенсивности проявления признаков, что можно проследить по таблице Кл2 и Кл5, в свою очередь, объединяет средняя выраженность практически всех оцениваемых признаков.

Описание особенностей комментаторов, объединяемых в кластеры, в терминах первичных и производных признаков приведено в табл. 2.

С точки зрения задачи определения иерархических отношений пользователей в сетевых дискуссиях полученные характеристики типичных представителей выявленных кластеров позволяют предполагать, что в роли ситуативных инфлюенсеров, влияющих на ход дискуссии своими комментариями, не могут выступать коммуниканты из категории малопишущих. По сути, у них отсутствует техническая возможность воздействовать на мнение других участников сетевого общения, поскольку они либо не создают развернутые комментарии (Кл0), либо крайне редко участвуют в дискуссиях (Кл3 и Кл7). Кроме того, для представителей самого многочисленного «малопишущего» Кл0 (к нему, как следует из данных табл. 1, относится 93% всех пользователей, чья активность анализировалась в ходе настоящего исследования) характерна позиция продуцента в эмоциональном обмене, что подразумевает скорее зависимое, чем лидерское поведение.

Перспективными для выявления текстовых признаков ситуативного инфлюенсерства в сетевом общении могут быть комментарии пользователей из категории многопишущих, одновременно с этим занимающих в эмоциональном обмене позицию реципиента, то есть представителей Кл4 и Кл1. К условно многопишущим относятся также представители Кл2. Количество относящихся

к перечисленным кластерам пользователей очень мало: Кл1 составляет всего 0,03%, Кл2 – 0,08%, Кл4 – 0,01% от общего числа комментаторов, учтенных в исследовании. Однако представители этих кластеров активны в плане текстовой деятельности, что может компенсировать их малочисленность при формировании корпусов комментариев для проведения текстового анализа.

2. Типологизация комментаторов городских сетевых сообществ

Данные об активности пользователей и психолингвистического анализа их комментариев позволяют реконструировать мотивацию участия посетителей страниц этих сообществ, создать «психологический портрет» пользователей, попавших в тот или иной кластер, выделить особенности оставляемых ими комментариев.

В табл. 3 представлены значения медиан параметров активности. Отметим, что исходные данные по составу кластеров были скорректированы путем удаления из них данных тех комментаторов, которые только ставили знак «лайк» и не оставили ни одного комментария. При этом самый многочисленный кластер из исходных данных «0» после такой коррекции значительно уменьшился, так как именно в нем оказалось собрана большая часть пользователей, не писавших комментариев.

С точки зрения мотивации участия в городских дискуссиях подавляющее число комментаторов в городских сообществах (Кл0, Кл3, Кл7), прокомментировавшие 1-2 дискуссии и редко оставлявшие лайки (от 1 до 7 лайков за все время

наблюдения), – это случайные, незаинтересованные посетители страниц горсообществ. Некоторые оставляют очень короткие комментарии в 4-6 слов, другие – длинный комментарий в 20-25 слов или даже очень длинный (около 200 слов), но т.к. это, как правило, один комментарий в несколько лет. Считать этих людей членами сообществ вряд ли возможно. Доля таких случайных комментаторов чрезвычайно высока – более 95%. Это значит, что существующие в этом городе новостные интернет-сообщества не очень популярны у пользователей – не на страницах этих сообществ происходит обсуждение волнующих горожан проблем.

Тем не менее для примерно 1 тысячи пользователей (Кл1, Кл2, Кл4, Кл5) эти дискуссионные площадки важны. Они прокомментировали множество постов (от 300 до 2,5 тыс.), оставили много комментариев (от 400 до 4,5 тыс.) и получили поддержку своим комментариям (их комментарии лайкали и под ними разворачивалась отдельная дискуссия).

Наконец, среди посетителей страниц городских интернет-сообществ есть пользователи, которых можно, видимо, отнести к группе заинтересованных наблюдателей (Кл6). Их втрое больше активных участников дискуссий – более 3 тыс. Их активность выражается чаще в проставлении лайков (постам и чужим комментариям), но все-таки они иногда оставляют и комментарии, вызывающие реакцию партнеров по общению (лайки и ответные комментарии).

Можно видеть, что в обследованном городе городские интернет-сообщества публикуют сообщения, привлекающие внимание значитель-

Табл. 3

Значения медиан параметров активности

Кластер	Количество комментариев	Число оставленных лайков	Число полученных лайков	Средняя длина комментария	Число начатых веток (тредов)	Число дискуссий, в которых принял участие пользователь	Размер кластера (число комментаторов)
0	2	7	1	5,75	2	2	39502
1	2405	2127,5	4633,5	12,2	1505,5	1454,5	62
2	1102	1724	1901	12,9	578	613	191
3	2	2	3	23,75	1	1	10596
4	4594	3593	10270	14,1	2639	2442	15
5	461	620	837	10,9	321	304	730
6	120	443	193	9,9	82	83	3191
7	2	0	6	189,5	2	2	33

ной части пользователей – посты этих сообществ лайкнуло около 220 тыс. пользователей. Однако подавляющее большинство этих пользователей – более 170 тыс. – не оставило ни одного комментария за все время наблюдения. Из тех же кто комментировал эти сообщения, более 95% делало это лишь один-два раза. Можно сказать, что сообщения городских интернет-сообществ служат жителям источником информации, но они (среди которых есть и страницы администрации города) не стали для горожан площадками общественных дискуссий. Тем не менее, потенциал этих страниц в таком качестве не исчерпан – исследование выделило группу пользователей (Кл4) в 15 человек, весьма заинтересованных в работе на страницах этих сообществ и де-факто выступающих инфлюенсерами для группы активистов (1 тыс. пользователей) и группы наблюдателей (3 тыс. пользователей).

Изучение текстового своеобразия комментариев пользователей, различающихся характером активности при комментировании материалов новостных городских сетевых сообществ, проводилось с опорой на данные инструмента интеллектуального анализа текста TITANIS. Все количественные текстовые маркеры нормализовались относительно количества комментариев у пользователя.

Сравнение критерием Манна – Уитни текстовых показателей корпусов комментариев показало следующее. Комментарии пользователей из Кл1, Кл2, Кл4 (268 человек из числа наиболее активных комментаторов) имеют лишь одно статистически значимое различие: для Кл1 и Кл2 по частоте упоминания природных и техногенных катастроф (*тематическая группа слов (ТГС) Катастрофы*). Однако в связи с незначительностью самой частоты употребления слов, относящихся к этой ТГС – 0,0009 для Кл1 и 0,0006 для Кл2, это различие можно трактовать как случайное. Таким образом, можно констатировать, что в психологическом и лингвистическом плане комментарии пользователей, вошедших по показателям активности в ядро активистов городских интернет-сообществ, различий нет.

Комментарии пользователей, входящих в Кл5 (730 человек из группы активистов) имеют одно статистически значимое отличие от комментариев пользователей из Кл4 – это отличие для редко встречающихся в обеих корпусах конструкций *Часть речи: часть композита – сложного слова* (0,0002 и 0,0001 соответственно). От комментариев пользователей Кл1 у этой группы есть отличия по четырем грамматическим показателям (*Прича-*

стия + деепричастия / Число предложений; Причастия; Часть речи: часть композита – сложного слова; Средняя длина предложения) и двум психолингвистическим (*ТГС Власти и ТГС ЖКХ*). Отличаются эти комментарии также более высокой *Долей прописных символов* (что может быть проинтерпретировано как массовая Большая аккуратность в письме, а может – как использование кем-то из пользователей капслока). От комментариев пользователей Кл2 у этой группы есть отличия по тем же четырем грамматическим показателям, что и при сравнении с Кл1, но есть и более значимые с психологической точки зрения различия. В комментариях этой группы реже используются слова, относимые к *Эмотивам Печали, Коэффициент Триггера, Коэффициент определенности действия, ТГС Когнитивная деятельность и коммуникация, ТГС Власти*, реже встречаются местоимения *Мы* и *ты*. Показатели длины реплик (*число предложений, символов, слов*) также снижены. Таким образом, можно сказать, что основная часть активистов городских интернет-сообществ не только дает менее развернутые реплики и делает это реже, чем представители большинства из «ядра», но и демонстрируют более спокойное отношение к обсуждаемым проблемам и реже апеллируют к групповой идентичности. Интересно, что эти более спокойные и выдержанные комментаторы похожи на тех, кто вошел в группу ситуативных инфлюенсеров (между корпусами комментариев пользователей Кл5 и Кл4 нет значимых различий).

В целом, можно отметить достаточное единодушие в группе активистов городских ВК-сообществ обследованного города. Активная тысяча пользователей примерно в равной мере заинтересована в обсуждении самых разных тем: власти как таковые (*ТГС Власти*) обсуждаются несколько чаще лишь 25% активистов (Кл1 и Кл2, 253 чел.), а проблемы ЖКХ (*ТГС ЖКХ*) – лишь 20% активистов (Кл2, 191 чел.). В отношении других тем и вовсе нет никаких различий. При этом обсуждение идет на приблизительно одном уровне эмоционального напряжения и этот уровень – средний для всех комментариев, оставленных на страницах городских сообществ самыми разными посетителями. Важно и то, что комментарии большинства активистов (Кл5) ничем в психологическом и тематическом плане не отличаются от комментариев инфлюенсеров (Кл4). Возможно, именно поддержка большинством активистов тех, кто пишет часто, но без излишней драматизации и агитации, выдвинула пользователей, вошедших в Кл4, в число инфлюенсеров при обсуждении городских проблем.

Противоположный полюс по активности составляют те, кого выше мы назвали случайными комментаторами (Кл0, Кл3, Кл7). Это люди, участвовавшие в одной-двух дискуссиях, оставившие в них один-два комментария, получившие незначительное число лайков (0–2) и сами оставившие их – немного (от 0 до 7). Формальные различия внутри этой группы связаны с длиной комментария, что находит отражение и в частеречных и грамматических показателях инструмента анализа текста TITANIS. Есть между этими группами и содержательные различия – психологически и тематически их комментарии имеют определенные особенности.

Комментарии представителей второго по численности Кл3 (10,5 тыс. пользователей) отличаются от комментариев людей из самого многочисленного Кл0 (39,5 тыс. пользователей) повышенным вниманием к действиям – повышенная доля глаголов любого лица, доля местоимений любого лица и числа, доля глаголов прошедшего времени, коэффициенты напряжения (КТ, КОД, КЛС) говорят о том, что эти комментарии содержат либо рассказы о чьем-то поведении, либо призывы к действию. При этом, повышенное внимание (по сравнению с текстами Кл0) в этих комментариях уделено темам власти, ЖКХ и силовых структур (ТГС *Власти*, ТГС *ЖКХ*, ТГС *Силовые структуры*), а пониженное – темам экономики и техногенных и природных катастроф (ТГС *Экономика*, ТГС *Катастрофы*). Учитывая то, что эти комментарии являются вдвое более длинными, чем «обычные» (т.е. комментарии тех, кто часто и много комментирует и лайкает чужие комментарии и посты сообществ – Кл1, Кл2, Кл4, Кл5), и вчетверо более длинными, чем комментарии основной массы случайных комментаторов (Кл0), можно предположить, что они содержат истории (возможно, заранее заготовленные) о поведении представителей власти, силовых структур и ЖКХ, требующем огласки и публичного осуждения. Интересно, что при этом комментаторы из Кл3 совершенно не склонны обсуждать эмоции, чужие или свои (сниженное число практически всех *эмотивов* по сравнению с корпусом комментариев Кл0), если это не эмоция гнева или удивления. При этом эти комментарии содержат повышенное (по сравнению с корпусом комментариев Кл0) количество *негативной эмоциональной лексики*. Итак, можно заключить, что Кл3 позволил выделить из общей массы случайных комментаторов тех, кто, попадая на страницы городских сообществ, реагирует стереотипно – рассказывает о поведении социально значимых акторов, при этом их рассказ содержит негативную оценку.

Отметим, что пользователи малочисленного Кл7 (33 пользователя), отличающиеся от пользователей Кл0 и Кл3 только многословностью (в среднем 190 слов на комментарий), продуцируют тексты, практически идентичные по своим психологическим и лингвистическим показателям комментариям представителей Кл3 и близкие по подавляющему числу показателей комментариям представителей Кл0. В связи с этим выделение такой группы пользователей можно признать нецелесообразным.

Рассмотрим в завершение особенности комментаторской продукции представителей Кл6 (3,2 тыс. комментаторов), который по маркерам активности занимает, в некотором смысле, промежуточное положение – этих людей нельзя назвать случайными комментаторами (около 80 дискуссий прокомментировано, написано в среднем по 120 комментариев), но их активность явно ниже, чем у группы активистов (где у большинства комментариев на порядок больше, а у некоторых – в 40 раз). Вот эти заинтересованные наблюдатели, чье число превышает группу активистов, публикуют комментарии более всего похожие на комментарии инфлюенсеров: содержательных различий между корпусами текстов этих двух групп немного, всего 6. При этом три показателя являются тематическими: инфлюенсеры чаще, чем наблюдатели поднимают в своих постах темы власти, силовых структур и катастроф (ТГС *Власти*, ТГС *Силовые структуры*, ТГС *Катастрофы*). Три показателя касаются демонстрируемых эмоций: в текстах инфлюенсеров больше *лексики негативных эмоций*, больше *эмотивов гнева* и *эмотивов отвращения*. Видимо, такой «тематическо-эмоциональный профиль» в сочетании со сверхвысоким уровнем комментаторской активности позволяет пользователю перейти из категории заинтересованного участника общественных дискуссий в категорию локального инфлюенсера.

Полученные результаты позволяют сделать следующие выводы:

1. Проведенная кластеризация комментаторов по маркерам активности дала содержательно интерпретируемые результаты.
2. Число групп комментаторов может быть сокращено на основе анализа текстовых показателей комментариев в зависимости от целей исследования до трех групп (случайные комментаторы, заинтересованные наблюдатели, активисты) или 6 групп (с выделением в группе активистов инфлюенсеров, взвинченных комментаторов и спокойных, а в группе случайных – нейтральных и осуждающих действия властей, силовых структур и работников ЖКХ).

3. В обследованном городе интернет-сообщества привлекли значительное внимание пользователей ВК и большое число т.н. случайных комментаторов (их число даже превышает численность всего населения города). При этом число активистов сетевых дискуссий невелико – 998 человек за весь период наблюдения. Тем не менее, происходящее на этих площадках сетевых дискуссий имеет, видимо, определенное влияние на жителей города – кроме 1 тыс. активистов еще 3 тыс. пользователей составляют круг заинтересованных наблюдателей, а комментарии 15 локальных инфлюенсеров собирают не менее двух лайков и более чем в половине случаев порождают отдельные «ветки» дискуссии (треды).
4. Активная тысяча пользователей примерно в равной мере заинтересована в обсуждении самых разных проблем и не отличается высокой политизированностью или сфокусированностью на одной теме (действия властей обсуждается несколько чаще лишь 25% активистов, а проблемы ЖКХ – лишь 20%). При этом обсуждение идет на приблизительно одном уровне эмоционального напряжения и этот уровень – средний для всех комментариев, оставленных на страницах городских сообществ самыми разными посетителями за 12 лет.
5. В группе активистов выделяются три подгруппы: ситуативные инфлюенсеры (15 чел.), взвинченные комментаторы (253 чел.) и спокойные комментаторы (730 чел.). Инфлюенсеры отличаются в основном своей активностью (в среднем приняли участие в 2,5 тыс. дискуссий, оставили более 4,5 тыс. комментариев), интересом к их комментариям со стороны других пользователей (получили в среднем более 10 тыс. «лайков») и относительно спокойным тоном обсуждения городских проблем. Взвинченные комментаторы демонстрируют далеко не такую высокую активность как инфлюенсеры, хотя они тоже довольно активны (комментируют примерно по 150 постов в год). Однако их комментарии чаще, чем комментарии других активистов, посвящены требованиям к властям. В них чаще выражается недовольство, обеспокоенность и сомнения, ярче выражена апелляция к групповой идентичности собеседников. Спокойные комментаторы отличаются как довольно ровным эмоциональным фоном своих высказываний, так и тем, что вообще реже вступают в дискуссии.

Заключение

Идеографический подход реализован в настоящем исследовании на двух уровнях. Представленное описание свидетельствует о том, что сама организация нашего исследования, способы сбора и обработки данных, а также процедура их интерпретации, отражают тот симбиоз номотетических и идеографических подходов, необходимость которого определяется сложной природой социо-психологических явлений как предмета изучения. Полученные качественные характеристики, определяющие специфику выявленных групп сетевых коммуникантов, как субъектов социального взаимодействия, базируются на результатах количественных измерений и математической обработке численных показателей. С другой стороны, значение и роль выявленных количественных закономерностей для характеристики различий в социо-психологической позиции представителей разных групп коммуникантов требует экспертного качественного анализа, учитывающего эксплицитные и имплицитные системные связи между отдельными разнородными проявлениями такой позиции.

Второй «уровень идеографичности» нашего исследования касается его общей схемы, не включающей в себя сравнения с другими подобными объектами. Выявленные характеристики относятся к поведению участников конкретных сообществ социальной сети ВКонтакте в их локальной и временной специфике. Полученный «портрет» не может быть распространен на другие сообщества и временные отрезки, но представляет интерес как некоторый социально-психологический феномен, описание которого может быть востребовано для таких управленческих практик, как организация на основе анализа сетевого контента мониторинга массовых настроений, выявление сетевых инфлюенсеров, формирование и реализация стратегий гармонизации сетевого общения и социальных отношений и т. п. Для решения подобных практических задач в большей степени ценно знание об уникальной ситуации «здесь и теперь», чем представление об общих закономерностях, по которым развиваются социальные отношения «как правило» и «в среднем».

Проведенное исследование показало, что в обследованном городе новостные городские сообщества ВК публикуют сообщения, привлекающие внимание значительной части пользователей. Сообщения городских интернет-сообществ служат жителям источником информации, но эти страницы не стали для горожан площадками общественных дискуссий – комментарии оставляет лишь четверть

посетителей, из них только 4 тыс. являются постоянными комментаторами. Тем не менее, потенциал этих страниц в таком качестве не исчерпан – исследование выделило группу пользователей в 15 человек, весьма заинтересованных в работе на страницах этих сообществ и де-факто выступающих инфлюенсерами в блогосфере города.

Литература

1. Черкашин А.К. Метатеоретическое семиотическое моделирование в науке и технике // Информационные и математические технологии в науке и управлении. 2022. № 2 (26). С. 5–23. DOI: 10.38028/ESI.2022.26.2.001
2. Антипов Г.А. Гуманитарные и социальные науки: теоретико-познавательные основания // Идеи и идеалы. 2024. Т. 16. № 2–1. С. 160–183. DOI: 10.17212/2075-0862-2024-16.2.1-160-183
3. Лебедев С.А. Современная наука как многомерная когнитивно-социальная структура // Современные философские исследования. 2023. № 2. С. 92–107. DOI: 10.18384/2310-7227-2023-2-92-107
4. Черников М.В. Возможна ли философия Правды в эпоху современности? // Свободная мысль. 2020. № 1 (1679). С. 197–210.
5. Кареев Н.И. Общая методология гуманитарных наук // Социология науки и технологий. 2020. Т. 11. № 2. С. 64–85. DOI: 10.24411/2079-0910-2020-12003
6. Липатов С.А., Нестерова Е.М. Интеграция качественных и количественных подходов как методологическая проблема социальной психологии // Национальный психологический журнал. 2024. № 3 (55). С. 81–90. DOI: 10.11621/prj.2024.0306
7. Ажимов Ф.Е. Антиметафизическое движение в контексте становления гуманитарных наук в XIX в. // Вестник МГУ. Серия 7. Философия. 2022. № 1. С. 9–22.
8. Калинин С.А. Уникальное и закономерное в современном понимании государства // Правовое государство: теория и практика. 2021. № 4 (66). С. 110–128. DOI: 10.33184/pravgos-2021.4.7
9. Линде А.Н. Значение личностного, понимающего направления в политической науке // Общество: политика, экономика, право. 2023. № 11. С. 35–49. DOI: 10.24158/per.2023.11.4
10. Глуханюк Н.С., Юртаева М.Н., Берсенева Е.И. Возможности метода семантического дифференциала для оценки персонала (на примере HR-брендинга организации) // Управление персоналом и интеллектуальными ресурсами в России. 2022. Т. 11. № 2. С. 83–88. DOI: 10.12737/2305-7807-2022-11-2-83-88
11. Еввалов В.В. Анализ социально-экономической деятельности тольяттинской городской администрации в период 1991–2005 гг. // Вестник Самарского университета. История, педагогика, филология. 2023. Т. 29. № 4. С. 53–60. DOI: 10.18287/2542-0445-2023-29-4-53-60.
12. Адеева Т.Н., Тихонова И.В. Опыт использования семантического анализа при исследовании психологических ресурсов подростков // Вестник Костромского государственного университета. Серия: Педагогика. Психология. Социокинетика. 2020. Т. 26. № 4. С. 111–119. DOI: 10.34216/1998-0817-2020-26-4-111-119.
13. Kuznetsova, Y.M., Smirnov, I.V., Stankevich, M.A. et al. Creating a Text Analysis Tool for Socio-Humanitarian Research. Part 2. The RSA Machine and the Experience in Using It. Sci. Tech. Inf. Proc. 2020. 47. P. 374–382. DOI: 10.3103/S0147688220060040
14. Григорьев О.Г., Кузнецова Ю.М., Никитина Е.Н., Станкевич М.А., Чудова Н.В. Каузально-эмотивный анализ. Часть II. Исследование реакции зрителей YouTube-каналов на пропаганду // Психологический журнал. 2022. Т. 43. № 4. С. 90–98. DOI: 10.31857/S020595920021484-8
15. Smirnov I., Stankevich M., Kuznetsova Y., Suvorova M., Larionov D., Nikitina E., Savelov M., & Grigoriev O. TITANIS: A tool for intelligent text analysis in social media // In Kovalev S.M., Kuznetsov S.O., Panov A.I. (eds.) Artificial Intelligence. The Nineteenth Russian Conference on Artificial Intelligence RCAI-2021, Taganrog, Russia, October 11-16, 2021 (RCAI 2021). Lecture Notes in Computer Science. Vol. 12948. Springer, Cham. P. 232–247. DOI: 10.1007/978-3-030-86855-0_16

Кузнецова Юлия Михайловна. Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва, Россия. Старший научный сотрудник. Кандидат психологических наук. Область научных интересов: психолингвистика, психосемантика, исследование социальных представлений, интеллектуальный анализ текста. E-mail: kuzjum@yandex.ru

Чудова Наталья Владимировна. Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва, Россия. Старший научный сотрудник. Кандидат психологических наук. Области научных интересов: психология агрессии, психология интернета, когнитивное моделирование, сетевая психодиагностика. E-mail: nchudova@gmail.com

Чуганская Анфиса Анваровна. Федеральный исследовательский центр «Информатика и управление» Российской академии наук», г. Москва, Россия. Научный сотрудник. Кандидат психологических наук. Области научных интересов: сетевая психодиагностика, социальная психология, коммуникация. E-mail: anfi.chuganskaya@yandex.ru (Ответственный за переписку)

The capacity of ideographic research of network communities by means of intellectual text analysis

Y.M. Kuznetsova, N.V. Chudova, A.A. Chuganskaya

Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia

Abstract. A description of the groups of participants in urban network communities is presented, based on the results of mathematical processing of data on the network activity of users and on the characteristics of their textual activity obtained using the TITANIS tool during network discussions. The results demonstrate the validity of conducting research for social monitoring tasks performed in a mixed nomothetic and ideographic manner, when the socio-psychological “portrait” of a large group is based on the methods of artificial intelligence and mathematical statistics.

Keywords: *automatic text analysis, psycholinguistic indicators, network community, ideographic approach.*

DOI: 10.14357/20790279250102 **EDN:** ZMUYAU

References

1. Cherkashin A.K. Metatheoretical semiotic modeling in science and technology. *Information and mathematical technologies in science and management*, 2022, 2 (26): 5–23. DOI: 10.38028/ESI.2022.26.2.001 (In Russian).
2. Antipov G. Humanities and social sciences: epistemological foundations. *Idey i idealy [Ideas and Ideals]*, 2024, 16, 2 (1): 160–183. DOI: 10.17212/2075-0862-202416.2.1-160-183 (In Russian).
3. Lebedev S.A. The modern science as the multidimensional cognitive-social structure. *Bulletin of Moscow Region State University. Series: Philosophy*, 2022, 2: 92–107. DOI: 10.18384/2310-7227-2023-2-92-107 (In Russian).
4. Chernikov M. The philosophy of Truth in the modern era. *Svobodnaya mysl' [Free Thought]*, 2020, 1 (1679): 197–210. (In Russian).
5. Kareev N.I. *Obshchaya metodologiya gumanitarnykh nauk [General Methodology of the Humanities]. Sotsiologia nauki i tehnologii [Sociology of Science and Technology]*, 2020, 11, 2: 64–85. DOI: 10.24411/2079-0910-2020-12003 (In Russian).
6. Lipatov S.A., & Nesterova E.M. Integration of qualitative and quantitative approaches as a methodological issue in social psychology. *National Psychological Journal*, 2024, 19 (3): 81–90. DOI: 10.11621/npj.2024.0306 (In Russian).
7. Azhimov F.E. History of philosophy anti-metaphysical movement in the making of humanities in the 19th century. *Moscow University Bulletin. Series 7. Philosophy*, 2022, 1: 9–22. (In Russian).
8. Kalinin S.A. Unique and naturally determined attributes of the modern understanding of a state. *The Rule of Law State: Theory and Practice*, 2021, 4 (66): 110–128. DOI: 10.33184/pravgos-2021.4.7 (In Russian).
9. Linde A.N. The meaning of personalistic, understanding school in political science. *Society: Politics, Economics, Law*, 2023, 11: 35–49. DOI: 10.24158/pep.2023.11.4 (In Russian).
10. Glukhanyuk N.S., Yurtaev M.N., & Berseneva E.I. Capabilities of the semantic differential method for personnel evaluation (using the example of HR branding of an organization). *Management of the Personnel and Intellectual Resources in Russia*, 2022, 11, 2: 83–88. DOI: 10.12737/2305-7807-2022-11-2-83-88 (In Russian).
11. Evpalov V.V. Analysis of social and economic activity by Togliatti local government in 1991-2005. *Vestnik Samarskogo universiteta. Istoriia, pedagogika, filologiya [Vestnik of Samara University. History, pedagogics, philology]*, 2023, 29, 4: 53–60. DOI: 10.18287/2542-0445-2023-29-4-53-60 (In Russian).

12. *Adeyeva T.N., & Tikhonova I.V.* Experience in using semantic analysis in the study of psychological resources of adolescents. *Vestnik of Kostroma State University. Series: Pedagogy. Psychology. Sociokinetics*, 2020, 26, 4: 111–119. DOI: 10.34216/1998-0817-2020-26-4-111-119 (In Russian).
13. *Kuznetsova Y.M., Smirnov I.V., Stankevich M.A. et al.* Creating a Text Analysis Tool for Socio-Humanitarian Research. Part 2. The RSA Machine and the Experience in Using It. *Sci. Tech. Inf. Proc.* 2020, 47, 374–382. DOI: 10.3103/S0147688220060040
14. *Grigoriev O.G., Kuznetsova Yu.M., Nikitina E.N., Stankevich M.A., Chudova N.V.* Causative-emotive analysis. Part II. A study of the reaction of YouTube channel viewers to propaganda // *Psychological Journal*. 2022. Vol. 43. No. 4. pp. 90-98. DOI: 10.31857/S020595920021484-8 (In Russian).
15. *Smirnov I., Stankevich M., Kuznetsova Y., Suvorova M., Larionov D., Nikitina E., Savelov M., & Grigoriev O.* TITANIS: A tool for intelligent text analysis in social media // In Kovalev S.M., Kuznetsov S.O., Panov A.I. (eds.) *Artificial Intelligence. The Nineteenth Russian Conference on Artificial Intelligence RCAI-2021, Taganrog, Russia, October 11-16, 2021 (RCAI 2021). Lecture Notes in Computer Science. Vol. 12948.* Springer, Cham. P. 232–247. DOI: 10.1007/978-3-030-86855-0_16

Kuznetsova Yulia M. PhD, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44/2 Vavilova str., Moscow, 119333, Russia. E-mail: kuzjum@yandex.ru

Chudova Natalia V. PhD, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44/2 Vavilova str., Moscow, 119333, Russia. E-mail: nchudova@gmail.com

Chuganskaya Anfisa A. PhD, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44/2 Vavilova str., Moscow, 119333, Russia. E-mail: anfi.chuganskaya@yandex.ru (Responsible for correspondence)