

УДК 159.9.07; 004.8

Воронин А. Н., Кубрак Т. А.

Институт психологии Российской академии наук
129366, г. Москва, ул. Ярославская, д. 13, к. 1, Российская Федерация

Смирнов И. В., Станкевич М. А.

Федеральный исследовательский центр «Информатика и управление» Российской академии наук
119333, г. Москва, ул. Вавилова, д. 44/2, Российская Федерация

ОЦЕНКА СУБЪЕКТНОСТИ СЕТЕВОГО СООБЩЕСТВА ПО ПОКАЗАТЕЛЯМ АВТОМАТИЧЕСКОГО РЕЛЯЦИОННО- СИТУАЦИОННОГО АНАЛИЗА ЕГО КОНТЕНТА

АННОТАЦИЯ

Цель. Представление исследования, посвящённого разработке регрессионных моделей субъектности сетевых сообществ по автоматически определяемым показателям реляционно-ситуационного анализа (РСА) его контента.

Процедура и методы. Для построения моделей был проанализирован контент 64 сетевых сообществ различной тематической направленности из открытого сегмента социальных сетей: Facebook, «ВКонтакте», «Одноклассники», Pikabu, Telegramm и др. Тексты сообществ были подвергнуты психолингвистическому анализу с помощью ранее разработанного перечня дискурсивных маркеров, по итогам которого были вычислены показатели субъектности. Автоматический реляционно-ситуационный анализ текстов осуществлялся с помощью машины РСА, разработанной в Институте системного анализа РАН.

Результаты. В результате проведённого анализа для всех показателей субъектности удалось построить исчерпывающие регрессионные модели с удовлетворительным качеством.

Теоретическая и/или практическая значимость. Использование полученных регрессионных моделей позволит осуществлять мониторинг различных секторов Рунета в автоматизированном режиме и проводить оценку субъектности контента.

КЛЮЧЕВЫЕ СЛОВА

интернет, субъектность сетевых сообществ, дискурсивные маркеры, цифровые следы, интеллектуальный анализ текстов, реляционно-ситуационный анализ текстов, регрессионные модели

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 18-00-00606 (К), 18-00-00233 и 18-00-00605

СТРУКТУРА

Введение

Процедура и методы исследования

Результаты и обсуждение

Выводы

A. Voronin, T. Kubrak

*Institute of Psychology of the Russian Academy of Sciences
13 korp. 1 Yaroslavskaya ul., Moscow 129366, Russian Federation*

I. Smirnov, M. Stankevich

*Federal Research Center "Computer Science and Management"
of the Russian Academy of Sciences
44/2 Vavilova ul., Moscow 119333, Russian Federation*

ASSESSMENT OF THE NETWORK COMMUNITY SUBJECTIVITY BY INDICATORS OF CONTENT AUTOMATIC RELATIONAL-SITUATIONAL ANALYSIS

ABSTRACT

Aim. Presentation of regression models of the subjectivity of network communities based on automatically determined indicators of the content relational situational analysis (RSA).

Methodology. To develop these models 64 network communities of various thematic focus from the open segment of social networks (Facebook, VKontakte, Odnoklassniki, Pikabu, Telegramm, etc.) were analyzed. The networks communities texts were subjected to psycholinguistic analysis using a previously developed list of discourse markers, and the results allowed to identify indicators of subjectivity. Automatic relational situational analysis of texts was performed using an RSA machine developed at the Institute for System Analysis of the Russian Academy of Sciences.

Results. Comprehensive regression models of satisfactory quality were constructed for all indicators of subjectivity.

Research implications. The use of the obtained regression models will allow to monitor various sectors of the Runet in an automated mode and to assess the subjectivity of the content.

KEYWORDS

internet, network community subjectivity, discourse markers, text mining, digital trails, relational-situational analysis of text, regression models

ACKNOWLEDGMENTS

The reported study was funded by RFBR under the research project № 18-00-00606 (K), 18-00-00233 и 18-00-00605

ВВЕДЕНИЕ

Современное состояние общества характеризуется перемещением значительной доли общественных отношений в интернет-пространство. В социальных сетях образуются новые связи, формируются группы, разворачивается внутригрупповое и межгрупповое взаимодействие. Интернет-контент воспроизводит существующие этические нормы и правила его создателей, отражает традиционные установки и предпочтения данного общества [27]. Группу людей, поддерживающих общение и ведущих совместную деятель-

ность при помощи социальных медиа, реализованных в рамках подхода Web 2.0., определяют как сетевое сообщество. Наука о сетях предлагает сужение определения сетевого сообщества до группы участников, связанных между собой большим числом связей, чем с участниками из других сообществ. При таком подходе сетевые сообщества определяются как локально плотно связанные подграфы в сети, где отдельные аккаунты – узлы, а их коммуникативные действия – рёбра графа. Объектом эмпирического исследования при изучении сетевых сообществ выступает пользовательский контент, генерируемый участниками сетевых сообществ, и сведения о характере взаимосвязей аккаунтов [1].

Наиболее адекватным психологическим подходом для изучения контента сетевых сообществ представляется дискурсивная парадигма, предусматривающая изучение реальной коммуникативной практики в различных ситуациях и социокультурных контекстах [11; 26]. Проблема исследования субъектности сетевого сообщества в рамках дискурса сводится к выделению адекватных *дискурсивных признаков*, связанных с теми или иными характеристиками субъектности. При этом возможно использование разноуровневых дискурсивных признаков (лексическо-грамматических показателей, процессуальных и семантических), связанных с различными признаками субъектности сообществ разного размера. Так, *маркерами субъектности* могут выступать многие лексические единицы, такие, как *лексика совместности, побудительные глаголы, парольные слова, жаргонная, обценная, эмоционально-оценочная лексика* и др. Крайне важны дискурсивные характеристики процесса коммуникации, разворачивающегося в ходе дискуссий: *нарастание сходства обсуждаемых тем, наличие цепочек комментариев, их длина, пространность* и пр. Среди процессуальных характеристик значимыми представляются такие, которые относятся к особенностям общения и выступают важнейшим проявлением субъектности сообщества. Не менее значимое место в перечне дискурсивных характеристик субъектности занимают семантические категории, которые не получают конкретной формы реализации, но могут быть надёжно идентифицированы экспертами по смыслу: *способность следовать организационным и управляющим воздействиям, отображение группового образа цели, информирование о текущем состоянии дела или проблемы* и пр.

Выбор фрагментов контента для анализа субъектности существенным образом зависит от того, как субъектность проявляется в дискурсе сетевого сообщества¹. Можно высказать несколько предположений о характере такого проявления субъектности в дискурсе сетевого сообщества [1].

¹ Подробное описание дискурсивных маркеров субъектности, способов их выделения и последующего структурирования в обобщённые показатели представлено в работе «Субъектность сетевого сообщества: сравнение психометрических моделей проявления дискурсивных маркеров в контенте» [15].

1. Предполагается, что субъектность – атрибут любого сетевого сообщества, и она так или иначе всегда проявляется в контенте.

2. Субъектность как феномен присуща сообществу лишь в определённые периоды его существования, и возможна фиксация лишь её «всплесков», поэтому следует оценивать максимальные проявления субъектности по разным дискурсивным признакам в разные периоды. Как вариант такого подхода можно рассмотреть субъектность сообщества в период максимальной спонтанной активности.

3. Субъектность проявляется как ответ сетевого сообщества на то или иное событие внешней социальной среды, либо событие «внутренней» жизни сетевого сообщества. Субъектность, в данном случае, рассматривается, как самопроизвольно разворачивающаяся активность, и её оценка напрямую связана с точной локализацией события и созданием выборки контента, инициированным в ответ на него.

Выделение дискурсивных маркеров открывает возможности оценки свойств сетевых сообществ, особенностей протекающих в них групповых процессов и их динамики [11; 28]. Социальная динамика в сетевой интернет-среде является сегодня предметом множества научных дисциплин, связанных с изучением человека информационной эпохи. Привлечение к социально-психологическим и дискурсивным исследованиям сетевых сообществ математических методов дает возможность осуществлять анализ больших данных, позволяющий обнаруживать ранее трудно выявляемые закономерности [12; 24].

При этом одним из приоритетных направлений исследований остаётся разработка методов анализа интернет-дискурса, осуществление которого возможно на разных уровнях и с разными задачами. Для выделения лингвистических маркеров психологических характеристик используются специальные инструменты, например, Linguistic Inquiry and Word Count [25] или Text Analysis, Crawling and Interpretation Tool [21]. По результатам, с помощью семантического анализа сообщений, интернет-запросов и постов в сетях с учётом «цифровых следов» были получены более точные оценки некоторых психологических характеристик человека, чем с помощью стандартизированных психологических опросников [18; 23], показаны возможности выявления личностных ценностей [31]. Технологии анализа цифровых следов как маркеров психологических характеристик разработаны с опорой преимущественно на компьютерную лингвистику при изучении сообщений блоггеров в Twitter [30; 32]. Сегодня анализ контента социальных сетей позволяет определять различные психические заболевания, суицидальные состояния, тяжёлые эмоциональные расстройства [17; 20; 22], в частности, выявлять уровень депрессии по текстовым сообщениям, как было сделано на материале, полученном из Facebook [29]. На основании используемой лексики возможно прогнозирование непо-

средственно поведения, такого, как, например, совершение повторных преступных действий [19].

Предлагаемый в данном исследовании подход автоматического анализа контента отличается от реализованных ранее подходов привлечением методов искусственного интеллекта, глубокого лингвистического анализа, преодолевающего рамки относительно легко формализуемых компонентов текста (преимущественно лексических, грамматических и морфологических), и частотного анализа, которые выступают в качестве основных средств извлечения информации. Для анализа семантико-синтаксической структуры контента в нашем исследовании применяется метод реляционно-ситуационного анализа (PCA) [8]. Реляционно-ситуационный анализ позволяет выявлять семантику текста, ставя в соответствие логической структуре действий, описанных в предложении, синтаксемную структуру предложения. Для выявления реляционно-ситуационной структуры высказываний выполняется семантико-синтаксический анализ текстов [13], в результате которого строится дерево полного синтаксического разбора: определяются предикаты и их аргументы, для последних устанавливаются семантические роли (синтаксические значения). Реляционно-ситуационное представление текста, учитывающее все уровни языка, показало свою эффективность в решении многих задач интеллектуального анализа текстов и информационного поиска. Так, установлено, что выявляемые с помощью метода PCA предикатно-синтаксические структуры с психологической точки зрения отражают глубинные структуры, соотносимые с языковой картиной мира личности, давая доступ к её смысловой сфере [14], определены текстовые проявления агрессивности и фрустрированности [7; 16].

В проводимых нами исследованиях объединение социально-психологического и дискурсивного подходов к изучению субъектности сетевого сообщества с математическими методами анализа (PCA текста) позволяло раскрыть психологические механизмы функционирования сетевых сообществ любой размерности, представляющих собой группы людей, объединённых особыми связями и отношениями, способных к совместной активности и саморефлексии, взаимодействие которых осуществляется в дискурсивном пространстве интернета и динамично меняется в режиме реального времени.

В ранее проведённых исследованиях было выявлено значительное количество (больше 250) соответствий между параметрами субъектности и показателями PCA, при этом часть из них была теоретически предсказуема и вполне объяснима в рамках семантического анализа при сопоставлении показателей PCA и психологических конструкторов субъектности [9]. В рамках данной статьи предпринята попытка дальнейшего исследования субъектности сетевых сообществ путём расширения эмпирической базы исследования до нескольких десятков сообществ, определения регрессионных

моделей субъектности по показателям PCA, что позволит в дальнейшем осуществлять автоматическую оценку субъектности сетевых сообществ различной направленности и тематики. Соответственно **целью данного исследования** является разработка регрессионных моделей субъектности по автоматически определяемым показателям PCA и определение наиболее значимых предикторов различных показателей субъектности сетевых сообществ.

ПРОЦЕДУРА И МЕТОДЫ ИССЛЕДОВАНИЯ

Для построения регрессионных моделей субъектности по показателям PCA был проанализирован контент 64 сетевых сообществ различной тематической направленности с различным уровнем проявления офлайн активности. Для исследования выбирались группы из открытого сегмента социальных сетей, в том числе:

- группа «Синие ведёрки» в Facebook (FB) – площадка для обсуждения проблем равноправия и нарушения прав граждан на российских дорогах;
- оппозиционный политический форум Politota;
- сообщество с лидером,
- страница в FB известного кинокритика;
- групповой чат в Telegramm «Прогрессоры», посвящённый отношениям между людьми и личностному росту;
- группа в VK «Страдающее Средневековье», площадка для обсуждения проблем «современной России» в саркастической форме;
- сообщество «Архнадзор» в FB, которое объединяет граждан, желающих способствовать сохранению культурного наследия Москвы;
- сообщество любителей игры «Detroit: Become Human»;
- сообщество мастеров кукол на BJDClub;
- сообщество «Крыса не подарок» на Pikabu;
- сообщество поиска пропавших с помощью нейросетевой программы «Лакмус» на LizaAlert и др.

Объём контента каждого из сообществ определялся обсуждением некоторой актуальной для сообщества проблемы и находился в пределах от 6040 до 30794 знаков. Все тексты подверглись психолингвистическому и реляционно-ситуационному анализу.

Психолингвистический анализ корпуса текстов сетевых сообществ и их кодирование осуществлялось с помощью перечня дискурсивных маркеров субъектности [15]. Тексты сетевых сообществ размечались 4 экспертами-психолингвистами путём выделения дискурсивных маркеров. Субъектность рассчитывалась как относительная частота дискурсивных маркеров, обнаруженных в текстах сетевых сообществ. В качестве единицы анализа использовались отдельные предложения. Показатели субъектности рассчитывались в соответствии с кумулятивно-аддитивной моделью с винзориро-

ванным усреднением с усечением 5% наибольших значений и последующей конволюцией среднего арифметического. Использование именно этой модели проявления конструкта «субъектность» в дискурсивных маркерах была признана наиболее адекватной [15]. Ранее было показано, что психометрика субъектности сетевого сообщества базируется на приоритете разнообразия различных её составляющих над уровнем их проявления [2; 15]. Вследствие этого для оценки субъектности использовались все возможные обобщенные характеристики субъектности сетевого сообщества: *девять* показателей субъектности первого уровня (*языковая и понятийная идентификация, готовность действовать, поддержка тематики сетевого сообщества, групповые нормы и ценности, планирование и результаты совместных действий, отстранение «других», защита целостности сообщества гражданская идентичность*), два показателя субъектности второго уровня (*обсуждение совместной деятельности, «свои» – «чужие»*) и показатель общей субъектности как среднее по всем показателям.

Автоматический реляционно-ситуационный анализ текстов сетевых сообществ осуществлялся с помощью машины РСА. Показатели РСА включали синтаксические, основные психолингвистические данные – семантические роли, семантические связи, показатели наличия в текстах сетевых сообществ лексики из словарей различной тематики – всего 202 показателя [13]. Поскольку тексты сообществ различались по объёму, то применялись относительные величины. В качестве единицы анализа использовалось предложение как наиболее адекватная единица среди множества возможных (реплики, клаузы, слова и др.) [9].

Для **статистических расчётов** использовался пакет PASW Statistics 18. Конкретные статистики и критерии приведены при описании результатов исследования.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для выявления предикторов субъектности был проведён множественный регрессионный анализ данных методом Stepwise. В качестве зависимой переменной последовательно выступали показатели субъектности сетевого сообщества, в качестве независимых переменных – показатели автоматического реляционно-ситуационного анализа. Так, для показателя «Общая субъектность» было подобрано 33 модели (табл. 1), при этом, начиная с 8 модели, скорректированный $R^2 > 0,8$, что говорит о её приемлемости и приемлемости всех последующих моделей с большим количеством независимых переменных [3].

Таблица 1 / Table 1

Сводная таблица регрессионных моделей общей субъектности /
Regression models of general subjectivity. Summary

Модель	R2	Скорректированный R2	Стд. ошибка оценки
8	0,875	0,856	0,005
9	0,909	0,894	0,004
10	0,93	0,917	0,004
11	0,942	0,93	0,003
12	0,96	0,951	0,003
13	0,977	0,971	0,002
14	0,982	0,977	0,002
15	0,986	0,982	0,002
16	0,991	0,987	0,001
17	0,993	0,991	0,001
18	0,996	0,995	0,001
19	0,998	0,997	0,001
20	0,999	0,999	0,000
21	1	0,999	0,000

Начиная с 21-ой модели R2 становится равным 1, и последующее увеличение количества независимых переменных не приводит к повышению точности регрессионной модели. Оценка эффекта мультиколлинеарности осуществлялась с помощью коэффициента разброса дисперсии (VIF). С 1-ой по 17-ую модель VIF для предикторов находится в пределах от 1.002 до 5.162, однако у моделей с большим количеством предикторов появляется коэффициент $VIF > 10$, что ограничивает количество приемлемых моделей до 17-ти. При этом наблюдается хорошее согласование гистограммы остатков с нормальным распределением у всех моделей из этого списка (на рис. 1 в качестве примера приведена гистограмма остатков для 17-ой модели). Автокорреляция остатков регрессионной модели, оценённая с помощью теста Дарбина-Ватсона (Durbin-Watson), составляет 2,436, что говорит об её отсутствии.

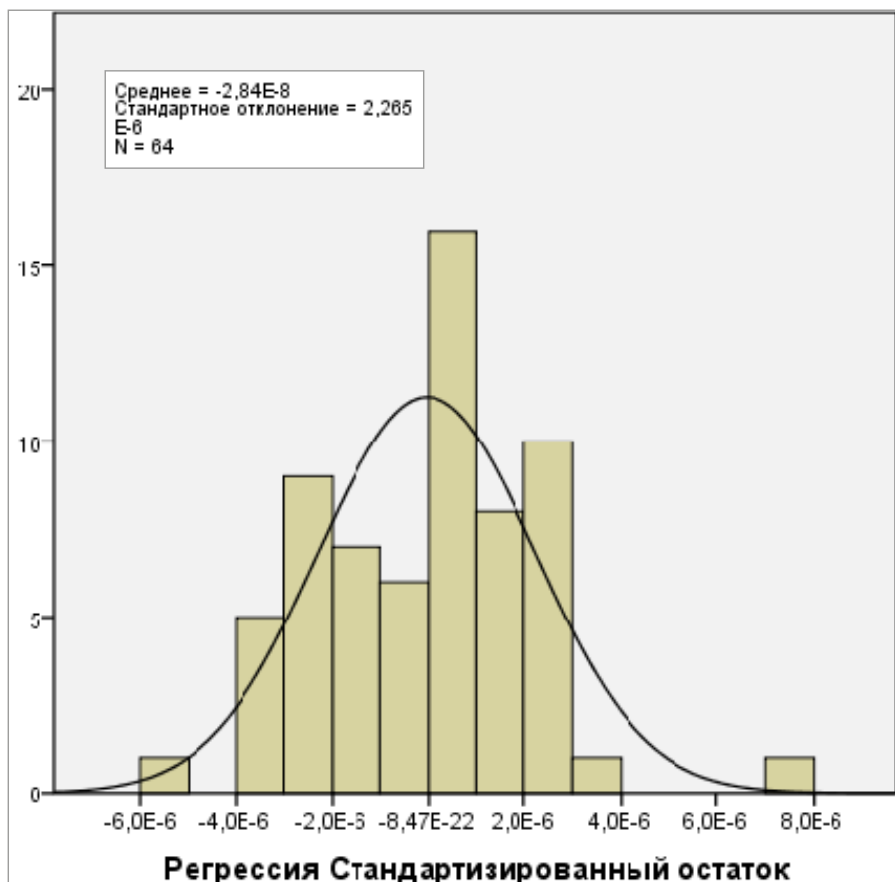


Рис. 1. / Fig. 1. Гистограмма остатков и кривая нормального распределения для 17-ой модели / Residual histogram and normal distribution curve for the 17th model

Примечание: изображение сгенерировано пакетом PASW Statistics 18 при проведении регрессионного анализа.

В результате проведения регрессионного анализа были выявлены основные предикторы *общей субъектности* и оценен вносимый ими вклад. В таблице 2 приведён перечень независимых переменных – показателей PCA – и указаны регрессионные коэффициенты, указывающие на важность этих переменных в регрессионном уравнении.

Таблица 2 / Table 2

Показатели PCA, выступающие предикторами общей субъектности /
Relational situational analysis indicators functioning as predictors of general
subjectivity

Модель 17. Скорректированный R ² = ,997		
	β	Знач.
Сем. роль: социальная_категория	-,678	,000
Сем. роль: объект	,651	,000
Словарь: Социальное неравенство и несправедливость	-,600	,000
Сем. связь: LOC	-,423	,000
Сем. роль: фабрикатив	,559	,000
Число символов	-,437	,000
Словарь: Тематическая_Власти	,375	,000
Сем. роль: квантитатив	,362	,000
Сем. роль: ликвидатив	,304	,000
Средняя длина слов (в количестве символов)	,302	,000
Сем. роль: оцениваемое_действие	,297	,000
Сем. связь: OBJ	-,273	,000
Число предложений	,253	,000
Словарь: Уголовный жаргон	,207	,000
Сем. роль: тематив	-,189	,000
Сем. роль: аблатив	,182	,000
Сем. роль: владелец_инф	-,059	,000

Анализ статистик согласия, в том числе скорректированного информационного критерия Акаике для ограниченных выборок (AICC) и информационного критерия Байеса (BIC), показывает, что они последовательно возрастают по мере увеличения количества предикторов в модели, соответственно, исходя из целей дальнейшего использования регрессионных моделей, имеет смысл ориентироваться либо на первую модель с приемлемым скорректированным R² (в нашем случае это модель 8), либо последнюю модель с допустимой мультиколлинеарностью (модель 17).

Поскольку целью исследования являлось предсказание субъектности сетевого сообщества по автоматически получаемым показателям PCA, вследствие наличия большого количества независимых переменных по сравнению с количеством анализируемых объектов (текстов сетевых сообществ), была проведена кросс-валидация данных по K блокам (K-fold cross-validation). Данные были разбиты случайным образом на 4 блока, 3 из которых путём перебора 4 раза выступали в качестве тренировочного набора и 1 оставшийся блок – в качестве контрольной выборки. В табл. 3 представ-

лены результаты оценки моделей при различных вариантах разбиения. Для всех моделей мультиколлинеарность и автокорреляция остатков находятся в допустимых пределах.

Таблица 3 / Table 3

Оценка качества модели при кросс-валидации данных / Assessing the quality of the model in data cross-validation

Показатели качества модели	Разбиение по блокам			
	1	2	3	4
Скорректированный R2	,860	,891	,781	,785
MSE	,002	,001	,004	,004
MAPE	2,813	1,258	8,531	9,297
Корреляция Спирмена между субъектностью по К-модели и уровнем субъектности по дискурсивным маркерам для контрольной выборкой К-разбиения	,732**	,750**	,483	,498*

**. Корреляция значима на уровне 0.01 (2-сторонняя).

*. Корреляция значима на уровне 0.05 (2-сторонняя).

Из табл. 3 следует, что для двух разбиений модель вполне адекватна и успешно предсказывает уровень субъектности сетевого сообщества при автоматическом анализе контента с помощью машины PCA ($R^2 > 0.8$, $MSE \leq 0.027$, $MAPE \leq 5\%$, значимый уровень корреляции). При двух других разбиениях качество модели находится на грани допустимого, и доля ошибок может достигать более 9%. Следует отметить, однако, что в трёх разбиениях из четырёх выявлена значимая корреляция в контрольной группе между уровнем субъектности, предсказанном по модели и показателем субъектности в ходе психолингвистического анализа по дискурсивным маркерам.

Аналогичным образом были рассчитаны регрессионные модели для *других показателей субъектности*. Для каждого из 11 оставшихся показателей субъектности удалось построить адекватные исчерпывающие регрессионные модели с R^2 , приближающимся к 1 с приемлемыми характеристиками качества. Обширность полученных результатов не позволяет в рамках статьи привести их все, вследствие этого укажем лишь наиболее популярные (встречающиеся во многих моделях) показатели религионно-ситуационного анализа, детерминирующие уровень субъектности сетевых сообществ. В первую очередь, к ним относится частота встречаемости специфической лексики: протестная лексика, молодежный жаргон, лексика, тематически связанная с «властью», лексика отрицательной рациональной оценки, уголовный жаргон, лексика «социального неравенства и несправедливости». Кроме того, в качестве предикторов часто выступают такие показатели реляционно-ситуационного анализа, как относительная частота встречаемости семантических ролей и связей (табл. 4).

Таблица 4 / Table 4

Наиболее значимые предикторы субъектности сетевых сообществ – показатели PCA (семантические роли и связи) / The most significant predictors of the subjectivity of network communities, i.e. the Relational situational analysis' indicators (semantic roles and connections)

Семантическая роль	Описание
аблатив	исходная точка движения
предикат	утверждение, высказанное о субъекте
каузат	объект каузирующего воздействия
субъект	компонент, которому приписывается предикативный признак
эстиматив	один из предикатов в авторизованных полипредикативных конструкциях с глаголами считать, называть, находить.
критерий_оценки	критерии оценки
поссесив	предикативный компонент со значением: а) поссесивного объекта или б) свойства, характеризующего субъект, при глаголе иметь.
тематив	компонент, называющий тему оцениваемой ситуации
темпоратив	компонент, выражающий временные характеристики
дименсив	размер, исчисляемая мера величины
социальная_категория	характеристика лица по принадлежности его к категории, группе лиц, обычно социально значимой)
объект	компонент с предметно-вещественным значением, подвергающийся действию или каузативному воздействию
квантитатив	количественный показатель изменения или соотношения признаков
ликвидатив	объект каузирующего воздействия; результатом каузирующего воздействия является прекращение существования объекта

Семантическая связь	Описание
TMP	темпоральная связь, в которой один компонент выражает временную локализацию признака, названного другим компонентом
ADR	адресатная связь, один компонент которой называет лицо или реже предмет, к которому обращено информативное, донативное или эмотивное действие лица, названного другим компонентом
LOC	локативная связь, один компонент которой называет местонахождение другого компонента
DLB	делиберативная связь, один компонент которой выражает содержание речемыслительного, социального действия или восприятия лица, названного другим компонентом
OBJ	объектная связь, один компонент которой называет объект действия субъекта, названного другим компонентом
SIT	ситуативная связь, в которой один компонент обозначает ситуацию, определяющую ситуацию или область действия другого компонента

Выявленные в результате регрессионного анализа предикторы субъектности согласуются с ранее описанными показателями субъектности сообществ. Они частично совпадают с характеристиками субъектности реальных социальных групп (например, совместная активность) [4; 6] и согласуются со специфичными для сетевых сообществ характеристиками: маркирование «чужих» и их отстранение, идентификация и поддержка «своих», гражданская идентичность [15]. Так, совместная активность, в различных формах которой, прежде всего, проявляется феномен субъектности [5], предполагает организацию и планирование деятельности, что релевантно *аблативу* и *каузату*, отражающему воздействие в дискурсе, *адресатной связи*, показывающей направленность действий, а также *темпоративу* и *темпоральной* связям, выражающим конкретизацию реальных планов. Кроме того, совместность выражается в языковой и понятийной идентификации участников, обнаруживающейся в специфической лексике контента сетевого сообщества. Реализовать совместную активность позволяет и взаимосвязанность участников сообщества, являющаяся базовой характеристикой субъектности. В контенте сообществ с более высоким уровнем субъектности реже встречается *тематив*, но чаще *критерий оценки*, что говорит о сформированности намерений действовать, так как уже известно, что следует делать, но важно определить, *как это нужно делать*. Совокупность перечисленных ролей позволяет говорить о некой гиперроли, характерной для контента сетевых сообществ с повышенной субъектностью – «оперативность действия».

Взаимосвязанность проявляется и в коммуникативной направленности собеседников друг на друга, что соответствует *адресатной связи*. В то же время выявление предиктора *критерий оценки*, а также *оцениваемое действие* и *лексики отрицательной рациональной оценки*, можно объяснить оценкой в дискурсе сетевых сообществ, не только обсуждаемых топиков, но и маркированием «своих» – «чужих», которое специфично именно для сетевых сообществ.

Ещё одной специфичной характеристикой субъектности сетевых сообществ является выражение в дискурсе гражданской идентичности. Декларации активной позиции и установок на изменение существующего положения, критике власти соответствуют в качестве предикторов субъектности показатели РСА, относящиеся как к специфической лексике, связанной с властью и протестом, отрицательной рациональной оценкой, так и к семантической роли – *ликвидатив*, отражающей радикальные действия. Получается, что именно направленность на активные преобразования, а не просто обсуждение общественно значимых тем, социального неравенства и несправедливости, способствует повышению субъектности сетевого сообщества. С определёнными оговорками совокупность этих предикторов складывается в ещё одну гиперроль – «протестные радикальные действия».

Обнаружение предиктора субъектности *прецедентные тексты*, *архаизмы*, *междометия* и др. можно интерпретировать по-разному: рассмотреть

его как показатель языковой и понятийной общности, проявляющейся в *прецедентных текстах* и *архаизмах*, или объяснить изменением тональности дискурса, выражающейся, в том числе и в *междометиях*, которые сопутствуют эмоциональному побуждению к действию или эмоциональной поддержке, представляющими собой дискурсивные маркеры субъектности. В связи с этим данный показатель РСА требует уточнения и дифференциации.

Наконец, большее *число предложений* наряду с меньшим *числом символов* свидетельствует о «насыщенности» дискурса сообществ с высоким уровнем субъектности и переходе от открытой полемики к «деловому» дискурсу с включением эллиптических конструкций, указывающих на слаженность, понимание с полуслова, повышенную активность участников коммуникации и их готовность к совместным действиям.

ВЫВОДЫ

Показатели РСА на прецедентных текстах позволяют исчерпывающим образом предопределять уровень субъектности сетевого сообщества. Для всех 12 показателей субъектности СС, описывающих субъектность сетевого сообщества на 3 уровнях, удалось построить регрессионные модели со скорректированным коэффициентом детерминации, приближающимся к 1.

Семантические связи и роли в контенте сетевых сообществ с высоким уровнем субъектности в совокупности описывают деятельностный характер отношений, складывающийся в этих сообществах. Другой характеристикой контента более субъектных сетевых сообществ является протестный характер. Указанные характеристики находят своё выражение в двух гиперролях таких сообществ: «оперативность действия» и «протестные радикальные действия».

При повышении субъектности сетевого сообщества повышается «насыщенность» контента и наблюдается переход к «деловому» дискурсу с элементами эллиптичности.

ЛИТЕРАТУРА

1. Воронин А. Н. Методологические проблемы исследования субъектности сетевых сообществ // Психология и Психотехника. 2019. № 3 [Электронный ресурс]. URL: https://nbpublish.com/ptmag/contents_2019_3.html#30388 (дата обращения 09.04.2020).
2. Воронин А. Н., Ковалева Ю. В. Изменение субъектности сетевого сообщества в процессе троллинга // Социальная и экономическая психология. 2019. Т. 4. № 3 (15). С. 25–61.
3. Вучков И., Бояджиева Л., Солаков Е. Прикладной линейный регрессионный анализ / пер. с болг. Ю. П. Адлера. М.: Финансы и статистика, 1987. 239 с.
4. Емельянова Т. П., Журавлев А. Л. Психология больших социальных групп как коллективных субъектов // Психологический журнал. 2009. Т. 30. № 3. С. 5–15.

5. Журавлев А. Л. Психология коллективного субъекта // Психология индивидуального и группового субъекта / К. А. Абульханова, В. А. Барабанщиков, А. В. Брушлинский и др. М.: ПЕР СЭ, 2002. С. 51–81.
6. Журавлев А. Л. Коллективный субъект как феномен и понятие в современной психологии // Разработка понятий современной психологии / отв. ред. А. Л. Журавлёв, Е. А. Сергиенко. М.: Институт психологии РАН, 2018. С. 116–161.
7. Методы выявления по тексту психологических характеристик автора (на примере агрессивности) / А. К. Ковалёв, Ю. М. Кузнецова, А. Н. Минин, М. Ю. Пенкина, И. В. Смирнов, М. А. Станкевич, Н. В. Чудова. Вопросы кибербезопасности. 2019. № 4 (32). С. 72–79.
8. Осипов Г. С., Смирнов И. В., Тихомиров И. А. Реляционно-ситуационный метод поиска и анализа текстов и его приложения // Искусственный интеллект и принятие решений. 2008. № 2. С. 3–10.
9. Оценка субъектности сетевых сообществ: сопоставление дискурсивных маркеров и показателей PCA / А. Н. Воронин, Н. Д. Павлова, Т. А. Гребенщикова, Т. А. Кубрак, И. В. Смирнов // Социальная и экономическая психология. 2020. Т. 5. № 2 (18). С. 330–364.
10. Павлова Н. Д. Интерактивный аспект дискурса: подходы к исследованию // Психологический журнал. 2005. Т. 26. № 4. С. 66–76.
11. Павлова Н. Д., Кубрак Т. А., Гребенщикова Т. А. Исследование динамики субъектности сетевых сообществ по её проявлению в дискурсе // Психологические исследования. 2020. Т. 13. № 70 [Электронный ресурс]. URL: <http://psystudy.ru/index.php/num/2020v13n72/1782-pavlova72.html?fontstyle=f-larger> (дата обращения: 09.04.2020).
12. Погорский Э. К. Особенности цифровых гуманитарных наук // Знание. Понимание. Умение. 2014. № 5 [Электронный ресурс]. URL: http://www.zpu-journal.ru/e-zpu/2014/5/Pogorskiy_Digital-Humanities (дата обращения: 09.04.2020).
13. Семантико-синтаксический анализ естественных языков. Часть II. Метод семантико-синтаксического анализа текстов / И. В. Смирнов, А. О. Шелманов, Е. С. Кузнецова, И. В. Храмоин // Искусственный интеллект и принятие решений. 2014. № 1. С. 11–24.
14. Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований / С. Н. Ениколопов, Ю. М. Кузнецова, И. В. Смирнов, М. А. Станкевич, Н. В. Чудова // Искусственный интеллект и принятие решений. Ч. 1. Методические и методологические аспекты. 2019. № 2. С. 28–38.
15. Субъектность сетевого сообщества: сравнение психометрических моделей проявления дискурсивных маркеров в контенте / А. Н. Воронин, Т. А. Гребенщикова, Т. А. Кубрак, Н. Д. Павлова // Вестник Московского государственного областного университета. Серия: Психологические науки. 2019. № 3. С. 6–24.
16. Текстовые проявления фрустрированности пользователя социальных сетей / Ю. М. Кузнецова, И. А. Курузов, И. В. Смирнов, М. А. Станкевич, Е. В. Старостина, Н. В. Чудова // Медиалингвистика. 2020. № 7 (1). С. 4–15.

17. Al-Mosaiwi M., Johnstone T. In an Absolute State: Elevated Use of Absolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation // *Clinical Psychological Science*. 2018. № 6 (4). P. 529–542.
18. Azucar D., Marengo D., Settanni M. Predicting the big 5 personality traits from digital footprints on social media: A meta-analysis // *Personality and Individual Differences*. 2018. V. 124. P. 150–159.
19. Drouin M., Boyd R. L., Greidanus Romanelli M. Predicting recidivism among internet child sex sting offenders using psychological language analysis // *Cyberpsychology, Behavior, and Social Networking*. 2018. Vol. 21. P. 78–83.
20. Grunebaum M. F. Suicidology meets “Big Data” // *Journal of Clinical Psychiatry*. 2015. Vol. 76. № 3. P. e383–e384.
21. TACIT: An open-source text analysis, crawling, and interpretation tool / M. Dehghani, K. M. Johnson, J. Garten, R. Boghrati, J. Hoover, V. Balasubramanian, A. Singh, Y. Shankar, L. Pulickal, A. Rajkumar, N. J. Parmar // *Behavior Research Methods*. 2017. № 49 (2). P. 538–547.
22. Tracking suicide risk factors through Twitter in the US / J. Jashinsky, S. H. Burton, C. L. Hanson, J. West, C. Giraud-Carrier // *Crisis*. 2014. Vol. 35. P. 51–59.
23. Lambiotte R., Kosinski M. Tracking the Digital Footprints of Personality // *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE)*. 2014. № 102 (12). P. 1934–1939.
24. Manovich L. Cultural Data: Possibilities and Limitations of Digitized Archives // *Museum and Archive on the Move. Changing Cultural Institutions in the Digital Era*. Berlin, Boston: De Gruyter, 2017. P. 259–276.
25. Pennebaker J. W. Campbell R. S. The effects of writing about traumatic experience // *Clinical Quarterly*. 2000. № 9. P. 17–21.
26. Potter J. Discursive psychology and the study of Naturally occurring Talk // *Qualitative Research*. SAGE Publications, 2011. P. 187–207.
27. Soffer O. The Internet and National Solidarity: A Theoretical Analysis // *Communication Theory*. 2013. Vol. 23. № 1. P. 48–66.
28. The Study of Network Community Capacity to be a Subject: Digital Discursive Footprints / A. N. Voronin, T. A. Grebenshikova, T. A. Kubrak, T. A. Nestik, N. D. Pavlova // *Behavioral Sciences*. 2019. № 9. URL: <https://doi.org/10.3390/bs9120119> (дата обращения: 09.04.2020).
29. Towards assessing changes in degree of depression through Facebook / A. H. Schwartz, J. Eichstaedt, M. Kern, G. Park, M. Sap, D. Stillwell, L. Ungar // *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology*. Association for Computational Linguistics. 2014. P. 118–125.
30. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text / F. Mairesse, M. A. Walker, M. R. Mehl, R. K. Moore // *Journal of Artificial Intelligence Research*. 2007. № 30. P. 457–500.
31. Wilson S. R. Natural language processing for personal values and human activities: a dis. ... Doctor of Philosophy. Michigan, 2019. 146 p. URL: <https://deepblue.lib.umich.edu/handle/2027.42/150025> (дата обращения: 09.04.2020).
32. Yarkoni T. Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers // *Journal of Research in Personality*. 2010. № 44. P. 363–373.

REFERENCES

1. Voronin A. N. [Methodological problems of studying the subjectivity of network communities]. In: *Psihologiya i Psihotekhnika* [Psychology and Psychotechnics], 2019, no. 3. Available at: https://nbpublish.com/ptmag/contents_2019_3.html#30388 (accessed: 09.04.2020).
2. Voronin A. N., Kovaleva Yu. V. [Changing the subjectivity of the network community in the process of trolling]. In: *Social'naya i ekonomicheskaya psihologiya* [Social and Economic Psychology], 2019, vol. 4, no. 3 (15), pp. 25–61.
3. Vuchkov I., Boyadzhieva L., Solakov E. *Prikladnoj linejnyj regressionnyj analiz* [Applied Linear Regression Analysis; transl. from Bulg. by Yu. P. Adler]. Moscow, Finance and statistics, 1987. 239 p.
4. Emel'yanova T. P., Zhuravlev A. L. [Psychology of large social groups as collective subjects]. In: *Psihologicheskij zhurnal* [Psychological Magazine], 2009, vol. 30, no. 3, pp. 5–15.
5. Zhuravlev A. L. [Psychology of collective subject]. In: *Psihologiya individual'nogo i gruppovogo sub'ekta* [Psychology of the individual and group subject; K. A. Abul'hanova, V. A. Barabanshchikov, A. V. Brushlinskij and etc.]. Moscow, *Per Se*, 2002, pp. 51–81.
6. Zhuravlev A. L. [The collective subject as a phenomenon and concept in modern psychology]. In: Zhuravlev A. L., Sergienko E. A., eds. *Razrabotka ponyatii sovremennoj psikhologii* [Development of concepts of modern psychology]. Moscow, Institute of Psychology RAS Publ., 2018, pp. 116–161.
7. Kovalyov A. K., Kuznecova Yu. M., Minin A. N., Penkina M. Yu., Smirnov I. V., Stankevich M. A., Chudova N. V. Methods for identifying psychological characteristics of the author in the text (on example of aggressiveness). In: *Voprosy kiberbezopasnosti* [Cybersecurity issues], 2019, no. 4 (32), pp. 72–79.
8. Osipov G. S., Smirnov I. V., Tihomirov I. A. [Relational-situational method of searching and analyzing texts and its applications]. In: *Iskusstvennyj intellekt i prinyatie reshenij* [Artificial intelligence and decision making], 2008, no. 2, pp. 3–10.
9. Voronin A. N., Pavlova N. D., Grebenshchikova T. A., Kubrak T. A., Smirnov I. V. [Assessment of subjectivity of network communities: comparison of discourse markers and PCA indicators]. In: *Social'naya i ekonomicheskaya psihologiya* [Social and Economic Psychology], 2020, vol. 5, no. 2 (18), pp. 330–364.
10. Pavlova N. D. [The Interactive Aspect of Discourse: Approaches to Research]. In: *Psihologicheskij zhurnal* [Psychological journal], 2005, vol. 26, no 4, pp. 66–76.
11. Pavlova N. D., Kubrak T. A., Grebenshchikova T. A. [Study of the dynamics of subjectivity of network communities by its manifestation in discourse]. In: *Psihologicheskie issledovaniya* [Psychological research], 2020, vol. 13, no 70. Available at: <http://psystudy.ru/index.php/num/2020v13n72/1782-pavlova72.html?fontstyle=f-larger> (accessed: 09.04.2020).
12. Pogorskiy E. K. [Features of the digital humanities]. In: *Znanie. Ponimanie. Umenie* [Knowledge. Understanding. Skill], 2014, no. 5. Available at: http://www.zpu-journal.ru/e-zpu/2014/5/Pogorskiy_Digital-Humanities (accessed: 09.04.2020).

13. Smirnov I. V., Shelmanov A. O., Kuznecova E. S., Hramoin I. V. [Semantic-syntactic analysis of natural languages. Part II. The method of semantic-syntactic analysis of texts]. In: *Iskusstvennyj intellekt i prinyatie reshenij* [Artificial intelligence and decision making], 2014, no. 1, pp. 11–24.
14. Enikolopov S. N., Kuznecova Yu. M., Smirnov I. V., Stankevich M. A., Chudova N. V. [Creation of a tool for automatic text analysis in the interests of socio-humanitarian research. Part 1. Methodological and methodological aspects]. In: *Iskusstvennyj intellekt i prinyatie reshenij* [Artificial intelligence and decision making], 2019, no. 2, pp. 28–38.
15. Voronin A. N., Grebenshikova T. A., Kubrak T. A., Pavlova N. D. The subjectness of the network community: a comparison of psychometric models of the discursive markers manifestation in content. In: *Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Seriya: Psihologicheskie nauki* [Bulletin of the Moscow Region State University. Series: Psychology], 2019, no. 3, pp. 6–24.
16. Kuznecova Yu. M., Kuruzov I. A., Smirnov I. V., Stankevich M. A., Starostina E. V., Chudova N. V. [Textual manifestations of frustration of a user of social networks]. In: *Medialingvistika* [Media linguistics], 2020. № 7 (1), pp. 4–15.
17. Al-Mosaiwi M., Johnstone T. In an Absolute State: Elevated Use of Absolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation. In: *Clinical Psychological Science*, 2018, no. 6 (4), pp. 529–542.
18. Azucar D., Marengo D., Settanni M. Predicting the big 5 personality traits from digital footprints on social media: A meta-analysis. In: *Personality and Individual Differences*, 2018, vol. 124, pp. 150–159.
19. Drouin M., Boyd R. L., Greidanus Romanelli M. Predicting recidivism among internet child sex sting offenders using psychological language analysis. In: *Cyberpsychology, Behavior, and Social Networking*, 2018, vol. 21, pp. 78–83.
20. Grunebaum M. F. Suicidology meets “Big Data”. In: *Journal of Clinical Psychiatry*, 2015, vol. 76, no. 3, pp. e383–e384.
21. Dehghani M., Johnson K. M., Garten J., Boghrati R., Hoover J., Balasubramanian V., Singh A., Shankar Y., Pulickal L., Rajkumar A., Parmar N. J. TACIT: An open-source text analysis, crawling, and interpretation tool. In: *Behavior Research Methods*, 2017, no. 49 (2), pp. 538–547.
22. Jashinsky J., Burton S. H., Hanson C. L., West J., Giraud-Carrier C. Tracking suicide risk factors through Twitter in the US. In: *Crisis*. 2014, vol. 35, pp. 51–59.
23. Lambiotte R., Kosinski M. Tracking the Digital Footprints of Personality. In: *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE)*, 2014, no. 102 (12), pp. 1934–1939.
24. Manovich L. Cultural Data: Possibilities and Limitations of Digitized Archives. In: *Museum and Archive on the Move. Changing Cultural Institutions in the Digital Era*. Berlin, Boston: De Gruyter, 2017. P. 259–276.
25. Pennebaker J. W. Campbell R. S. The effects of writing about traumatic experience. In: *Clinical Quarterly*, 2000, no. 9, pp. 17–21.
26. Potter J. Discursive psychology and the study of Naturally occurring Talk. In: *Qualitative Research*. SAGE Publications, 2011, pp. 187–207.
27. Soffer O. The Internet and National Solidarity: A Theoretical Analysis. In: *Communication Theory*, 2013, vol. 23, no. 1, pp. 48–66.

28. Voronin A. N., Grebenshchikova T. A., Kubrak T. A., Nestik T. A., Pavlova N. D. The Study of Network Community Capacity to be a Subject: Digital Discursive Footprints. In: *Behavioral Sciences*, 2019, no. 9. Available at: <https://doi.org/10.3390/bs9120119> (accessed: 09.04.2020).
29. Schwartz A. H., Eichstaedt J., Kern M., Park G., Sap M., Stillwell D., Ungar L. Towards assessing changes in degree of depression through Facebook. In: *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology. Association for Computational Linguistics*, 2014, pp. 118–125.
30. Mairesse F., Walker M. A., Mehl M. R., Moore R. K. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. In: *Journal of Artificial Intelligence Research*, 2007, no. 30, pp. 457–500.
31. Wilson S. R. Natural language processing for personal values and human activities: a dis. ... Doctor of Philosophy. Michigan, 2019. 146 p. University of Michigan. 2019. Available at: <https://deepblue.lib.umich.edu/handle/2027.42/150025> (accessed: 09.04.2020).
32. Yarkoni T. Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. In: *Journal of Research in Personality*, 2010, no. 44, pp. 363–373.

ДАТА ПУБЛИКАЦИИ

Статья поступила в редакцию: 23.08.2020

Статья размещена на сайте: 28.09.2020

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Воронин Анатолий Николаевич – доктор психологических наук, профессор, главный научный сотрудник лаборатории психологии речи и психолингвистики Института психологии Российской академии наук; e-mail: voroninan@bk.ru

Кубрак Тина Анатольевна – кандидат психологических наук, научный сотрудник лаборатории психологии речи и психолингвистики Института психологии Российской академии наук; e-mail: kubrak.tina@gmail.com

Смирнов Иван Валентинович – кандидат физико-математических наук, доцент, заведующий отделом Федерального исследовательского центра «Информатика и управление» Российской академии наук; e-mail: ivs@isa.ru

Станкевич Максим Алексеевич – инженер-исследователь Федерального исследовательского центра «Информатика и управление» Российской академии наук; e-mail: stankevich@isa.ru

Anatoly N. Voronin – Dr. Sci. (Psychology), Prof., Chief Researcher, Laboratory of Psychology of Speech and Psycholinguistics, Institute of Psychology, Russian Academy of Sciences; e-mail: voroninan@bk.ru

Tina A. Kubrak – Cand. Sci. (Psychology), Researcher, Laboratory of Psychology of Speech and Psycholinguistics, Institute of Psychology, Russian Academy of Sciences; e-mail: kubrak.tina@gmail.com

Ivan V. Smirnov – Cand. Sci. (Phys.-Math.), Assoc. Prof., Departmental Head, Federal Research Center “Computer Science and Management” of Russian Academy of Sciences; e-mail: ivs@isa.ru

Maxim A. Stankevich – Research Engineer, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences; e-mail: stankevich@isa.ru

ПРАВИЛЬНАЯ ССЫЛКА НА СТАТЬЮ / FOR CITATION

Оценка субъектности сетевого сообщества по показателям автоматического реляционно-ситуационного анализа его контента / А. Н. Воронин, Т. А. Кубрак, И. В. Смирнов, М. А. Станкевич // Вестник Московского государственного областного университета (электронный журнал). 2020. № 3. URL: www.evestnik-mgou.ru.

Voronin A. N., Kubrak T. A., Smirnov I. V., Stankevich M. A. Assessment of the Network Community Subjectivity by Indicators of Content Automatic Relational-Situational Analysis. In: *Bulletin of Moscow Region State University (e-journal)*, 2020, no. 3. Available at: www.evestnik-mgou.ru.