

МЕТОДЫ ВЫЯВЛЕНИЯ ПО ТЕКСТУ ПСИХОЛОГИЧЕСКИХ ХАРАКТЕРИСТИК АВТОРА (НА ПРИМЕРЕ АГРЕССИВНОСТИ)¹

Ковалёв А.К.², Кузнецова Ю.М.³, Минин А.Н.⁴, Пенкина М.Ю.⁵, Смирнов И.В.⁶, Станкевич М.А.⁷, Чудова Н.В.⁸

Цель исследования: проверка возможностей нового инструмента автоматического анализа текста в задаче выявления текстовых параметров, специфичных для людей с определёнными психологическими особенностями; получение данных о признаках, отличающих тексты людей с высокой личностной агрессивностью.

Метод: применялся инструмент корпусных лингво-статистических исследований, опирающийся на использование реляционно-ситуационного анализа, психолингвистических показателей и словарей, охватывающих лексику эмоциональной и рациональной оценки; для получения данных об уровне агрессивности испытуемых применялся опросник Басса-Перри; при обработке данных использовались алгоритмы бинарной классификации – метод опорных векторов (SVM) и случайный лес (Random Forest).

Полученные результаты: созданное средство когнитивного анализа текста позволило получить достаточное количество дифференцирующих признаков в задаче оценки психологических особенностей автора текста; разработанная процедура обработки данных привела к улучшению качества классификации; выявлены синтаксические, семантические и лексические признаки, отличающие тексты, написанные людьми с высокой личностной агрессивностью разных форм.

Ключевые слова: когнитивный анализ текста, личностная агрессивность, методы бинарной классификации

DOI: 10.21681/2311-3456-2019-4-72-79

1. Введение

Для социо-гуманитарных исследований в современных условиях большое значение приобретает анализ дискурсивного пространства социальных медиа, в том числе, анализ текстовой продукции людей с различными психологическими особенностями. Одной из важнейших тем современных исследований медиа является распространение агрессивных установок в СМИ и социальных сетях. В настоящей работе представлены результаты использования в оценке личностной агрессивности автора по его текстам разрабатываемого в ФИЦ ИУ РАН инструмента корпусных лингво-статистических

исследований, опирающегося на использование реляционно-ситуационного анализа, психолингвистических показателей и словарей, охватывающих лексику эмоциональной и рациональной оценки [5]. Применение машинного обучения к данным психодиагностического обследования испытуемых, чьи тексты исследовались с помощью нового инструмента когнитивного анализа текста, привело к выделению текстовых характеристик, позволяющих отнести автора текста к группе людей с высокой или низкой агрессивностью.

1 Работа выполнена при частичной финансовой поддержке РФФИ (проект № 17-29-02247 «Создание методов диагностики распространения фрустрации в сетевых дискуссиях» и проект № 18-00-00606 (18-00-00233) «Методы комплексного интеллектуального анализа информации различных типов для социо-гуманитарных исследований в социальных медиа»).

2 Ковалев Алексей Константинович, младший научный сотрудник, Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва, Россия. E mail: alexeykkov@gmail.com

3 Кузнецова Юлия Михайловна, кандидат психологических наук, старший научный сотрудник, Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва, Россия. E mail: kuzjum@yandex.ru

4 Минин Алексей Николаевич, старший преподаватель Курганского государственного университета (ГБОУ ВПО КГУ), г. Курган, Россия. E mail: arni_arv@yahoo.com

5 Пенкина Марина Юрьевна, старший преподаватель Московского городского психолого-педагогического университета (ГБОУ ВПО МГППУ), г. Москва, Россия. E mail: mpenkina@mail.ru

6 Смирнов Иван Валентинович, кандидат физико-математических наук, заведующий отделом, Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва, Россия. E mail: ivs@isa.ru

7 Станкевич Максим Алексеевич, младший научный сотрудник, Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва, Россия. E mail: maxastan95@gmail.com

8 Чудова Наталья Владимировна, кандидат психологических наук, старший научный сотрудник, Федеральный исследовательский центр «Информатика и управление» РАН, г. Москва, Россия. E mail: nchudova@gmail.com

2. Обзор

Развитие информационно-коммуникативного пространства, опосредованного цифровыми технологиями, сопровождается не только расширением предоставляемых ими возможностей, но и возникновением специфических рисков и угроз. Негативные феномены виртуальной среды Т.В. Владимирова объединяет в несколько групп: 1) угрозы безопасности личности, связанные с расширением возможностей манипулирования сознанием человека, информационной перегрузкой, с ростом интернет-зависимости и развитием форм психосоциальной депривации; 2) угрозы использования во вред персональных данных, расширение возможностей скрытого сбора персональной информации; 3) информационные угрозы, связанные с расширением масштабов манипуляции общественным мнением, появлением возможностей эффективной организации деструктивных процессов в ценностных системах общества; 4) информационные угрозы безопасности личности, общества, государства, связанные с эффективностью сетевых структур отечественной и международной преступности и терроризма; 5) угрозы неконтролируемого распространения информационно-психологического оружия и его применения в информационных войнах; 6) угрозы стабильности существующих политических режимов власти: системные и периферийные, а также обусловленные сетевой логикой многих социальных процессов в обществе [3, с. 177-178]. Как отмечает Л.М. Мкртчян, «систематическая незащищенность человека перед угрозами, вызванными динамикой и плотностью коммуникативного пространства, нерациональным использованием достижений информационно-коммуникативных технологий, отсутствием самоконтроля, информационной культуры пользователя и культуры безопасности обостряют проблему выработки механизмов по обеспечению социальной безопасности в сети» [8].

В психологическом отношении на первый план среди угроз, возникающих в сетевой среде, выдвигается вербальная агрессия, формами проявления которой выступают троллинг, кибербуллинг, недобросовестная социальная инженерия, экстремистская пропаганда и манифестации психически нездоровых коммуникантов [11]. Отмечаемые исследователями свойства современного сетевого коммуникативного пространства – провокативность и принудительность коммуникации [10] – создают предпосылки для роста уровня агрессивности и вообще деструктивных форм он-лайн и офф-лайн поведения. Так, в ходе проведенного в США в 2013 г. онлайн-опроса было выявлено, что 78% респондентов старше 18 лет отмечают «спад в уровне цивилизованности общения» и возрастающую грубость в Facebook и Twitter, что в некоторых случаях приводило к ссорам между членами семьи и друзьями из реальной жизни [7]. Все более значимой становится проблема распространения в сети эмоций и настроений, и в первую очередь негативных, «заражение» которыми значительного числа коммуникантов является опасным с точки зрения индивидуального и социального психологического здоровья [16, 18], причем, как показывают данные двадцатилетнего лонгитюдного исследования, эффект от

распространения по социальным сетям эмоциональных состояний имеет долговременный характер [19].

Возникающие в сетевой среде угрозы закономерным образом стимулируют разработку исследовательского аппарата для изучения их факторов, механизмов и последствий. Эмоциональная природа тех процессов, которые преимущественно определяют деструктивный характер воздействия на человека негативных сетевых феноменов, задает в качестве приоритетного направления создание средств автоматического анализа, позволяющих оценивать выраженность в сетевом общении признаков аффективных состояний участников коммуникации, то есть методов сантимерт-анализа в широком смысле слова.

В работе [21] описаны три подхода к решению задачи автоматического распознавания эмоций в тексте: 1) подход, основанный на правилах, при котором выявляются ключевые слова и их сочетания со словами, имеющими явное эмоциональное значение; при этом используются лексические ресурсы, содержащие лексику с эмоциональным значением, такие, как WordNet-Affect и SentiWordNet; 2) машинное обучение, подразумевающее анализ n-грамм, пунктуации, эмотиконов, хэштегов, средств отрицания и т.п. с последующим использованием классификаторов на основе деревьев решений, наивного байесовского классификатора, SVM и др.; 3) глубокое машинное обучение на основе нейронных сетей – рекуррентных (LSTM, BiLSTM), сверточных и т.д. Авторы работы [17] противопоставляют подход, основанный на машинном обучении в различных его вариантах, лексическому.

Имея достаточно высокие показатели (так, инструмент SentiStrength позволяет распознавать в текстах на MySpace положительные эмоции с точностью 60,6%, а отрицательные – с точностью 72,8% [24]), лексический подход ограничен тем, что направлен на выявление только вербально выраженных эмоций. Позволяющие преодолевать данное ограничение методы идентификации скрытого сантимерта подразумевают использование баз знаний, содержащих понятия, ассоциируемые носителями языка с определенными эмоциями [15]. В настоящее время ведутся также исследования, направленные на расширение способов определения тональности текстов на стадии обучения; например, предлагается альтернатива традиционной экспертной разметке в виде автоматической классификации на основе включенных в высказывание маркеров, имеющих доказанную дифференцирующую силу (например, предварительно типологизированных смайликов [21]). Еще одним направлением совершенствования инструментария является расширение перечня идентифицируемых эмоций; например, в работе [14] предлагается метод классификации, основанный на алгоритме предсказания по частичному совпадению (PPM), который позволяет с достаточной степенью точности выявлять в текстах LiveJournal выраженность шести базовых эмоций: гнева, отвращения, страха, радости, печали и удивления.

Определение эмоциональной составляющей текста используется для исследования различных форм и проявлений он-лайн агрессивности: речевой враждебно-

сти [22], троллинга [20], кибербуллинга [26], разных видов манипулирования [12, 13], разжигания розни [27] и пр. Сетевая составляющая такой формы офф-лайн агрессии, как массовый протест, является предметом комплексного метода киберметрии, или киберметрического анализа, описанного в [1, 2]. Киберметрия применяется в целях сегментирования информационных потоков на основе поисковых запросов и словарей маркеров; всего в исследовании фигурируют словари маркеров 14 типов социально-медийных документов по степени радикальности выраженных в них протестных установок и объектов сообщений. Сопоставление долей активно-нигилистических и пассивно-нигилистических массивов сообщений позволяет оценивать динамику реальных массовых протестов, а также осуществлять поиск факторов и триггеров, провоцирующих массовый протест. Кроме того, киберметрия включает в себя проведение анализа облаков тегов, геолокационного анализа сообщений социальных медиа, анализ демографических характеристик авторов сообщений, анализ публикационной активности лидеров общественного мнения в социальных медиа и выявление структуры сетевых социальных связей.

С точки зрения структуры сетевого общения анализ эмоциональной составляющей может проводиться на уровне отдельного высказывания, отдельного поста или комментария, а также на уровне пользователя [25]. В последнем случае агрессивность выступает не как характеристика текстов, а как личностная особенность автора, поэтому для ее оценки применяются психодиагностические методы [6, 23].

3. Метод

В Институте проблем искусственного интеллекта ФИЦ ИУ РАН разрабатывается система когнитивного анализа текста, основанная на методах реляционно-ситуационного анализа, лексико-частотного анализа и психолингвистического анализа русскоязычных текстов научного, публицистического и бытового жанров – «Машина РСА». Система работает на основе процессора Exactus Expert и в настоящий момент позволяет искать в тексте показатели трёх типов:

- предикатно-синтаксемные структуры, отражающие представления автора об объектах и событиях, а также о принимаемых и отвергаемых им сценариях поведения
- показатели частотности в тексте лексических единиц, относящихся к определенным тематическим группам слов (ТГС),
- психолингвистические показатели, отражающие эмоциональное состояние автора.

Метод РСА направлен на выявление предикатно-аргументной семантики, которая представляет высказывание на естественном языке в виде действия, события или ситуации, которые выражаются предикатами. Такая модель в принципе не зависит от предметной области и способна описать почти любое высказывание на естественном языке. В РСА семантика высказываний и текста, состоящего из высказываний, представляется в виде семантической сети, в узлах которой находятся предикатные слова и синтаксемы (Г.А.Золотова), а в

качестве ребер выступают синтаксические связи и семантические связи (Г.С.Осипов). Так организованная модель высказывания позволяет выявлять представление автора об упоминаемых им в тексте предметах, событиях или ситуациях. РСА работает на основе Словаря предикатов (2,7 тыс. глаголов, причастий, деепричастий и девербативов) и Определителя семантических ролей (75 семантических отношений) [9].

Под тематическими группами слов (ТГС) в лингвистике понимается объединение слов на основе классификации предметов и явлений в соответствии с экстралингвистическим принципом сопряженности с определенной темой, так что в ТГС объединяются слова вне зависимости от их частеречной принадлежности и связанные друг с другом различными парадигматическими и синтагматическими отношениями. Для составления ТГС применяется способ сплошной выборки лексических единиц с соответствующей целям исследования семантикой из наиболее полных по объему лексического материала словарей, и привлечение материала из специальных тематических словарей. Благодаря этому ТГС представляет собой относительно полный список существующих в конкретном языке средств выражения, используемых носителями для речевого общения на определенную тему. В лексике русского языка нами были выделены группы с тематикой напряжения и фрустрации и в настоящий момент используются словари с ТГС четырёх типов: лексика экспрессии, лексика отрицательной эмоциональной оценки, лексика отрицательной рациональной оценки, лексика социального стресса. Всего в состав созданных 30 словарей входит более 51 тыс. лексических единиц.

В разрабатываемой системе также используется 27 психолингвистических и синтаксических показателей (средняя длина слова, количество безличных глаголов и т.п.), продемонстрировавших в нашем пилотажном исследовании статически значимые связи с более чем 50 показателями психологических опросников [4].

В данной работе была предпринята попытка найти текстовые показатели, используемые в Машине РСА, указывающие с определенной мерой достоверности на особенности личностной агрессивности автора.

4. Процедура исследования

Было проведено психодиагностическое обследование (487 испытуемых – студенты и взрослые г.Москвы и г.Кургана), включающее в себя изучение личностной агрессивности с помощью опросника Басса-Перри ВРАQ. Опросник содержит три шкалы – шкала склонности к физической агрессии, шкала склонности к гневу, шкала враждебности (подозрительность, обидчивость, недоброжелательность). Кроме того, испытуемые писали эссе на тему «Я, другие, мир» объемом 0,5-1 стр. Тексты эссе были обработаны с помощью инструмента «Машина РСА», так что текст каждого автора оказался численно охарактеризован по 177 параметрам.

5. Процедура обработки данных

Данные представляют собой набор из 177 признаков для 487 испытуемых. Признаки разбиты на шесть групп: 1. Психолингвистические показатели (PI) – 27 признаков;

- 2. Семантические роли (SR) – 75 признаков;
- 3. Семантические связи (SL) – 32 признака;
- 4. Словари оценки и состояния (AS) – 20 признаков;
- 5. Тематические словари (SD) – 9 признаков;
- 6. Части речи (PS) – 14 признаков.

В качестве целевых признаков выступали физическая агрессия (Physical Aggression), гневливость (Anger) и враждебность (Hostility). Для каждого целевого признака высчитывались значения первого и третьего квартиля. Испытуемые, у которых значение целевого признака больше первого квартиля, но меньше третьего квартиля исключались и выборки. При этом значение больше третьего квартиля считались высокими и таким испытуемым присваивалась метка класса 1, значения меньше третьего квартиля считались низкими и таким испытуемым присваивалась метка класса 0. Таким образом задача предсказания значения целевого признака сводилась к задаче классификации. В таблице 1 представлены некоторые статистики полученных выборок.

Так же, признаки, которые принимали значение «ноль» в более чем 70% случаев, бинаризовались – значения, не равные нулю, приравнивались к единице и играли роль индикатора.

В качестве алгоритмов бинарной классификации использовались метод опорных векторов (SVM) и случайный лес (Random Forest), реализованные в библиотеке scikit-learn, метрика качества – доля верных ответов (accuracy). Настройка гиперпараметров производилась подбором по сетке с 5-кратным скользящим контролем.

Эксперименты проводились отдельно для каждой группы признаков. С помощью оценки важности признаков в алгоритме случайного леса, в каждой группе отбирались 5 признаков с наибольшим значением важности. Совокупность этих признаков по всем группам составила группу наиболее важных признаков (TOP 5). Также были проведены эксперименты на всех признаках (ALL). Результаты экспериментов представлены в таблице 2.

Таблица 1

	Physical Aggression	Anger	Hostility
Всего	259	265	289
Низкое/высокое значение признака	136 / 123	137 / 128	146 / 143
Возраст	21.12 ± 7.15	20.70 ± 6.48	20.99 ± 6.60
Пол	Женщины – 145 Мужчины – 45 Не указан – 69	Женщины – 148 Мужчины – 43 Не указан – 74	Женщины – 156 Мужчины – 57 Не указан – 76

Таблица 2

	Physical Aggression	Anger	Hostility
SVM			
PI	0.529 ± 0.071	0.536 ± 0.055	0.533 ± 0.052
SR	0.525 ± 0.005	0.552 ± 0.102	0.553 ± 0.051
SL	0.548 ± 0.051	0.517 ± 0.004	0.502 ± 0.096
AS	0.540 ± 0.017	0.536 ± 0.035	0.526 ± 0.046
SD	0.590 ± 0.035	0.551 ± 0.049	0.505 ± 0.061
PS	0.525 ± 0.005	0.517 ± 0.004	0.519 ± 0.029
ALL	0.549 ± 0.068	0.596 ± 0.023	0.568 ± 0.025
TOP 5	0.544 ± 0.027	0.623 ± 0.065	0.591 ± 0.060
PI+AS+PS	-	-	0.591 ± 0.060
Random Forest			
PI	0.462 ± 0.078	0.493 ± 0.070	0.513 ± 0.051
SR	0.505 ± 0.076	0.487 ± 0.060	0.536 ± 0.042
SL	0.501 ± 0.079	0.438 ± 0.064	0.502 ± 0.078
AS	0.448 ± 0.042	0.565 ± 0.048	0.512 ± 0.026
SD	0.475 ± 0.059	0.517 ± 0.041	0.436 ± 0.042
PS	0.425 ± 0.070	0.483 ± 0.078	0.571 ± 0.050

	Physical Aggression	Anger	Hostility
ALL	0.517 ± 0.056	0.520 ± 0.088	0.544 ± 0.047
TOP 5	0.621 ± 0.035	0.581 ± 0.044	0.564 ± 0.038
PI+AS+PS	-	-	0.578 ± 0.045

Как видно из результатов отбор наиболее важных признаков каждой группы (использование показателя TOP 5) приводит к улучшению качества классификации. Для враждебности был проведен дополнительный эксперимент, в котором классификация делалась только по психолингвистическим показателям, словарям оценки и состояния и частям речи, что позволило незначительно улучшить качество классификации при использовании случайного леса.

6. Результаты

Проведённая обработка позволила выявить текстовые признаки, наличие которых позволяет отличить тексты, написанные людьми с высокой личностной агрессивностью разных форм. В таблице 3 приведены данные о 10 наиболее важных признаков для текстов людей физически агрессивных, гневливых или враждебных.

В задачу настоящей работы не входит содержательная интерпретация полученных данных, которая объясняла бы более частую встречаемость в текстах людей с разной формой агрессивности тех или иных синтаксических или семантических конструкций и определённой лексики, однако отметим две важные особенности полученных данных. Во-первых, можно выделить признаки, характерные для текстов людей с любой формой агрессивности: семантическая роль «объект» (компо-

нент с предметно-вещественным значением, подвергающийся воздействию), лексика социального разобщения и лексика положительной рациональной оценки и ментальных действий, прилагательные. Во-вторых, можно указать на текстовые параметры, уникальные для каждой формы агрессивности. Для текстов людей с высокой склонностью к физической агрессии это частое употребление слов в семантической роли «адресат» (лицо или предмет, к которому обращено информативное, донативное или эмотивное действие), большая средняя глубина синтаксического дерева, высокая доля глаголов 1го лица, использование лексики страдания. Тексты людей, склонных к гневу, отличаются высокой долей знаков препинания, причастий и деепричастий, большим количеством частиц, а также использованием лексики отрицательной рациональной оценки. Тексты, написанные высоковраждебными людьми, содержат повышенное количество слов в семантической роли «предикат» (признак, приписываемый субъекту), лексику мотивации, деятельности и напряжения и лексику положительной эмоциональной оценки, а также имеют повышенную тональность. Таким образом, полученные данные позволяют как выделять среди текстов те, которые вероятно были написаны людьми с повышенной личностной агрессивностью, так и те, которые могут принадлежать людям с одной из интересующих исследователя форм агрессивности.

Таблица 3

	Physical Aggression	Anger	Hostility
1	Семантическая роль «адресат»	Число знаков препинания / число слов	Число прилагательных
2	Семантическая роль «объект»	Лексика социального разобщения	Семантическая роль «объект»
3	Лексика социального разобщения'	Компьютерный жаргон	Число местоимений-существительных
4	Средняя глубина синтаксического дерева	Лексика отрицательной рациональной оценки	Лексика социального разобщения
5	Доля глаголов 1го лица	Число отрицательных приставок и форм (не, ни, бес, без) / Число слов	Лексика мотивации, деятельности и напряжения
6	Лексика страдания	Семантическая роль «объект»	Лексика положительной эмоциональной оценки
7	Число местоимений-существительных	Число частиц	Число отрицательных приставок и форм (не, ни, бес, без) / Число слов
8	Лексика положительной рациональной оценки и ментальных действий	Лексика положительной рациональной оценки и ментальных действий	Семантическая роль «предикат»
9	Компьютерный жаргон	Число прилагательных	Тональность слов
10	Число прилагательных	Число причастий и деепричастий / число предложений	Лексика положительной рациональной оценки и ментальных действий

7. Выводы

Результаты, полученные с помощью нового инструмента автоматического анализа текста, психодиагностического обследования и применения методов машинного обучения, позволяют говорить о том, что созданное средство когнитивного анализа текста может применяться в задачах, требующих оценки психологических особенностей автора текста. Выявленные в исследовании текстовые характеристики позволяют с определённой долей уверенности отнести автора текста к группе людей с высокой или низкой агрессивностью.

Необходимо в завершение подчеркнуть ограничения, существующие при использовании как наших результатов, так и любых результатов в данной области исследований. Во-первых, оценка психологических особенностей человека по созданному им тексту не может рассматриваться как результат психодиагностического обследования и применяться в задачах постановки

диагноза. Полученные с помощью автоматического анализа текста оценки могут применяться в широком круге задач информационно-психологической безопасности, выделения группы риска по социальным и психологическим стрессорам, социотерапии. Во-вторых, следует учитывать, что помимо стабильных личностных особенностей автора на текст оказывает влияние эмоциональное состояние пишущего и та задача, которую он решает при создании текста. В нашем исследовании испытуемые находились в спокойном состоянии и создавали текст в жанре эссе. Соответственно, перенос полученных данных на другие задачи, т.е. использование выявленных текстовых признаков агрессивности для оценки личности пишущего, возможен для ситуаций, где люди в спокойной обстановке высказывают свои суждения о жизни и делятся своими мировоззренческими установками.

Литература

1. Бродовская Е.В., Домбровская А.Ю., Карзубов Д.Н. и др. Развитие методологии и методики интеллектуального поиска цифровых маркеров политических процессов в социальных медиа // Мониторинг общественного мнения: Экономические и социальные перемены. 2017. № 5. С. 79-104. DOI: 10.14515/monitoring.2017.5.06.
2. Бродовская Е.В., Домбровская А.Ю., Карзубов Д.Н. и др. Состояние и динамика русскоязычных потоков о межнациональных и межрелигиозных отношениях в Крыму и Севастополе: киберметрия и дискурс-анализ сообщений в социальных медиа Рунета // Гуманитарные науки. Вестник Финансового университета. 2018. № 4(34). С. 23-35.
3. Владимирова Т.В. Социальная природа информационной безопасности. М.: АНО Изд. Дом «Науч. обозрение», 2014. 239 с.
4. Воронцова О.Ю., Ениколопов С.Н., Кузнецова Ю.М., Чудова Н.В. и др. Лингвистические характеристики текстов психически больных и здоровых людей // Психологические исследования. 2018. Т. 11. № 61. URL: <http://psystudy.ru/index.php/pum/2018v11n61/1622-enikolopov61.html>
5. Ениколопов С.Н., Кузнецова Ю.М., Смирнов И.В., Станкевич М.А., Чудова Н.В. Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований. Ч.1. Методические и методологические аспекты // Искусственный интеллект и принятие решений. 2019. № 2 (в печати).
6. Кузнецова Ю.М. О существующих подходах к автоматическому анализу текстов в целях психодиагностики // Филология в XXI веке. 2019. Вып. 1 (в печати)
7. Марченко Ф.О., Маховская О.И. Психология сетевой агрессии (кибербуллинга) во время эпидемии нарциссизма // Человек: Образ и сущность. Гуманитарные аспекты. 2018. № 4(35). С. 100–119.
8. Мкртчян Л.М. Риски и угрозы социальной безопасности личности в сетевом коммуникативном пространстве: к постановке проблемы // Приоритетные научные направления: от теории к практике. 2013. № 8. С. 149–155.
9. 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.
10. Сеницкая А.В. Дискурс «жертвы» и границы публичной рефлексии в сетевых сообществах // Международный журнал исследований культуры. 2017. № 4(29). С. 120–127.
11. Шумилова А.А. Информационно-коммуникативные риски виртуального пространства // Международный журнал социальных и гуманитарных наук. 2016. Т. 8. № 1. С. 160–165.
12. Adewole, K.S., Anuar, N.B., Kamsin, A., et al. Malicious accounts: dark of the social networks // Journal of Network and Computer Applications, 2017, 79: 41-67.
13. Ahmad, M., Agarwal, N., Jabin, S., & Hussain, S.Z. Analyzing Real and Fake users in Facebook Network based on Emotions // 11th International Conference on Communication Systems & Networks (COMSNETS). 2019. DOI: 10.1109/COMSNETS.2019.8711124
14. Almahdawi, A., & Teahan, W.J. Automatically Recognizing Emotions in Text Using Prediction by Partial Matching (PPM) Text Compression Method // Al-mamory S., Alwan J., Hussein A. (eds) New Trends in Information and Communications Technology Applications. NTICT 2018. Communications in Computer and Information Science, 2018, 938: 269–283. URL: https://doi.org/10.1007/978-3-030-01653-1_17
15. Balahur, A., Hermida, J.M., & Montoyo, A. Detecting implicit expressions of sentiment in text based on commonsense knowledge // Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis. 2011. URL: <https://www.researchgate.net/publication/262388481>
16. Bond, R.M., Fariss, C.J., Jones, J.J., et al. A 61-million-person experiment in social influence and political mobilization. Nature. 2012, 489(7415): 295–298. doi: 10.1038/nature11421 PMID: 22972300
17. Du, P., & Nie, J.-Y. Mutux at SemEval-2018 Task 1: Exploring Impacts of Context Information On Emotion Detection // Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018). New Orleans, Louisiana, June 5–6. 2018: 345–349.

18. Ferrara, E., & Yang, Z. Measuring Emotional Contagion in Social Media // PLoS ONE, 2015, 10(11): e0142390. doi:10.1371/journal.pone.0142390
19. Fowler, J.H., Christakis, N.A., et al. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study // BMJ. 2008; 337: a2338. doi: 10.1136/bmj. a2338 PMID: 19056788
20. Galán-García, P., Puerta, J.G.D.L., Gómez, C.L., Santos, I., & Bringas, P.G. Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying // Logic Journal of the IGPL, 2016, 24(1): 42-53.
21. Gupta, U., Chatterjee, A., Srikanth, R., & Agrawal, P. A Sentiment-and-Semantics-Based Approach for Emotion Detection in Textual Conversations // Neu-IR: Workshop on Neural Information Retrieval, SIGIR 2017, ACM. URL: arXiv:1707.06996
22. Malmasi, S., & Zampieri, M. Detecting Hate Speech in Social Media. 2017. URL: <https://www.researchgate.net/publication/321902238>
23. Pabian, Sara, De Backer, C.J.S., & Vandebosch, H. Dark Triad personality traits and adolescent cyber-aggression // Personality and Individual Differences, 2015, 75: 41-46.
24. Stieglitz, S., & Dang-Xuan, L. Emotions and information diffusion in social media – Sentiment of microblogs and sharing behavior // Journal of Management Information Systems. 2013, 29(4): 217–248.
25. Schwartz, H.A. Predicting individual well-being through the language of social media // Pacific Symposium on Biocomputing, 2016: 516-527.
26. Tahmasbi, N., & Rastegari, E. A Socio-Contextual Approach in Automated Detection of Public Cyberbullying on Twitter // ACM Transactions on Social Computing – Special Issue on HICSS 2018, 1, 4. doi>10.1145/3290838
27. Wani, M.A., Agarwal, N., Jabin, S., & Hussain, S.Z. User emotion analysis in conflicting versus non-conflicting regions using online social networks. Telematics and Informatics. 2018. DOI: 10.1016/j.tele.2018.09.012

TEXT ANALYSIS APPROACH FOR IDENTIFYING PSYCHOLOGICAL CHARACTERISTICS (WITH AGGRESSIVENESS AS AN EXAMPLE)

***Kovalev A.K.⁹, Kuznetsova Y.M.¹⁰, Minin A.N.¹¹, Penkina M.Y.¹², Smirnov I.V.¹³,
Stankevich M.A.¹⁴, Chudova N.V.¹⁵***

The purpose of the study: to test the capabilities of a new automatic text analysis tool for identifying text parameters specific to people with certain psychological characteristics; to obtain data on the signs that distinguish texts of people with high personal aggressiveness.

Method: a corpus linguistic-statistical research tool was applied, based on the relational-situational analysis, psycholinguistic indicators and dictionaries covering the vocabulary of emotional and rational evaluation; to assess the level of aggressiveness, the Bass-Perry questionnaire was used; the data were processed by binary classification algorithms: support vector machine (SVM) and Random Forest.

Results: thanks to new tool, several textual signs of the authors' psychological characteristics were identified; classification was improved through the use of our data processing method; and some syntactic, semantic and lexical features of texts of highly hostile persons have been identified.

Keywords: cognitive text analysis, personal aggressiveness, binary classification methods

9 Alexey Kovalev, Associate Scientist in Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, Moscow, Russia. E mail: alexeykkov@gmail.com

10. Julia Kuznetsova, Ph.D., Senior Research Fellow in Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, Moscow, Russia. E mail: kuzjum@yandex.ru

11 Alexey Minin, Senior Lecturer of Kurgan state University, Kurgan, Russia. E mail: arni_arv@yahoo.com

12 Marina Penkina, Senior Lecturer of Moscow city University of psychology and education, Moscow, Russia. E mail: mpenkina@mail.ru

13 Ivan Smirnov, Ph.D., head of a department of the Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, Moscow, Russia. E mail: ivs@isa.ru

14 Maxim Stankevich, Associate Scientist in Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, Moscow, Russia. E mail: maxastan95@gmail.com

15 Natalia Chudova, Ph.D., Senior Research Fellow in Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, Moscow, Russia. E mail: nchudova@gmail.com

References

1. Brodovskaya E.V., Dombrovskaya A.Y., Karzubov D.N. i dr. Razvitie metodologii i metodiki intellektual'nogo poiska cifrovyyh markerov politicheskikh processov v social'nykh media // Monitoring obshchestvennogo mneniya: Ekonomicheskie i social'nye peremeny. 2017. № 5. S. 79-104. DOI: 10.14515/monitoring.2017.5.06.
2. Brodovskaya E.V., Dombrovskaya A.Y., Karzubov D.N. i dr. Sostoyanie i dinamika russkoyazychnyyh potokov o mezhnatsional'nykh i mezhreligioznykh otnosheniyah v Krymu i Sevastopole: kibernetriya i diskurs-analiz soobshcheniy v social'nykh media Runeta // Gumanitarnye nauki. Vestnik Finansnogo universiteta. 2018. № 4(34). S. 23-35.
3. Vladimirova T.V. Social'naya priroda informatsionnoy bezopasnosti. M.: ANO Izd. Dom «Nauch. obozrenie», 2014. 239 s.
4. Voroncova O.Y., Enikolopov S.N., Kuznecova Y.M., Chudova N.V. i dr. Lingvisticheskie karakteristiki tekstov psicheskii bol'nykh i zdorovykh lyudej // Psihologicheskie issledovaniya. 2018. T. 11. № 61. URL: <http://psystudy.ru/index.php/num/2018v11n61/1622-enikolopov61.html>
5. Enikolopov S.N., Kuznecova Y.M., Smirnov I.V., Stankevich M.A., Chudova N.V. Sozdanie instrumenta avtomaticheskogo analiza teksta v interesakh socio-gumanitarnyyh issledovaniy. CH.1. Metodicheskie i metodologicheskie aspekty // Iskustvennyy intellekt i prinyatie resheniy. 2019. № 2 (v pechati).
6. Kuznecova Y.M. O sushchestvuyushchih podhodah k avtomaticheskomu analizu tekstov v celyakh psihodiagnostiki // Filologiya v XXI veke. 2019. Vyp. 1 (v pechati)
7. Marchenko F.O., Mahovskaya O.I. Psihologiya setevoy agressii (kiberbullinga) vo vremya epidemii narcissizma // Chelovek: Obraz i sushchnost'. Gumanitarnye aspekty. 2018. № 4(35). S. 100–119.
8. Mkrtchyan L.M. Riski i ugrozy social'noj bezopasnosti lichnosti v setevom kommunikativnom prostranstve: k postanovke problemy // Prioritetnye nauchnye napravleniya: ot teorii k praktike. 2013. № 8. S. 149–155.
9. Osipov G.S., Smirnov I.V., Tihomirov I.A. Relyacionno-situatsionnyy metod poiska i analiza tekstov i ego prilozheniya // Iskustvennyy intellekt i prinyatie resheniy. 2008. № 2. S. 3–10.
10. Sinickaya A.V. Diskurs «zhertvy» i granicy publichnoy refleksii v setevykh soobshchestvakh // Mezhdunarodnyy zhurnal issledovaniy kul'tury. 2017. № 4(29). S. 120–127.
11. Shumilova A.A. Informatsionno-kommunikativnye riski virtual'nogo prostranstva // Mezhdunarodnyy zhurnal social'nykh i humanitarnyyh nauk. 2016. T. 8. № 1. S. 160–165.
12. Adewole, K.S., Anuar, N.B., Kamsin, A., et al. Malicious accounts: dark of the social networks // Journal of Network and Computer Applications, 2017, 79: 41-67.
13. Ahmad, M., Agarwal, N., Jabin, S., & Hussain, S.Z. Analyzing Real and Fake users in Facebook Network based on Emotions // 11th International Conference on Communication Systems & Networks (COMSNETS). 2019. DOI: 10.1109/COMSNETS.2019.8711124
14. Almahdawi, A., & Teahan, W.J. Automatically Recognizing Emotions in Text Using Prediction by Partial Matching (PPM) Text Compression Method // Al-mamory S., Alwan J., Hussein A. (eds) New Trends in Information and Communications Technology Applications. NTICT 2018. Communications in Computer and Information Science, 2018, 938: 269–283. URL: https://doi.org/10.1007/978-3-030-01653-1_17
15. Balahur, A., Hermida, J.M., & Montoyo, A. Detecting implicit expressions of sentiment in text based on commonsense knowledge // Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis. 2011. URL: <https://www.researchgate.net/publication/262388481>
16. Bond, R.M., Fariss, C.J., Jones, J.J., et al. A 61-million-person experiment in social influence and political mobilization. Nature. 2012, 489(7415): 295–298. doi: 10.1038/nature11421 PMID: 22972300
17. Du, P., & Nie, J.-Y. Mutux at SemEval-2018 Task 1: Exploring Impacts of Context Information On Emotion Detection // Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018). New Orleans, Louisiana, June 5–6. 2018: 345–349.
18. Ferrara, E., & Yang, Z. Measuring Emotional Contagion in Social Media // PLoS ONE, 2015, 10(11): e0142390. doi:10.1371/journal.pone.0142390
19. Fowler, J.H., Christakis, N.A., et al. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study // BMJ. 2008; 337: a2338. doi: 10.1136/bmj. a2338 PMID: 19056788
20. Galán-García, P., Puerta, J.G.D.L., Gómez, C.L., Santos, I., & Bringas, P.G. Supervised machine learning for the detection of troll profiles in twitter social network: Application to a real case of cyberbullying // Logic Journal of the IGPL, 2016, 24(1): 42-53.
21. Gupta, U., Chatterjee, A., Srikanth, R., & Agrawal, P. A Sentiment-and-Semantics-Based Approach for Emotion Detection in Textual Conversations // Neu-IR: Workshop on Neural Information Retrieval, SIGIR 2017, ACM. URL: arXiv:1707.06996
22. Malmasi, S., & Zampieri, M. Detecting Hate Speech in Social Media. 2017. URL: <https://www.researchgate.net/publication/321902238>
23. Pabian, Sara, De Backer, C.J.S., & Vandebosch, H. Dark Triad personality traits and adolescent cyber-aggression // Personality and Individual Differences, 2015, 75: 41-46.
24. Stieglitz, S., & Dang-Xuan, L. Emotions and information diffusion in social media – Sentiment of microblogs and sharing behavior // Journal of Management Information Systems. 2013, 29(4): 217–248.
25. Schwartz, H.A. Predicting individual well-being through the language of social media // Pacific Symposium on Biocomputing, 2016: 516-527.
26. Tahmasbi, N., & Rastegari, E. A Socio-Contextual Approach in Automated Detection of Public Cyberbullying on Twitter // ACM Transactions on Social Computing – Special Issue on HICSS 2018, 1, 4. doi>10.1145/3290838
27. Wani, M.A., Agarwal, N., Jabin, S., & Hussain, S.Z. User emotion analysis in conflicting versus non-conflicting regions using online social networks. Telematics and Informatics. 2018. DOI: 10.1016/j.tele.2018.09.012