

ПОДХОД К АНАЛИЗУ И УПРАВЛЕНИЮ КОНТЕНТОМ В ВИРТУАЛЬНОЙ СРЕДЕ

А.О. Исхакова

Институт проблем управления им. В.А. Трапезникова РАН

Россия, 117997, Москва, Профсоюзная ул., 65

E-mail: iao@ipu.ru

Ключевые слова: интернет-контент, социальные сети, социо-киберфизическая система, интеллектуальный анализ данных, большие данные, искусственный интеллект, искусственные нейронные сети, звуковое воздействие, психологическое давление, агрессия в интернете.

Аннотация: Научные исследования в области анализа Интернет-контента являются актуальными в последние годы, это связано с возрастающей и укрепляющейся тенденцией к взаимодействию человека с виртуальной реальностью и вовлечением пользователя в виртуальную среду. Проводимые исследования чаще всего имеют междисциплинарный характер, так как в данной области имеются как социальная, лингвистическая направленность, так и медицинская составляющая, а также математические и технические аспекты. В работе рассмотрены особенности анализа электронных материалов, основные подходы, которые могут составлять решение в данной области. Автором также приведены некоторые свойства элементов Интернет-контента, исследование которых позволит принять решение о классификации их как вредоносных.

1. Введение

Сеть Интернет изначально была задумана как способ передачи данных, многие государства развивали компьютерные сети исключительно в военных целях. Однако с течением времени и развитием многих сопутствующих информационных технологий Интернет переродился в мощнейший инструмент связи с общественностью, став своего рода мостом для проведения регулярного диалога в различных форматах между множеством его пользователей. Сегодня этот уникальный социальный феномен способствует формированию принципиально новой среды общения между участниками процесса. Социальная составляющая влияния глобальной сети на различные сферы человеческой жизни настолько значительна, что наше сегодняшнее общество принято называть информационным, что говорит о высокой роли информации, знаний и информационных технологий в социуме.

Развитие виртуальных способов коммуникации, популяризация социальных сетей и иных форм общения через Интернет-сервисы фактически переводит плоскость социального взаимодействия на другой уровень. Удобство использования, высокая скорость передачи данных, доступность и разнообразие форм общения играют решающую роль в привлечении пользователя. На сегодняшний день можно констатировать, что подавляющее большинство населения так или иначе использует Интернет-сервисы для ежедневного общения и взаимодействия. Стремительное развитие и популярность подобных текст-ориентированных средств коммуникации позволяет злоумышленнику использовать уязвимость и доверчивость человека для достижения своих неправомерных целей: от навязчивого спама и распространения недостоверной информации до целена-

правленной пропаганды различных идей и мошенничества. Анонимность, характерная для сети Интернет, обостряет степень проблематики как для пользователя, так и для владельца информационной системы, через которую осуществляется коммуникация.

На сегодняшний день проблема кибербезопасности в аспекте воздействия и давления на пользователей крайне актуальна. Об этом свидетельствует множество событий и действий, направленных на ее решение. Начиная с блокирования вредоносного контента, которое уже несколько лет определено на законодательном уровне [1], государство продолжает стратегию по поддержке перспективных разработок по интеллектуальному анализу и краткосрочному выявлению материалов с элементами деструктивного информационного воздействия на личность и общество. Ведутся работы по созданию научно-технических решений, позволяющих детектировать факты использования информационных технологий для пропаганды и осуществления противоправных деяний [2]. При этом статистические исследования лишь подтверждают актуальность вопроса и свидетельствуют о том, что большинство пользователей так или иначе ощущали негативное влияние Интернет-контента. По результатам опросов [3], 73% взрослых пользователей всемирной сети были свидетелями фактов проявления агрессии в Интернете и 40% испытали ее лично на себе, несмотря на то что интернет-взаимодействие не подразумевает непосредственного контакта злоумышленника и пострадавшего пользователя (встречи участников коммуникации).

2. Анализ и управление Интернет-контентом

Данное исследование направлено на развитие функционирования социокриберфизических систем в части анализа разнородного интернет-контента: текстовых данных, изображений, видеоматериалов. Предполагается объединение воедино знаний о воздействии на состояние человека звуков определенных частот, тестовых постов, призывов с элементами агрессии и т.п. для решения единой комплексной задачи. Основным этапом решения указанного проблемного вопроса планируется внедрение технологий интеллектуального анализа с элементами динамического обучения на основе применения искусственных нейронных сетей для выявления признаков проявления агрессии, психологического давления, деструктивного воздействия на индивидуальное и групповое сознание пользователей.

Учитывая особенности процессов анализа и управления контентом в виртуальной среде, связанными с большим потоком данных, необходимостью потоковой обработки в реальном времени и сложностью формализации критериев классификации, были выделены основные подходы к выбору методов и инструментов обработки данных, анализа и принятия решений для автоматизации данных процессов:

1. Так как понятие деструктивного влияния Интернет-материалов на состояние пользователя достаточно широко и многозначно, в рамках исследования оно определяется как явления, характеризующиеся попытками манипуляции, изменения состояния человека, его мнения, настроения, мировоззрения за счет распространения в виртуальной среде специально созданных материалов с элементами агрессии, навязывания, пропаганды, смысловая составляющая которых искажена и не соответствует реальности.

2. Формализация задачи определения вредоносного Интернет-контента в рамках данного исследования сводится к изучению контента различных классов, определению признаков пространства для классификации, выбору вычислительного аппарата и разработке необходимого методического и алгоритмического обеспечения для автоматизации целевого процесса [4].

3. Для формирования признакового пространства предлагается исследовать следующие свойства анализируемого контента (в формате текстового, графического, аудио- и видеоматериалов):

- рубрикация материалов и анализ соответствия рубрики более ранним материалам автора или ресурса;
- анализ наличия ключевых слов, выделенных для искомым материалов, частоты их упоминания, соответствия рубрике;
- оценка проявления эмоциональных компонент в материале, в том числе резко негативных, резко позитивных, агрессивных, оскорбительных и т.д.;
- выявление массово порожденных текстов, сгенерированных автоматически с применением специальных алгоритмов и программных средств;
- анализ параметров медиаматериалов – звукового ряда, видеоряда, изображений – для определения наличия встроенных шумов, использования неблагоприятных для человека шумов, неприятных для глаза изображений и пр.

4. Двойственный характер подхода к классификации обусловлен двумя заключениями:

- деструктивное воздействие есть следствие одного или нескольких способов воздействия на Интернет-контент. Исследования таких воздействий позволяют сделать заключения о формализуемости идентификации подобных материалов и ее эффективности (например, исследование автоматических рубрикаторов [5], ключевых слов [6], искусственно созданных текстов [7-9], шумов в медиафайлах [10, 11] и т.д.);
- на основании того факта, что даже сам пользователь, непосредственно испытывающий неблагоприятное воздействие «на себе», не всегда готов с точностью заключить, что контент является ложным, пропагандистским, искаженно отражающим действительность, излишне эмоционально «окрашен» и вызывает неоправданные изменения собственного мнения, поступков, настроения, можно сделать вывод, что выделенный класс контента обладает слабо формализуемым набором свойств и требует применения более глубокого анализа.

Схема исследования предполагает два направления: с применением методов классического обучения на размеченном материале с выделением уникального инварианта для каждого класса, а также с применением искусственных нейронных сетей для решения слабоформализуемой задачи классификации. Предполагается, что анализ эффективности обоих подходов позволит сформировать комплексный подход для решения задачи автоматизации поиска вредоносных материалов в онлайн среде.

3. Заключение

Рассмотренная в работе проблема создания эффективных методов интеллектуального анализа и выявления материалов с элементами деструктивного информационного воздействия на личность и общество, а также использования информационных технологий для пропаганды и осуществления противоправных деяний крайне актуальна на сегодняшний день. Автором предлагается набор подходов для формирования комплексного решения, также выделяются свойства материалов для исследования. Новые методы анализа и управление виртуальным контентом позволят создать более благоприятную и безопасную среду для информатизации и коммуникации посредством сети Интернет, снизят эффективность ее использования в неправомерных целях за счет более своевременного и комплексного подхода к детектированию и тем самым защитят пользователей от деструктивного воздействия.

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научного проекта № 18-29-22104.

Список литературы

1. Федеральный закон от 27 июля 2006 г. №149-ФЗ «Об информации, информационных технологиях и о защите информации».
2. Основные направления научных исследований в области обеспечения информационной безопасности Российской Федерации (утверждены Секретарем Совета Безопасности Российской Федерации Н.П. Патрушевым 31 августа 2017 г.)
3. Duggan M. Online Harassment // Pew Research Center. Internet and Technology. <http://www.pewinternet.org/2014/10/22/online-harassment/>
4. Исхакова А.О., Исхаков А.Ю., Мещеряков Р.В. Интеллектуальный анализ электронных текстовых сообщений и интернет-контента // Сборник трудов Международной научной конференции «Актуальные проблемы прикладной математики, информатики и механики» (Воронеж, 2018). Воронеж: Научно-исследовательские публикации, 2018. С. 1453-1461.
5. Dli M., Bulygina O., Kozlov P. Development of Multimethod Approach to Rubrication of Unstructured Electronic Text Documents in Various Conditions // 2018 International Russian Automation Conference (RusAutoCon). Sochi, 2018. P. 1-5.
6. Солошенко А.Н., Орлова Ю.А., Розалиев В.Л. Автоматизированный анализ новостного потока из сети интернет с учетом семантики текстов // Известия ВолгГТУ. 2014. Т. 22, № 25 (152). С. 80-85.
7. Исхакова А.О. Модель процесса формирования инвариантов классов текстов // Доклады ТУСУРа. 2016. № 3. С. 65-69.
8. Шумская А.О. Метод определения искусственных текстов на основе расчета меры принадлежности к инвариантам // Труды СПИИРАН. 2016. № 6 (49). С. 124-140.
9. Shumskaya A.O. Using Euclidean and Mahalanobis distances while solving the problem of the text origin identification, scientific paper // Interactive systems: Problems of Human-Computer Interaction: Collection of scientific papers. 2015. P. 211-217.
10. Касперски К. Тайные рычаги подсознания. Методы психовизуальной атаки // Хакер. 2007. № 9 (105). С. 134-137.
11. Крапивенко А.В. Технологии мультимедиа и восприятие ощущений. М.: Бином. Лаборатория знаний, 2009. 271 с.