УДК 316.454: 004.622
DOI: 10.17223/1998863X/59/15

С.А. Кузнецов, А.Ю. Карпова, А.О. Савельев

АВТОМАТИЗИРОВАННОЕ ОБНАРУЖЕНИЕ ПЕРЕКРЕСТНЫХ СВЯЗЕЙ ПОЛЬЗОВАТЕЛЕЙ УЛЬТРАПРАВЫХ СООБЩЕСТВ В СОЦИАЛЬНОЙ СЕТИ

Исследование выполнено при финансовой поддержке РФФИ и ЭИСИ в рамках научного проекта № 20-011-31583.

В статье представлен анализ современных исследований и методов, направленных на автоматизированное обнаружение экстремистских сообществ, в частности – автоматизированное обнаружение перекрестных связей сообществ социальной сети. В качестве перспективного метода автоматизированного обнаружения перекрестных связей пользователей ультрарадикальных сообществ, в частности ультраправых, предлагается использование существующих парсинговых сервисов.
Ключевые слова: *радикализация, социальная сеть, ультраправое сообщество, парсинг, web mining.*

В последнее десятилетие охват социальных сетей в интернете радикальными группами расширился и предоставляет воинствующим экстремистам широкие возможности для рекрутинга новых адептов, выстраивания цепочек взаимодействий, распространения противоправного контента [1]. В частности, экстремистские организации все чаще используют технологии таргетинга – совокупность подходов, направленных на выделение целевой аудитории внутри сообщества, для привлечения новых членов [2]. Недавние исследования показывают, что инструменты вовлечения наиболее эффективны в начале экстремистской деятельности, в фазах радикализации и рекрутинга [3]. Радикальные группы используют свободную и открытую природу интернета для формирования онлайн-сообществ и распространения литературы, учебных материалов в социальных сетях, для которых, в отличие от традиционных средств массовой коммуникации, не существует эффективного инструментария регулирования содержимого. Экстремистские организации занимаются направленным таргетингом, рекрутируя новых участников на социальных сайтах, таких как Facebook, «ВКонтакте», и радикализированных веб-форумах, в том числе в рамках отдельных сообществ YouTube, 4chain, 8chan, Gab [4].

Анализ открытых источников по тематике исследования выявил устойчивый, растущий с течением времени интерес исследователей к автоматизации процесса сбора, предварительной обработки, анализа и интерпретации содержания интернет-ресурсов для изучения процесса радикализации. Следует выделить следующие проекты в рассматриваемой области:

а) общетематические исследования процесса радикализации и создания моделей радикализации [5–12];

б) узкотематические исследования различных идеологических платформ радикалов, индивидуальной и групповой динамики мобилизации на совершение насильственных инцидентов [13–20];

в) создание баз данных по инцидентам, связанным с активностью радикалов и экстремизмом PIRUS [21–23];

г) изучение инструментов анализа данных социальных сетей [24–27].

Таким образом, имеет место факт растущего с течением времени интереса исследователей к автоматизации процесса сбора, предварительной обработки, анализа и интерпретации содержания интернет-ресурсов с целью прогнозирования процесса распространения деструктивного, противоправного контента и радикализации молодежи; прогнозирования инцидентов экстремистской и террористической направленности. Противодействие распространению идеологии терроризма, насильственного экстремизма, а также выявление противоправного контента является ключевой задачей нового Комплексного плана противодействия идеологии терроризма в Российской Федерации на 2019–2023 гг. Несмотря на значительный прогресс в области контент-анализа интернет-ресурсов, отсутствуют единый комплексный подход к реализации превентивных мер по пресечению распространения деструктивного контента в социальных медиа, и инструменты, позволяющие оперативно выявлять такого рода контент. В частности, инструменты для выявления информации распространяемой ультраправыми радикальными организациями.

Методы автоматизированного обнаружения ультрарадикальных сообществ и перекрестных связей между ними будут полезны для предварительной проверки текстовых документов, уменьшая нагрузку на аналитиков в сфере информационной безопасности [28]. Перекрестные связи – пользователи, являющиеся участниками одновременно хотя бы одного радикального и студенческого сообщества. Использование автоматизированных или полуавтоматических инструментов сбора данных, основанных на анализе содержания текстовых сообщений форумов или веб-сайтов, где обмениваются крайними мнениями, потенциально позволяет оперативно выявить планирование инцидента. Такого рода инструменты полезны аналитикам правоохранительных органов при поиске «цифровых следов» радикалов.

Ультраправые, ультралевые, исламистские онлайн-сообщества представляют собой распространенные типы насильственных экстремистских групп, которые действуют с конкретной целью – воздействие на общественное мнение или разжигание политической нестабильности [29]. Радикальная группа, организующая подстрекательские, но мирные протесты, или политически мотивированный человек, участвующий в гражданском неповиновении, не считаются насильственными экстремистами. Чтобы классифицировать их как насильственных экстремистов, надо идентифицировать их намерения, такие как пропаганда, продвижение, мобилизация на совершение конкретных насильственных инцидентов. Насильственный экстремизм рассматривается как «поощрение, оправдание или поддержка совершения насильственного действия для достижения политической цели, идеологических, религиозных, социальных или экономических целей» [30]. Основная опасность онлайн-радикализации заключается в ее способности быстро «заражать» большие онлайн-сообщества деструктивным контентом.

Анализ современного состояния исследований в данной области позволяет сделать следующие выводы:

1. Большая часть исследований посвящена тому, как воинствующие экстремистские группы используют сайты социальных сетей, а также онлайн-

дискуссионные форумы для рекрутинга и мобилизации на совершение насильственных действий.

2. Компьютерный анализ социальных сетей является одной из наиболее актуальных областей исследований. Основной исследовательской задачей является выявление организационной структуры сетей.

3. Попытки профилировать отдельных пользователей с использованием методов интеллектуального анализа текста являются довольно успешными, примеры тому репозитории, созданные на базе консорциума START, университета Мэриленда: PIRUS, BAAD, IVEO, TEVUS и др. [31].

Потребность в инструментах автоматизированного обнаружения деструктивного контента и радикальных сообществ привела к росту интереса к методам анализа так называемого контента «темной паутины» (Dark Web). Под «темной путиной» понимается информация, полученная, как правило, с интернет-источников, в рамках которых взаимодействуют экстремисты [32]. Проблема в том, что «темные сети» имеют динамическую природу, и соответствующим образом усложняются задачи по сбору, предварительной обработке, хранению и анализу данных.

Задача сбора информации из социальных сетей, даже при наличии ограничений, устанавливаемых владельцами ресурсов, является к настоящему моменту частично решенной, благодаря разработкам и совершенствованию методов web mining. Нерешенными остаются задачи повышения качества формализации поисковых параметров, например словаря триггерных слов и выражений, под решаемые задачи и точности установления корреляционной зависимости между значениями различных параметров [33].

Наличие некачественных метаданных увеличивает сложности и технические проблемы для сбора данных и лингвистического анализа, увеличивает количество «ложных» срабатываний. Это связано с тем, что пользовательские данные содержат «шум» – флуктуацию значений признаков, описывающих процесс, неправильную грамматику, слова с ошибками, интернет-сленг, аббревиатуры, многоязычный текст, неформальные языковые выражения и т.п.

Применение методологии оценки сходства профилей для выявления экстремистских групп основано на линейном и обобщенном регрессионном моделировании и представляет собой набор инструментов для применения к данным, которые получены в форме наблюдений (именованных групп) по переменным (признакам и поведению групп), чтобы выявить сеть отношений между группами на основе их атрибутов и поведенческого сходства.

Был поставлен исследовательский вопрос – каким образом можно в оперативном режиме проводить мониторинг радикализации студентов, не обладая собственной инфраструктурой мониторинга социальных сетей. Изучая вопрос автоматизации обнаружения ультраправых сообществ в социальных сетях, была сформулирована гипотеза о том, что существуют автоматизированные парсинговые сервисы, позволяющие сотрудникам службы безопасности университета проводить мониторинг социальной сети для определения количества студентов из официальных студенческих сообществ, подписанных на радикальные сообщества, а также использование данных сервисов позволит корректировать программу профилактических мероприятий управления по режиму безопасности университета в рамках выполнения Комплексного плана противодействия экстремизму и терроризму на 2019–2023 гг. В данной

работе рассматривается только информация, находящаяся в свободном доступе в сети интернет. В качестве экспериментальной площадки была выбрана социальная сеть «ВКонтакте». Фокус нашего внимания направлен на ультраправые сообщества. Ультраправые (alt-right) движения, сообщества в зарубежной исследовательской практике чаще всего рассматривают как политический блок, который стремится объединить деятельность нескольких различных экстремистских движений или идеологий. Несмотря на то, что alt-right распространился в международном масштабе, его центр влияния находится в Соединенных Штатах и берет свое начало от Ричарда Спенсера, известного белого националиста, который ввел в оборот термин «alt-right». Ряд различных типов ультраправых отождествляют себя, идентифицируют или мимикрируют под белых националистов. Региональные движения могут иметь совершенно разные приоритеты и острые проблемы, но в то же время разделяют основные идеи белых националистов. Всплеск инцидентов, совершенных ультраправыми группами, наблюдается в последние несколько лет, и исследователи фиксируют экспоненциальный рост количества сообществ такого рода [34, 35].

Парсер – это сервис, который позволяет выявлять и анализировать целевую аудиторию, как правило, для интернет-коммерции. Например, можно найти промопосты, которые конкуренты используют в рекламе, посмотреть, какие из них наиболее популярны. С помощью парсера можно найти тех, кто разместил целевые аудио- и видеофайлы, или получить информацию о количестве лайков, комментариев, репостов и т.п. При этом имеются различные фильтры для сортировки сообществ или пользователей: по интересам, вероисповеданию, количеству контента.

В рамках исследования были рассмотрены 13 общедоступных парсеров для социальной сети «ВКонтакте»: «Гамаюн», CleverTarget, ТаргетоLOG, MarkFinder, BorstchРетаргет, «Лимботаргет», «Барков», PepperNinja, RetargetSexy, SegmentoTarget, TargetHunter 3.0, «ЦереброТаргет», Starcomment.

Специфика работы указанных парсеров в следующем:

– наличие возможности поиска сообществ с аналогичной целевому сообществу аудиторией;

– количество подписчиков и количество пересечений с исходным сообществом социальной сети отображается в абсолютных величинах и процентном соотношении;

– наличие возможности формирования целевой аудитории (список пользователей, состоящих в трех и более целевых сообществах);

– определение лидеров мнений в рамках целевых сообществ – пользователей, обладающих наибольших количеством подписчиков;

– наличие возможности параметризованного поиска целевой аудитории: по ключевым словам; являющихся участниками нескольких аналогичных сообществ; активных по отдельным постам и т.д.;

– привязка геолокационных данных основана на многоступенчатой оценке данных из разных источников, а не только информации из профиля пользователя;

– категория интересов в профиле пользователя формируется механизмами социальной сети на основе того, куда пользователь заходит, что кликает и т.д.

Для решения поставленной задачи были проанализированы вовлеченность сообщества в радикальные ультраправые группы, а также перекрестные связи между ними и студенческими сообществами, выбранными в качестве экспериментальных. На первом этапе работы был проведен экспертный отбор сообществ с выраженной ультраправой идеологической платформой. В выборку попало тридцать сообществ. На втором этапе с помощью парсинговых сервисов «Барков», Peper.ninja и Target Hunter выявлены пересечения между сообществами по подкатегориям:

– поиск пересечений всех ультраправых групп (выбранных для эксперимента) с отдельно взятым студенческим сообществом;

– поиск пересечения выявленных участников ультраправых сообществ со студенческими.

Обобщенная схема автоматизированного обнаружения перекрестных связей сообществ социальной сети приведена на рис. 1. По этическим соображениям радикальные и студенческие сообщества обозначены как РС и СС соответственно. На третьем этапе работы были сегментированы выявленные перекрестные связи и проведено аналогичное тестирование через другие парсеры для сравнения результатов.

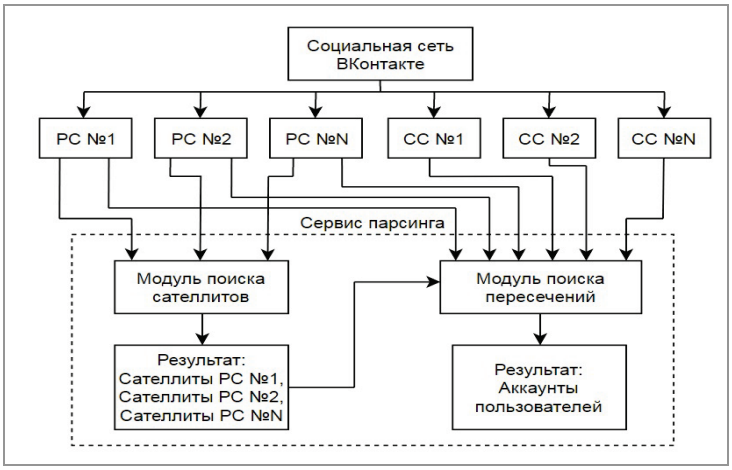

Рис. 1. Обобщенная схема автоматизированного обнаружения перекрестных связей пользователей ультраправых сообществ социальной сети

Чтобы дать некоторое представление о том, как проводилось исследование, и сравнить производительность каждого парсера, опишем этапы сегментирования и перекрестной проверки. Статистика, представленная в этом разделе, основана на агрегированных результатах.

Последовательность поиска и обнаружения перекрестных связей:

1. Через сервис «Барков» в разделе «Состоящие в нескольких группах» вводятся группа СС № 1 и все выбранные ультраправые группы. Задается значение параметра «состоящих в не менее чем N сообществах одновременно», равное 2.

2. Предыдущую операцию продолжаем для разных количеств сообществ, увеличивая каждый раз на одно, и записываем каждый результат в отдельный файл.

3. Далее через сервис TargetHunter в разделе «Инструменты – Пересечение баз», вводим Аудитория-1: СС № 1, Аудитория-2 – список ссылок пользователей из каждого файла по пересечению. Получаем результат – участники, состоящие в группе СС № 1 и в группах ультраправых с количеством пересечений не менее заданного числа. Выполняем данную операцию для каждого файла.

4. Так как сервис «Барков» позволяет искать пересечения пользователей, состоящих в количестве сообществ «не менее чем», необходимо обработать полученный результат, убрав пользователей, находящихся одновременно в разных пересечениях. Правильный результат у большего количества пересечений.

5. Предыдущие операции повторяем для сообщества СС № 2.

6. Повторяем те же операции в другом парсинговом сервисе. Сравниваем результат. По итогам тестирования – результат сравнения обнаруживает набор со значительной точностью.

В качестве эксперимента были выбраны два крупных студенческих сообщества СС № 1 и СС № 2 с количеством подписчиков 22 000 и 1 600 соответственно, наиболее полно представляющих студенческую аудиторию университета. Экспертом были выбраны наиболее активные радикальные ультраправые сообщества, их количество составило 30. Фильтрация от ботов предусматривается средствами парсингового сервиса.

Далее по предложенному алгоритму были найдены пользователи, состоящие в радикальных ультраправых сообществах. Также были найдены пользователи, состоящие одновременно в нескольких радикальных ультраправых сообществах. Визуализация полученных результатов приведена на рис. 2.

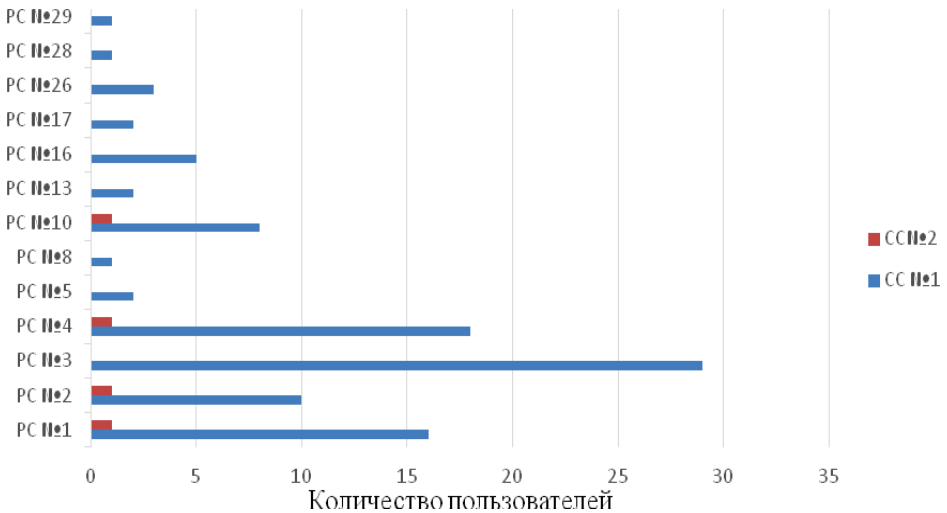

Рис. 2. Гистограмма пересечения сообществ СС и РС

Для радикальных ультраправых сообществ, которые не указаны на рис. 2, пересечения не обнаружены.

Также были найдены пользователи, состоящее одновременно в нескольких радикальных ультраправых сообществах. Результаты показаны на рис. 3.

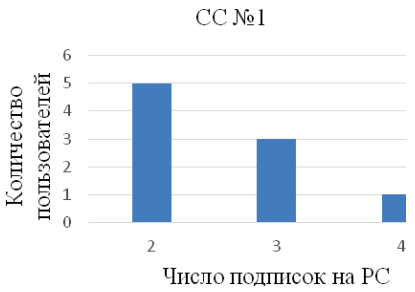

Рис. 3. Гистограмма одновременного нахождения пользователей в разных РС

На рис. 3 отображена информация для пользователей, находящихся одновременно в двух и более радикальных ультраправых сообществах. В сообществе СС № 2 пользователей, состоящих в более чем в 1 РС, на момент проведения исследования не выявлено.

Заключение

Результаты, представленные в этой статье, подтверждают вывод о том, что автоматизированное обнаружение перекрестных связей пользователей радикальных ультраправых сообществ в социальной сети является достижимой целью.

Научная новизна данного исследования заключается в том, что данный подход расширяет возможности исследователей, не имеющих собственной инфраструктуры мониторинга и анализа социальных сетей с целью изучения процесса радикализации.

Проверив эффективность нашей гипотезы через парсинговые сервисы для реальных условий, мы продемонстрировали, что такие инструменты автоматизированого обнаружения будут четко вписываться в рабочий процесс групп по анализу информации.

Однако исследование показало, что различные парсинговые сервисы на аналогичные запросы предоставляют отличные друг от друга результаты. Поскольку внутренняя логика работы этих сервисов является скрытой, однозначно указать причину несовпадения результатов не представляется возможным. Также следует отметить, что, хотя существующие решения могут быть использованы в качестве инструментов выявления целевых сообществ в социальных сетях, предоставляемые ими функции не являются достаточными при решении задач выявления деструктивных идеологических платформ и могут применяться лишь на начальных этапах соответствующих социологических исследований.

Таким образом, актуальной и все еще нерешенной является задача разработки программного инструментария автоматизации социологических исследований на основе данных, извлекаемых из социальных сетей. В настоящее время научная группа ТПУ проводит тестирование разработанного специализированного парсера под вышеобозначенные задачи.

Литература

1. *Overbey L.A, McKoy G., Gordon J., McKitrick S.* Automated sensing and social network analysis in virtual worlds // Intelligence and Security Informatics (ISI). IEEE, Vancouver, BC, Canada, 2010. P. 179–184.

2. *Lakomy M.* Let's Play a Video Game: Jihadi Propaganda in the World of Electronic Entertainment, Studies in Conflict & Terrorism. URL: https://www.tandfonline.com /doi/full/10.1080/1057610X.2017.1385903?scroll=top&needAccess=true#aHR0cHM6Ly93d3cudGFuZGZvbmxpbmUuY29tL2RvaS9wZGYvMTAuMTA4MC8xMDU3NjEwWC4yMDE3LjEzODU5MDM/bmVlZEFjY2Vzcz10cnVlQEBAMA== (accessed: 30.10.2019).

3. *Torok R.* Developing an explanatory model for the process of online radicalization and terrorism // Security Informatics. 2013. Vol. 2, № 1. P. 1–10. DOI: 10.1186/2190-8532-2-1

4. *The* New York Times. Zeynep Tufekci. How Everyday Social Media Users Become Real-World Extremists. March 10, 2018. URL: https://www.nytimes.com/2018/03/10/ opinion/sunday/youtube-politicsradical.html?action=click&module=RelatedLinks&pgtype= Article (accessed: 28.07.2019).

5. *Borum R.* Radicalization into Violent Extremism I: A Review of Social Science Theories // Journal of Strategic Security. 2011. Vol. 4, № 4. P. 7–36.

6. *McCauley C., & Moskalenko S.* Understanding political radicalization: The two-pyramids model // American Psychologist. 2017. Vol. 72, № 3. P. 205–216.

7. *Neumann P.R.* Old and New Terrorism. Cambridge: Polity Press, 2009.

8. *Braniff W.* Recasting and Repositioning CVE as a Grand Strategic Response to Terrorism. START (November). 2017. URL: https://www.start.umd.edu/news/recasting-and-repositioning-cve-grand-strategic-response-terrorism (accessed: 18.12.2018).

9. *Hamm M., Spaaij R.* Lone wolf terrorism in America: Using knowledge of radicalization pathways to forge prevention strategies. Final grant report to NIJ. 2015. URL: www.ncjrs.gov/pdffiles1/nij/grants/248691.pdf (accessed: 11.05.2018).

10. *Raab J., Milward H.B.* Dark Networks as Problem // Journal of Public Administration research and Theory. 2003. Vol. 13, № 4. P. 413–439.

11. *Gerdes L.M.* Illuminating Dark Networks: The Study of Clandestine Groups and Organizations (New York: Cambridge University Press, 2015);

12. *Krebs V.* Mapping Networks of Terrorist Cells // Connections. 2002. Vol. 24. P. 43–52.

13. *Fact* Sheet: Far-Right Fatal Ideological Violence against Religious Institutions and Individuals in the United States: 1990-2018. URL: https://www.start.umd.edu/pubs/START_ECDB_FarRightFatalIdeologicalViolenceAgainstReligiousTargets1990-2018_Oct2018.pdf (accessed: 05.03.2019).

14. *Application* of a Profile Similarity Methodology for Identifying Terrorist Groups that Use or Pursue CBRN Weapons. Social Computing, Behavioral-Cultural Modeling and Prediction. URL: https://link.springer.com/chapter/10.1007/978-3-642-19656-0_5 (accessed: 04.02.2019).

15. *Kruglanski A.W., Fernandez J.R., Factor A.R., Szumowska E.* Cognitive mechanisms in violent extremism. 2018. Elsevier. URL: https://doi.org/10.1016/j.cognition.2018.11.008 (accessed: 11.02.2019).

16. *Kruglanski A.W.* Violent radicalism and the psychology of prepossession. Social Psychological Bulletin, 2018. 13(4), Article e27449. URL: https://doi.org/10.32872/spb.v13i4.27449 (accessed: 25.01.2019).

17. *Langman P.* Different Types of Role Model Influence and Fame Seeking Among Mass Killers and Copycat Offenders // American Behavioral Scientist. 2018. Vol. 62, № 2. P. 210–228.

18. *Towers S., Gomez-Lievano A., Khan M., Mubayi A., Castillo-Chavez C.* Contagion in Mass Killings and School Shootings // PLoS ONE. 2015. Vol. 10, № 7. P. e0117259. URL: https://doi.org/10.1371/journal.pone.0117259 (accessed: 25.01.2019).

19. *Caiani M., Wagemann C.* Online networks of the italian and german extreme right // Information, Communication & Society. 2009. Vol. 12:1. P. 66–109. DOI: 10.1080/13691180802158482

20. *Southern* Poverty Law Center. Hate& Extremism. URL: https://www.splcenter.org/ (accessed: 08.03.2019).

21. *START* 2017. Overview: Profiles of Individual Radicalization in the United States-Foreign Fighters. URL: https://www.start.umd.edu/ (accessed: 15.02.2019).

22. *Project* 2018–2020 «A Multi-Level Approach to the Study of Violent Extremism», Investigators: Gary LaFree, Michael Jensen, STAR (accessed: 02.03.2019).

23. *Project* 2018–2019 «Social Media Influencers: Re-domaining Fashion Industry Forecasting to Anticipate Online Extremist Radicalization», Investigators: Barnett S. Koven, Ramon F. Brena, START (accessed: 02.03.2019).

24. *Cohen K., Johansson F., Kaati L., Mork J.C.* Detecting Linguistic Markers for Radical Violence in Social Media // Terrorism and Political Violence. 2014. Vol. 26. P. 246–256.

25. *Xie D., Xu J., Lu T.-C.* Automated Classification of Extremist Twitter Accounts Using Content-Based and Network-Based Features, 2016. IEEE International Conference on Big Data. P. 2545–2549.

26. *Gilani Z., Kochmar E., Crowcroft J.* Classification of Twitter Accounts into Automated Agents and Human Users // Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017. P. 489–496.

27. *Pennebaker J.W., Booth R.J., Boyd R.L., Francis M.E.* Linguistic Inquiry and Word Count: LIWC2015. 2015. Austin, TX: Pennebaker Conglomerates (www.LIWC.net) (accessed: 23.01.2019).

28. *Yang M., Kiang M., Chen H., Li Y.* Artificial immune system for illicit content identification in social media // J. Am. Soc. Inf. Sci. Technol. 2012. Vol. 63, № 2. P. 256–269. DOI: 10.1002/asi.21673

29. *START* Research. Where the Extremes May Touch: The Potential for Collaboration Between Islamist, Right- and Left-Wing Extremists. URL: https://www.start.umd.edu/research-projects/where-extremes-may-touch-potential-collaboration-between-islamist-right-and-left (accessed: 28.07.2019).

30. *Borum R.* Radicalization into Violent Extremism I: A Review of Social Science Theories // Journal of Strategic Security. 2011. Vol. 4, № 4. P. 7–36.

31. *Карпова А.Ю.* Механизмы индивидуальной радикализации в процессе самоорганизации молодежи // Молодежь и молодежная политика: новые смыслы и практики: сборник / под ред. С.В. Рязанцева, Т.К. Ростовской. Сер.: Демография. Социология. Экономика. М., 2019. С. 69–81.

32. *Fu T., Abbasi A., Chen H.* A focused crawler for dark web forums // J. Am. Soc. Inf. Sci. Technol. 2010. Vol. 61, № 6. P. 1213–1231.

33. *Карпова А.Ю. Савельев А.О., Вильнин А.Д., Чайковский Д.В.* Изучение процесса онлайн-радикализации молодежи в социальных медиа (междисциплинарный подход) // Мониторинг общественного мнения: экономические и социальные перемены. 2020. № 3 (157). С. 159–181.

34. *START* Fact Sheet. October 2018. Far-Right Fatal Ideological Violence against Religious Institutions and Individuals in the United States: 1990–2018. URL: https://www.start.umd.edu/pubs/START_ECDB_ FarRightFatalIdeologicalViolenceAgainstReligiousTargets1990-2018_Oct2018.pdf

35. *Smith A.G.* How Radicalization to Terrorism Occurs in the United States: What Research Sponsored by the National Institute of Justice Tells Us. National Institute of Justice / NIJ.ojp.gov (accessed: 28.07.2019).

Sergey A. Kuznetsov, Tomsk Polytechnic University (Tomsk, Russian Federation).
E-mail: ksa11@tpu.ru
Anna Yu. Karpova, Tomsk Polytechnic University (Tomsk, Russian Federation).
E-mail: belts@tpu.ru
Aleksey O. Savelev, Tomsk Polytechnic University (Tomsk, Russian Federation).
E-mail: sava@tpu.ru

Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya – Tomsk State University Journal of Philosophy, Sociology and Political Science. 2021. 59. pp. 156–166.
DOI: 10.17223/1998863X/59/15

AUTOMATED DETECTION OF ULTRA-RIGHT COMMUNITIES' CROSS-LINKS IN A SOCIAL NETWORK

Keywords: radicalization; social network; ultra-right community; parsing; web mining.

The reported study was funded by RFBR and EISR, project number 20-011-31583.

In the last decade, the coverage of social networks on the Internet by radical groups has expanded and provided militant extremists with many opportunities to recruit new adherents, build chains of interactions, and distribute illegal content. Extremist organizations engage in targeting, recruiting new members on social sites such as Facebook, VKontakte, and on radicalized web forums, including within individual communities. The main danger of online radicalization lies in its ability to quickly "infect" large online communities with destructive content. On the other hand, the Internet facilitates the study of extremist views. The use of automated or semi-automatic data collection tools based on the analysis of the website text content where extreme opinions are potentially distributed allows identifying incident planning quickly. As part of the study, we formulate an assumption on the effectiveness of using existing parsing services to automate the detection of cross-links between ultra-right community users. In this article, we consider only information that is freely available on the

Internet. As an experimental site, we chose the social network VKontakte. To solve the problem, we analyzed the involvement of the community in radical far-right groups, as well as cross-links between them and student communities selected as experimental. Two student communities SC 1 and SC 2 were selected with the number of subscribers 22,000 and 1,600, respectively. The expert selected 30 radical ultra-right communities. According to the proposed algorithm, we found users who are members of a radical far-right community, as well as users who are simultaneously members of several radical far-right communities. The study showed that automated detection of cross-links between users of radical far-right communities in a social network is an achievable goal. However, various parsing services for similar requests provide results different from each other. Thus, the challenge remains to develop software tools for automating sociological research based on data retrieved from social networks.

References

1. Overbey, L.A, McKoy, G., Gordon, J. & McKitrick, S. (2010) Automated sensing and social network analysis in virtual worlds. *Intelligence and Security Informatics (ISI)*. IEEE, Vancouver, BC, Canada, 2010. pp. 179–184.

2. Lakomy, M. (n.d.) *Let's Play a Video Game: Jihadi Propaganda in the World of Electronic Entertainment, Studies in Conflict & Terrorism*. [Online] Available from: https://www.tandfonline.com /doi/full/10.1080/ 1057610X.2017.1385903?scroll=top&needAccess=true#aHR0cHM6Ly93d3cudGFuZGZvbmxpbmUuY29tL2RvaS9wZGYvMTAuMTA4MC8xMDU3NjEwWC4yMDE3LjEzODU5MDM/bmVlZEFjY2Vzcz10cnVlQEBAMA== (Accessed: 30th October 2019).

3. Torok, R. (2013) Developing an explanatory model for the process of online radicalization and terrorism. *Security Informatics*. 2(1). pp. 1–10. DOI: 10.1186/2190-8532-2-1

4. Tufekci, Z. (2018) How Everyday Social Media Users Become Real-World Extremists. *The New York Times*. 10th March. [Online] Available from: https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politicsradical.html?action=click&module=RelatedLinks&pgtype= Article (Accessed: 28th July 2019).

5. Borum, R. (2011) Radicalization into Violent Extremism I: A Review of Social Science Theories. *Journal of Strategic Security*. 4(4). pp. 7–36. DOI: 10.5038/1944-0472.4.4.1

6. McCauley, C. & Moskalenko, S. (2017) Understanding political radicalization: The two-pyramids model. *American Psychologist*. 72(3). pp. 205–216. DOI: 10.1037/amp0000062

7. Neumann, P.R. (2009) *Old and New Terrorism*. Cambridge: Polity Press.

8. Braniff, W. (2017) *Recasting and Repositioning CVE as a Grand Strategic Response to Terrorism*. [Online] Available from: https://www.start.umd.edu/news/recasting-and-repositioning-cve-grand-strategic-response-terrorism (Accessed: 18th December 2018).

9. Hamm, M. & Spaaij, R. (2015) *Lone wolf terrorism in America: Using knowledge of radicalization pathways to forge prevention strategies*. Final grant report to NIJ. [Online] Available from: www.ncjrs.gov/pdffiles1/nij/grants/248691.pdf (Accessed: 11th May 2018).

10. Raab, J. & Milward, H.B. (2003) Dark Networks as Problem. *Journal of Public Administration research and Theory*. 13(4). pp. 413–439.

11. Gerdes, L.M. (2015) *Illuminating Dark Networks: The Study of Clandestine Groups and Organizations*. New York: Cambridge University Press.

12. Krebs, V. (2002) Mapping Networks of Terrorist Cells. *Connections*. 24. pp. 43–52.

13. Start.umd.edu. (n.d.) *Fact Sheet: Far-Right Fatal Ideological Violence against Religious Institutions and Individuals in the United States: 1990–2018*. [Online] Available from: https://www.start.umd.edu/ pubs / START_ECDB_FarRightFatalIdeologicalViolenceAgainstReligious Targets1990-2018_Oct2018.pdf (Accessed: 5th March 2019).

14. Breiger, R.L. et al. (2011) Application of a Profile Similarity Methodology for Identifying Terrorist Groups That Use or Pursue CBRN Weapons. In: Salerno, J., Yang, S.J., Nau, D. & Chai, SK. (eds) *Social Computing, Behavioral-Cultural Modeling and Prediction*. SBP 2011. Lecture Notes in Computer Science. Vol 6589. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-642-19656-0_5

15. Kruglanski, A.W., Fernandez, J.R., Factor, A.R. & Szumowska, E. (2018) Cognitive mechanisms in violent extremism. *Cognition*. 188. pp. 116–123. DOI: 10.1016/j.cognition.2018.11.008

16. Kruglanski, A.W. (2018) Violent radicalism and the psychology of prepossession. *Social Psychological Bulletin*. 13(4). Article e27449. DOI: 10.32872/spb.v13i4.27449

17. Langman, P. (2018) Different Types of Role Model Influence and Fame Seeking Among Mass Killers and Copycat Offenders. *American Behavioral Scientist*. 62(2). pp. 210–228. DOI: 10.1177/0002764217739663

18. Towers, S., Gomez-Lievano, A., Khan, M., Mubayi, A. & Castillo-Chavez, C. (2015) Contagion in Mass Killings and School Shootings. *PLoS ONE*. 10(7). Article e0117259. DOI: 10.1371/journal.pone.0117259 (Accessed: 25th January 2019).

19. Caiani, M. & Wagemann, C. (2009) Online networks of the italian and german extreme right. *Information, Communication & Society*. 12(1). pp. 66–109. DOI: 10.1080/13691180802158482

20. Southern Poverty Law Center. *Hate& Extremism*. [Online] Available from: https://www.splcenter.org/ (Accessed: 8th March 2019).

21. Start.umd.edu. (2017) *Overview: Profiles of Individual Radicalization in the United States-Foreign Fighters*. [Online] Available from: https://www.start.umd.edu/ (Accessed: 15th February 2019).

22. Start.umd.edu. (n.d.) *Project 2018–2020 "A Multi-Level Approach to the Study of Violent Extremism"*. Investigators: Gary LaFree, Michael Jensen. (Accessed: 2nd March 2019).

23. Start.umd.edu. (n.d.) *Project 2018–2019 "Social Media Influencers: Re-domaining Fashion Industry Forecasting to Anticipate Online Extremist Radicalization"*. Investigators: Barnett S. Koven, Ramon F. Brena. (Accessed: 2nd March 2019).

24. Cohen, K., Johansson, F., Kaati, L. & Mork, J.C. (2014) Detecting Linguistic Markers for Radical Violence in Social Media. *Terrorism and Political Violence*. 26. pp. 246–256. DOI: 10.1080/09546553.2014.849948

25. Xie, D., Xu, J. & Lu, T.-C. (2016) Automated Classification of Extremist Twitter Accounts Using Content-Based and Network-Based Features. *IEEE International Conference on Big Data*. pp. 2545–2549.

26. Gilani, Z., Kochmar, E. & Crowcroft, J. (2017) Classification of Twitter Accounts into Automated Agents and Human Users. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. pp. 489–496.

27. Pennebaker, J.W., Booth, R.J., Boyd, R.L. & Francis, M.E. (2015) *Linguistic Inquiry and Word Count: LIWC2015. Austin, TX: Pennebaker Conglomerates (www.LIWC.net)*. (Accessed: 23rd January 2019).

28. Yang, M., Kiang, M., Chen, H. & Li, Y. (2012) Artificial immune system for illicit content identification in social media. *Journal of American Society for Information Science and Technology*. 63(2). pp. 256–269. DOI: 10.1002/asi.21673

29. Start.umd.edu. (n.d.) *START Research. Where the Extremes May Touch: The Potential for Collaboration Between Islamist, Right- and Left-Wing Extremists*. [Online] Available from: https://www.start.umd.edu/research-projects/where-extremes-may-touch-potential-collaboration-between-islamist-right-and-left (Accessed: 28th July 2019).

30. Borum, R. (2011) Radicalization into Violent Extremism I: A Review of Social Science Theories. *Journal of Strategic Security*. 4(4). pp. 7–36. DOI: 10.5038/1944-0472.4.4.1

31. Karpova, A.Yu. (2019) Mekhanizmy individual'noy radikalizatsii v protsesse samoorganizatsii molodezhi [Mechanisms of individual radicalization in the process of self-organization of youth]. In: Ryazantsev, S.V. & Rostovskaya, T.K. (eds) *Molodezh' i molodezhnaya politika: novye smysly i praktiki* [Youth and youth policy: new meanings and practices]. Moscow: Ekon-infom. pp. 69–81.

32. Fu, T., Abbasi, A. & Chen, H. (2010) A focused crawler for dark web forums. *Journal of American Society for Information Science and Technology*. 61(6). pp. 1213–1231. DOI: 10.1002/asi.21323

33. Karpova, A.Yu. Saveliev, A.O., Vilnin, A.D. & Chaykovsky, D.V. (2020) Studying online radicalization of youth through social media (interdisciplinary approach). *Monitoring obshchestvennogo mneniya: ekonomicheskie i sotsial'nye peremeny – Monitoring of Public Opinion: Economic and Social Changes*. 3(157). pp. 159–181. (In Russian). DOI: 10.14515/monitoring.2020.3.1585

34. Start.umd.edu. (2018) Far-Right Fatal Ideological Violence against Religious Institutions and Individuals in the United States: 1990–2018. *START Fact Sheet*. October 2018. [Online] Available from: https://www.start.umd.edu/pubs/ START_ECDB_ FarRightFatalIdeologicalViolenceAgainstReligiousTargets1990-2018_Oct2018.pdf

35. Smith, A.G. (n.d.) *How Radicalization to Terrorism Occurs in the United States: What Research Sponsored by the National Institute of Justice Tells Us*. National Institute of Justice / NIJ.ojp.gov (Accessed: 28th July 2019).