ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН

ҚазҰТЗУ ХАБАРШЫСЫ

ВЕСТНИК КазНИТУ

VESTNIK KazNRTU

№ 5 (141)

АЛМАТЫ 2020 ОКТЯБРЬ

ISSN 1680-9211

Учредитель:

Казахский национальный исследовательский технический университет
имени К.И. Сатпаева

Регистрация:

Министерство культуры, информации и общественного согласия
Республики Казахстан № 951 – Ж "25" 11. 1999 г.

Основан в августе 1994 г. Выходит 6 раз в год

Адрес редакции:

г. Алматы, ул. Сатпаева, 22,
каб. 609, тел. 292-63-46
Nina. Fedorovna. 52 @ mail.ru

ӘДЕБИЕТТЕР

[1] Statistics. Most popular social networks worldwide. The Statistics Portal. 2020. Available online // https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ (accessed on 2 May 2020).

[2] Azizan, S.A.; Aziz, I.A. Terrorism Detection Based on Sentiment Analysis Using Machine Learning. J. Eng. Appl. Sci. 2017, 12, 691–698.

[3] Yadron, D. Twitter Deletes 125,000 ISIS Accounts and expands anti-Terror Teams. Available online // https://www.theguardian.com/technology/2016/feb/05/ twitter-deletes-isis-accounts-terrorism-online.

[4] M. Viktor, K. Cukier. Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, 2013.

[5] Agarwal, S., Sureka, A.: A focused crawler for mining hate and extremism promoting videos on youtube. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media, pp. 294–296 (2014). ACM.

[6] Sureka, A., Agarwal, S.: Learning to classify hate and extremism promoting tweets. In: Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint, pp. 320–320 (2014). IEEE

[7] Ahmad, S., Asghar, M.Z., Alotaibi, F.M. et al. Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. Hum. Cent. Comput. Inf. Sci. 9, 24 (2019) // https://doi.org/10.1186/s13673-019-0185-6.

[8] Көленкелі Веб және геосаяси веб-зерттеулер // https://ai.arizona.edu/research/dark-web-geo-web. Қаралған күні: 05.11.2017.

[9] М.А. Болатбек, Ш.Ж. Мусиралиева ЭКСТРЕМИСТІК МӘТІНДЕРДІ МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ АНЫҚТАУ // ВЕСТНИК КазНИТУ. - 2018г.. - №№6 (130). - С. 300-304.

[10] Kaggle URL: https://www.agkgle.com/fifthtribe/isis-religious-texts (дата обращения: 15.02.2020).

Мусиралиева Ш.Ж., Омаров Б.С., Медетбек Ж.Б., Караман Ғ.Р., Бекетова А.К.

Идентификация сообщений, содержащих элементы экстремизма в социальных системах с использованием методов машинного обучения

Резюме: Вместе с ростом социальных сетей растет количество религиозной ненависти и расизма в сети. Кроме того, активность радикальных групп в Интернете, призывающих к насилию и экстремизму, является одним из важнейших вопросов общественной безопасности. Потому что для таких структур основным инструментом для обмена информацией, найма и продвижения является Интернет, в частности, является веб-ресурсы, социальные сети, социальные мессенджеры и т. д. В связи с этим необходимо выявить отдельных пользователей, группы и интернет-сообщества, которые создают и распространяют террористическую и экстремистскую информацию в Интернете, а также предотвращать распространение экстремистских материалов.

Данная работа посвящена изучению и разработке методов машинного обучения, направленных на решение проблемы выявления экстремистского текста в социальных системах. Кроме того, представлены модели и методы выявления экстремистского текста, которые используются для углубленного лингвистического анализа и статистической обработки текстов. Чтобы классифицировать текст как экстремистский или неэкстремистский на основе сообщений в социальных сетях, оставленных пользователем в Интернете, мы создаем систему классификации текста с использованием методов анализа настроений на основе машинного обучения.

Ключевые слова: социальная сеть, онлайн-экстремизм, текст с радикальным содержанием, лингвистический анализ, машинное обучение, большие данные, векторная модель, логистическая регрессия, наивный Байес, глубокое обучение.

Sh.Zh. Mussiraliyeva, M.Y. Aidyn, R.K. Ospanov
(Al-Farabi Kazakh National University, Almaty, Kazakhstan
Email: {mussiraliyevash, aidynme, ospanov.ruslan.k} @gmail.com)

USER IDENTIFICATION METHOD BASED ON FRIENDSHIP AND DEMOGRAPHIC ATTRIBUTES IN SOCIAL NETWORKS

Abstract. Nowadays, social networks are a platform for a lot of information. With the exception of useful information, social media has become a convenient platform for illegal activities. Suspicious activity is often overshadowed by the lack of threat detection and analysis systems on social media. This article provides a brief overview of approaches for analyzing information from a user profile. Methods of programming interface of social network for intelligence based on open data are considered. The method of identifying the user of the profile based on the analysis of data of friendships and attributes has been tested.

Key words: social network, Vkontakte, data analysis, OSINT, virtual connections.

Ш.Ж. Мусиралиева, М.Е. Айдын, Р.К. Оспанов
(Al-Farabi Kazakh National University, Almaty, Kazakhstan
Email: {mussiraliyevash, aidynme, ospanov.ruslan.k} @gmail.com)

МЕТОД ИДЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЯ НА ОСНОВЕ ДРУЖЕСКИХ СВЯЗЕЙ И ДЕМОГРАФИЧЕСКИХ АТРИБУТОВ В СОЦИАЛЬНЫХ СЕТЯХ

Аннотация. В настоящее время социальные сети являются площадкой большого количества информации. За исключением полезной информации социальные сети стали удобной платформой для противозаконных действий. Зачастую подозрительная активность остается в тени ввиду отсутствия систем обнаружения и анализа угроз в социальных сетях. В статье представлен краткий обзор подходов для анализа информации из профиля пользователя. Рассмотрены методы интерфейса программирования социальной сети для разведки на основе открытых данных. Опробован метод идентификации пользователя профиля, основанный на анализе данных дружеских связей и атрибутов.

Ключевые слова: социальная сеть, в контакте, анализ данных, OSINT, виртуальные связи.

Введение

В последние годы социальные сети, такие как Facebook, Twitter и Google+, Вконтакте привлекли миллионы пользователей. Наиболее широко используемой социальной сетью на территории СНГ является социальная сеть "ВКонтакте" которая насчитывает 100 миллионов активных пользователей [1]. Среди жителей Казахстана на долю данной социальной сети приходится 7,2 миллиона пользователей [2]. Социальные сети привлекли внимание всего мира из-за их возможности обращаться к миллионам пользователей и строить коммуникативные связи. Потенциальные возможности социальных сетей часто используются злоумышленниками, которые извлекают конфиденциальную информацию неосведомленных пользователей или используют социальную сеть в качестве платформы для противоправных действий. Одним из наиболее распространенных способов создания анонимности для противозаконных действий является использование поддельных профилей, когда злонамеренные пользователи представляют себя в профилях, выдавая себя за фиктивных или реальных людей. Основная цель этого исследования провести анализ профилей пользователей в социальной сети на основе заполненных данных и дружеских связей пользователя. Для этого мы воспользовались инструментами для разведки на основе открытых источников и произвели сбор данных, было проанализировано взаимодействие между профилями. В результате нашей работы был рассмотрен метод идентификации профиля пользователя в социальной сети на основе демографических атрибутов дружественных связей пользователя.

Обзор литературы

В последние годы в социальных сетях наблюдается экспоненциальный рост в взаимодействиях между пользователями. Быстрый рост интересов параллельно вызвал резкий рост противоправных действий. В социальных сетях достаточно много поддельных профилей которые превратились в удобный инструмент для злоумышленников. Исследования сообщают о подозрительной активности поддельных аккаунтов на страницах политиков, знаменитостей и медийных личностей [3]. По результатам исследования поддельные профили могут быть использованы социальными ботами или злоумышленниками для раскрутки ложных новостей, разжигания межнациональной розни или распространения иной информации [4]. Информационная безопасность и конфиденциальность являются одними из основных требований пользователей социальных сетей, поддержание и обеспечение этих требований повышает надежность сети и, следовательно, уровень доверенности к социальной сети.

В 2012 году Facebook заметил злоупотребление на их платформе, включая публикацию ложных новостей, разжигания ненависти, сенсаций и поляризации, и других [5]. Это явление подняло вопрос о необходимости новых методов обнаружения таких действий и их предотвращения.

В 2015 году Facebook подсчитал, что почти 14 миллионов его активных пользователей ежемесячно являются нежелательными, представляя вредоносные поддельные учетные записи, которые были созданы в нарушение условий обслуживания веб-сайтов [6]. Facebook впервые опубликовал отчет за первый квартал 2018 года, в котором показаны их внутренние руководящие принципы, используемые для обеспечения соблюдения стандартов сообщества применительно к их усилиям в период с октября 2017 года по март 2018 года. Этот отчет иллюстрирует количество

нежелательного контента, который был удален Facebook и охватывает шесть категорий: сцены насилия, нагота взрослых, сексуальная активность, террористическая пропаганда, ненавистнические высказывания, спам [7].

В нынешних условиях можно отметить две задачи, решение которых можно найти с помощью анализа определенных параметров в социальной сети. Задачи выявления лидеров в определенной группе пользователей и идентификация профилей.

Разработка системы определения связей и верификации профилей является одной из сложных задач на сегодняшний день.

Зачастую социальные сети используют поведенческий анализ, благодаря которому удается обнаружить пользователей, нарушающих тем или иным образом пользовательское соглашение [8]. Подозрительная активность ведет к появлению механизма верификации Captcha.

Следующий подход представляет собой методы статического или семантического анализа текстов. В случае если определенное сообщество пользователей ведет обсуждение определенной темы или профили пользователей управляются одним лицом. При использовании семантического анализа используется корпус текстов для присвоения определенного идентификатора.

Третий подход заключается в проведении анализа связей пользователя. Данный метод основан на анализе количества тех или иных атрибутов связей пользователя.

В качестве комбинированного метода можно использовать несколько методов одновременно, что позволяет на основе многих факторов получать результаты с высокой достоверностью. Данный метод строится на основе машинного обучения, где ключевыми задачами являются подбор классификатора и обучение модели.

Методы и инструменты

Зачастую при заполнении профиля в социальной сети пользователи преднамеренно либо по ошибке указывают ложную информацию о фактах биографии. Некорректно указанные данные затрудняют решение задачи идентификации пользователя. Однако возможность анализа демографических атрибутов профиля пользователя и профилей друзей пользователя предоставляет возможность частичной идентификации пользователя. В работе мы рассматриваем метод предназначенный для разведки на основе открытых источников, а именно использование инструмента верификации профиля пользователя в социальной сети ВКонтакте для последующего анализа. Для получения предполагаемых данных был применен алгоритм вычисления среднего значения атрибутов. Для анализа были выбраны следующие атрибуты:

1. Часто встречающийся пол друзей пользователя.
2. Возраст друзей.
3. Города друзей.
4. Увлечения пользователя на основе групп пользователя.

Для получения необходимых данных из страницы пользователя были использованы следующие инструменты:

1. Интерфейс программирования API Вконтакте по адресу: https://vk.com/dev
2. Язык программирования Python 3.6
3. Библиотека requests для HTTP запросов в Python
4. Встроенный модуль Collections для языка программирования Python

API Вконтакте - это интерфейс, который позволяет получать информацию из базы данных vk.com с использованием http-запросов к серверу социальной сети. Синтаксис запросов и тип возвращаемых ими данных строго определены на стороне сервера.

Для получения необходимой информации из профиля пользователя использовались методы (API) Вконтакте указанные в таблице - 1.

Таблица - 1. Методы API Вктонтакте

Метод	Ключевой параметр	Результат
groups.get - возвращает список сообществ указанного пользователя.	user_id - идентификатор пользователя, информацию о сообществах которого требуется получить.	Метод возвращает объект, содержащий число результатов в поле count и массив идентификаторов сообщества в поле items.
groups.getByid - возвращает информацию о заданном сообществе или о нескольких сообществах.	group_id - идентификатор или короткое имя сообщества.	Возвращает массив объектов, описывающих сообщества.
utils.resolveScreenName - определяет тип объекта (пользователь, сообщество, приложение) и его идентификатор по короткому имени screen_name.	screen_name - короткое имя пользователя, группы или приложения.	После успешного выполнения возвращает объект, который содержит следующие поля: 1. Type - тип объекта. Возможные значения (user, group, application). 2. object_id - идентификатор объекта.
friends.get - возвращает список идентификаторов друзей пользователя или расширенную информацию о друзьях пользователя (при использовании параметра fields)	user_id - идентификатор пользователя, для которого необходимо получить список друзей. Если параметр не задан, то считается, что он равен идентификатору текущего пользователя (справедливо для вызова с передачей access_token).	После успешного выполнения возвращает список идентификаторов (id) друзей пользователя. При использовании параметра fields возвращает список объектов пользователей, но не более 5000.

Для анализа и сортировки полученных значений были использован простой алгоритм для определения медианы. Центральную тенденцию данных можно рассматривать не только, как значение с нулевым суммарным отклонением или максимальную частоту, но и как некоторую отметку, делящую ранжированные данные (отсортированные по возрастанию или убыванию) на две равные части, в таком случае первая половина исходных данных меньше данной отметки, вторая половина больше по значению, данное значение является медианой.

Медиана, как и среднее значение, необходима для определения типичного значения в наборе данных. В общем случае для нахождения медианы, данные нужно расположить в порядке возрастания или убывания. В случае если выявлено чётное число элементов, медиана может быть не определена однозначно: для числовых чаще используют полсуммы двух соседних значений.

Для практической реализации вычисления медианного значения был использован встроенный модуль collections для Python. Модуль collections предоставляет специализированные типы данных, на основе словарей, кортежей, множеств и списков. Из типов данных модуля collections был использован тип данных Counter. Вид словаря collections.Counter позволяет считать количество неизменяемых объектов. Пример использования модуля указан на рисунке - 1.

```
>>> import collections
>>> c = collections.Counter()
>>> for heading in ['spam', 'music', 'programming', 'Books', 'Books', 'Books']:
        c[heading] += 1


>>> print(c)
Counter({'Books': 3, 'spam': 1, 'music': 1, 'programming': 1})
>>> print(c ['Books'])
3
>>>
```

Рис. 1. Пример использования словаря Counter

Для определения частоты элементов, расположенных в словаре, используется метод most_common([n]). Данный метод возвращает n наиболее часто встречающихся элементов, в порядке убывания встречаемости. В случае если n не указано, будут возвращены все элементы.

Применение инструмента к случайной выборке пользователей ВКонтакте из списка друзей.

Для построения исходного набора данных был произведен сбор данных пользователей с разными демографическими атрибутами в случайном порядке. Были проанализированы профили 10 пользователей, результаты указаны в таблице - 2. Результаты точности могут меняться в зависимости от объема подаваемых на вход данных.

Таблица - 2. **Результаты точности атрибутов**

Атрибут	Точность в процентах
Соотношение мужчин и женщин	95%
Предполагаемый пол	70%
Предполагаемый возраст	80%
Возможный город	90%
Возможные увлечения	90%

В результате обработки инструмент генерирует отчет с подробной информацией, указанной ниже:

1. Данные друзей пользователя (ФИО, идентификатор пользователя, дата рождения, возраст, пол, город, мобильный телефон)
2. Количество представителей полов в друзьях у пользователя
3. Предположительный пол владельца таргетированного профиля на основе анализа соотношения представителей разного пола в друзьях
4. Предполагаемый возраст на основе среднего значения возраста друзей
5. Предполагаемый город на основе часто встречающихся городов у друзей

На рисунке - 2 представлен экземпляр отчета сформированный после таргетированного выбора одного из друзей в социальной сети ВКонтакте. В качестве таргетированного профиля был выбран профиль знакомого человека, с которым ранее познакомились на курсах по программированию. Отчёт показывает достаточно хороший результат, следует отметить атрибут возможных увлечений и возможный город.

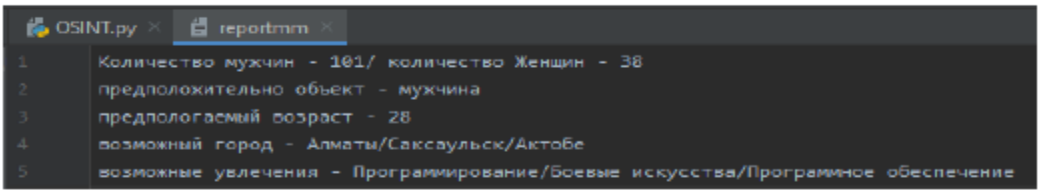

Рис. 2. Экземпляр отчета

Было установлено, что уровень ложных результатов и ошибок зависит также от настроек приватности страницы пользователей в социальной сети. Наличие в списке друзей профиля поддельных аккаунтов или профилей с неполными или некорректными анкетными данными ухудшает точность полученных результатов. Однако в случае высокой интеграции аккаунта в дружеские связи и сообщества инструмент выдает наиболее верные атрибуты для идентификации.

Заключение

В рамках данного исследования был рассмотрен метод анализа демографических атрибутов. Был протестирован инструмент для разведки на основе открытых источников. Рассмотрены методы получения исходных данных путем обращения к интерфейсу программирования социальной сети. Полученные результаты подтверждают корректность работы инструмента и достоверность результатов со средним значением 85%. Метод может быть использован в качестве одного из элементов для системы мониторинга и сбора данных из социальных сетей. Данный метод и инструмент могут быть также полезны для набора определенных датасетов с целью последующего анализа.

Работа выполнена в рамках проекта «Разработка моделей, алгоритмов семантического анализа веб-контента для определения экстремистской направленности и создание инструментария кибер-криминалистики», IRN AP06851248, договор №4, финансирование осуществлено Министерством цифрового развития, оборонной и аэрокосмической промышленности Республики Казахстан.

ЛИТЕРАТУРА

[1] VK Statistics and Facts (2020) | By the Numbers. Режим доступа: https://expandedramblings.com/index.php/vk-statistics-facts/ — Загл. с экрана. — Яз. анг. (дата обращения: 01.07.2020).

[2] В Казахстане открылось представительство ВКонтакте. Режим доступа: https://vk.com/press/kazakhstan-start — Загл. с экрана. — Яз. рус. (дата обращения: 01.07.2020).

[3] Cresci S., Di Pietro R., Petrocchi M., Spognardi A., Tesconi M., Fame for sale: Efficient detection of fake Twitter followers. [Текст]. / Decision Support Systems — 2015

[4] Ferrara E., Varol O., Davis C., Menczer F., and Flammini A., The Rise of Social Bots. [Текст]. / Communications of the ACM — 2016

[5] Facebook shares drop on news of fake accounts. Режим доступа: https://www.cbc.ca/news/technology/facebook-shares-drop-on-news-of-fake-accounts-1.1177067 — Загл. с экрана. — Яз. анг. (дата обращения: 01.07.2020).

[6] Ripeanu, K. Beznosov, and E. Santos-Neto, "Thwart-ing fake osn accounts by predicting their victims," inProceedings ofthe 8th ACM Workshop on Artificial Intelligence and Security. [Текст]. /ACM — 2015

[7] Facebook Publishes Enforcement Numbers for the First Time. Режим доступа: https://about.fb.com/news/2018/05/enforcement-numbers/ — Загл. с экрана. — Яз. анг. (дата обращения: 01.07.2020).

[8] Чесноков В.О., Применение алгоритма выделения сообществ в информационном противоборстве в социальных сетях. [Текст]. / Вопросы кибербезопасности No1(19) — 2017

Мусиралиева Ш.Ж., Айдын М.Е., Оспанов Р. Қ.

Әлеуметтік желідегі достық байланыс және демографикалық атрибуттарға негізделген пайдаланушыны анықтау әдісі

Түйіндеме: Бүгінгі таңда әлеуметтік желілер көптеген ақпаратқа қол жеткізуге арналған алаң болып табылады. Пайдалы ақпаратты қоспағанда, әлеуметтік заңсыз әрекеттер үшін ыңғайлы алаңға айналды. Күдікті әрекеттер көбінесе әлеуметтік желілерде қауіптерді анықтау және талдау жүйелерінің жоқтығымен байланысты. Бұл мақалада пайдаланушы профиліндегі ақпаратты талдау тәсілдеріне қысқаша шолу жасалды. Ашық деректерге негізделген барлау әдістері және әлеуметтік желінің бағдарламалау интерфейсінің әдістері қарастырылған. Достық қарым қатынас пен қосымша атрибуттар деректерін талдау негізінде профиль пайдаланушысын анықтау әдісі сыналды.

Түйінді сөздер: әлеуметтік желі, Вконтакте, деректерді талдау, OSINT, виртуалды байланыстар

UTC: 62:004.9
УОК: 62:004.9

D. Mukhammedzhanova
(Information Systems, Almaty University of Energy and Communications, Almaty, Kazakhstan.
E-mail: dinargul_97@mail.ru)

PERSONAL DATA PROTECTION IN INFORMATION SYSTEMS BY METHOD OF DE-IDENTIFICATION

Abstract. The rapid development of computer technology has led to the expansion of personal data storage capacity. The vast amount of information that governments and businesses collect from individuals has become a cause for concern. The collection of personal data violates human rights, especially if it violates confidentiality or the right to control information about yourself; there is disclosure of personal facts; and the information can be used in a way that