ВЛАДИМИР АЛЕКСАНДРОВИЧ МИНАЕВ,
доктор технических наук, профессор,
профессор кафедры специальных
информационных технологий
Московский университет МВД России
имени В. Я. Кикотя
E-mail: m1va@yandex.ru

АНДРЕЙ НИКОЛАЕВИЧ РАБЧЕВСКИЙ,
кандидат технических наук, доцент,
доцент кафедры информационной
безопасности и систем связи
Пермский государственный национальный
исследовательский университет
E-mail: ran@psu.ru

ЕВГЕНИЙ АНДРЕЕВИЧ РАБЧЕВСКИЙ,
Генеральный директор компании СЕУСЛАБ
E-mail: e.rabchevskiy@seuslab.ru

УДК 623.624

Моделирование активности ботов в ходе информационных атак

Аннотация

В статье представлены и исследованы на эмпирических материалах методики детектирования активности ботов, создаваемых для реализации информационных атак в социальных сетях. С целью выявления признаков и активности информационных атак разработаны нейросетевые модели применительно к информационным потокам в сети ВКонтакте. Учтено, что в ней имеются открытые и закрытые профили пользователей. Проверка нейросетевых моделей на тестовых данных показала следующие результаты: для открытых профилей – 95 % точности выявления ботов, для закрытых – 86 %. С помощью созданной программы «Детектор ботов» проанализированы сопровождавшиеся атаками информационные поводы – «Дворец Путина» и «Специальная военная операция». Анализ показал, что в среднем 24 % постов по инфоповоду «Дворец Путина» создано ботами, по инфоповоду «Специальная военная операции» данный показатель оказался равным 67 %. Выявлена стабилизация структуры постов в социальной сети при их последо-

вательной генерации в виде «Пользователи → Боты», а также при параллельной генерации теми же источниками.

Ключевые слова и словосочетания: *информационная атака; детектирование; бот; нейросетевая модель; динамика постов.*

Платформы социальных сетей позволяют для распространения информации прибегать к использованию так называемых ботов, способных в автоматическом режиме выполнять определенные действия, например, публикацию целенаправленных сообщений, акцентированное комментирование постов.

Злоумышленники нередко используют ботов при проведении целевых информационных атак в социальных сетях. Под информационной атакой понимается спланированное, специально организованное, массированное информационно-психологическое воздействие на конкретных индивидов, определенные группы социума, все население с целью формирования общественного мнения и массового поведения в соответствии с задачами злоумышленников [2]. Именно боты часто выступают эффективным инструментом для манипуляции мнением общества, распространения дезинформации и создания фейковых новостей в процессе совершения информационных атак. В этой связи использование ботов при распространении информации может служить явным признаком проведении информационных атак в социальных сетях. Поэтому анализ применения ботов важен для выявления такого рода атак и реализации стратегий противодействия им.

Боты в социальных сетях

Ботом называется *компьютерная программа, автоматически решающая определенные типовые задачи* [3]. Такие программы классифицируют по различным типам.

Одним из наиболее распространенных типов являются *чат-боты*. Чатботы используют технологии обработки естественного языка и машинного обучения для предоставления пользователям актуальной информации или услуг. Такие боты применяются для автоматизации взаимодействия с пользователями, предоставления информации при решении таких задач, как поддержка клиентов и онлайн-продажи.

Другим распространенным типом ботов являются *веб-скраперы*, наделенные алгоритмами для автоматического сбора и анализа данных с различных веб-сайтов.

Авторами изучались *социальные боты*, под которыми подразумевается программное обеспечение, которое автоматически генерирует контент и взаимодействует с пользователями социальных сетей, чтобы повлиять на их поведение [4]. Такие боты автоматически публикуют сообщения, отвечают на вопросы пользователей и даже проводят опросы, давая возможность компаниям поддерживать активное присутствие в социальных сетях без постоянного участия в них реальных сотрудников. При этом боты могут помогать в создании персонализированного контента, что делает взаимодействие более эффективным.

Боты могут быть использованы для создания фальшивых аккаунтов, которые имитируют действия реальных пользователей. Подобные аккаунты могут участвовать в обсуждениях, подстрекая к разжиганию различного рода конфликтов и тем самым создавая токсичную атмосферу, чувство незащищенности пользователей в социальных сетях.

Еще одним негативным аспектом является использование ботов для распространения спама, когда ими массово рассылаются рекламные сообщения, ссылки на мошеннические сайты или нежелательный контент. Все это затрудняет поиск полезной информации, снижает доверие к платформам социальных сетей со стороны пользователей.

Боты в социальных сетях играют большую роль в продвижении информации из ненадежных источников, активизируя распространение подобного контента на ранних этапах, а также целенаправленно взаимодействуя с пользователями, обладающими большим количеством подписчиков.

Отметим, что распространение ложной информации особенно опасно в условиях политических выборов или во время кризисов, когда достоверность информации имеет решающее значение [1]. Так, перед выборами в Европарламент в 2019 году опубликованы тысячи постов с дезинформацией по вопросам иммиграции, преступности и национальной безопасности. Это значительно повлияло на выборы, так как взгляды многих избирателей изменились на противоположные [6].

Социальные боты способствуют распространению только некоторых идей из их множества. Таким образом, боты могут быть использованы для создания фильтров. Продвигая определенные идеи и игнорируя альтернативные мнения, боты могут активно создавать иллюзию консенсуса. Такой искусственный консенсус опасен возможностью поляризации общества. Согласно работе [8], даже незначительное присутствие ботов в онлайн-пространстве

может кардинально изменить восприятие и взгляды реальных пользователей.

Для выявления ботов используются различные методы, сравнительный анализ которых приводится в [9–10]. Самым эффективным признан метод машинного обучения, позволяющий быстро анализировать большие объемы данных и обнаруживать ботов в реальном времени. Тем не менее и у этого метода есть недостатки. Так, его эффективность**17** в значительной степени зависит от качества и объема входных данных. Если обучающая выборка содержит недостаточно примеров или искаженные данные, это может привести к неверным итоговым результатам. Кроме того, алгоритмы, реализующие метод, ориентируются на особенности конкретных социальных сетей, что усложняет использование единого программного решения для различных платформ.

Для обнаружения ботов с помощью методов машинного обучения можно использовать два подхода. Первый заключается в анализе контента, который публикует пользователь. Основное внимание в данном случае уделяется текстам сообщений, изображениям и видео, которые могут указывать на автоматизированные действия. Второй подход основывается на изучении метаданных аккаунта пользователя, таких как имя, статус, количество друзей, количество публикаций и др. Ключевой задачей здесь является выбор и извлечение значимых характеристик, поскольку для достижения высокой точности анализа важно корректно определить, какие параметры являются наиболее информативными.

Детектирование ботов

Целью авторов в настоящей работе выступило исследование применения ботов в информационных атаках, проводимых в социальной сети ВКонтакте. Для детектирования ботов применялись нейросети как эффективный способ анализа большого числа аккаунтов пользователей.

После изучения аккаунтов, выделены 29 параметров в качестве входных данных для нейронной сети. Среди них: количество дней с последнего захода, возраст аккаунта в днях, количество подписок, полнота заполнения профиля, среднее количество постов в день, количество друзей, отношение постов и репостов, среднее количество комментариев на постах, среднее количество лайков на постах, среднее количество просмотров на постах, количество подписчиков, количество фото профилей, количество видео, количество аудио, количество групп, количество городов в группах, а также другие параметры.

Учтено, что в сети ВКонтакте имеются открытые и закрытые профили пользователей, отличающиеся числом доступных параметров. Поэтому созданы две нейросетевые модели – для поиска ботов среди открытых профилей и для поиска ботов среди закрытых профилей.

Обе нейронные сети представляют собой многослойный персептрон с одним скрытым слоем и одним выходным нейроном, который выдает значения от 0 до 1, обозначающие вероятность того, что данный аккаунт является ботом. Аккаунты ботов получены с помощью сайта https://botnadzor.org [11]. В процессе разработки нейросетевых моделей проанализирована значимость входных параметров.

Проверка полученных нейросетевых моделей на тестовых данных показала следующие результаты: для открытых профилей – 95% точности выявления ботов, для закрытых – 86%. После успешного тестирования нейросетевых моделей создана программа «Детектор ботов» [12], которая классифицирует профили пользователей ВКонтакте на две категории: «Боты» и «Не боты».

Экспериментальные результаты

С помощью программы проанализированы два повода, в связи с которыми проводились информационные атаки в социальной сети ВКонтакте – «Дворец Путина» и «Специальная военная операция».

Анализ показал, что 24 % постов по инфоповоду «Дворец Путина» создано ботами. Для инфоповода «Специальная военная операции» данный показатель почти в три раза выше – 67 %.

Экспериментальная часть исследования показала следующее.

Во-первых, в рамках, рассмотренных инфоповодов найдены и изучены потоки распространения информации (рис. 1), когда публикацию сначала создает реальный человек (группа людей), затем ее начинают массово распространять другие пользователи, а позже подключаются боты, число которых также возрастает.

Во-вторых, выявлены информационные потоки, когда изначально несколько постов создаются ботами, а затем их начинают распространять как пользователи, так и боты (рис. 2).

На рис. 1 стабилизация происходит на уровне 20–21 %. Синяя кривая отражает асимптотический выход на определенный уровень. Красная – рост общего числа постов. Оранжевая – распределенное во времени отношение числа постов со стороны ботов к общему числу постов в течение часа. Из рисунка видно, что в итоге, с течением времени отношение постов, генерируемых ботами, к общему числу публикаций стабилизируется.

Рис. 1. Стабилизация структуры постов в социальной сети при их последовательной генерации «Пользователи → Боты»

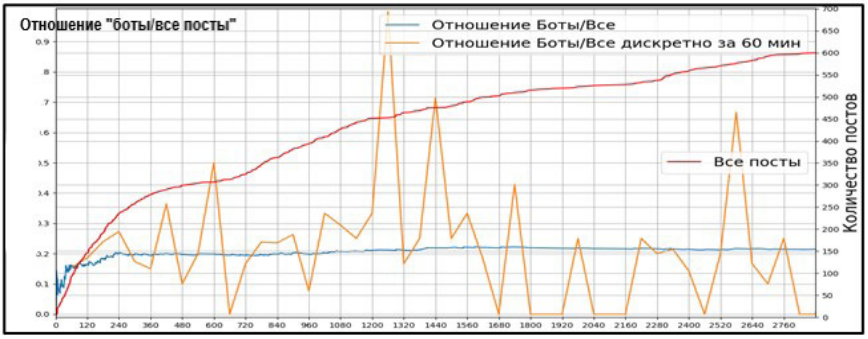

Рис. 2. Стабилизация структуры постов в социальной сети при их параллельной генерации «Пользователи + Боты»

На рис. 2 синяя кривая отражает стабилизацию на уровне 30 %. Красная описывает рост выхода общего числа постов на асимптотический уровень. Оранжевая – распределенное во времени отношение числа постов со стороны ботов к общему числу постов в течение часа.

Предложенная и успешно апробированная нейронная модель позволила провести еще ряд очень полезных экспериментов с данными из социальной сети ВКонтакте. В частности, подтверждены две исследовательские гипотезы относительно закономерностей в динамике постов, созданных ботами и реальными пользователями.

Первая из них сформулирована следующим образом – со временем доля постов, созданных ботами, в их общем количестве начинает снижаться, поскольку реальные пользователи сами начинают более активно делиться полученной информацией.

Вторая гипотеза, напротив, отражает рост доли постов, распространяемых ботами. Это происходит в случае, когда пользователи не проявляют интереса к информации и не делятся ею самостоя-

тельно с другими. Тогда боты «возлагают» на себя задачу распространения информации с целью ее донесения до как можно более широкой аудитории.

Обе гипотезы нашли свое подтверждение при анализе с помощью нейронной модели различных практических вариантов распространения информации в социальной сети, отражающих информационные атаки с применением ботов [5].

Результаты проведенных исследований демонстрируют, что злоумышленники активно создают и применяют ботов для осуществления информационных атак в социальных сетях. Такие автоматизированные системы дают возможность эффективно распространять дезинформацию и манипулировать общественным мнением, что делает задачу обнаружения Интернет-ботов весьма актуальной.

Очевидно, что наибольшую опасность как для отдельных индивидов, социальных групп, так и общества в целом представляют социальные боты, выступающие в руках злоумышленников инструментом дестабилизации общества, деструктивной трансформации в нем политических, социально-экономических отношений и морально-этических норм [7].

Изучение эмпирических данных из социальной сети ВКонтакте с помощью разработанных нейросетевых моделей свидетельствует о том, что они позволяют эффективно выявлять структурные и динамические особенности в системе постов, генерируемых ботами и пользователями, в условиях реализации информационных атак.

При этом возможно моделирование не только специфических характеристик открытых и закрытых профилей пользователей, как это сделано авторами, но и воспроизведение более глубоких количественно-качественных параметров информационных потоков, характеризующих латентные, то есть скрытые аспекты данных об информационных атаках. Это весьма важно для надежного обоснования мероприятий по противодействию таким атакам, особенно, если они связаны с деструктивными воздействиями злоумышленников с применением ботов.

Список литературы:

1. *Коцюбинская Л. В.* Информационная атака: понятие и онтологические свойства // Политическая лингвистика. 2017. №16. С. 106–111.

2. Что такое боты – определение и описание. URL: https://www.kaspersky.ru/resourcecenter/definitions/what-are-bots (дата обращения: 26.04.2025).

3. *Faris R. et al.* Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election. Ed. Berkman Klein Center for Internet & Society at Harvard University. Cambridge, Massachusetts, 2017. P. 142.

4. *Pierri F., Artoni A., Ceri S.* Investigating Italian Disinformation Spreading on Twitter in the Context of 2019 European Elections // PLoS One. 2020. Vol. 15. № 1.

5. *Aldayel A., Magdy W.* Characterizing the Role of Bots' in Polarized Stance on Social Media // Social Network Analysis and Mining. 2022. Vol. 12. № 30. Pp. 1–24.

6. *Hayawi K. et al.* Social Media Bot Detection with Deep Learning Methods: Systematic Review // Neural Computing and Applications. 2023. Vol. 35. Pp. 8903–8918.

7. Ботнадзор. URL: https://botnadzor.org/ (дата обращения: 26.04.2025).

8. Программа «Детектор ботов». Свидетельство о регистрации программы для ЭВМ. Регистрационный номер: 2024617353 от 1 апреля 2024 г.

9. *Минаев В. А. [и др.].* Противодействие экстремистской идеологии в социальных медиа: математические модели и методы. Хабаровск, 2023. 232 с.

10. *Минаев В. А., Рабчевский А. Н., Мустакимова Я. Р.* Выявление информационных операций в социальных сетях на их ранних стадиях // Информация и безопасность. 2022. № 4 (25). С. 485–494.