УДК 004.056.53

Смолич Д.С., студент гр. 125м-16-1
Научный руководитель: Кручинин Александр Владимирович, ст. преп. кафедры безопасности информации и телекоммуникаций
(Государственное ВУЗ «Национальный горный университет», г. Днепр, Украина)

АНАЛИТИЧЕСКИЙ ОБЗОР МЕТОДОВ ДЕТЕКТИРОВАНИЯ ВРЕДОНОСНЫХ АККАУНТОВ В СОЦИАЛЬНЫХ СЕТЯХ

В данной работе рассматриваются виды вредоносных аккаунтов, риски и последствия их использования в социальных сетях. Также представлены результаты проведенного анализа существующих методов обнаружения и детектирования «социальных ботов».

Ключевые слова: боты, детектирование, обнаружение вредоносных аккаунтов, кластеризация, Captcha, Sms-verification, Ratelimit.

ВВЕДЕНИЕ

В настоящее время социальные сети – неотъемлемая часть большинства сфер жизни человека. Простое общение, поиск людей и необходимой информации, продвижение товаров и услуг, реклама, проведение бизнес-конференций и многое другое возможно в режиме онлайн, В социальных сетях ежедневно общается 40% населения планеты. Здесь сконцентрировано огромное количество информации и денег. Естественно, что некоторые стремятся использовать такие безграничные возможности для наживы и достижения своих далеко не благородных целей.

Одна из основных угроз на сегодняшний день – так называемые «социальные боты». [1] Это вредоносные программы, поддельные аккаунты, способные имитировать поведение людей. Как правило, боты используются для:

- организации информационных вбросов;
- массового хищения персональных данных;
- ухудшения доверия в соц.сетях;
- создания ложных новостей и голосований;
- как средство для легального бизнеса киберпреступности;
- создания проблем в социальном маркетинге.

Так, к примеру, при SMM продвижении в Вконтакте либо Facebook часть целевой аудитории это «мертвые» аккаунты или аккаунты-двойники, [4]

Угрожающие масштабы использования социальных ботов требуют создания эффективных алгоритмов их детектирования. Решением проблемы является развитие методов сетевого анализа, которые предназначены для выявления и кластеризации сообществ в соц.сетях, оценки их связности, степени доверия.

ВИДЫ «СОЦИАЛЬНЫХ БОТОВ»

Виртуальная (онлайн) социальная сеть – структура Интернет-среды, представленная электронными порталами, такими как «Одноклассники», «Вконтакте», «Twitter». «Facebook» и др. Узлами сети являются отдельные люди или организации, связанные корпоративными, политическими, служебными, семейными, дружескими взаимодействиями. Внутри сети образуются группы и сообщества с сильными стабильными связями, некая иерархия (рисунок 1), что позволяет их легко идентифицировать. [9]

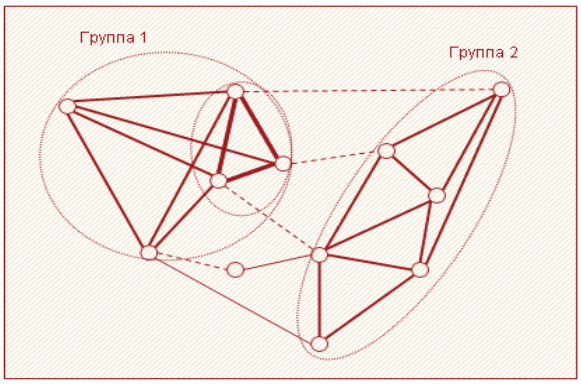

Рис 1. Иерархия групп в соц. сетях.

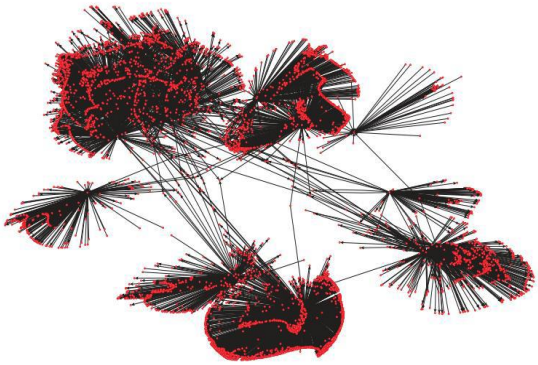

Рис 2. График взаимосвязей пользователей в соц.сети.

Социальные боты классифицируются следующим образом;

1. Аватары – профили с хорошо налаженными социальными связями. [13]
2. Новостные боты – аккаунты, распространяющие как реальные, так и фейковые новости.
3. Боты для организации информационных вбросов, распространяя ложные сообщения реальных людей.
4. Игровые боты используются в играх для накрутки упоминаемости конкретных приложений.
5. Боты, участвующие в накручивании репутации для заработка денег на репостах.
6. Аккаунты, публикующие чужие новости и перепосты.

Основная сложность при детектировании вредоносных аккаунтов – отличить реальных ботов от людей, похожих по поведению на боты.

МЕТОДЫ ДЕТЕКТИРОВАНИЯ БОТОВ

Основой в разработке современных методов выявления вредоносных аккаунтов стал принцип распознавания образов. Научная дисциплина позволяет классифицировать объекты, основываясь на «прецедентах» - образах, правильная классификация которых известна. В зависимости от наличия или отсутствия прецедентной информации различают классификацию с обучением и без обучения (кластеризация).

Существующие алгоритмы по борьбе с ботами являются скорее профилактическими:

- Captcha – автоматизированный тест Тьюринга, определяющий, является пользователь компьютером или человеком.
- Sms-verification – проверка пользователя посредством отправки смс с кодом подтверждения.
- Ratelimit – временное ограничение числа запросов к системе.

Задача классификации – разбиение множества объектов или наблюдений на априорно заданные группы, называемые классами. Внутри каждой из которых они предполагаются похожими друг на друга, имеющими одинаковые признаки и свойства. Если количество классов не более двух, имеет место бинарная классификация. Для классификации сложных моделей применяются традиционные алгоритмы.

К примеру, для исследования проблемы обнаружения спам-ботов в Twitter разработчики применили нейронные сети, деревья решений, машины опорных векторов, наивный байесовский классификатор. [6] и [14]. В качестве признаков использовалось количество подписчиков и читаемых, а также граф-ориентированные взаимосвязи. (рис 2.)

Общий подход при создании алгоритма детектирования социальных ботов (проектирование фреймворка) [7] включает в себя:

- сбор данных;

- выявление признаков (метрик);
- выборка классифицированных данных для обучения классификатора;
- обучение классификатора на данной выборке.

Проблема различения аккаунтов на человека, бота и киборга [13] решается с помощью создания обширной базы аккаунтов и классификации по различным признакам: поведению, содержанию публикаций, характеристик профиля. В результате формируется система, состоящая из 4 частей:

- компонент энтропии (entropycomponent) – анализирует интервал между сообщениями аккаунта;
- компонент обнаружения спама (spamdetectioncomponent) – анализирует содержание сообщений по заданным шаблонам;
- компонент анализа профиля (accountpropertiescomponent) – анализирует признаки профиля;
- компонент классификации (decisionmarkercomponent) – классификация аккаунта на результатах предыдущих компонентов.

Одними из эффективных классификаторов являются:

- алгоритм приманки (Honeypot) [12];
- анализ маштабограмм (scalogram) [15];
- математический алгоритм Байеса, классификатор, основанный на теореме Байеса[18].

Современные разработчики алгоритмов детектирования ботов используют методики, учитывающие особенности контента, краудсорсинг, технологии сетей [17] и многие другие характеристики и признаки аккаунтов социальных сетей, создавая все более универсальные классификаторы.

ВЫВОДЫ

В ходе проведения обзора методологии детектирования социальных ботов были проанализированы наиболее эффективные способы обнаружения, классификации и борьбы с вредоносными аккаунтами. Рассмотрены основные принципы распознавания образов, достоинства и недостатки популярных классификаторов. Увы, имеют скорее профилактический характер. Современные тенденции широчайшего использования социальных сетей в различных сферах требуют создания более гибких и универсальных моделей для распознавания вредоносных программ.

ПЕРЧЕНЬ ИСТОЧНИКОВ

1. Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, Matei Ripeanu. The Socialbot Network: When Bots Socialize for Fameand Money / University of British Columbia Vancouver, Canada, 2011. – 1 с.

2. Korrespondent.net – Продажа ботов в Twitter стала многомиллионным бизнесом – исследование. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: https://korrespondent.net/business/web/1540808-prodazha-botov-v-twitter-stala-mnogomillionnym-biznesom-issledovanie (дата обращения 15.11. 2017)

3. Carlo De Micheli, Andrea Stroppa. Twitter and the underground market / 11th Nexa Lunch Seminar, May, 2013.— 5–9 с.

4. Topmarketing.by – Честное SMM-продвижение, выявляем аккаунты-боты в социальных сетях. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://topmarketing.by/internet-marketing/chestnoe-smm-prodvizhenie-vyyavlyaem-akkaunty-boty-v-socialnyx-setyax.html (дата обращения 14.11. 2017)

5. Cossa.ru – Использование ботов (технических аккаунтов) в работе с отзывами. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://www.cossa.ru/155/12121/ (дата обращения 12.11. 2017)

6. Alex Hai Wang. Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach / College of Information Sciences and Technology The Pennsylvania State University, USA, 2010. — 2–10 с.

7. R. Nithin Reddy, Nitesh Kumar. Automatic Detection of Fake Proles in Online Social Networks / Department of Computer Science and Engineering National Institute of Technology Rourkela, Orissa, India, May, 2012. — 10–15 с.

8. Inosmi.ru – Бот в твиттере оказался настолько убедительным, что люди начали сочувствовать «ей». [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://inosmi.ru/world/20120627/194149517.html (дата обращения 15.11. 2017)

9. Basegroup.ru, BasegroupLabs – Введение в SocialMining. [Электронный ресурс]: [Веб-сайт]. – Режим доступа: http://www.basegroup.ru/library/web_mining/introduction_in_social_mining (дата обращения 16.11. 2017)

10. Wikipedia.org. Сводная энциклопедия – Кластерный анализ. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://ru.wikipedia.org/wiki/Кластеризация (дата обращения 15.11. 2017)

11. Wikipedia.org. Сводная энциклопедия – Задача классификации. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://ru.wikipedia.org/wiki/Задача_классификации (дата обращения 15.11. 2017)

12. Securitylab.ru SecurityLab – Технология Honeypot. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://www.securitylab.ru/analytics/275420.php (дата обращения 13.11. 2017)

13. Zi Chu, Steven Gianvecchio, Haining Wang, SushilJajodia. Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg? / IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, 2012. – 2-10 с.

14. Alex Hai Wang. DON'T FOLLOW ME: SPAM DETECTION IN TWITTER / College of Information Sciences and Technology, The Pennsylvania State University, Dunmore, USA, 2012. – 2-7 с.

15. Wikipedia.org. Сводная энциклопедия – Scaleogram. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://ru.wikipedia.org/wiki/Scaleogram (дата обращения 11.11. 2017)

16. Wikipedia.org. Сводная энциклопедия – JSON. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://ru.wikipedia.org/wiki/JSON (дата обращения 11.11. 2017)

17. J. Ratkiewicz, M. D. Conover, M. Meiss, B. Gonc¸alves, A. Flammini, F. Menczer. Detecting and Tracking Political Abuse in Social Media / Center for Complex Networks and Systems Research School of Informatics and Computing Indiana University, Bloomington, IN, USA, 2011. — 2-5 с.

18. Wikipedia.org. Сводная энциклопедия – Теорема Байеса. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://ru.wikipedia.org/wiki/Теорема_Байеса (дата обращения 12.11. 2017)

19. Ibm.comПрограммное обеспечение IBM – Анализ социальных сетей – техническая публикация. [Электронный ресурс]: [Веб-сайт]. – Режим доступа: http://public.dhe.ibm.com/software/dw/ru/download/ZZW03070.pdf (дата обращения 15.11. 2017)

20. Haijian Shi. Best-first Decision Tree Learning / The university of Waikato, Hamilton, NewZealand, 2007. — 5 с.

21. Wikipedia.org. Сводная энциклопедия – Дерево принятия решений. [Электронный ресурс] : [Веб-сайт]. – Режим доступа: http://ru.wikipedia.org/wiki/ Дерево_принятия_решений (дата обращения 12.11. 2017)