Polytechnic Journal. 2022. 12(2): 1-7
ISSN: 2313-5727
http://journals.epu.edu.iq/index.php/polytechnic

POLYTECHNIC JOURNAL

RESEARCH ARTICLE

Social Bot Detection using Machine Learning Algorithms: A Survey and Research Challenges

Kayhan Zrar Ghafoor

Department of Software Engineering, Salahaddin , University-Erbil, Iraq.

***Corresponding Author:**

Kayhan Zrar Ghafoor,
Department of software Engineering, Salahaddin University Erbil-Iraq.

E-mail:
mrkayhanz@gmail.com

Received:
Accepted: 02/3/2022
Published:1/2/2023

DOI
10.25156/ptj.v12n2y2022.pp219-2

ABSTRACT

In the past decade social media platforms growing rapidly and they are part of our routine life. Each platform has its own specification which uses for specific purposes. After this widely spread, those SMPs were targeted by the cybercriminals to cast their malicious activities. There are many different malicious activities in SMPs such as spamming, phishing, fake account. In these papers, Bots activities in SMPs one of those threats which include fake accounts, fake friends/followers, spreading misinformation by purpose, and many more. At the beginning of our work, we explain all terminology related to this topic to have a clear understanding of what is going on now. Then we reviewed the recent papers about this topic. We found out different models suggested by the researchers for recognizing those malicious activities. Until now most of the work focusing on Twitter as a platform, English as a language, and machine learning as a detection method but there are many gaps in this research area because Twitter is the 17th most used SMPs in 2020, also there are many malicious actions in other languages, and detection method needs lots of improvement in reliability, accuracy, real-time detection, and performance area. As a result, we are at the beginning of the game and we need lots of improvement for controlling the bot's activities. Besides all technical term also people awareness has a big impact on controlling a bot because most of the times the botmaster use people ignorance to make their actions easy.

Key Words: Social bots, Social media, malicious activities, Bot detection

1-INTRODUCTION

Social media platforms are affecting all aspects of our life. Moreover, the number of users on those platforms increased rapidly day by day. With an extraordinary number of users, they have many advantages/disadvantages, some of their advantages are making the world a small town, you can buy, communicate, learn, teach, and help others very easily. Besides their advantages, they are the right place for the attacker to spamming, phishing, Click Fraud, and other differences of attacks. Also, many young people are addicted to those social media platforms, They spend most of their life there[Richards D, Caldwell PHY and Go H, 2015]. One of the new term related to those social media platform is a social bot, those social bots are born with social platforms but they are not popular until the last decade, they used for spreading all kind of attack, also they affect people mind by spreading misinformation news about a recent topic.

Those social bots have an enormous effect on our lives and they are a source of many malicious activities in social media. Due to their rapid increase. In the first quarter of 2018 583 million fake accounts were removed by Facebook, 837 million spam shared, 81 million unacceptable content removed. Still, they expect nearly 4% of remaining accounts are fake.[Karunakar E, Pavani VDR, Priya TNI, Sri V and Tiruvalluru K, 2020][Ram A and Galav RK, 2020] Some reports show that 9% to 15% of the active accounts on Twitter are social bots. [Richards D, Caldwell PHY and Go H, 2015] Likely, the fake news reached 100,000 users and false news 70% more retweets than true news on Twitter. [Pulido CM, Villarejo-Carballido B, Redondo-Sama G and Gómez A, 2020] moreover in the US, the EU, Israel, and Canada enormous bot activities were reported during a Parliament and presidential elections.
[Hanouna S, Neu O, Pardo S, Tsur O and Zahavi H, 2019] recently all the papers [Gallotti R, Valle F, Castaldo N, Sacco P and De Domenico M, 2020][Singh L, Bansal S, Bode L, Budak C, Chi G , Kawintiranon K, et al, 2020] related to COVID-19 information and spread explain the huge impact of social bots.

With this huge spreading and influence of socials bots. Unfortunately, this spreading is increased in future years. [Cresci S, 2020] And the reason is social media like a big free cake everyone wants a piece of it. there are many reasons behind this spreading such as advertising for business, The politician sharing their last believes, celebrities sharing their life with their fans, etc… this variety and amount make social media the perfect place for the cybercriminal to spread the malicious activities and gain.
benefit from it. Also, it is a difficult challenge for researchers in the industry and science side to make social media a safe place for ordinary people. In this work, we focus on the scientific attempt and detection method for controlling these malicious activities.

In this paper, we describing related terms and methods related to social media bots to having a clear view of the state of the art on this topic. We focus on two criteria of the problem: the datasets and machine learning detection methods. From our understanding clarifying those two areas clear the way for upcoming researchers. The paper is organized as follows: In section 2 we review all works related to this topic. In section 3 we highlight the definition and term related to this topic. In section 4 we discuss the machine learning techniques for detecting social bots. In section 5 we focus on the datasets used by the papers for detecting bots. In section 6 we explain the open problems of the area. Lastly, we discuss the paper conclusion and revealing the comparison table between the papers.

2-RELATED WORK

There are other works related to this topic. Some of them choose a couple of papers for review. Other use the systemic literature review (SLR) which is a more organized and clear scope review. [Okoli C and Schabram K, 2010] we describe some of the works that release in 2016 or later. In [Adewole KS, Anuar NB, Kamsin A, Varathan KD and Razak SA, 2017] work they focus on detecting spam, fake, compromised, and phishing accounts on social platforms. Also, they classify the features into Behavioral, Profile, and Network each of those has sub-features, furthermore, they classify detection methods into Crowdsourcing, graph-base, and machine learning which contain Supervised, Unsupervised, Semi-supervised subcategories. Another work is [Ferrara BYE, Varol O, Davis C, Menczer F and Flammini A, 2014] which specially focuses on social bots. Categorize the detection methods into three classes, graph-based, crowd source, and machine-learning

Table 1: The previous survey for bots topic.

Ref	Year	Platform	Detection Method	SLR	# Paper	Focus On
[Ferrara E et al, 2016]	2016	All SM	Graph, ML, Crowdsourcing	no	---	Detecting
[Adewole KS et al, 2017]	2016	All SM	Graph, ML, Crowdsourcing	no	---	Detecting Fake, spam, phishing, compromised accounts
[Stieglitz S et al, 2017]	2017	All SM	---	yes	103	Types and behavioural
[Karataş A et al, 2017]	2017	All SM	Graph, ML, Crowdsourcing	no	---	Detecting
[Alothali E et al, 2018]	2018	Twitter	Graph, ML, Crowdsourcing	no	---	Detecting, Datasets
[da Silva F et al, 2019]	2019	Facebook, Twitter	ML	yes	169	Fake news
[Orabi M et al, 2020]	2020	All SM	Graph, ML, Crowdsourcing, Anomaly	yes	55	Detecting
[Cresci S, 2020]	2020	All SM	---	no	---	Bots nature
OURS	2021	Twitter	ML	no	20	All activities related to bots

without any sub categorization. Also, they categorize the twitters features into Network, User, Friends, Timing, Content, and Sentiment.

In [Grimme C, Preuss M, Adam L and Trautmann H, 2017] paper the focus more on "what is the bots", the difference between good and bad bots, the potential influence of bots on social media also they do some experiment depend on [Ferrara BYE, Varol O, Davis C, Menczer F and Flammini A, 2014],[Davis CA, Varol O, Ferrara E, Flammini A and Menczer F, 2016] features classifications. In [Karataş A and Şahin S, 2017] work which is classify the method detection the same as previous work also describes the malicious influence of Social bots. In [Stieglitz S, Brachten F, Ross B and Jung AK, 2017] work that uses SLR techniques, They review 103 different papers from those databases: Scopus, ScienceDirect, and AISEL. they focus on social bot intention on social platforms. Likewise, they differentiate between social bots and other bots.

In [Alothali E, Zaki N, Mohamed EA and Alashwal H, 2018] work they focus on the Twitter platform precisely. Their focus is on papers datasets and detection methods. They show the numbers of tweets and accounts that are used in each reviewed paper. Additionally, they define the most repeated features used in the papers.

Finally, they categorize the detection method as the same as previous work but they focus on machine learning approaches, they show all the different methods and the papers that adopted them. In [Cardoso Durier da Silva F, Vieira R, Garcia AC,2019] work which SLR paper, they review around 170 papers in the last five years. they define all terminology related to social bots then they focus on platforms that are used by those papers. In the last section of their work they focus on all related topic of machine learning algorithm such as Datasets, Preprocessing, Features, etc In [Cresci S, 2020] work, they differentiate between old and new bots, they find out the new generation bots created by using more advance techniques and detecting them are not the easy ways. Additionally, they describe individual and group account detection. Finally, in Orabi M, Mouheb D, Al Aghbari Z and Kamel I, 2020] which is also SLR, it is recent and very well organized work. They collected all papers from popular databases for the last ten years. Their main focus on detection methods besides the previous categorizes they add a new detection category called Anomaly-based detection. However, they differentiate between Behavior/content detection methods for supervised and unsupervised machine learning.

In our work, we try to review those papers that release in 2020. Besides that, we review older papers if needed. We try to add extra new information for previous work related to bots detection in social media. One of our future works is converting this work to SLR and review all the papers related to COVID-19 bots detection.

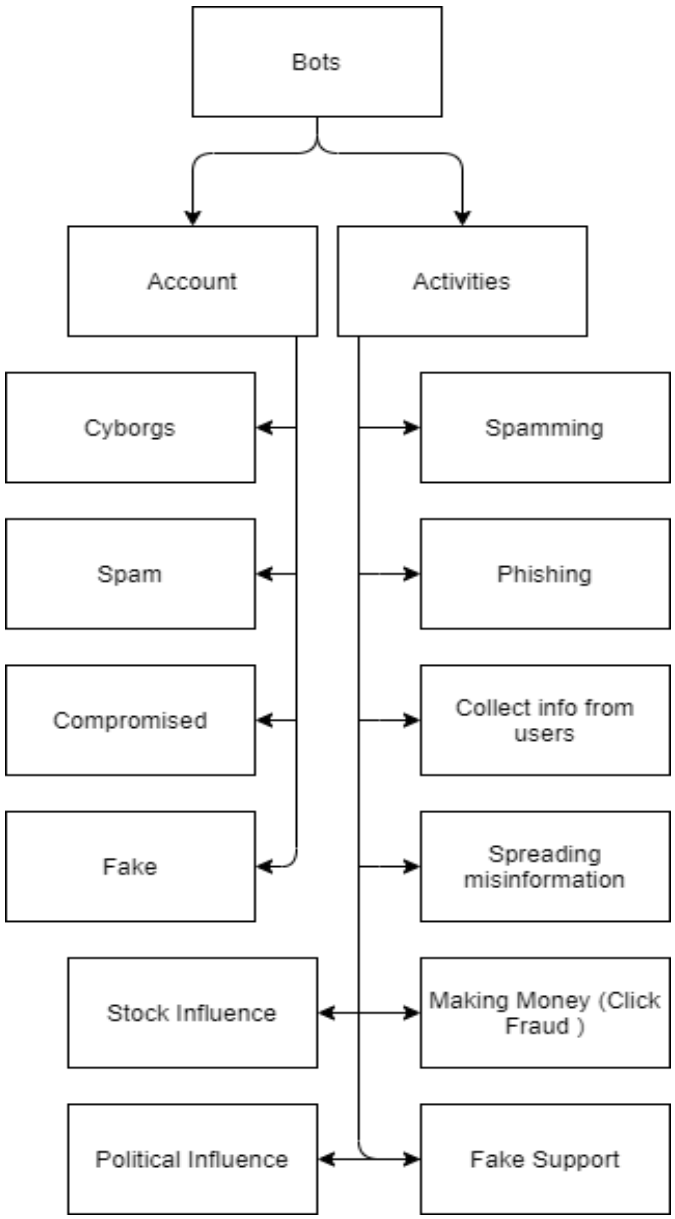

Fig. 1. Different bot accounts and activities

3-DETECTION METHOD

As we discussed shortly in previous chapters, there are many different ways of detecting SMBs. We don't have a standard way for the classifying detection method. Each reviewer is classified in their ways. But generally, all of them [Adewole KS, Anuar NB, Kamsin A, Varathan KD and Razak SA, 2017], [Karataş A and Şahin S. A, 2017], [Alothali E, Zaki N, Mohamed EA and Alashwal H, 2018], [Orabi M, Mouheb D, Al Aghbari Z and Kamel I,2020] classify the methods into three different categories, Graph-based detection, Crowdsourcing detection, Machine Learning-based detection, and ML is divided into Supervised, Unsupervised, Semi-supervised ML. What we noted in the previous works is the categorization has the same base and this base will expand year by year due to finding new algorithms.

But the mean expansion on machine learning-based because most of the researchers go to this path. Also, most of the

methods applied for detecting the Bots on Twitter. In our work, we review all the categories of machine-learning-based in detail.

Machine learning (ML) is the way of teaching machines to do, recognize, detect the things around them by learning it. Machine learning algorithms having a huge role in the new technology due to their nature of working with big data, good performance, and accurate results.[Daya AA, Salahuddin MA, Limam N and Boutaba R, 2020] Because of that, they are the most adapted methods in the articles for detecting social bots which is very suitable for such kinds of problems. ML can be categorized into three sub-categories which are supervised ML (SML), Unsupervised ML (UML), Semisuperviser ML (SSML). IN SML the model learns from very big classified data which is called a dataset in a process called the training process.
[Daya AA, Salahuddin MA, Limam N and Boutaba R, 2020] then it goes through two other processes called validation and test afterward the model can predict the result. UML is not classifying the dataset instead we divide the data into different groups depending on their similarity. In another word, we are clustering the data by their similarity.
[Daya AA, Salahuddin MA, Limam N and Boutaba R, 2020] SSML simply combines both SML and UML to get rid of their disadvantages and observe their strength.
[Thesis M , 2018] In **Fig. 2.** explain the categorization and give some algorithms for the machine learning approaches.

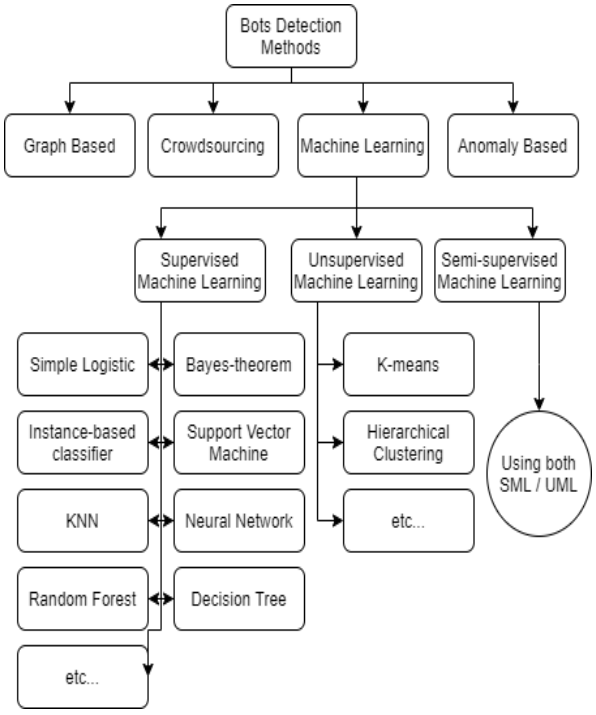

Fig. 2. This figure explains the different algorithms for detecting bots.
In our survey, we will only focus on machine learning approaches.

3.1 Supervised machine learning

In [Khalil A, Hajjdiab H and Al-Qirim N , 2017] work, they focus on detecting fake followers on the Twitter platform they use 18k accounts as the dataset and focusing on 6 different features. They use three supervised machine learning algorithms which are SVM, Simple Logistic, and Instance-based classifier using 1 nearest neighbor. By using the Weka tool with 10-fold cross-validations and default configuration for the algorithm they get those accuracy results: SVM 60.48%, Simple Logistic 90.02%, and Instance-based classifier 98.74%.
In [Kudugunta S and Ferrara E, 2018] work, they worked on detecting fake accounts by using a CNN based on contextual LSTM architecture. They test their algorithm upon a big dataset using 15 different features for detecting fake accounts. They worked in two-level the account-level for detecting fake accounts depends on account metadata and tweet-level for detecting depends on tweets content. Six of the features belong to content features other nine features are belonging metadata features. For account-level, they get 99% accuracy, and for tweet-level, they get 96% accuracy. It's worth mentioning that besides DNN they used Logistic regression, SGD, Randon forest classifier, AdaBoost algorithms (all of them belong to SML) for their test but DNN achieve the best result.

3.2 Unsupervised machine learning

In [Abu-El-Rub N and Mueen A, 2019] work, they create a pubic bot detector. They investigate the 2016 US presidential campaign after the election was ended. They work on the Twitter platform by collecting 75M tweets from 6M account for 60 days. by working on 15 different features that belong to the context, temporal, sentiment features. They used unsupervised machine learning for clustering their dataset. They create five different graphs for mention, media, hashtag, temporal, and retweets using Louvain clustering, afterward by using Boosted decision tree classifier with 10-fold cross-validation for evaluating the interaction of the previous step. As a result, they detect five bot groups who interact during the campaigns, three of them supporting a specific candidate and the two of the remaining focusing on gaining human interaction by tweeting the recent events.
In [Hanouna S, Neu O, Pardo S, Tsur O and Zahavi H, 2019] work, which goals are detecting those account that suspiciously participates in political activities. They used six different datasets contain the tweets related to the US, Canada, France, Israel election with more than 1 billion data. They use the k-means algorithm which is unsupervised ML to clustering their datasets. They focus only on behavior, content features from the datasets by extracting more than 28 different features. As their final result, they find out there are many suspicious accounts involving the Canadian politician and election but the paper goals are finding evidence for involving bots account on smoking gun scandal in Canada which they fail and they didn't find any direct relationship.

3.3 Semi-supervised machine learning

In [Dorri A, Abadi M and Dadfarnia M, 2018] article, they use

Semi-supervised ML for detecting an account if it is a spammer or not in social media. They combine both social interaction graphs with users' behavior for detecting bots they called their algorithm SocialBotHunter. They used an 11k dataset for testing the algorithm they focus on ten different features from behavioral, context, URL, temporal features categories. By using a variety of features it increases the detection rate of the algorithm. Their work looking for the similarity between the features In the first step it creates a social graph for a small group of the real user then trains their classifier depends on this small user group. Then they use Markov random field for detecting the bots on the graph. Finally, they compare their result with other works that used supervised ML binary classifiers with the same datasets. They did the comparison in two-level: Behavior-based and Structural-based in both cases the proposed method has better results.

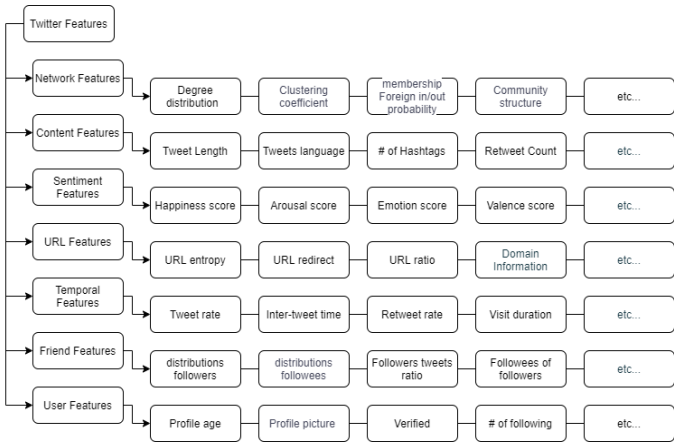

Fig. 3. Twitter features by categories

4-DATASETS

Dataset is a set of data that usually used for testing accuracy, performance, etc… of a program or an algorithm. There are many different ways of collecting those data, Such as coming from a user form submission, User purchases, provided APIs by an organization, and many other ways. For bots detection in social media, the mean sources for getting our datasets are using the APIs provided by those social platforms. If we look at Twitter as an example, there are different ways of collecting data, use Streaming API for collecting 1% unfiltered real-time tweets (public), and Search API for collecting historical data.
[Hanouna S, Neu O, Pardo S, Tsur O, Zahavi H, 2019],[Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M, 2020] during retrieving, we can use several filtering for catching those features that we want. But the number of filtering is limited we can use more than 400 filters for one API request.
[Hanouna S, Neu O, Pardo S, Tsur O and Zahavi H, 2019] but the data collection process is a long process and has many limitations. For instance, in the case of [Singh L, Bansal S, Bode L, Budak C, Chi G, Kawintiranon K, et al, 2020], it took two months to collect around 20M of tweets. The same thing happens to [Gallotti R, Valle F, Castaldo N, Sacco P and De Domenico M, 2020] they need around two months to collect their data after the increasing attention of COVID-19 they can't collect all tweets related to this subject because Twitter does not allow to retrieving more than 4.5 million of data per day.
Some other researchers such as [Battur R and Yaligar N , 2019][Pozzana I and Ferrara E, 2020] used the public datasets which is the dataset collected by the previous researchers. This approach is used to reduce the needed time for data collection but it isn't the right way for all the cases because old datasets reduce the accuracy of the algorithms due to Bots nature that changes very quickly.

Finally, in the case of [Khalil A, Hajjdiab H and Al-Qirim N, 2017], they use the university accounts as legitimate followers and bought some fake followers for 5$. The datasets containing data called features, the variety of data depend on the platforms. For example, on Twitter, we can retrieve more than 1000 features contain almost everything (public data) about the users and tweets. In [Davis CA, Varol O, Ferrara E, Flammini A and Menczer F, 2016] work they extract more than 1000 features from Twitter APIs and create an online service for detecting Twitter bots. Besides their detection method, they classify those features into 6 groups which is Network, User, Friends, Temporal, Content, and Sentiment. Those classifiers showed in **Fig. 3.** However we add 4 features for each classifier just for simplicity. We noted that in some papers they work on tweets URL and detect the Bots using this feature because of that we also add the URL feature to **Fig. 3**. Specifying the features in this way makes the work easier for the researchers they can focus on specific features for their work. In **Table 2** we review the datasets of some papers, most of the papers focus on user features and behavior features. Furthermore, we can notice that the number of features used by the papers is different. This number is very crucial because a small number affects accuracy and a big number affects performance. **Table 2** also shows the number of datasets, tweets, account, and availability of papers data.

Ref	Flatforms	NoD	Availability	TPoT	Total Account	Behavior	USF	NF	TF	UF	S	EF
[Ram A. et al, 2020]	Twitter	1	private	-	10K*	✖	✓*	✖	✖	✖	✖	-
[Karunaka E.et al, 2020]	Twitter	1	private	-	54.5K	✖	✓*	✖	✖	✖	✖	9
[AMUD. et al, 2020]	Twitter	1	public	5K	-	✓	✖	✖	✖	✖	✖	100
[Ferrara E. et al, 2020]	Twitter	2	private	43M	-	✓	✓	✖	✖	✖	✖	6
[Yang KC. et al, 2020]	Twitter	2	public	-	-	✖	✖	✖	✖	✓	✖	-
[Gallotti R. et al, 2020]	Twitter	1	public	100M	-	✖	✓	✖	✖	✓	✖	-
[Hanouna S. et al, 2019]	Twitter	6	public	828M	53M	✓	✓	✖	✖	✖	✖	>28
[Sahoo SR. et al, 2020]	Facebook	8	private	-	120K	✓	✓	✖	✖	✖	✖	13
[Chu Z et al, 2012]	Twitter	1	public	1k	-	✓	✖	✖	✖	✖	✖	-
[Pozzana I et al, 2020]	Twitter	3	public	27.8 M	5.4M	✓	✖	✖	✓	✖	✖	4
[Yang KC et al,2019]	Twitter	14	public	-	250K	✖	✓	✖	✖	✖	✖	20
[Iş H et al, 2019]	Twitter	1	public	2.8M	4.2K	✓	✓	✖	✓	✖	✖	10
[Battur R et al, 2019]	Twitter	1	public	-	-	✖	✓	✖	✖	✖	✖	6
[Singh L et al, 2020]	Twitter	4	public	21M	-	✓	✓	✖	✖	✓	✖	-
[Kudugunta Set al, 2018]	Twitter	4	public	11.8 M	8.3K	✓	✓	✖	✖	✖	✖	15
[Khalil A et al,2017]	Twitter	1	private	-	18K	✖	✓	✓	✓	✖	✖	6
[Abu-El-Rub et al, 2019]	Twitter	1	public	75M	6M	✓	✖	✖	✓	✖	✓	15
[Dorri A et al, 2018]	Twitter	1	public	-	11k	✓	✓	✖	✓	✓	✖	>10

If you put * symbol inside any cells, It means this data is not mentioned directly on the paper instead we expected this data by analyzing the paper.

NoD: Number of Datasets, TPoT: Total Post or tweets, SF: Sentiment Features, UF: URL Features, NF: Network Features, USF: User Features, EF: Extracted Features, TF: Temporal Features

Table 2 : The features used by reviewed papers

5- FUTURE WORKS

The bot detection is similar to the cat mouse game, it is an ongoing game. Bot detection is a new area for researchers besides those great works that researchers already did. Many open challenges have not been handled yet. We summarize some of the open problems as below:

- Most of the released papers are focusing on detecting bots on Twitter. If we look at **Table 3,** it shows that Twitter at the bottom of the popular platform just 353M people using it around the globe. Of course, there are lots of bots on other platforms detecting those bots is one of the future works.

- As we discussed in the dataset section(we show Twitter as an example), most of the platforms have many restrictions for collecting data due to user privacy violations [Adewole KS, Anuar NB, Kamsin A, Varathan KD and Razak SA, 2017]. If growing social bots continue in such a way those social platforms can make APIs with less restriction to make the data collection process easier for the researcher.

- As we discussed in the dataset section, collecting, pre-processing data is consumable work and it needs lots of time and hard work. using public datasets not always the right choice because they are not up to date. Because of that creating huge organized multi-feature datasets is one of the future works in this area.

- A machine-learning algorithm uses by most of the papers has good accuracy but we should be considering the performance of the algorithm especially for lots of features because in supervised ML we need to teach the algorithm to classify between Bots and human and also adding new data to the dataset is important hence recognizing new bots with SML algorithm which trained by old data is difficult. Likewise, an unsupervised machine learning algorithm that uses clustering also a consumable process and affecting the performance of the algorithm if defined well.

- Balancing between the number of features used by the algorithm on one hand and the speed and scalability of the algorithm on another hand is one of the future works by the researcher. (regardless of algorithm category)

- Focusing on real-time detection and preparing for big events is more necessary from now on [Orabi M, Mouheb D, Al Aghbari Z and Kamel I, 2020] Because until now the attacker (botmasters) leading this game. They prepare themself for big events around the globe, afterward, the researchers search for what they did.

6-CONCLUSIONS

Using social media platforms increasing day by day, they connected into many areas of our life. Besides their benefit usage, they have many disadvantages. Cybercriminals using those platforms to attacking society in different ways. In this paper, we focus on spreading misinformation to change crowd options about those topics that cybercriminals are interested in or they paid for. At the beginning of our work, we define all terminology related to this research area to giving a clear view and the state of the art to the new researchers. Then we review different papers using a different detection method, datasets, and detection topic. We can conclude our work as its:

Detecting social bots is a new and challenging research topic. Until now most of the workaround detecting bots on Twitter due to the nature of the platform. Also, most of the works focusing on detecting English tweets but also there are a few papers that focusing on bots in other platforms or detecting tweets with the languages. Because of that, one of the future challenges in front of the researchers is detecting bots from other platforms.

The nature of the bots is another challenge due to continuous improvement and making sophisticated bots by the cybercriminal which is detecting them is very difficult. To overcome this issue the researchers should work on new datasets because detecting bots with old datasets does not have any benefit and when applying the algorithm in real life it will be useless. However controlling those black markets that selling bots accounts, bot followers by the authority will be a big step because the bots business have a very good outcome and the cybercriminals have a reason to continue growing and developing a new type of bots.

Detection methods can be categorized into machine learning and non-machine learning. With NML most of the papers focusing on detecting a small amount of account with few detection features by using human effort, mathematical equation, and statistical algorithm. This type of detection will be suitable in case of ignoring algorithm performance due to their cost and the accuracy rate which is not very high most of the time. However, using ML algorithms is more promised. There are different ML algorithms for detecting bots and we can say most of them have a better result than the NML algorithm. But as we discussed above detection method, not the only factor that affects the whole process performance, reliability, new datasets, etc… all together defining the good detection method. As the result, this topic is a new and promising area going forward. there are lots of extra work and improvement that need to be done in the future.

Ref	Detection Category	Method	Result	Weaknesses	Strengths	Social Platform
[Karunakar et al,2020]	Fake account	**SVM** and **Bayesian**	Not clear	• Small dataset and no clue about features • Using old dataset is release in 2014	• Besides accuracy, also work on performance	Twitter
[AMUD. et al, 2020]	COVID-19 fake tweets	**LR, NB, SVM,** Decision Trees (**DT**)	• LR **98.3%** • NB **97.23%** • SVM **98.2%** • DT **98.53%**	• Small dataset, it is just 5k. • The filtering tweets are expensive. they hire three people to do this job	• Make dataset public • Using feature engineering	Twitter
[Gallotti R. et al, 2020]	COVID-19 fake tweets	• Deep neural network (**DNN**)	• DNN **95%**	• Detect the fake news depend on the shared URL but maybe the legitimate user share a fake URL	• Collect tweets from 64 different languages • Using a very big dataset • make dataset public • filtering data depend on geolocated	Twitter
[Sahoo SR. et al, 2020]	Detecting spam account	**Weka** software with different ML classifier	• The result between **98.01%** and **99.51%**	• Create a new dataset depend on categories but in reality there is no categorization when detecting spam bots	• Use Particle Swarm Optimization for selecting the best features in the datasets	Facebook
[Pozzana I. et al, 2020]	Bots behaviour	• **DT**, Extra Trees (**ET**), Random Forest (**RF**), and Adaptive Boosting (**AB**).	• Between **84%** and **97%** AUC with a session. • **83%** AUC without a session	• The work depends on account session It will be difficult to detect those Bots that act like a human (In session).	• Test the algorithm with 25M tweets	Twitter
[Iş H. et al, 2019]	Fake account base on behavioural analyses	• **SVM, KNN,** and **ANN**	• SVM **94.17%** • KNN **96.81%** • ANN **92.33%**	• The collection process is slow, it takes nearly 8 months • Using 10 features to classification maybe those filters affect performance for the big dataset	• They classify their users into three different classes (most of the other papers just classify them into two classes)	Twitter
[Battur R. et al, 2019]	Fake account	• **DT**, **RF**, and Naïve Bayes (**NB**)	• DT **87.85%** • NB **69.76%** • RF **86.19%**	• The number of accounts in the dataset did not mention anywhere in the paper	• Using the public dataset and remove unnecessary columns • They test their algorithm with life data from Twitter	Twitter
[Kudugunta. et al, 2018]	Fake account	• **DNN** based on LSTM	• Single tweet **96%** • Account-level **99%**	• Pre-process data before using DNN. It affects performance for big dataset	• Using both tweet content and metadata to detecting bots with a big dataset	Twitter
[Khalil A. et al, 2017]	Fake followers	• **SVM,** Simple Logistic, and Instance-based classifier	• SVM **60.48%** • SL **90.02%** • IBC **98.74%**	• The dataset is small • Manually verify legitimate users • Not mention anything about performance	• Using different attribute selection during classification	Twitter
[Hanouna et al, 2019]	Political fake account	• **K-means**	The focus on datasets rather than improving the algorithm	• The algorithm will be slow for big datasets because it tests many parameters	• Test the algorithm with 4 different datasets • Test the algorithm with around 1 billion tweets	Twitter
[Dorri A. et al, 2018]	spammer account	emi-supervised ML	• **99.11%**	• The dataset not enough for a real-world bot detector • Due to the algorithm	• Use more than 10 features. • Detecting by four categories aset public	Twitter

Table 3 : Comparison table of existing researches

7-REFRENCES

Richards D, Caldwell PHY, Go H. Impact of social media on the health of children and young people. J Paediatr Child Health 2015;51:1152–7. https://doi.org/10.1111/jpc.13023.

Karunakar E, Pavani VDR, Priya TNI, Sri V, Tiruvalluru K. Ensemble Fake Profile Detection Using Machine Learning (ML) n.d.;10:1071–7. (Volume 10 Issue 3 - 2020)

Ram A, Galav RK. Detection and identification of bogus profiles in online social network using machine learning methods. Eur J Mol Clin Med 2020;7:395–400.

Pulido CM, Villarejo-Carballido B, Redondo-Sama G, Gómez A. COVID-19 infodemic: More retweets for science-based information on coronavirus than for false information. Int Sociol 2020;35:377–92. https://doi.org/10.1177/0268580920914755.

Hanouna S, Neu O, Pardo S, Tsur O, Zahavi H. Sharp power in social media : Patterns from datasets across electoral campaigns. Aust New Zeal J Eur Stud 2019;11:95–111.

Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat Hum Behav 2020. https://doi.org/10.1038/s41562-020-00994-6.

Ferrara E. What Types of Covid-19 Conspiracies Are Populated By Twitter Bots? ArXiv 2020. https://doi.org/10.5210/fm.v25i6.10633.

Khanday AMUD, Khan QR, Rabani ST. Identifying propaganda from online social networks during COVID-19 using machine learning techniques. Int J Inf Technol 2020. https://doi.org/10.1007/s41870-020-00550-5.

Singh L, Bansal S, Bode L, Budak C, Chi G, Kawintiranon K, et al. A first look at COVID-19 information and misinformation sharing on Twitter. ArXiv 2020.

Cresci S. Detecting malicious social bots: Story of a never-ending clash. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2020;12021 LNCS:77–88. https://doi.org/10.1007/978-3-030-39627-5_7.

Okoli C, Schabram K. Working Papers on Information Systems A Guide to Conducting a Systematic Literature Review of Information Systems Research n.d.;10.

Adewole KS, Anuar NB, Kamsin A, Varathan KD, Razak SA. Malicious accounts: Dark of the social networks. J Netw Comput Appl 2017;79:41–67. https://doi.org/10.1016/j.jnca.2016.11.030.

Ferrara BYE, Varol O, Davis C, Menczer F, Flammini A. The Rise of Social Bots n.d. 2014

Grimme C, Preuss M, Adam L, Trautmann H. Social Bots: Human-Like by Means of Human Control? Big Data 2017;5:279–93. https://doi.org/10.1089/big.2017.0044.

Davis CA, Varol O, Ferrara E, Flammini A, Menczer F. BotOrNot: A System to Evaluate Social Bots 2016:14–6. https://doi.org/10.1145/2872518.2889302.

Karataş A, Şahin S. A Review on Social Bot Detection Techniques and Research Directions. Proc Int Secur Cryptol Conf Turkey 2017:156–61.

Stieglitz S, Brachten F, Ross B, Jung AK. Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts. ArXiv 2017:1–11.

Alothali E, Zaki N, Mohamed EA, Alashwal H. Detecting Social Bots on Twitter: A Literature Review. Proc 2018 13th Int Conf Innov Inf Technol IIT 2018 2019:175–80. https://doi.org/10.1109/INNOVATIONS.2018.8605995.

Cardoso Durier da Silva F, Vieira R, Garcia AC. Can Machines Learn to Detect Fake News? A Survey Focused on Social Media. Proc 52nd Hawaii Int Conf Syst Sci 2019;6:2763–70. https://doi.org/10.24251/hicss.2019.332.

Orabi M, Mouheb D, Al Aghbari Z, Kamel I. Detection of Bots in Social Media: A Systematic Review. Inf Process Manag 2020;57:102250. https://doi.org/10.1016/j.ipm.2020.102250.

Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. Commun ACM 2016;59:96–104. https://doi.org/10.1145/2818717.

Orabi M, Mouheb D, Aghbari Z Al, Kamel I. Detection of Bots in Social Media : A Systematic Review. Inf Process Manag 2020;57:102250. https://doi.org/10.1016/j.ipm.2020.102250.

Daya AA, Salahuddin MA, Limam N, Boutaba R. BotChase: Graph-Based Bot Detection Using Machine Learning. IEEE Trans Netw Serv Manag 2020;17:15–29. https://doi.org/10.1109/TNSM.2020.2972405.

Thesis M. Fraudulent social media users detection by a supervised machine learning technique 2018.

Khalil A, Hajjdiab H, Al-Qirim N. Detecting fake followers in twitter: A machine learning approach. Int J Mach Learn Comput 2017;7:198–202. https://doi.org/10.18178/ijmlc.2017.7.6.646.

Kudugunta S, Ferrara E. Deep neural networks for bot detection. Inf Sci (Ny) 2018;467:312–22. https://doi.org/10.1016/j.ins.2018.08.019.

Abu-El-Rub N, Mueen A. BotCamp: Bot-driven interactions in social campaigns. Web Conf 2019 - Proc World Wide Web Conf WWW 2019 2019:2529–35. https://doi.org/10.1145/3308558.3313420.

Dorri A, Abadi M, Dadfarnia M. SocialBotHunter: Botnet detection in twitter-like social networking services using semi-

supervised collective classification. Proc - IEEE 16th Int Conf Dependable, Auton Secur Comput IEEE 16th Int Conf Pervasive Intell Comput IEEE 4th Int Conf Big Data Intell Comput IEEE 3 2018:496–503. https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00097.

Battur R, Yaligar N. Twitter Bot Detection using Machine Learning Algorithms 2019;8:304–7.

Yang KC, Torres-Lugo C, Menczer F. Prevalence of Low-Credibility Information on Twitter During the COVID-19 Outbreak. ArXiv 2020. https://doi.org/10.36190/2020.16.

Pozzana I, Ferrara E. Measuring Bot and Human Behavioral Dynamics. Front Phys 2020;8:1–11. https://doi.org/10.3389/fphy.2020.00125.

Sahoo SR, Gupta BB. Popularity-based detection of malicious content in facebook using machine learning approach. Adv Intell Syst Comput 2020;1045:163–76. https://doi.org/10.1007/978-981-15-0029-9_13.

Chu Z, Gianvecchio S, Wang H, Jajodia S. Detecting automation of Twitter accounts: Are you a human, bot, or cyborg? IEEE Trans Dependable Secur Comput 2012;9:811–24. https://doi.org/10.1109/TDSC.2012.75.

Yang KC, Varol O, Hui PM, Menczer F. Scalable and generalizable social bot detection through data selection. ArXiv 2019. https://doi.org/10.1609/aaai.v34i01.5460.

Iş H, Tuncer T. Interaction-based behavioral analysis of twitter social network accounts. Appl Sci 2019;9. https://doi.org/10.3390/app9204448.