Fake Social Media Account Detection System and Analysis

Anand Kumar
Computer Science and Engineering
Galgotias University
Greater Noida, India
kumaranand09861@gmail.com

Shubham Singh
Computer Science and Engineering
Galgotias University
Greater Noida, India
shubham171515@gmail.com

***Abstract*—**There are a lot of fake accounts on social-media now a days which may of our family, friends, neighbor's, school/class mates etc. This may be harmful for many people as there are many crimes which takes place using fake social- media accounts like terrorism, online fraud, phishing, spam messages, etc. this can sometimes turn into very dangerous situation and can form some hilarious crimes like murder, kidnapping, terrorist threats, robbery, etc. So, to avoid such things, "Fake Social- Media Detection System" can be used to detect all the fake social media of a particular person, track their activeness, can also be known how many times the account was active, active time, etc. The project can be created using python (AIML), HTML, CSS, JS. Python is used for the main algorithm design in AIML. HTML, CSS, JS is for designing the front end of the system. In turn it will show all the fake social media accounts, their activity log in the form of graph, which will show how many times the account was active and when the account was active. At last, i can conclude by saying that fake account detection system is will be very helpful in detecting the fake social media accounts and keeping their logs of activity. So, it will help in deactivating those accounts so, it could not be used for any fraud means.

Keywords—fake, social media, html, CSS, js, aiml, python

I. About Project

There are a lot of fake accounts on social-media now a days which may of our family, friends, neighbor's, school/class mates etc. This may be harmful for many people as there are many crimes which takes place using fake social-media accountslike terrorism, online fraud, phishing, spam messages, etc. This can sometimes turn into very dangerous situation and can form some hilarious crimes like murder, kidnapping, terrorist threats, robbery, etc. So, to avoid such things, "Fake Social-Media Detection System" can be used to detect all the fake social media of a particular person, track their activeness, can also be known how many times the account was active, active time, etc. The project can be created using python (AIML), HTML, CSS, JS. Python is used for the main algorithm design in AIML. HTML, CSS, JS is for designing the front end of the system. In turn it will show all the fake social media accounts, their activity log in the form of graph, which will show how many times the account was active and when the account was active. At last, i can conclude by saying that fake account detection system is will be very helpful in detecting the fake social media accounts and keeping their logs of activity. So, it will help in deactivating those accounts so, it could not be used for any fraud means.

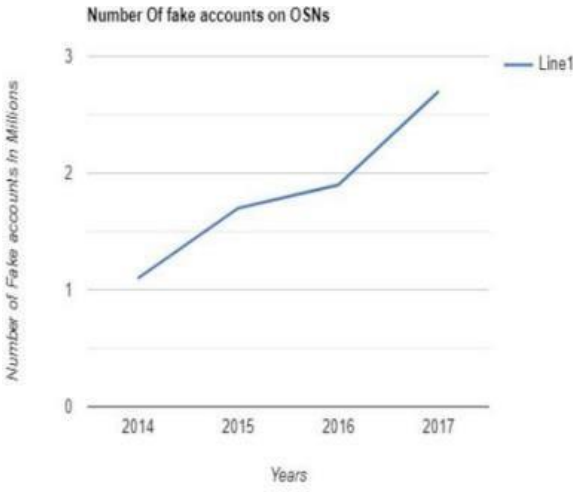

Fig 1. Graph Showing increase in number of Fake accounts over the years

II. Literature Review

In [1][18] authors proposed a new algorithm namely SVM-NN which provide efficient detection for fake Twitter accounts and bots, four feature selection and dimension reduction techniques were applied. Three machine learning classification algorithms were used to decide the target account's identity real or fake, those algorithms were support vector machine (SVM), neural Network (NN), and our newly developed algorithm, SVM-NN, that uses a smaller number of features, while still being able to correctly classify about 98% of the accounts of our training dataset. A similar type of data was also presented by the authors in [2] the authors studied various machine learning algorithms that are used in detecting the fake accounts which are Random Forest, Support Vector Machines and Neural Network. [3] consists an article on human or bot which states the average daily creation of fake accounts related data and the various case historied which states that fake accounts are dangerous in what ways. [4][7] stated that in present world the social media platforms are being used on daily basis and has become an important part of our lives. The number of peoples on social media platforms are incrementing at a greater level for malicious use. There are numerous cases where produced accounts have been effectively distinguished utilizing machine adapting techniques however the amount of research work is very low to recognize counterfeit characters made by people. For bots the ML models used various features to calculate the no. of followers to the no. of friends that an account has on

social media platforms (social media PLATFORMSs). The no. of friends to the no. of followers of any account are easily available in the account profiles and no rights are violated of any accounts. In order to accomplish the task of detecting, identifying and eliminate the fake accounts we establish a forged human account. [5][6] consists of techniques used for detecting spammers on Twitter. Moreover, a taxonomy of the Twitter spam detection approaches is presented that classifies the techniques based on their ability to detect: (i) fake content, (ii) spam based on URL, (iii) spam in trending topics, and (iv) fake users. The presented techniques are also compared based on various features, such as user features, content features, graph features, structure features, and time features. We are hopeful that the presented study will be a useful resource for researchers to find the highlights of recent developments in Twitter spam detection on a single platform. [8] is a study of methods which were used to detect fake accounts that could mislead people. For this purpose, the dataset generated was pre-processed and fake accounts were determined by machine learning algorithms. Decision trees, logistic regression and support vector machines algorithms are used for the detection of fake accounts. Classification performances of these methods are compared and the logistic regression proved to be more successful than the others. The authors in [11] used the study of [9] to do the research which states Machine learning is a fundamental way that enable the computer to have the intelligence; Its application which had been used mainly the method of induction and the synthesis, rather than the deduction has already reached many fields of Artificial Intelligence. Mean while the authors of [10] proposed some thing called IASS, integrated anti-spam system, which adopts machine learning to filter spam in a intelligent, flexible, precise, and self-adaptive way. The methods of linear classification based on optimal separating hyperplane and K-means clustering are used in action recognition layer. The method of improved naive Bayes is used in content analysis layer. The application of machine learning helps improve the performance of IASS. [12] is an investigation done on a possible approach of fake profiles on social media and to mitigate the same. Meanwhile [13] was a research report on a workshop done on fake social media accounts. Authors of [14] published the research on detecting fake accounts on twitter but if we read the research of [15] then we can see that the authors really developed something new. The authors in [15] made the research that have been developed for detecting malicious content, primarily considered the characteristics of user profile. Most of the existing techniques lack comprehensive evaluation. In this work we propose new model using machine learning and NLP (Natural Language Processing) techniques to enhance the accuracy rate in detecting the fake identities in online social networks. We would like to apply this approach to Facebook by extracting the features like Time, date of publication, language, and geo position. [16] is the research of detecting compromised or hacked social media accounts which states by hijacking control of a popular media or business account, attackers can distribute their malicious messages or disseminate fake information to a large user base. The impacts of these incidents range from a tarnished reputation to multi-billion-dollar monetary losses on financial markets. Authors of [17] and [15] have the same study as if we summaries. In [19] the authors introduced a new system called SybilRank. It relies on social graph properties to rank users according to their perceived likelihood of being fake (Sybils). SybilRank is computationally efficient and can scale to graphs with hundreds of millions of nodes, as demonstrated by our Hadoop prototype. We deployed SybilRank in Tuenti's operation center. We found that ~90% of the 200K accounts that SybilRank designated as most likely to be fake, actually warranted suspension. On the other hand, with Tuenti's current user-report-based approach only ~5% of the inspected accounts are indeed fake. The authors of [20] proposed the research to evaluate the proposed model, all steps were implemented on the Twitter dataset. It was found that the Medium Gaussian SVM algorithm predicts fake accounts with high area under the curve=1 and low false positive rate=0.02. An attempt has been made in [21] to use a hybrid model based on machine learning and skin detection algorithms to detect the existence of fake accounts. The experimentation process clearly brought out the strength of the proposed scheme in terms of detecting fake accounts with high accuracy. [22] is a work with research which proposes a feature set, capturing the user social interaction behavior to identify fraud. The problem being solved is one of the characteristics that lead to fraud rather than detecting fraud. In [23] it is the research which is to improve detection of spammers, integrated approach is proposed which combines the advantages of two learning algorithms. Each learning algorithm performance is measured based on spammers detection accuracy and non-spammers detection accuracy. The integrated approach that combines both algorithms performance better than other approaches in terms of overall accuracy and detect non spammers with 99% accuracy with an overall accuracy of 94.1%. the authors in [24] proposed a system using a pattern-matching algorithm on screen-names with an analysis of tweet update times, a highly reliable subset of fake user accounts was identified. Analysis of profile creation times and URLs of these fake accounts revealed distinct behavior of the fake users relative to a ground truth data set. In [25] the authors proposed a system a technique to build extensive datasets of impersonation attacks in current social networks and we gather 16,572 cases of impersonation attacks in the Twitter social network. Our analysis reveals that most identity impersonation attacks are not targeting celebrities or identity theft. Instead, we uncover a new class of impersonation attacks that clone the profiles of ordinary people on Twitter to create real-looking fake identities and use them in malicious activities such as follower fraud. We refer to these as the doppelgänger bot attacks. Our findings show (i) that identity impersonation attacks are much broader than believed and can impact any user, not just celebrities and (ii) that attackers are evolving and create real-looking accounts that are harder to detect by current systems. We also propose and evaluate methods to automatically detect impersonation attacks sooner than they are being detected in today's Twitter social network. At the end I will conclude that all the authors proposed same research and techniques and algorithms and we are also working on the same algorithms.

III. EXISTING SYSTEM

The existing systems use very fewer factors to decide whether an account is fake or not. The factors largely affect the way decision making occurs. When the number of factors is low, the accuracy of the decision making is reduced significantly. There is an exceptional improvement in fake account creation, which is unmatched by the software or application used to detect the fake account. Due to the advancement in creation of fake account, existing methods have turned obsolete.

The most common algorithm used by fake account detection Applications is the Random Forest algorithm. The algorithm has few downsides such as inefficiency to handle the categorical variables which has different number of levels. Also, when there is an increase in the number of trees, the algorithm's time efficiency takes a hit.

IV. PROPOSED SYSTEM

The existing system uses random forest algorithm to identify the fake account. It is efficient when it has the correct inputs and when it has all the inputs. When some of the inputs are missing it becomes difficult for the algorithm to produce the output. To overcome such difficulties in the proposed systems we used a gradient boosting algorithm. Gradient boosting algorithm is like random forest algorithm which uses decision trees as its main component. We also changed the way we find the fake accounts i.e., we introduced new methods to find the account. The methods used are spam commenting, engagement rate and artificial activity. These inputs are used to form decision trees that are used in the gradient boosting algorithm. This algorithm gives us an output even if some inputs are missing. This is the major reason for choosing this algorithm. Due to the use of this algorithm, we were able to get highly accurate results.

A. METHODOLOGY

Proposed system is equipped with various Machine Learning tasks and the architecture followed is as shown below. The proposed system collects the dataset which are pre-processed by providing a framework of algorithms using which we can detect fake profiles in Facebook by comparing the accuracy of three machine learning algorithms and the algorithm with very high efficiency is found for the given dataset.

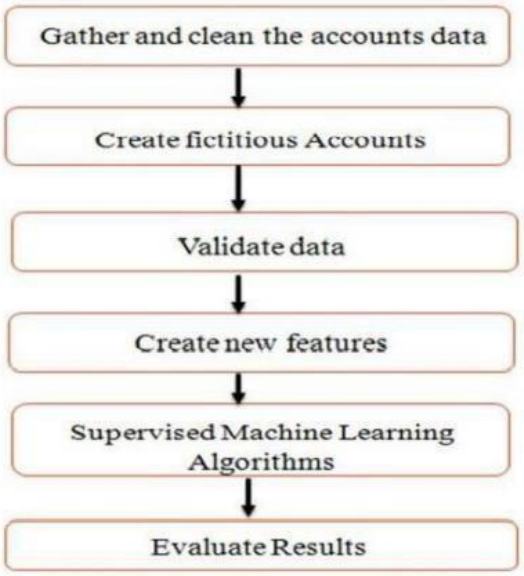

Fig 2. Proposed Methodology

The different ways in which an algorithm can model a problem is based on its interaction with the experience or environment for the model preparation process that helps in choosing the most appropriate algorithm for the given input data in order to get the best result.

1. Support Vector Machine (SVM): Support-vector machines (SVMs, also support-vector networks) are the supervised learning models with associated learning algorithms that analysed data used for classification and regression analysis. For the given labelled training data (supervised learning), the algorithm outputs an optimal hyper plane which categorizes new examples.

2. Neural Networks: A neural network is a network or circuit of neurons, or in a modern sense, an artificial neural network, composed of artificial neurons or nodes. A neural network (NN), in the case of artificial neurons is an interconnected group of natural or artificial neurons that uses a mathematical model for information

3. Random Forest: Random Forest algorithm is a supervised classification algorithm. As the name suggest, this algorithm creates the forest with a number of trees. In general, the more trees in the forest the more robust the forest looks like. In the same way in the random forest classifier, the higher the number of trees in the forest gives the high accuracy results.

4. Explanation of attributes: Attribute importance is a supervised function that identifies and ranks the attributes that are most important in predicting a target attribute. Raw machine learning data contains a mixture of attributes, some of which are relevant to making predictions.

Attribute Name	Description
ID	The unique ID given to the account holder
NAME	The name given to the account holder
SCREEN_NAME	The pseudonym given to the account holder
CREATED_AT	The date when the account is created
FRIENDS_COUNT	The number of friends for the account
STATUSES_COUNT	The number of statuses posted from the account
FOLLOWRES_COUNT	The number of followers for the account
LISTED_COUNT	The number of groups the account belongs to
URL	The URL of the account
TIMEZONE	The time zone of the account holder
UTC_OFFSET	The UTC offset, given TIMEZONE
LOCATION	The location of the account holder
GEO_ENABLED	This field must be true for the current user to attach geographic data when using POST statuses / update.
VERIFIED	When true, indicates that the user has a verified account.

Table 1: Attributes that define a Dataset

A . 1 DETECTION STRATEGY

In Our Research, we define an account as fake when it doesn't meet the minimum engagement rate, have artificial activities or when the account has a history ofspam comments.

A. Web Scraper

Web Scraper is used to extract data from a website. When a user pastes a link of a social media Account, Using outwit hub, a Web scraper tool, we extract necessary pieces of information from thesocial media site. We extract data such as login activity, Total Likes, Total Comments, Number of posts, Number of followings and Number of followers.

B. Calculation of Engagement Rate

An engagement rate is a metric that measures the level of engagement of a Post or Story received on social media. It is the percentage by which the audience interact with a post. By checking the number of interactions with the number of followers we can evaluate the engagement rate. Interactions can be of likes, comments, and shares. Most Fake accounts will boast of 1000s of followers and a very minimum number of likes. Since the engagement rate is relatively calculated, comparisons between popular accounts and semi- popular accounts are comparatively easy. This metric is one of the most vital ones because lesser audience engagement signifies that the account is fake.

$$\text{Engagement rate percentage} = \left(\frac{\text{Total number of interactions}}{\text{Total number of followers}} \right) \times 100$$

Fig 3. Engagement rate calculation

C. Spam Comments

BOT comments are always known to be very Genericand often lack Substance. At this stage, the comments made from the account will be gone through in a detailed manner. Total number of comments by the user made since the creation of the account will be compared with average comments of users in that particular OSN's. If there is a big difference the account may be considered fake. Commenting links will lead to the account being termed as Fake account. Same or Similar type of comments will also be considered as spam comments.

D. Detection of Fake Accounts

In this step, we combine all the data we extracted from the website. In this paper we mainly focus on engagement rate, artificial activity and spam comments. The data collected using web scraper is used to compute the values for the factors mentioned above. Using these factors different decision tress is formed. Using gradient boosting algorithm and with the formed decision trees fake accounts are detected.

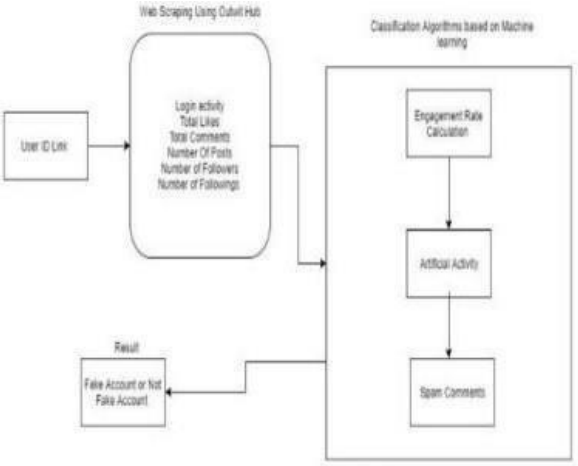

Fig 4. Architecture diagram

B. Algorithm

Input: training set $\{(x_i, y_i)\}_{i=1}^{n}$ a differentiable loss function L(y,F(x)), number of iterations *M*.

Algorithm:

I. Initialize model with a constant value:
$F_0(x) = argmin \sum_{i=1}^{n} L(y_i, \gamma)$.

II. For $m = 1$ to M:

1. Compute so-called *pseudo-residuals*:
$$r_{im} = -\left[\frac{\partial(L(y_i, F(x_i))}{\partial(F(x_i)}\right]_{F(x)=F_{m-1}(x)}$$
For i =1, . . . , n.
2. Fit a base learner (e.g.tree) $h_m(x)$ to pseudo-residuals, i.e. train it using the training set $\{(x_i, y_i)\}_{i=1}^{n}$.
3. Compute multiplier γ_m by solving the following one-dimensional optimization problem:
$$\gamma_m = argmin \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_{i)})).$$
4. Update the model:
$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x).$$

III. Output $F_m(x)$.

$$a + b = \gamma$$

V. EXPERIMENTAL RESULT AND DISCUSSION

1. **Performance of model using Random Forest Algorithm:** The random forest is a model made up

of many decision trees. When training the model using Random Forest algorithm, each tree in a random forest learns from a random sample of the data points and the samples drawn with replacement are known as bootstrapping in which some samples will be used multiple times in a single tree.

2. **Performance of model using Support Vector Machine Algorithm:** In many supervised learning tasks, labeling instances to create a training set is time consuming and costly; thus, finding ways to minimize the number of labeled instances is beneficial. The Support Vector Machine algorithm is used to minimize the instances by improving efficiency. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. We then perform the detection of fake accounts through classification technique by finding the hyper-plane that differentiate the two classes very well (look at the below snapshot).
3. **Performance of model using Neural Networks Algorithm:** Neural networks (NNs) can be defined as "The algorithms in machine learning are implemented by using the structure of neural networks. These neural networks model the data using artificial neurons. Neural networks thus mimic the functioning of the brain." The 'thinking' or processing that a brain carries out is the result of these neural networks in action. The Neural networks algorithm tries to improve the performance of the model by using smart computational methods to create new and better performing types of prediction and detection model.

VI. CONCLUSION

The existing system uses random forest algorithm to identify the fake account. It is efficient when it has the correct inputs and when it has all the inputs. When some of the inputs are missing it becomes difficult for the algorithm to produce the output. To overcome such difficulties in the proposed systems we used a gradient boosting algorithm. Gradient boosting algorithm is like random forest algorithm which uses decision trees as its main component. We also changed the way we find the fake accounts i.e., we introduced new methods to find the account. The methods used are spam commenting, engagement rate and artificial activity. These inputs are used to form decision trees that are used in the gradient boosting algorithm. This algorithm gives us an output even if some inputs are missing. This is the major reason for choosing this algorithm. Due to the use of this algorithm, we were able to get highly accurate results.

REFERENCES

[1] S. Khaled, N. El-Tazi and H. M. O. Mokhtar, "Detecting Fake Accounts on Social Media," 2018 IEEE International Conference on Big Data (Big Data), Seattle,WA, USA, 2018, pp. 3672- 3681.

[2] Rao, K. Sreenivasa, N. Swapna, and P. Praveen Kumar. "Educational data mining for student placement prediction using machine learning algorithms." Int. J. Eng. Technol. Sci 7.1.2 (2018): 43-46.

[3] (2018) Human or 'bot'? doubts over italian comic beppegrillo's twitter followers. Internet draft. [Online]. Avail-able: https://www.telegraph.co.uk/technology/twitter/9421072/Human-or-bot-Doubts-over-Italian-comic-Beppe-Grillos-Twitter-followers.html

[4] N. Singh, T. Sharma, A. Thakral and T. Choudhury, "Detection of Fake Profile in Online Social Networks Using Machine Learning," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), Paris, 2018, pp. 231-234.

[5] Sreeram Gutha, Bandrapalli Deevena Raju and K Sreenivasa Rao "Detecting Fake Account On Social Media Using Machine Learning Algorithms" 2020 International Journal of Control and Automation 13(1s(2020)):95-100.

[6] FAIZA MASOOD1, GHANA AMMAD1, AHMAD ALMOGREN 2, (Senior Member, IEEE), ASSAD ABBAS 1, HASAN ALI KHATTAK 1, (Senior Member, IEEE), IKRAM UD DIN 3, (Senior Member, IEEE), MOHSEN GUIZANI 4, (Fellow, IEEE), AND MANSOUR ZUAIR5, "Spammer Detection and Fake User Identification on Social Networks", IEEE Access, 2019.

[7] "Detection of fake profile in online social networks using Machine Learning" Naman singh, Tushar sharma,Abha Thakral, Tanupriya Choudhury.

[8] "Detection of Fake Twitter accounts with Machine Learning Algorithms" Ilhan aydin, Mehmet sevi, Mehmetumut salur.

[9] "a new heuristic of the decision tree induction" ning li, li zhao, ai-xia chen, qing-wu meng, guo-fang zhang.

[10] "statistical machine learning used in integrated anti-spam system" peng-fei zhang, yu-jie su, cong wang

[11] " a study and application on machine learning of artificial intellligence" ming xue, changjun zhu.

[12] Conti, M., Poovendran, R., Secchiero, M., 2012. FakeBook: Detecting Fake Profiles in On-Line Social Networks, in: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), ASONAM '12. IEEE Computer Society, Washington, DC, USA, pp. 1071– 1078. doi:10.1109/ASONAM.2012.185.

[13] Douceur, J.R., 2002. The Sybil Attack, in: Revised Papers from the First International Workshop on Peer- to-Peer Systems, IPTPS '01. Springer-Verlag, London, UK, UK, pp. 251–260.

[14] "Twitter fake account detection", Buket Ersahin, Ozlem Aktas, Deniz kilinc, Ceyhun Akyol.

[15] Kedir Lemma Arega, "Social Media Fake Account Detection for Afan Oromo Language using Machine Learning" in New Media and Mass Communication ISSN 2224-3267 (Paper) ISSN 2224-3275 (Online) Vol.90, 2020.

[16] Egele, M., Stringhini, G., Kruegel, C., Vigna, G., 2015. Towards Detecting Compromised Accounts on Social Networks. IEEE Trans. Dependable Secure Comput. PP, 1–1. doi:10.1109/TDSC.2015.2479616.

[17] Srinivas Rao Pulluri1, J. G. (2017). A Comprehensive Model for Detecting Fake Profiles in Online Social Networks. International Journal of Advanced Research in Science and Engineering, 1-10. www.facebook.com. (2019). Fake account . (MAU) on Facebook .

[18] S. P. Maniraj, H. K. (2019). Fake Account Detection using Machine Learning and Data Science. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 583-585.

[19] Pregueiro, Q. C. (n.d.). Aiding the Detection of Fake Accounts in Large Scale Social Online Services. 1-14.

[20] Mohammadreza Mohammadrezaei, 1. M. (August 2018). Identifying Fake Accounts on Social Networks Based on. WILEY HINDAWI, 1-9.

[21] M. Smruthi, N. H. (, February 2019). A Hybrid Scheme for Detecting Fake Accounts in Facebook. International Journal of Recent Technology and Engineering (IJRTE), 213-217.

[22] Kunal Goswami, Y. P. (2017). Impact of reviewer social interaction. Springer Journal of Big Data, 1-19.

[23] Hao, K. (2020, March 4). Hao, Karen Archive Page. Retrieved from https://www.technologyreveiw.com K Subba Reddy, D. E. (2017). An Efficient Methodology to Detect Spam. International Journal of Computer Science and Information Security (IJCSIS),, 151-158.

[24] S. Gurajala, J. S. White, B. Hudson, J. N. Matthews, ―Fake Twitter accounts: Profile characteristics obtained using an activity-based pattern detection approachl, SMSociety '15, July 27 - 29, 2015, Toronto, Canada © 2015 ACM. ISBN 978-1- 4503- 3923-0/15/07

[25] Goga, G. Venkatadri, K. P. Gummadi, ―The Doppelgänger Bot Attack: Exploring Identity Impersonation in Online Social Networks‖, IMC‘15, October 28–30, 2015, Tokyo, Japan, c 2015 ACM. ISBN 978-1-4503-3848-6/15/10.