Detection of Fake and Clone accounts in Twitter using Classification and Distance Measure Algorithms

Sowmya P and Madhumita Chatterjee

***Abstract*—Online Social Network (OSN) is a network hub where people with similar interests or real world relationships interact. As the popularity of OSN is increasing, the security and privacy issues related to it are also rising. Fake and Clone profiles are creating dangerous security problems to social network users. Cloning of user profiles is one serious threat, where already existing user's details are stolen to create duplicate profiles and then it is misused for damaging the identity of original profile owner. They can even launch threats like phishing, stalking, spamming etc. Fake profile is the creation of profile in the name of a person or a company which does not really exist in social media, to carry out malicious activities. In this paper, a detection method has been proposed which can detect Fake and Clone profiles in Twitter. Fake profiles are detected based on set of rules that can effectively classify fake and genuine profiles. For Profile Cloning detection two methods are used. One using Similarity Measures and the other using C4.5 decision tree algorithm. In Similarity Measures, two types of similarities are considered – Similarity of Attributes and Similarity of Network relationships. C4.5 detects clones by building decision tree by taking information gain into consideration. A comparison is made to check how well these two methods help in detecting clone profiles.**

***Index Terms*—Clone, C4.5, Fake, Identity Theft, Online Social Networks, OSN**

I. Introduction

ONLINE Social Networks (OSN) like Facebook, Twitter, LinkedIn, Instagram etc are used by billions of users all around the world to build network connections. The ease and accessibility of social networks have created a new era of networking. OSN users share a lot of information in the network like photos, videos, school name, college name, phone numbers, email address, home address, family relations, bank details, career details etc. This information if put into hands of attackers, the after effects are very severe. Most of the OSN users are unaware of the security threats that exist in the social networks and easily fall prey to these attacks. The risks are more dangerous if the victims are children. In Profile Cloning attack, the profile information of existing users are stolen to create duplicate profiles and these profiles are misused for spoiling the identity of original profile owners[1-6]. There are two types of Profile Cloning namely - Same Site and Cross Site Profile Cloning[1,7-9].

If user credentials are taken from one Network to create a clone profile in same Network then it is called Same Site profile cloning[1,10-12]. In Cross Site profile cloning, attacker takes the user information from one Network to create a duplicate profile in other Network in which the user is not having any account[1,13-15].

As the registration process in social networks have become very simple in order to attract more and more users, the creation of fake profiles are also increasing in an alarming rate. An attacker creates a fake profile in order to connect to a victim to cause malicious activities. And also to spread fake news and spam messages.

The paper organized as below. Section II describes the literature survey. Section III explains the proposed methodology. Section IV discusses the results. At last, Section V concludes the paper with the conclusion.

II. Literature Survey

Today, Fake and Clone profiles have become a very serious threat in social networks. So, a detection method is very much necessary to find these frauds who use people's faith to gather private information and create duplicate profiles. Many authors have worked in this area and have proposed methods to identify these type of profiles in social networks. Some of these methods are discussed below.

Georgios Kontaxis, Iasonas Polakis, Sotiris Ioannidis and Evangelos P Markatos [2] have proposed a prototype to check whether the users have become victim to cloning attack or not. Information is extracted from user profile and a search is made in OSN to find profiles which match to that of user profile and a similarity score is calculated based on commonality of attribute values. If the similarity score is above the threshold value then the particular profile is termed as clone.

Sowmya P is with the Department of Computer Engineering, Pillai College of Engineering, University of Mumbai, Maharashtra, India (e-mail: sowmya@mes.ac.in).

Madhumita Chatterjee is with the Department of Computer Engineering, Pillai HOC College of Engineering and Technology, University of Mumbai, Maharashtra, India (e-mail: mchatterjeee@mes.ac.in).

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

Brodka, Mateusz Sobas and Henric Johnson in their paper [3] have proposed two novel methods for detecting cloned profiles. The first method is based on the similarity of attribute values from original and cloned profiles and the second method is based on the network relationships. A person who doubts that his profile has been cloned will be chosen as a victim. Then treating name as primary key, a search is made for profiles with the same name as that of victim, using query search. Potential clone (Pc) and the Victim profile (Pv) are compared and similarity S is calculated. If S(Pc, Pv) > Threshold, then profile is suspected to be a clone. In the verification step, the user does it manually as he knows which is his original profile and which one is a duplicate.

Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M [4], in their paper have reviewed some of the most relevant existing features and rules (proposed by Academia and Media) for fake Twitter accounts detection. They have used these rules and features to train a set of machine learning classifiers. Then they have come up with Class A classifier which can effectively classify original and fake accounts.

Ahmed El Azab, Amira M Idrees, Mahmoud A Mahmoud, Hesham Hefny [5], have proposed a classification method for detecting fake accounts on Twitter. They have collected some effective features for the detection process from different research and have filtered and weighted them in first stage. Various experiments are conducted to get minimum set of attributes which gives accurate results. From 22 attributes, only seven attributes were selected which can effectively detect fake accounts and have applied these factors on classification techniques. A comparison of the classification techniques based on results are made and the one which provides most accurate result is selected.

III. Proposed System

Fake and clone profiles have become a very serious social threat. As information like phone number, email id, school or college name, company name, location etc are readily exposed in social networks, hackers can easily hack this information to create fake or clone profiles. They then try to cause various attacks like phishing, spamming, cyberbullying etc. They even try to defame the legitimate owner or the organisation. So, a detection method has been proposed which can detect both fake and clone profiles in order to make the social life of the users more secure. The architecture of proposed system is as shown in Fig. 1.

The proposed architecture consists of modules for Fake Profile detection and Clone Profile detection.

A. *Fake Profile Detection*

This module is used to detect fake Twitter profiles. Here fake profiles are detected based on rules that effectively distinguish fake profiles from genuine ones. Some of the rules that are used to detect fake profiles are - usually fake profiles do not have profile name or image. They do not include any description about the account. The geo-enabled field will be false as they do not want to expose their location in tweets. They usually make large number of tweets or sometimes the profiles would not have made any tweets etc. The rules are applied on the profile, for each matching rule, a counter is incremented, if the counter value is greater than pre-defined threshold, then the profile is termed as fake.

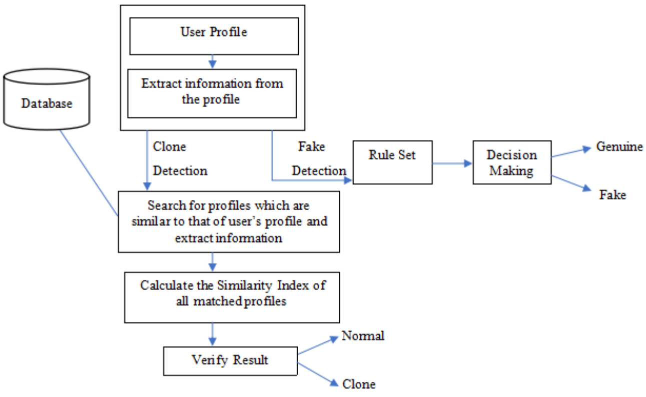
User Profile
Extract information from the profile
Database
Clone Detection
Fake Detection
Rule Set
Decision Making
Genuine
Fake
Search for profiles which are similar to that of user's profile and extract information
Calculate the Similarity Index of all matched profiles
Verify Result
Normal
Clone

Fig. 1. Architecture of proposed system.

B. *Clone Profile Detection using Similarity Measures*

This module detects clones based on Attribute and Network similarity. User profile is taken as input. User identifying information are extracted from the profile. Profiles which are having attributes matching to that of user's profile are searched. Similarity index is calculated and if the similarity index is greater than the threshold, then the profile is termed as clone, else normal[1].

i) Attribute Similarity

Attribute similarity is calculated based on the similarity of attribute values between the profiles. The attributes that are considered for similarity measurement are Name, ScreenName, Language, Location and Time_zone. Two similarity measures are used to measure the similarity between the attributes – Cosine similarity and Levenshtein distance. Cosine similarity is used to find similarity between words and Levenshtein distance is used to find similarity between two sequences.

Cosine similarity formula is given by equation (1)

$$\cos(\theta) \;=\; = \frac{\sum_{i=1}^{n} A_i . B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\;\sqrt{\sum_{i=1}^{n} B_i^2}} \qquad (1)$$

where Ai and Bi are two non-zero vectors [1].

Two vectors have a cosine similarity of 1 if they are with the same orientation; have a similarity of 0 if they are at 90° and -1 if they are diametrically opposed [1]. Levenshtein distance is a similarity measuring metric to find similarity between two sequences.

If two sequences are given, the Levenshtein distance between them is the minimum number of insert, delete or substitution operations required to change one sequence into another. Mathematically, the Levenshtein distance between two strings a, b of length i and j respectively is given by equation (2)

$$
lev_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0 \\ \min \begin{cases} lev_{a,b}(i-1,j) + 1 \\ lev_{a,b}(i,j-1) + 1, \\ lev_{a,b}(i-1,j-1) + 1_{a_i \neq b_j} \end{cases} & \text{otherwise} \end{cases} \qquad (2)
$$

ii) Network Similarity

Network similarity is calculated based on network relationships[1]. Here, Followers_ids attribute is used to find the network similarity between the profiles. Followers_ids gives the list of accounts which follows the user. The clone profile always try to connect to same set of users as that of legitimate owner in order to show that it is genuine one. So, by comparing the Followers_ids of two profiles, we can find whether they are similar with respect to network relationships or not.

Network similarity is calculated as given in equation (3)

NetSim (Pv, Pc) = (|MFF vc|)/(√|Fv|.|Fc|) (3)

where:[3]

NetSim - Network Similarity

Pv - Profile of victim

Pc - Profile of clone

MFFcv - Set of matching Followers_ids of Pv and Pc

Fv - Set of Followers_ids of Pv

Fc - Set of Followers_ids of Pc

If the NetSim value is greater than the threshold, then the profile is treated as clone, else normal[1].

C. Clone Profile Detection using C4.5 algorithm

In this module, C4.5 algorithm is used to detect whether the given profile is a clone or not. C4.5 is a decision tree algorithm used for classification. It builds a decision tree based on given data. At each node of tree, the attribute that most effectively splits the sample sets into subsets is chosen.

The splitting factors used in C4.5 are information gain and entropy. The attribute with highest information gain is chosen to make decision and then it re-curses over the partitioned sub-trees. The information gain as shown in equation (4)

Info (D) = - ∑_(i=1)^n 〖Pi log2 Pi〗 (4)

where, Pi refers to probability.

C4.5 algorithm find the similarity between the attributes by building a tree-like structure. The given profile is compared against the profiles which are already in the database. If the given profile matches with any of the profiles in database, then the profile is termed as clone, else normal.

IV. Experiments and Results

A. Datasets Used

The datasets used in the experiment are collected from MIB projects. It consists of Genuine and Fake Twitter datasets. The Genuine accounts dataset contains accounts of people who came forward to be part of academic study for detecting fake accounts on Twitter and it is mostly a mixture of accounts of researchers, social experts and journalists from Italy, US and other European countries[4]. The fake accounts were purchased from three different Twitter online markets namely fastfollowerz.com, intertwitter.com and twittertechnology.com [4].

B. Evaluation Metrics

In order to evaluate the performance of the system, various evaluation metrics are used based on following four standard indicators

- True Positive (TP): True positives are records that are correctly detected with expected vectors.
- True Negative (TN): True negatives are records correctly detected expected as Neutral.
- False Positive (FP): False positives are records that were detected by the system as expected but actually are listed in the other vectors.
- False Negative (FN): False negatives are records not detected by the system.

The evaluation metrics considered are

1. Accuracy which gives the ratio of number of correct results to the total number of inputs
2. Precision which gives the proportion of positive detection that was actually correct
3. Recall which gives the proportion of actual positives that was detected correctly
4. F1 Score which takes into account both precision and recall to compute the score. F1-score is given by harmonic mean of precision and recall. If F1-score is 1, then it is best value and worst is 0.

TABLE I
PERFORMANCE EVALUATION OF FAKE DETECTION

Total no. of records checked	2200
No. of fake records detected by rule as fake (TP)	990
No. of genuine records detected by rule as fake (FP)	105
No. of fake records detected by rule as genuine (FN)	110
No. of genuine records detected by rule as genuine (TN)	995

TABLE II
PERFORMANCE EVALUATION OF CLONE DETECTION USING SIMILARITY MEASURES

Total no. of records checked	800
No. of normal records detected by system as normal (TN)	769
No. of normal records detected by system as clone (FN)	11
No. of clone records detected by system as normal (FP)	2
No. of clone records detected by system as clone (TP)	18

For detection of fake profiles, a total of 2200 accounts were fed into the system in which 1100 were genuine and 1100 were fake. The rule set worked fine and was able to classify genuine and fake accounts with an accuracy of 90.2% shown in Fig. 2. Table I gives the performance evaluation of fake detection module.

For detection of clone profiles, 780 normal profiles along with 20 artificially generated clone profiles were fed to the modules to check how accurately it detects clone profiles from the given set. The modules worked fine and was able to detect clones with good accuracy. Table II and Table III gives the performance evaluation of Clone Detection using Similarity Measures and using C4.5 respectively.

TABLE III
PERFORMANCE EVALUATION OF CLONE DETECTION USING C4.5

Total no. of records checked	800
No. of normal records detected by system as normal (TN)	765
No. of normal records detected by system as clone (FN)	15
No. of clone records detected by system as normal (FP)	4
No. of clone records detected by system as clone (TP)	16

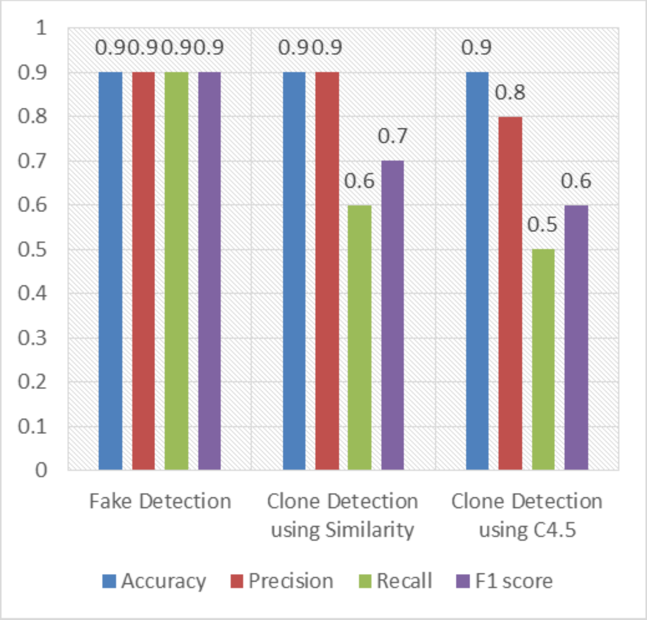

Fig. 2. Performance Evaluation Result.

Results of Table II and Table III shows that 18 out of 20 clones were detected using similarity measures whereas only 16 clones were detected using C4.5 classification algorithm. So it can be concluded that clone detection using similarity measures gives better results as compared to that of using C4.5 classification algorithm.

V. CONCLUSION

Fake and clone profiles have become a very serious problem in online social networks. We hear some or the other threats caused by these profiles in everyday life. So a detection method has been proposed which can find both fake and clone Twitter profiles. For fake detection, a set of rules were used which when applied can classify fake and genuine profiles. Clone detection was carried out using Similarity Measures and C4.5 algorithm and a comparison was made to check the performance. Clone detection using Similarity Measures worked better than C4.5 and was able to detect most of the clones which were fed into the system. In this work we have considered only the profile attributes for fake and clone detection. In future this work can be extended by taking tweets also into consideration by applying some NLP techniques.

REFERENCES

[1] Sowmya P and Madhumita Chatterjee ," Detection of Fake and Cloned Profiles in Online Social Networks", Proceedings 2019: Conference on Technologies for Future Cities (CTFC)

[2] Georgios Kontaxis, Iasonas Polakis, Sotiris Ioannidis and Evangelos P. Markatos, "Detecting Social Network Profile Cloning", 2013

[3] Piotr Bródka, Mateusz Sobas and Henric Johnson, "Profile Cloning Detection in Social Networks", 2014 European Network Intelligence Conference

[4] Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angello Spognardi, Maurizio Tesconi, "Fame for sale: Efficient detection of fake Twitter followers", 2015 Elsevier's journal Decision Support Systems, Volume 80

[5] Ahmed El Azab, Amira M Idrees, Mahmoud A Mahmoud, Hesham Hefny, "Fake Account Detection in Twitter Based on Minimum Weighted Feature set", World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering Vol:10, 2016

[6] M.A.Devmane and N.K.Rana, "Detection and Prevention of Profile Cloning in Online Social Networks", 2014 IEEE International Conference on Recent Advances and Innovations in Engineering

[7] Kiruthiga. S, Kola Sujatha. P and Kannan. A, "Detecting Cloning Attack in Social Networks Using Classification and Clustering Techniques" 2014 International Conference on Recent Trends in Information Technology

[8] Buket Erşahin, Ozlem Aktaş, Deniz Kilinç, Ceyhun Akyol, "Twitter fake account detection", 2017 International Conference on Computer Science and Engineering (UBMK)

[9] Arpitha D, Shrilakshmi Prasad, Prakruthi S, Raghuram A.S, "Python based Machine Learning for Profile Matching", International Research Journal of Engineering and Technology (IRJET), 2018

[10] Olga Peled, Michael Fire, Lior Rokach, Yuval Elovici, "Entity Matching in Online Social Networks", 2013 International Conference on Social Computing

[11] Aditi Gupta and Rishabh Kaushal, "Towards Detecting Fake User Accounts in Facebook", 2017 ISEA Asia Security and Privacy (ISEASP)

[12] Michael Fire, Roy Goldschmidt, Yuval Elovici, "Online Social Networks: Threats and Solutions", JOURNAL OF LATEX CLASS FILES, VOL. 11, NO. 4, DECEMBER 2012, IEEE Communications Surveys & Tutorials

[13] Ashraf Khalil, Hassan Hajjdiab and Nabeel Al-Qirim, "Detecting Fake Followers in Twitter: A Machine Learning Approach" 2017 International Journal of Machine Learning and Computing

[14] Mohammad Reza Khayyambashi and Fatemeh Salehi Rizi, "An approach for detecting profile cloning in online social networks" 2013 International Conference on e-Commerce in Developing Countries: with focus on e-Security

[15] Mauro Conti, Radha Poovendran and Marco Secchiero, "FakeBook: Detecting Fake Profiles in On-line Social Networks", 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining