International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Special Issue: | Apr 2023 www.irjet.net p-ISSN: 2395-0072

International Conference on Trends in Engineering & Technology- 2023 (ICRTET)
Organised by: VSM College of Engineering, Ramachandrapuram

DETECTION OF FAKE AND CLONE ACCOUNTS IN TWITTER BY USING CLASSIFICATION AND DISTANCE MEASURE ALGORITHMS

B.Prasanna Rani[1], P. Devi Mounika[2], V.TejaSri[3], A.S.P.Madhuri[4], A.Kishore[5], K.Ajay[6]

[1]Associate Professor,[2-6]Students B.Tech. Computer Science Engineering, V. S. M. College of Engineering, Ramachandrapuram, A.P, India

---

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 bytaking information gain into consideration. A comparison is made to check how well these two methods help in detecting clone profiles.

Key Words: fake account, clone account, Detection, Algorithims, Existing User, Social Network.

1. INTRODUCTION

Online Social Networks (OSN) like Face book, 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, bankdetails, career details etc. This information if put into hands ofattackers, 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. There are two types of Profile Cloning namely - Same Site and Cross Site Profile Cloning. If user credentials are taken from one Network to create a clone profile in same Network then it is called Same Site profile cloning. 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 isnot having any account. 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 alsoincreasing 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.

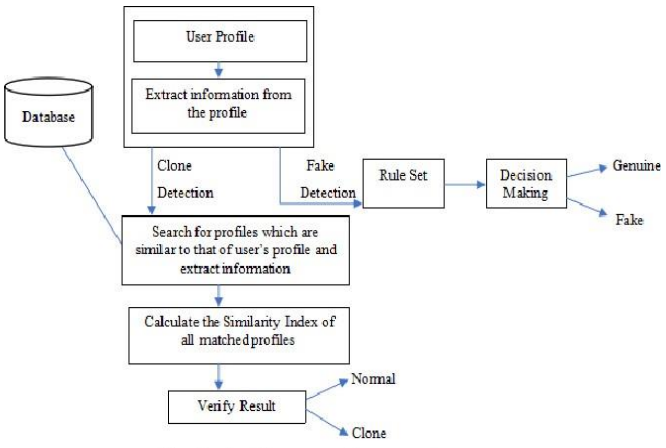

Fig. 1. Architecture of proposed system.

IDENTIFY, RESEARCH AND COLLECT IDEA

In [1],Georgios Kontaxis, Iasonas Polakis, Sotiris Ioannidis and Evangelos P Markatos 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.

In [2], Brodka, Mateusz Sobas and Henric Johnson in their paper 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, 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 classifier which can effectively classify original and fake accounts.

In [3], Chandy and Abraham proposed a random forest classifier in extracting the features for data processing using cloud computing. The extracted features are requesting number, user identification, expiry time, time of arrival and memory requirement. After feature extraction, the predictionof work load is done by using the trained data that has been perceived from the learning stage that allows to learn the details of the extracted features from user's request.

In [4], Ahmed El Azab, Amira M Idrees, Mahmoud A Mahmoud, Hesham Hefny, 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.

2. PROPOSED APPROACH

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 eventry 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 the proposed system. 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.

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 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 Timezone. 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)

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

BIOGRAPHIES

Team Guide ,B.Prasanna Rani is Computer Science and Engineering student in VSM College of Engineering, Ramachandrapuram.

Team Leader, P. Devi Mounika is Computer Science and Engineering student in VSM College of Engineering, Ramachandrapuram.

Team Member, V. Tejasri is Computer Science and Engineering student in VSM College of Engineering, Ramachandrapuram.

Team Member,
A.S.P Madhuri is Computer Science and Engineering student in VSM College of Engineering, Ramachandrapuram.

Team Member,
hore is ComputerScience and
Engineering student in VSM College of Engineering, Ramachandrapuram.

Team Member,
K. Ajay is Computer Science and Engineering student in VSM College of Engineering, Ramachandrapuram.