DETECTION OF FAKE AND CLONE ACCOUNTS IN TWITTER

SHRUTHI M T[1], SUGANDHA M S[2]

[1]Assistant professor, Department of MCA, BIET, Davanagere
[2]Student,4th Semester MCA, Department of MCA, BIET, Davanagere

ABSTRACT

OSN's growing popularity is accompanied with a rise in security and privacy concerns. One significant risk is the creation of duplicate user profiles through the theft of existing user information, which is subsequently exploited to damage the original profile owner's identity. Phony profiles are identified using a set of rules that can discern between authentic and fraudulent ones. There are two methods for identifying Cloning a profile. Similarity measures are used in one, and the C4.5 decision tree technique is used in the other. The two categories of parallels that are considered in Similarity metrics include network and attribute similarity. connections. Clone detection is done using C4.5.

1. INTRODUCTION

What Is a Social Network? "The building and verifying of online social networks for communities of people who share interests and activities, or who are interested in exploring the interests and activities of others, and which necessitates the use of software," constitutes how Wikipedia defines a social network service.

"Websites Social networking sites like Facebook, Mixi, and Myspace are primarily created to enable interaction between users who share interests, attitudes, and activities.

sites are defined in an OCLCreport. What Uses Are Possible for Social Networks Members of an organization can benefit from a variety of social networks. Social networking sites like Facebook, Mixi, and Myspace are primarily created to enable interaction between users who share interests, attitudes, and activities. Professional networks can flourish with the help of social media platforms. Social interaction: Although it may give rise to ethical dilemmas, the passive use of social media platforms can provide valuable business information and feedback on institutional services. Simple information and app access Typical interface: One potential benefit of social networks is their capacity to facilitate communication across boundaries, both social and professional. Users may be

familiar with the interface and functioning of these services since they are frequently used for personal purposes, which lessens the requirement for training and support.

2. LITERATURE REVIEW

Online social network fake and clone profile detection Websites for online social networking are becoming more and more popular every day[1]. Popular websites like Google+, Facebook, LinkedIn, and Twitter have millions of users worldwide.

Due to their widespread use, Online Social Networks (OSN) users face several security and privacy risks, including virus assaults, viral marketing, privacy breaches, structural attacks, and profile cloning. the several survey results as well as an experiment we conducted on Facebook, LinkedIn, and Google+ to identify instances of profile cloning using a methodology we created. We recommended a few

Identifying Cloned Social Network Profiles With millions of users worldwide, social networking is one of the most well-liked online pastimes[2].

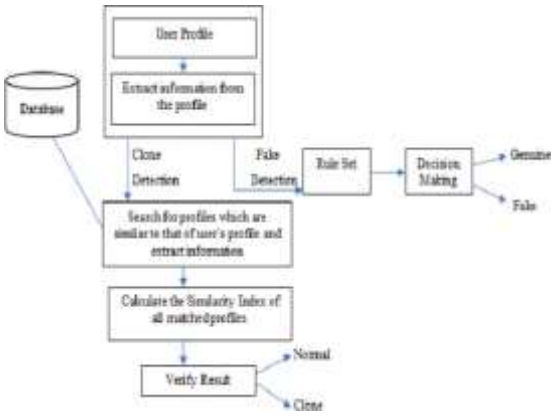
Attackers have the ability to replicate a user's online appearance on one or more social networks, deceiving other users into establishing trustworthy social relationships with the false profile. Through misuse of the implicit trust derived from the idea of relationships in the real world, they can conduct phishing scams, obtain private user data, or negatively impact the legitimacy of the authentic profile.

the social network and beyond, as they might change ideas like influence and popularity in the Twittersphere, having an effect on society, politics, and the economy. We make several contributions in this paper along various aspects.

Next, we establish a baseline dataset comprising authenticated human and fraudulent follower accounts. The scientific community has public access to this baseline dataset.

As a result of utilizing the less expensive features, the new Class A classifier is lightweight, capable of preventing overfitting, and can accurately identify over 95% of the original training set's accounts. Finally, we carry out a sensitivity analysis based on information fusion.

3. METHODOLOGY



Detecting fake and clone accounts on Twitter involves a systematic approach combining data collection, feature extraction, and classification using machine learning techniques. Below is a detailed methodology typically followed for such a task:

Data Collection

API Usage: Utilize the Twitter API to collect user profiles, tweets, follower/following relationships, and other relevant data.

Data Sources: Gather data from other sources if available, such as public datasets or previously identified fake and clone accounts for training purposes.

Data Preprocessing

Cleaning: Remove irrelevant data, handle missing values, and standardize formats (e.g., date and time).

Normalization: Normalize numerical features to make sure they are of a similar size.

Feature Extraction

Profile Features: Extract features from user profiles, such as the number of followers, count, and bio information.

Content Features: Analyze tweets for linguistic patterns, sentiment analysis, posting frequency, and content similarity.

Network Features: Examine the social graph, including followers, followings, and interaction patterns. Use network analysis metrics like centrality and community detection.

Behavioral Features: Track user behaviors, such as login times, tweet times, retweet ratios, and hashtag usage.

Machine learning algorithms

Rule-Based Classification for Fake Profiles:

Relies on predefined rules based on profile attributes and behaviors.Effective for detecting obvious fake profiles using domain-specific knowledge.

Similarity Measure Algorithms for Clone Profiles:

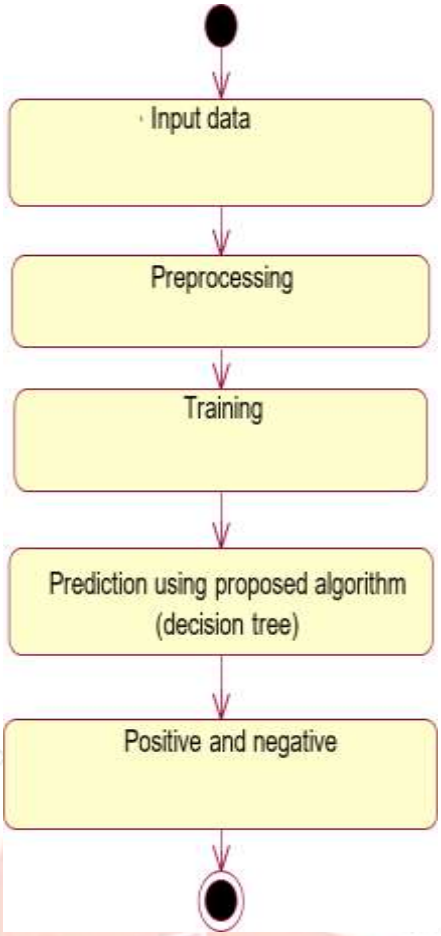
Uses similarity measures to compare profile attributes and network relationships.

C4.5 Decision Tree for Clone Profiles:

Builds a decision tree based on information gain to classify profiles.

Provides a structured and interpretable model for detecting clone profiles.

By combining these approaches, the proposed method leverages both rule-based heuristics and machine learning techniques to effectively detect fake and clone profiles on Twitter, addressing.



4. RESULT

An output is deemed high quality if it successfully communicates the information and satisfies the needs of the end user. Any system's outputs are the means by which processing results are conveyed to users and other systems. How data is transported for both the hard copy output and immediate usage is determined by the output design. For the user, it is the most significant and easily available source of information. An efficient and well-planned output design facilitates user decision- making and enhances system integration.

1. Decide how information will be presented.

2. Generate documents, reports, or other formats that the system generates that contain data. The output form of an information. Disseminate details concerning previous actions, the situation as it is, or future expectations.

- Indicate significant occasions, chances, issues, or cautions.

- Set off an event.

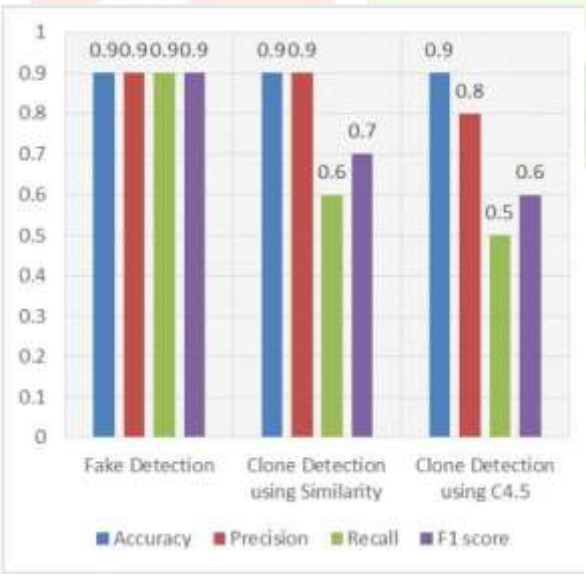
• Verify an activity.

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

Tables 1 demonstrate that similarity metrics were used

CONCLUSION



In online social networks, fake and duplicate profiles have grown to be a serious problem. Every day, we become more conscious of the various threats that these profiles present. As a result, a detection method that might detect fake and cloned Twitter profiles has been introduced. For fake detection, a set of rules that, when put into practice, may discern between authentic and fraudulent profiles were utilized. The C4.5 algorithm and Similarity Measures were used to detect clones, and the results were compared to assess performance. Similarity metrics-based clone identification performed better than C4.5 and identified the majority of the clones given into the system. Only the profile attributes for clone and fake detection have been taken into consideration in this work.

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