International Journal of Research in Engineering, Science and Management
Volume 6, Issue 9, September 2023
https://www.ijresm.com | ISSN (Online): 2581-5792

Spammer Identification Using Novel Machine Learning Algorithm in Industrial Mobile Cloud Computing

F. N. Afrah Fathin[1*], A. Rizwana[2], J. Afrah[3], Alagesan[4]
[1,2,3]Student, Department of Information Technology, Aalim Muhammed Salegh College of Engineering, Avadi, India
[4]Assistant Professor, Department of Information Technology, Aalim Muhammed Salegh College of Engineering, Avadi, India

***Abstract*: Social media networks are gaining more traction in a variety of industries throughout the world and have emerged as one of the most used and well-liked digital marketing platforms for monitoring societal trends and better understanding consumer preferences. False social profiles are multiplying quickly and disseminating false information and news through this expanding channel. Online impersonation and fraudulent accounts are a major problem on the social media network, which is a vital part of our lives. Intruders frequently utilize false personas to engage in illegal activity on Online Social Networks (OSN), including injuring others, identity theft, and invasions of privacy. Consequently, determining whether an account is real or false is one of the major issues with OSN. Many classification algorithms, including the Advanced Support Vector Machine Learning algorithm and deep machine learning networks, are proposed in this paper. This research offers a Spammer Identification strategy based on Gaussian Mixture Model (SIGMM) for industrial mobile networks in order to solve this issue. It delivers accurate spammer identification without relying on fluid and erratic interactions. SIGMM integrates the data presentation with the model generation process, which assigns a class to each user node.**

***Keywords*: accounts, fake identities, social media, data science, friends, followers, fake profiles.**

1. Introduction

Spammers are difficult to eradicate. Spam emails continue to circulate online despite the fact that email providers like Google mail, Microsoft, Facebook, and others have been effectively detecting them for years. According to these services, email spam accounts for up to 90 to 95 percent of all email exchanges. Companies are unable to halt spammers even after successful spam detection, which guarantees the financial gain spammers receive when they trick a user into clicking on a spam link. With the development of online social networks, the severity of the threat posed by spam has grown. Twitter is one of the most well-known online social networks that has been severely impacted by spam. Twitter spam is more dangerous because it targets Twitter's trending topics and is therefore a little simpler to access, thanks in part to the hash-tag operator. Users of Twitter, an online social network (OSN), can post whatever they wish, including news, commentary, and other material.

The fact that debates may be carried on a range of subjects and in a variety of moods makes twitter a more accessible and lucrative target for spammers. Twitter users come from all walks of life, including the general public, celebrities, politicians, educators, and even customers and clients. They are of various ages, but the 55 to 64 age group utilizes twitter the most frequently. About 60% of Twitter users use their phones to access the service. With 288 million active users per month, Twitter is a rapidly expanding social networking service. There are over 400 million tweets written every day, with each user posting an average of 208 tweets each account. There are six basic types of spam, and each one affects Internet users in a different way, such as (1) spam in e-mails, (2) spam in comments, (3) spam in instant messengers, (4) spam in unsolicited text messages, and (5) spam in social networking sites. Various anti-spam techniques have been presented in cutting-edge research to decrease or eradicate spam. [2] Anti-spam strategies were divided into three groups by Heymann et al: (1) prevention-based, (2) detection-based, and (3) demotion-based. There are several anti-spam tactics, such as content-based, link-based, graph analysis, and clocking schemes, however despite the variety of anti-spam strategies, there are still a number of unresolved issues that need to be resolved. In the study, some of these are emphasized.

2. Literature Survey

An architecture designed specifically for brief spam content on SMPs like Twitter was put out by Gauri Jain et al. [1].

In contrast, earlier extended spam emails were detected and deleted. Using the fundamental settings of results using CNN algorithms demonstrated that the suggested model was effective when using Twitter and SMS text datasets. A model for spam detection has been put forth by Thayakorn Dangkesee et al. [3] by exploiting spam word lists utilizing a billboard URL-based security tool. The data have been analyzed using the Naive Bays algorithm using both general and specialized data types. The spam detector is operating more effectively than normal as a result. One can demonstrate that their approach matches the outcomes of experiments.

In order to categorize unlabeled tweets using a different

*Corresponding author: afrahfathin444@gmail.com

dataset, Rutuja Katpatal [4] has created an extra input training dataset. The author has put up a plan that modifies training data sets. By discarding too-old samples after a set amount of time, unexpected information has been removed, conserving space.

3. Major Issue for Spam Detection

Due to the following negative effects, spam identification is becoming a significant difficulty for service providers [2].

- Spam degrades search result quality and deprives legitimate websites of money.
- Spam has an economic impact since it increases online traffic because websites with high rankings receive a lot of free advertising.
- Because there is no cost involved in switching from one search provider to another, it erodes users' trust in search engine providers, which is a particularly noticeable problem.
- Spam websites are platforms for the distribution of malware, adult content, and fishing attacks.
- Spam compels a search engine provider to squander a sizable amount of compute and storage capacity.
- Finding the best tags for the content at hand and getting rid of the spam tags are two significant challenges in tagging.

4. Spam Detection Techniques Content Based Analyses

Methods for analyzing aspects of material like word count, linguistic models, and content duplication.

Web spam pages may be recognized via statistical analysis, according to a theory put forth by Fetterly et al. Spam pages contain some peculiar characteristics, including: (1) URLs with unusually high numbers of dots, dashes, and digits; (2) very little word count variation across spam pages hosted on the same server; and (3) content that changes very quickly. In order to identify script-generated spam pages, T. Urvoy et al. created features based on HTML page structure [5]. In this preprocessing, all content is removed and only the page's layout is taken into account. They used the finger printing technique with later clustering to identify collections of spam pages with similar structural features [5]. An area of research on language modeling for spam identification was suggested by G. Mishne et al.

A. Group Analyzes

A different team examines link-based data, such as neighbor graph connectivity. based on the de-weighting of questionable nodes and linkages after they have been identified. Each node's link-based attributes are extracted, and various machine learning algorithms are used to find spam. Methods of graph regularization for the identification of spam. For all web pages, global relevance scores are calculated in this link using information.

Page Pi's trust in Page PJ is evidenced by the link between the two pages. The algorithm is based on the repeated improvement concept, where the true score is calculated as the point at which an iterative updating process converges [5].

These algorithms carry out standard classification or clustering analysis and represent pages as feature vectors. to do website categorization based on functionality, research link-based characteristics. They make the supposition that websites with comparable structural patterns, such as average page levels or the quantity of outbound connections on each leaf page, play similar functions on the internet. For instance, spam sites have a specific topology designed to maximize Page Rank boost and show high content duplication, while web directories typically consist of pages with high ratios of out links to in links, form a tree-like structure, and the number of out links increases with the depth of pages. Overall, each website is represented as a vector of 16 connection attributes, and cosine is used as a similarity metric to achieve clustering [4].

B. Algorithms that Exploit Click Stream

Data and information about user activity, HTTP session metadata, and query popularity. Since click spam attempts to introduce "malicious noise" into a query log with the objective of contaminating the data used to construct the ranking function, the majority of countermeasures focus on ways to make learning algorithms resilient to this noise. The understanding of the economic aspects supporting the spammers' ecology serves as the driving force behind additional anti-click fraud techniques. The proposed solution to stop click spam is interesting. To stop click fraud manipulation, the author advises utilizing individualized ranking functions because they are more reliable [3].

5. Challenges in Spam Detection Techniques

Support A binary classification system called the Vector Machine determines the greatest hyperplane of separation between two classes. It is a supervised learning algorithm that classifies fresh cases, splits two classes rather effectively, and provides adequate training examples. Because of its mathematical basis in statistical learning theory, it offers a general approach to machine learning problems [10]. The weighted sum of SVs, which make up only a portion of the training input, is how SVM generate their answer. When the number of dimensions exceeds the number of provided samples, it works well considering dynamics of trust would lead to better modeling in real world application.

- Most of the existing approaches based on text information assuming monolingual environment.
- However social network services are used by people from various countries, so various languages simultaneously appears in tags and comment. In such cases some text information may be regarded as wrong or considered as spam due to language spam. Therefore, incorporating multilinguism in trust modeling would solve this problem.
- It is observed that interaction across social network become popular. For e.g., users can use their Face book accounts to log in some other social network services. Thus, future challenge is to investigate how trust model across domains can be effectively connected and shared.

- Trust modeling most of the current techniques for noise and spam reduction focus only on textual tag processing and user profile analysis while audio and visual content features of multimedia content can also provide useful information about the relevance of the content and content tag relation. In future challenge could be to combine multimedia content analysis with the conventional tag processing and user profile analysis.

6. Support Vector Machine Analysis

Support A binary classification system called the Vector Machine determines the greatest hyperplane of separation between two classes. It is a supervised learning algorithm that classifies fresh cases, splits two classes rather effectively, and provides adequate training examples. Because of its mathematical basis in statistical learning theory, it offers a general approach to machine learning problems [10]. The weighted sum of SVs, which make up only a portion of the training input, is how SVM generate their answer. When the number of dimensions exceeds the number of provided samples, it works well.

A. Unsupervised Machine Learning

According to Miller et al. [6], supervised machine learning models need a label present in the corpus to forecast the desired result. In unsupervised machine learning, the data are not labeled and are sorted according to how similar they are. Searching via the phony account class is not realistic. A one class support vector machine is often trained using data from the minority class.

B. Advanced Machine Learning

Garadi et al. [7] assess the viability of employing freely available and engineered features for the successful detection of fraudulent identities manufactured by bots or computers using machine learning algorithms. It is done by taking into account how identical qualities can act as a catalyst for exposing human identity fraud on online social networks.

A reinforcement proof-of-concept model that rewards itself for effectively identifying bots was presented by Venakatesan et al. For reinforcement machine learning models to adapt and advance, the environment must provide feedback. There are no easy ways to access this on social media.

C. Filtering and Identification

A sender will be added to a blacklist whenever a new threat has been found and verified [9]. Similar approaches to combating spam have been put forth on Twitter, where known harmful URL content has been blacklisted and known bots have been quarantined.

D. Random Forest Method

Random Forest is a flexible method that may be used for both regression and classification tasks [8]. Its hyper parameters are almost identical to those of a decision tree or a bagging classifier. It produces a wide variety of trees. Identity deception will be predicted using the best result. Each classification result reflects a separate branch of a tree.

E. Suspicious Behavior

Jiang et al.'s research study used Catch Sync to identify suspicious Twitter activity based on synchronized and unusual user activity. They were able to demonstrate that their strategy produced a high efficiency of false account detection on Twitter.

Table 1
Attributes used in previous study

ATTRIBUTE	DESCRIPTION
HAS_IMAGE	Whether the account has a profile image
HAS_NAME	Whether the account has a profile name
HAS_PROFILE	Whether the account has a profile description
PROFILE_HAS_URL	Whether the profile description contains URL
FF_RATIO	Friends-to-Followers ratio

7. Proposed Work

This paper proposes the detection process starts with the selection of the profile that needs to be tested. After selection of the profile the suitable attributes i.e., features are selected on which the classification algorithm is being implemented, the attributes extracted is passed to the trained classifier

Classification is the process of learning a target function f that maps each record consulting of set of attributes to one of the predefined classes models from an input data set. Classification technique is a approach of building classification models from an input data set. This technique uses a learning algorithm to identify a model that best fits the relationship between the attribute set and class label of the training set.

The model generated by the learning algorithm should both fit the input data correctly and correctly predict the class labels of the learning algorithm is to build the model with good generality capability. Different steps are executed to classify an account as fake or genuine profiles. They are: Data set of both fake and genuine profiles with various attributes like number of friends, followers, status count. Dataset is divided into training and testing data. Classification algorithm are trained using training dataset and testing data set is used to determine the efficiency of algorithm. From the dataset used 80% of both (real and fake) are used to prepare a training data set and 20% of both profiles are used to prepare a testing dataset.

Table 2
Features extracted

S.No.	Features
1	Number of friends
2	Number of
3	Followers
4	Favorite Count
5	Languages
6	Known Status Count

After selection of attributes, the dataset of profiles that are already classified as fake or genuine are needed for the training purpose of the classification algorithm. We have used a publicly

available dataset of 1337 fake users and 1481 genuine users consisting of various attributes including listed count, status count, number of friends, followers count, favourites, languages known, sex code. Classification is the process of categorizing a data object into categories called classes based upon features/attributes associated with that data object. Classification uses a classifier, an algorithm that processes the attributes of each data object and outputs a class based upon this information. In this project, we use Support Vector Machine as a classifier.

Support Vector Machine is an elegant and robust technique for classification on a large data set not unlike the data sets of Social Network with several millions of profiles. Algorithm used for classification are Support Vector Machine, Random Forest and Deep Neural Networks.

Confusion Matrix is a technique for describing the performance of a classification algorithm. Confusion Matrix is used to give you a better idea of what your classification model is getting right and what types of errors it is making. All the algorithm results are plotted in confusion matrix to know where the error has occurred.

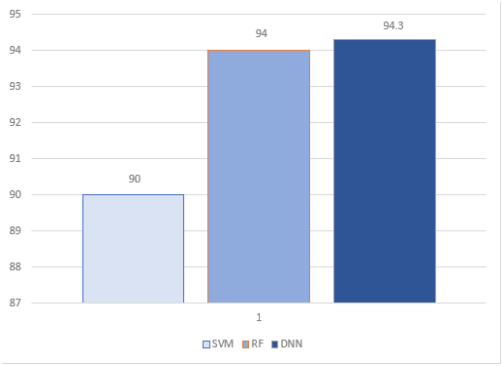

Fig. 1. Results

8. Modules

A. *Data Description*

Our data contain the following contents, user's ID, the relationship with other users, the time-stamped post record, and the activity in the past three months. From the post record, we calculate the frequency of posting and proportion of posts with URL or @, and the average similarity among the user's posts. The activity reflects whether the account is normal or not. It indicates the frequency of following others, which is necessary because spammers tend to follow others all the time.

B. *Feature Scaling*

The data we obtained have the following two constraints. First, the labeled data are far fewer than the unlabeled data which severely decreases the precision of training. Second, there is large data noise that may cause incorrect factors in the parameters of the model. Data points that do not belong to any class are defined as data noise. The values of some data may greatly differ from the mean of samples. SIGMM reduces data noise by calculating the similarity among users to increase the precision of training.

C. *Feature Grouping*

The multidimensional feature is divided into three parts to indicate three user perspectives, which are basic features, content features and network features. The basic features include the number of fans, following users, posts, and the frequency of following. Previous studies show that spammers tend to follow a large number of users, and their fans are rare. The proportion of fans to following is particularly low. These characteristics reflect whether the user is normal or not. Spammers' frequency in following others is further higher than that of normal users.

The content features mainly reflect the characteristics of the information sent by a user in the most recent three months. The user's activity can be analyzed by the content features.

The network feature mainly describes user characteristics in an industrial mobile network. The number and proportion of following each other represent the degree of intimacy between users. Spammers usually follow a large number of normal users to attack. Therefore, their proportion of following each other is lower than that of normal users.

9. SIGMM Model

The SIGMM mechanism fits the behavior data of normal users and spammers, where the behavior data of normal users and spammers are mixed random sampling. The SIGMM mechanism learns the parameters of the two distributions (normal users and spammers) to obtain the classification model. The SIGMM instrument fits the conduct information of ordinary clients and spammers, where the conduct information of typical clients and spammers are blended arbitrary examining. The SIGMM component learns the parameters of the two disseminations (ordinary clients and spammers) to acquire the characterization. A boss among the most generally utilized comfortable get-together tallies is the Fuzzy C recommends gathering (FCM) Algorithm. The FCM computation attempts to section a restricted assembling of parts into a gathering of c feathery groups with respect to some given principle. Reinforce learning estimations are absolutely reasonable for comprehending how to control an expert. They proposed an approach of spam detection in blogs by comparing the language models for blog comments and pages linked with these comments. They use KL divergence as a measure of discrepancy [7]. In other work by M. Sydow linguistic features were analyzed for web spam detection by considering lexical validity and content diversity, syntactical diversity and entropy, usage of active and passive voices and various other NLP features.[9] this paper proposes a spammer identification scheme based on Gaussian mixture model (SIGMM) that utilizes machine learning for industrial mobile networks. It provides intelligent identification of spammers without relying on flexible and unreliable relationships. SIGMM combines the presentation of data, where each user node is classified into one class in the construction process of the model. We validate the SIGMM by comparing it with the reality mining algorithm and hybrid fuzzy c-means (FCM) clustering algorithm using a mobile network dataset from a cloud server. Simulation results

show that SIGMM outperforms these previous schemes in terms of recall, precision, and time complexity. in light of gaussian mixture model, it proposes an attestation procedure named the SIGMM appear for depiction without depending upon clients' scrappy affiliations. [4] SIGMM can name information typically, which builds the precision of model by extending the status set. it denotes a ton of unlabeled data subject to two or three named data and deals with the issue of the disproportion between named data and unlabeled data. they use a cutting-edge versatile framework dataset from a cloud server to perform reenactments.

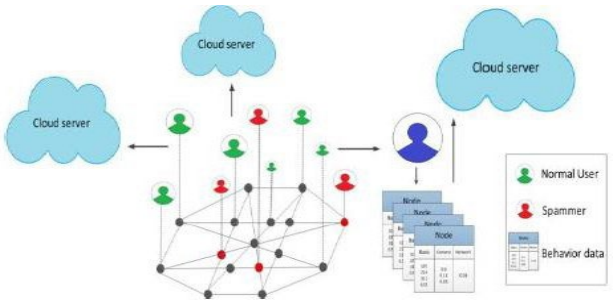

Fig. 2. System architecture

10. Conclusion

In this project, we have presented a machine learning pipeline for detecting fake accounts in online social networks. Rather than making a prediction for each individual account, our system classifies clusters of fake accounts to determine whether they have been created by the same actor. Our evaluation on both in-sample and out- of-sample data showed strong performance, and we have used the system in production to find and restrict more than 250,000 accounts. In this work we evaluated our framework on clusters created by simple grouping based on registration date and registration IP address. In future work we expect to run our model on clustering that are created by grouping on other features, such as ISP and other time periods, such as week or month. Another promising line of research is to use more sophisticated clustering algorithms such as k-means or hierarchical clustering. While these approaches may be fruitful, they present obstacles to operating at scale: k-means may require too many clusters (i.e., too large a value of k) to produce useful results and clustering of data may be too intensive for classifying millions of accounts in Online Social Network.

In order to solve the malicious attack problem in industrial mobile networks and reduce the computational complexity of using large cloud server datasets, this paper proposes SIGMM, a spammer identification model based on the Gaussian Mixture Model. We extract features related to labels from originally labeled data in a given dataset containing both labeled and labeled data, and visualize the data to add labels to then labeled data. According to the characteristics of data presentation, each user data belongs to one distribution. Multidimensional features are divided into three groups, and SIGMM separates the two distributions based on these features.

References

[1] Estee Van Der Walt and Jan Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," in *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 1, no. 1, pp. 61–71 March 2018.

[2] Loredana Caruccio, Domenico Desiato and Giuseppe Polese, "Fake Account Identification in Social Networks," in *IEEE International Conference on Big Data.*, vol. 9, no. 6, pp. 811–824, 2018.

[3] SarahKhaled, Neamat El-Tazi and Hoda M. O. Mokhtar, "Detecting Fake Accounts on Social Media," in *IEEE International Conference on Big Data*, vol. 6, pp. 101-110, 2018.

[4] Suyash Somani and Somya Jain, "Resolving Identities on Facebook and Twitter," in *Tenth International Conference on Contemporary Computing (IC3)*, 10-12 August 2017.

[5] Francesco Buccafurri, Gianluca Lax, Denis Migdal, Serena Nicolazzo, Antonino Nocera and Christophe Rosenberger, "Contrasting False Identities in Social Networks by Trust Chains and Biometric Reinforcement," in *International Conference on Cyberworlds*, vol. 5, 2017.

[6] Mohamed Torky, Ali Meligy and Hani Ibrahim, "Recognizing Fake Identities in Online Social Networks Based on a Finite Automaton Approach," in *International Journal of Computer Applications*, 2016.

[7] Supraja Gurajala, Joshua S White, Brian Hudson, Brian R. Voter, and Jeanna N. Matthews, "Profile characteristics of fake Twitter accounts," in *Big Data & Society*, July–December 2016: 1–13.

[8] Simranjit. Kaur. Tuteja, "A survey on classification algorithms for email spam filtering," in *International Journal Eng. Sci.*, vol. 6, no. 5, pp. 5937–5940, 2016.

[9] M. A. Devmane and N. K. Rana, "Detection and Prevention of Profile Cloning in Online Social Networks," in *IEEE International Conference on Recent Advances and Innovations in Engineering*, May 09- 11, 2014.

[10] Sara Keretna, Ahmad Hossny and Doug Creighton, "Recognising User Identity in Twitter Social Networks via Text Mining," in *IEEE International Conference on Systems, Man, and Cybernetics,* 2013.

[11] B. Erçahin, Ö. Aktaş, D. Kilinç, and C. Akyol, "Twitter fake account detection," in *Proc. Int. Conf. Comput. Sci. Eng. (UBMK)*, Oct. 2017, pp. 388–392, 2017.

[12] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on Twitter, " in *Proc. Collaboration, Electron. Messaging, Anti- Abuse Spam Conf. (CEAS)*, vol. 6, pp. 12. Jul. 2010

[13] S. Gharge, and M. Chavan, "An integrated approach for malicious tweets detection using NLP," in *Proc. Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, pp. 435–438, 2017.

[14] T. Wu, S. Wen, Y. Xiang, and W. Zhou, "Twitter spam detection: Survey of new approaches and comparative study," in *Comput. Secur.*, vol. 76, pp. 265–284, Jul. 2018.

[15] S. J. Soman, ‘‘A survey on behaviors exhibited by spammers in popular social media networks," in *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, pp. 1–6, 2016.

[16] A. Gupta, H. Lamba, and P. Kumaraguru, "Analyzing fake content on Twitter," in *Proc. eCrime Researchers Summit (eCRS)*, pp. 1–12, 2013.

[17] F. Concone, A. De Paola, G. Lo Re, and M. Morana, "Twitter analysis for real-time malware discovery," in *Proc. AEIT Int. Annu. Conf.*, pp. 1–6, 2017.

[18] N. Eshraqi, M. Jalali, and M. H. Moattar, "Detecting spam tweets in Twitter using a data stream clustering algorithm," in *Proc. Int. Congr. Technol., Commun. Knowl. (ICTCK)*, pp. 347–351, 2015.

[19] C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, "Statistical features-based real-time detection of drifted Twitter spam," in *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 4, pp. 914–925, Apr. 2017.