Global Journal of Engineering Innovations & Interdisciplinary Research

Original Article

Spammer Detection and Fake User Identification On Social Network

S Ruksana, B.Rushi kumar, KM Nikhil, R Sujatha, K Sai kumar, A Vinod kumar

Computer Science and Engineering, Gates Institute of Technology, Gooty, AP, India

Correspondence

S Ruksana

Department of CSE, Gates Institute of Technology, Gooty, AP, India

- Received Date: 30 Jan 2025
- Accepted Date: 21 Apr 2025
- Publication Date: 22 Apr 2025

Abstract

Social networking sites engage millions of users around the world. The users' interactions with these social sites, such as Twitter and Facebook have a tremendous impact and occasionally undesirable repercussions for daily life. The prominent social networking sites have turned into a target platform for the spammers to disperse a huge amount of irrelevant and deleterious information. Twitter, for example, has become one of the most extravagantly used platforms of all times and therefore allows an unreasonable amount of spam. Fake users send undesired tweets to users to promote services or websites that not only affect legitimate users but also disrupt resource consumption. Moreover, the possibility of expanding invalid information to users through fake identities has increased that results in the unrolling of harmful content. Recently, the detection of spammers and identification of fake users on Twitter has become a common area of research in contemporary online social Networks (OSNs). In this paper, we perform a review 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.

Introduction

What Can Social Networks Be Used For? Social networks can provide a range of benefits to members of an organization: Support for learning: Social networks can enhance informal learning and support social connections within groups of learners and with those involved in the support of learning. Support for members of an organization: Social networks can potentially be used my all members of an organization, and not just those involved in working with students. Social networks can help the development of communities of practice. Engaging with others: Passive use of social networks can provide valuable business intelligence and feedback on institutional services (although this may give rise to ethical concerns). Ease of access to information and applications: The ease of use of many social networking services can provide benefits to users by simplifying access to other tools and applications. The Facebook Platform provides an example of how a social networking service can be used as an environment for other tools. student's record management [1]. Block chain, the technology underpinning the Bit coin currency, is a decentralized sharing ledger that records data from the various parties.

Participating in the Bit coin network's transactions. The Bit coin network, in particular, uses the Block chain to store the history of transactions as well as other transaction related information, such as the time that the transaction was completed, the sender's (or spender's) address, and the receiver's address. It will assist the spenders in avoiding double-spending. To secure the Block chain's privacy, all of the information is encrypted. The Block chain can also be defined as a shared ledger since it holds all of the information about all Bit coin transactions [2]. The world of education is transitioning into the modern age. Indeed, technology and education are an excellent match that has grown in popularity in recent years. As a result, educational his work ―Blueprint for a new economy‖, says that the technology has become a worldwide phenomenon. However, we growth of block chain technologies could be divided into cannot discuss the use of technologies without discussing the three generations – 1) Block chain 1.0 2) Block chain 2.0 3) issue of protection. Failure to adhere to adequate protection Block chain 3.0. Block chain 1.0 relates to development of procedures will result in increased financial and human resource crypto currencies. Block chain 2.0 widens its scope to use. Researchers

Citation: Ruksana S, Rushi kumar B, Nikhil KM, Sujatha R, Sai kumar K, Vinod kumar A. Spammer Detection and Fake User Identification On Social Network . GJEIIR. 2025;5(2):49.

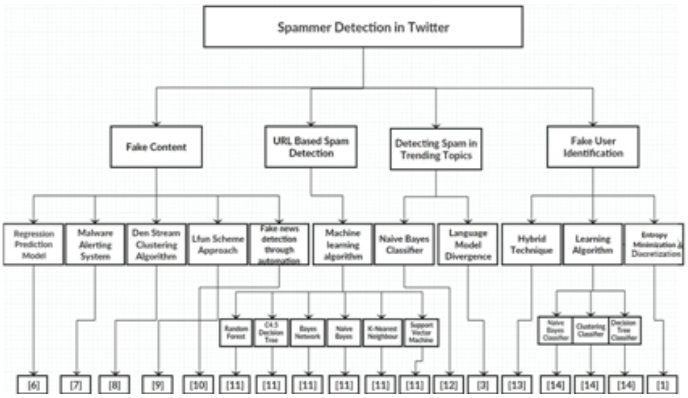

Figure 1. *Spammer Detection Overview*

and practitioners have proposed various include applications based on loans, smart contacts, recommendations, approaches, and strategies that help the property and bonds. Block chain 3.0 explores the possibility decision-making process on the security steps to be adopted after of leveraging block chain technologies for applications approaches, and strategies that help the decision-making process other than finance, such as healthcare, governance and on the security steps to be adopted after the early implementation education [8].

Background of blockchain

The project on spammer detection and fake user identification in social networks addresses the growing concern of malicious activities that undermine user trust and platform integrity. With the rise of social media, platforms like Twitter have become targets for spammers and fake accounts, which can spread misinformation, promote scams, and disrupt genuine interactions. The challenge lies in developing effective algorithms that can accurately identify and differentiate between legitimate users and fraudulent accounts. This involves analyzing user behavior, content patterns, and network interactions to detect anomalies indicative of spam or deceitful practices. The project aims to create a robust framework that automates the detection process, enhancing user safety and improving the overall quality of social interactions. By leveraging machine learning techniques and data mining, the system will continuously adapt to evolving spam tactics, ensuring timely responses to emerging threats. Furthermore, the implementation of this project could lead to better user experiences, reduce unwanted content, and foster a more authentic online community. Ultimately, the goal is to contribute to a safer digital environment where users can engage without fear of deception or harassment.

Blockchain is the core technology used by cryptocurrencies like Bitcoin. It maintains immutable distributed ledgers across thousands of nodes. As defined by Satoshi Takemoto [2], blockchain is a single list of chained blocks, where each block contains transactions or data, its own hash value, and the hash value of the previous block. Any alteration to a block changes its hash value, making tampering virtually impossible. Blockchain also includes concepts like distributed consensus, privacy and security protection, peer-to-peer (P2P) communication, network protocols, and smart contracts [4]. It has the potential to transform the Internet from an "Internet of Information Sharing" to an "Internet of Value Exchange" [5]. The technology has attracted attention in finance, healthcare, governance, and business due to its transparency, decentralization, and security. There are two types of blockchain: Permissionless (open to anyone) and Permissioned (restricted to selected users). Based on their needs, organizations can choose from Public, Private, Consortium, and Hybrid blockchain types [6,7].

Methodology

This section outlines the methodology adopted for detecting spammers and fake users on Twitter, as well as for implementing blockchain technology to ensure the authentication of educational credentials.

Data Collection and Preprocessing

The User Management Module is designed to manage various types of users, including the public and administrative officials. It has different components, such as the Public User Interface, which allows the public to upload images of suspected missing children along with relevant details like location, time, and remarks.

- **Twitter Spam Detection:**
  - A large dataset of tweets is collected, consisting of both spam and non-spam tweets.
  - Tweets are labeled based on predefined criteria (e.g., fake content, malicious URLs, trending topic spam, fake user profiles).
  - The dataset is cleaned by removing stop words, links, and special characters, and converted into a structured format suitable for analysis.
- **Blockchain Document Authentication:**
  - User-submitted applications are collected, including personal information, educational records, and credentials.
  - These are digitized and securely submitted through a blockchain-based application portal.

Feature Extraction

- For each tweet, various features are extracted, which include:
  - User features: Account age, followers/following ratio, verification status.
  - Content features: Presence of hashtags, URLs, mentions, sentiment polarity.
  - Graph features: Retweet and mention network analysis.
  - Time features: Posting frequency and activity timelines.
- For document authentication, metadata such as issue date, institution ID, and certificate hash is extracted and stored.

Classification and Detection Model

- A supervised machine learning approach is adopted using models such as:
  - Logistic Regression
  - Random Forest
  - Support Vector Machine (SVM)
  - Deep Learning (LSTM for sequential tweet data)
- The L-Fun Scheme is implemented to dynamically update the classifier by discovering and retraining with drifted spam patterns from unlabeled data (as per C. Chen et al.).
- The classifier is trained on balanced labeled datasets and evaluated using accuracy, precision, recall, and F1-score

Smart Contract and Blockchain Integration

- A consortium blockchain is used to record verified student applications and educational documents.
- Smart contracts operate on an "if-this-then-that" logic to automatically:
  - Validate application data.
  - Check digital credentials.
  - Record valid credentials as immutable entries on the blockchain.
- These credentials are stored in a tamper-proof manner and can be verified by institutions or employers without

System Workflow

1. User registers and submits an application with credentials.
2. Application is hashed and submitted to the blockchain.
3. Smart contract validates the data.
4. Tweet data is processed in real time and passed through the trained spam classifier.
5. Detected spam accounts are flagged and handled by the system admin.
6. Validated documents are stored on the blockchain for future verification.

Evaluation Metrics

- **Spam Detection:**
  - Accuracy, precision, recall, F1-score
  - Performance before and after applying the L-fun scheme
- **Blockchain Authentication:**
  - Verification time
  - Data integrity checks
  - Resistance to tampering and forgery

Conclusion

In this paper, we performed a review of techniques used for detecting spammers on Twitter. In addition, we also presented a taxonomy of Twitter spam detection approaches and categorized them as fake content detection, URL based spam detection, spam detection in trending topics, and fake user detection techniques. We also compared the presented techniques based on several features, such as user features, content features, graph features, structure features, and time features. Moreover, the techniques were also compared in terms of their specified goals and datasets used. It is anticipated that the presented review will help researchers find the information on state-of-the-art Twitter spam detection techniques in a consolidated form.

Despite the development of efficient and effective approaches for the spam detection and fake user identification on Twitter , there are still certain open areas that require considerable attention by the researchers. The issues are briefly highlighted as under: False news identification on social media Networks is an issue that needs to be explored because of the serious repercussions of such news at individual as well as collective level. Another associated topic that is worth investigating is the identification of rumors sources on social media. Although a few studies based on statistical methods have already been conducted to detect the sources of rumors, more sophisticated approaches, e.g., social network based approaches, can be applied because of their proven effectiveness. This research focuses on detecting Parkinson's disease using image and speech data by implementing and assessing three machine learning models: Support Vector Machine (SVM), Random Forest, and Decision Tree. The objective was to accurately classify individuals with Parkinson's disease based on features extracted from medical imaging and speech recordings. Among the three models, Random Forest showed the highest accuracy and reliability in detecting Parkinson's disease. However, further improvements with deep learning could yield even better results. Random Forest showed strong performance by handling feature variability and reducing overfitting. It provided better feature importance insights, making it useful for understanding key contributors to Parkinson's detection. In future exploring deep learning models (CNN, LSTMs, or hybrid models) could enhance performance. Increasing dataset size and diversity can improve model generalization. Implementing real-time detection using IoT-enabled systems for early Parkinson's diagnosis.

References

1. B. Enchain, Ö. Aktaş, D. Killing, and C. Akyol, "Twitter fake account detection," *Proc. Int. Conf. Comput. Sci. Eng. (UBMK)*, Oct. 2017, pp. 388–392.
2. F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on Twitter," *Proc. Collaboration, Electron. Messaging, Anti-Abuse Spam Conf. (CEAS)*, vol. 6, Jul. 2010, p. 12.
3. S. Gharge and M. Chavan, "An integrated approach for malicious tweets detection using NLP," *Proc. Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, Mar. 2017, pp. 435–438.
4. T. Wu, S. Wen, Y. Xiang, and W. Zhou, "Twitter spam detection: Survey of new approaches and comparative study," *Comput. Secur.*, vol. 76, pp. 265–284, Jul. 2018.
5. S. J. Soman, "A survey on behaviours exhibited by spammers in popular social media networks," *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, Mar. 2016, pp. 1–6.
6. A. Gupta, H. Lamba, and P. Kumaraguru, "$1.00 per RT #BostonMarathon #prayforboston: Analysing fake content on Twitter," *Proc. eCrime Researchers Summit (eCRS)*, 2013, pp. 1–12.
7. F. Concone, A. De Paola, G. Lo Re, and M. Morana, "Twitter analysis for real-time malware discovery," *Proc. AEIT Int. Annu. Conf.*, Sep. 2017, pp. 1–6.
8. N. Eshraqi, M. Jalali, and H. Moattar, "Detecting spam tweets in Twitter using a data stream clustering algorithm," *Proc. Int. Congr. Technol., Commun. Knowl. (ICTCK)*, Nov. 2015, pp. 347–351.
9. C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, "Statistical features-based real-time detection of drifted Twitter spam," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 4, pp. 914–925, Apr. 2017.
10. C. Buntain and J. Golbeck, "Automatically identifying fake news in popular Twitter threads," *Proc. IEEE Int. Conf. Smart Cloud (SmartCloud)*, Nov. 2017, pp. 208–215.
11. C. Chen, J. Zhang, Y. Xie, Y. Xiang, W. Zhou, M. M. Hassan, A. AlElaiwi, and M. Alrubaian, "A performance evaluation of machine learning-based streaming spam tweets detection," *IEEE Trans. Comput. Social Syst.*, vol. 2, no. 3, pp. 65–76, Sep. 2015.
12. G. Stafford and L. L. Yu, "An evaluation of the effect of spam on Twitter trending topics," *Proc. Int. Conf. Social Computers.*, Sep. 2013, pp. 373–378.
13. M. Mateen, M. A. Iqbal, M. Aleem, and M. A. Islam, "A hybrid approach for spam detection for Twitter," *Proc. 14th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Jan. 2017, pp. 466–471.
14. A. Gupta and R. Kaushal, "Improving spam detection in online social networks," *Proc. Int. Conf. Cogn. Compute. Inf. Process. (CCIP)*, Mar. 2015, pp. 1–6.