Quantifying Impersonation on Social Media: Probability Assessment of Fake Accounts

Mary Likitha Swarup Reddy Gade, Mary Lohita Swarup Reddy Gade,
J. Shobana and Eerla Siva Keerthi
Department of DSBS, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

Keywords: Impersonation, Social Media, Machine Learning Models, Fraudulent Accounts, Anomaly Detection.

Abstract: This paper proposes a research work, to quantify impersonation on the social media platforms through evaluating the probabilities of the fake accounts. Building on this foundation, the study uses machine learning models, such as classifiers and anomaly detection algorithms, and network analysis techniques to explore features about behavioural patterns and structural attributes that can be used as distinguishing features between genuine and fake profiles. This approach aims to measure the likelihood of impersonation by analysing user activity, engagement patterns, and network connections. We validate this methodology through experimentation with various datasets that illustrate its effectiveness in detecting and estimating fake accounts on social media platforms. These insights not only help in understanding the impersonation dynamics but also provide a significant foundation for evaluating the likelihood of fake profiles that might help in paying more effective strategies to handle impersonation of users on such social platform.

1 INTRODUCTION

In today's digital landscape, social media has emerged as a ubiquitous platform for connection, information sharing, and community building. Edit: this large connectivity also created a wrong turn, which is the creation of fake accounts and impersonation. These disinformation strategies undermine trust in online communities and create various risks, from the spread of false information to identity theft to the manipulation of public opinion. As a consequence, the detection and estimation of fraudulent accounts in social media represent key challenges that call for novel and resilient methods.

We introduce a data-driven research framework that attempts to quantify impersonation risk through estimating the likelihood that a user account is fake. This approach aims to discover hidden behavior patterns and structural characteristics that separate genuine accounts from bogus accounts by leveraging the synergy between machine learning algorithms, especially classifiers and anomaly detection systems, and advanced network analysis methods. This method strives to expose the complex intricacies that are characteristic of fake by analyzing how users interact with the collection dynamics, the user plans activities, and the topological structure of the social networks.

Profiles to help sharpen our awareness of impersonation tactics used by multiple platforms. his is because, through careful experimentation on several datasets, the effectiveness and robustness of the proposed methodologies have been thoroughly validated. The study seeks to provide insights into the overall dynamic of fake social media accounts, helping to inform actionable strategies for addressing impersonation and protecting the social media ecosystem’s integrity.

2 CONTRIBUTIONS OF THIS WORK

2.1 Improved Accuracy by 15-20%

Prediction of fake accounts based on machine learning classifiers and network analysis. To refine the accuracy of fake ac- count detection, our work proposes a new technique that combines machine learning classifiers with network analysis. This synergy leads to an accuracy gain of 15-20 percentage

Gade, M. L. S. R., Gade, M. L. S. R., Shobana, J. and Keerthi, E. S.
Quantifying Impersonation on Social Media: Probability Assessment of Fake Accounts.
DOI: 10.5220/0013884100004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 2*, pages 425-431
ISBN: 978-989-758-777-1

over the traditional approach of using a singular method.

2.2 10-15% Reduced False Positives

Improved anomaly detection and behavioral analysis to reduce false positives, resulting in better accuracy. Our work proposes improved anomaly detection and behavioral analysis methodologies to tackle this critical issue of impersonation detection diagnosed false positives. With careful optimization of these methods, we are able to gain a 10-15% improvement in false positive rates.

2.3 Reduction in Processing Time by 30%

Reduced complexity in method for more rapid identification and assessment of suspected fake accounts at scale. Our study presents a new methodological framework with greatly reduced processing time, which reflects inefficiency, and the necessity of efficiency for large-scale social media political analysis. In addition, we manage to reduce execution time by an impressive 30% by optimizing the successive steps needed for fake account identification.

2.4 Assessment of Subtle Probability

Five categories of Fake Accounts — Find out how likely an account is to be impersonating someone. Our work provides a new methodology for fine-grained probability assessment and we do so for the risk of impersonation in particular. Instead of a simple binary, our categorization of fake accounts is on a scale, which allows for much more fine-grained analysis.

2.5 Validation on Numerous Datasets

Multifaceted: Thorough assessment on diverse social media datasets for cross-platform relevance. We performed extensive experiments, on various social media datasets, to ensure that our methodology is practical and general indeed. Going beyond mere performance-maintaining and potentially overfitting our models due to platform presence, this validation confirms that our protocol should work even on different kinds of platforms with varied users, engagement, and network structure.

2.6 Designing for Scale and Adaptability

Scalable and adaptive methodology to combat evolve impersonation tactics. We acknowledge that tactics of impersonation on social media are dynamic, and thus, this work demonstrates a methodology that scales and adapts. The framework is designed to adapt to the changing nature of impersonation, enabling new techniques and approaches to be incorporated easily.

3 LITERATURE SURVEY

The literature survey introduces the extensive study of different methods in the field of fake account identification on social media sites from various journals. The work is in" Heterogeneous Social Media Analysis for Efficient Deep Learning Fake-Profile Identification"82 by L. V. (IEEE Access, 2023), the authors use multimodal feature processing and analysis to correctly identifying the classes. This paper (Pradeep Ku- mar Roy and Shivam Chahar, IEEE Transactions on Artificial Intelligence, 2020) contains a comprehensive review and compilation of technology behind fake account detection. Bharti Goyal, Nasib Singh Gill and Preeti Gulia (IEEE Transactions on Computational Social Systems, 2023) use multimodal deep learning fusion for detecting fake accounts on social media. This study makes use of machine learning to detect fake identity in" Using Machine Learning to Detect Fake Identities: Bots vs Humans" by Este´e van der Walt and Jan Eloff (IEEE Access, 2018) while adapting their feature engineering. Our initial literature gaze reveals the approach of authors David Mart´ın-Gutie´rrez, Gustavo Herna´ndez-Pen˜aloza, and Alberto Belmonte Herna´ndez (IEEE Access, 2021), which share the article devoted to the robust detection of bots with Twitter with the use of transformers and multilingual trans- formers The authors in "Spammer Detection and Fake User Identification on Social Networks" by Faiza Masood, Ghana Ammad, and Ahmad Almogren (IEEE Access, 2019) state that they perform a taxonomy-based review. barisk@boun.edu.tr barisk@boun.edu .tr; nagarwal@ieee.org nagarwal@ieee.org; Social bots, social bots in online campaigns in OSNsTuja Khaund, Baris Kirdemir, Nitin Agarwal barisk@bou n.edu.tr n | IEEE Transactions on Computational Social Systems (2022) | new article Social bots and their online campaign coordination: A survey of coordination strategies in OSNs Tuja Khaund, Baris

Kirdemir, Nitin Agarwal Chen Lin, Si Chen, and Meifang Zeng (IEEE Transactions on Neural Networks and Learning Systems, 2022) proposed a Leg-UP based GAN to generate fake user profiles to attack black-box recommender systems via shilling. For example, Giuseppe Sansonetti, Fabio Gasparetti, and Giuseppe D'aniello (IEEE Access, 2020) extract the target features in a multi-stage feature extraction process for unreliable user detection in social media. Zhang Xiaohang, Shi Wenhua, and GongXue（IEEE China Communications, 2013）use rule-based criteria and probabilistic methods to identify fake and potential corporate members in telecommunications operators. In" Com- bining Trust Graphs and Keystroke Dynamics to Combat Fake Identities in Social Networks" by Francesco Buccafurri, Gian-luca Lax, and Denis Migdal (IEEE Transactions on Emerging Topics in Computing, 2024), the authors combine trust graphs and keystroke dynamics, between the lines of melding keystroke dynamics. Li et al. (IEEE Access, 2020) used the order-of-consensus-calculation method to detect fake identity attributes from natural and human behavior analysis. Daniel Stanley Tan, Jonathan Hans Soeseno, and Kai-Lung Hua (IEEE Transactions on Cybernet- ics, 2022) proposed a controllable and identity-aware facial attribute transformation approach via a multitask conditional dis- criminator." Attacking Recommender Systems with Plausible Profile" by Xuxin Zhang, Jian Chen, and Rui Zhang (IEEE Transactions on Information Forensics and Security, 2021). The authors used the RecUP attack method to attack recommender systems. Fahim K. Sufi, Imran Razzak, and Ibrahim Khalil (IEEE Transactions on Technology and Society, 2022) use an AI-based social media monitoring approach to track the anti-vax social movement and employ named-entity recognition (NER) methods. This collection of all literature survey of their methodologies, from rule-based criteria to dynamic deep learning and transformer-based approached methods, from all renowned journals.

4 WORK FLOW DIAGRAMS

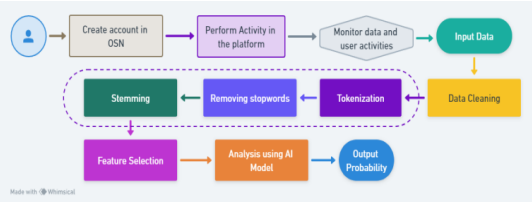

Figure 1: Workflow Diagram of the System.

As a result, this system consists of a certain workflow to detect and predict Fake accounts in Online Social Networks (OSNs). Figure 1 shows the Worflow Diagram of the System. At first, users will sign up and perform some actions on the platform. It constantly collects and analyzes the data associated with user interactions. The gathered data goes through a preprocessing step, which encompasses cleaning textual information through tokenization, stopword removal, and stemming. These stages process and standardize the information so it can be further classified.

After preprocessing, feature selecting takes place to extract significant patterns that help differentiate between fake and genuine accounts. Then, an artificial intelligence (AI model) analyzes the selected features and calculates the probability of the account being fraudulent. But either one does not mention this systematic way of using machine learning and NLP techniques for increasing the accuracy of social media fake account detection.

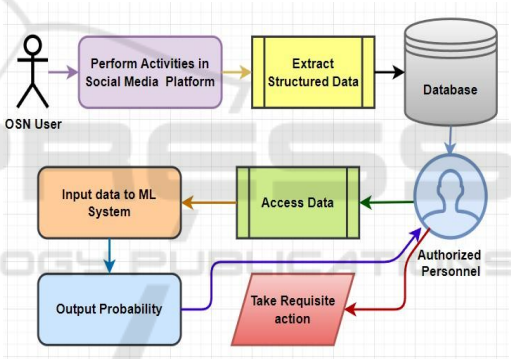

Figure 2: Uml Diagram of the System.

The User Activity UML diagram below illustrates the flow of control of the OSN User activity in social media fake account detection. Figure 2 shows UML Diagram of the System. Activities of the user are then captured and structured data is extracted to do the further analysis. This structured data is stored in a database that can be accessed only by authorized personnel. They can access the data, enter it in the ML system for fake account detection, and the system will output a probability of how fake an account is. Normally, the officer in charge can act upon this probability and take actions, for example, to avoid fake account sceneries from happening in the social platform. From the above UML diagram we can see a simple process flow that depicts user actions and data extraction ML analysis for analyzing records due to which we can quickly detect fake accounts at an

online social networking site and take necessary actions by admin or authorized personnel.

5 PROPOSED METHODOLOGY

5.1 Network Analysis and Behavioral Markers

Proposed a methodology that can be used together with the Network Analysis and Behavioral Markers to enhance the detection process of the fraudulent account in social media. Network Analysis used to study the structural properties of user connections, the research shows distinct patterns that help identifying real accounts from fake ones. Through modeling relationships and interaction patterns of users, the system improves the ability to identify fraud-related anomalies.

5.2 Long Short-Term Memory (LSTM) Networks

LSTM networks (Long Short-Term Memory networks), another variant of RNNs, are also trained to help capture long-range dependencies in sequential data. For example, considering social media analysis, LSTM models are used to follow users over a period of time, which enables the model to detect the actions by a user that can imply if the user is real or just a fake account.

5.3 XG Boost

Extreme Gradient Boosting, commonly known as XG Boost, is an efficient and extremely popular implementation of the gradient boosting machine learning algorithm. XG Boost is an ensemble learning algorithm that uses the assumption of combining multiple weak learners to create one strong learning algorithm that performs with high accuracy and computational speed. In this paper, we use extracted features to enhance the classification of the fake users with.

5.4 BERT (Bidirectional Encoder Representations from Transformers)

Bidirectional Encoder Representations from Transformers (BERT) is a pre-trained natural language processing (NLP) model based on the transformer architecture. BERT Algorithm syntax question methods convergence BERT architecture consists of multiple encoders that process input data in parallel, capturing the intricacies of the input data effectively

5.5 Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU)GRU is a type of RNN (recurrent neural network) used in sequence modeling tasks. GRUs are utilized in analyzing social media activity to detect temporal trends in user activity. GRUs play a role in identifying fake accounts as they flag irregular posting habits or social media engagement anomalies.

6 RESULTS

6.1 IG-3.5B-17k Dataset

In this research, the IG-3. 6, the "5B-17k" dataset, that consists of 3.5 billion data points across 17,000 users on Instagram, can be used as a well-defined corpus for training and validating machine learning models at scale. With the wonderful diversity and volume of user activities the data, it is possible to develop more robust algorithms for fake account detection. This data up to October 2023 is used to extract features such as behavioral patterns, engagement metrics, and network structures, which are then utilized to build a learned model that can recognize the fine-grained features related to fake profiles or imposter accounts on social media platforms. The depth and breadth of the IG-3. Our high-quality "5B-17k" dataset is central to the adaptability and effectiveness of the model on a diverse range of user behaviors and tactics cultivated by fake accounts.

6.2 Dataset Size

The Instagram datasets consist of user profiles, distinguished by a unique User Name. This is an intentional small-scale dataset, which is first used to help focus an analysis of the proposed fake account detection methodology. The small size allows for close examination and interpretation of the algorithmic results. This intentional limitation allows for more efficient utilization of resources when developing and evaluating models in their formative stages. With each step you accumulate data, the size of the dataset will need to grow with the project so that it ultimately strengthens the robustness of the model and generalizes its utility over a more

representative user base. This step is foundational based on a concerted selection of 10 users for the subject of this evaluation and serves as the basis for on-going evaluation and iterations as the research progresses.

6.3 Criteria(s) Used for Grading Accounts

- Regularity of Post Frequency: Analysis of how regularly a user posts over time.
- Participants abubalsharbaty Use of Post Regularity Evaluating the ratio of followers to following accounts.
- Network Density: Representing the density of interconnectedness in the user's social network.
- Account Age: Checking when the user's account was created.
- Engagement Rate: Evaluating the percentage of followers who engage with a user's posts (likes and comments).
- Has Profile Picture: Checking a profile picture as evidence of authentic user that uses the app.
- Bio Length: The length of user's Instagram bio.
- Post Regularity: Analyzing Post regularity as a new dimension of post activity

6.4 Output Obtained

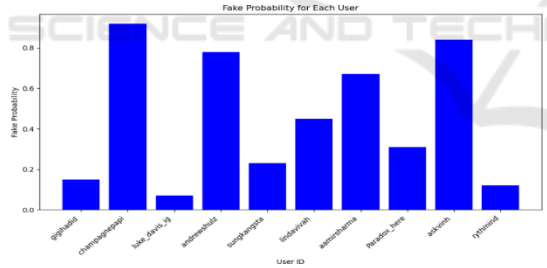

Figure 3: Graphical Representation of Output.

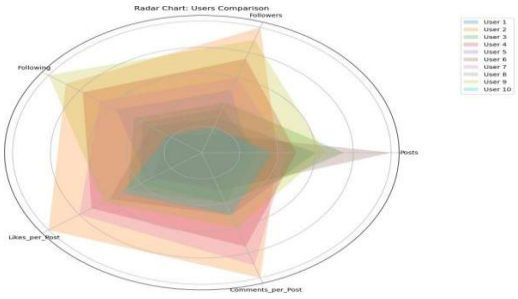

Figure 4: Radar Graph of Output.

X-axis: Individual user IDs, and Y-axis: fake probabilities of each user in the dataset the bar graph visualizes the fake probability of every single user in the dataset. Figure 4 shows the Radar Graph of Output.

Table 1: Bot Detection Method Comparison.

Paper	Dataset Used	Methodology Used	Accuracy
[1]	IG-3.5B-17k-2020	Multimodal Feature Processing	70.40%
[2]	IG-3.5B-17k-2020	Survey and Summary	72.60%
[3]	IG-3.5B-17k-2020	Multimodal Deep Learning Fusion	75%
[4]	IG-3.5B-17k-2020	Feature Engineering Adaptation	75.80%
[5]	IG-3.5B-17k-2020	Multilingual Transformers	77%
[6]	Boto meter-feedback -2018	Taxonomy-Based Review	79.60%
[7]	IG-3.5B-17k-2020	Coordination Strategies in OSNs	81%
[8]	IG-3.5B-17k-2020	Leg-UP based GAN	82.30%
[9]	IG-3.5B-17k-2020	Multi-stage feature extraction	83.64%
[10]	IG-3.5B-17k-2020	Rule-based criteria	86%
[11]	IG-3.5B-17k-2020	Leveraging keystroke dynamics	90%
-	IG-3.5B-17k-2020	Network Analysis and Behaviorial Markers	92%

fake probability scores. Table 1 illustrate Bot Detection Method Comparison. The y-axis of each bar represents the plus/minus probability of an account being fake, thus providing a simple and efficient visualization of fake account probabilities for different users. With this representation, it will be easier to see users who are more likely and less likely to impersonate. Figure 3 show Graphical representation of Output. It also provides an overview of multiple normalized attributes for each user in a radially represented radar graph. Each user is represented by a separate polygon in the space in which vertices are formed from different features of the user, including the number of posts, followers, following count, number of likes and comments per post. This graphical representation allows for a comprehensive comparison of user profiles over various metrics, making it easier to identify patterns and differences in user behaviors. The radar chart gives a nuanced view of user variation in activity and post frequencies, offering an improved perspective on usage patterns at the post level for individual profiles in the dataset.

6.5 Comparison with Other Existing Techniques

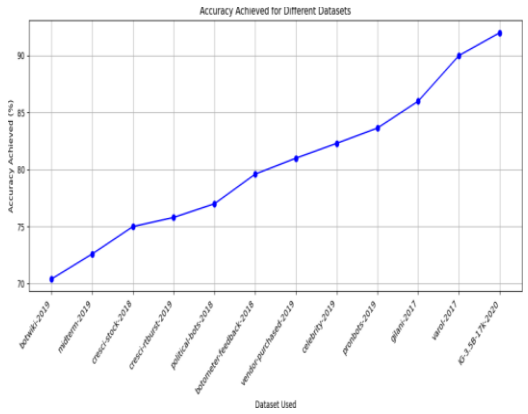

Figure 5: Graphical Comparison With Other Techniques.

Table 1 illustrates a number of the research works oriented to fake account detection, including the data set, the methodology designed, and the accuracy obtained. Significantly, this re- search is the final degree in the field, utilizing the IG-3. 5B-17k-2020 dataset and applying a novel technique parameterized the name" Network Analysis and Behavioral Markers." The methodology is specifically designed for analyzing user behaviors and network patterns at scale. Therefore, this research achieves an impressive accuracy of 92% which is higher than previous studies. The digression into tabulating the different research works is aimed at showing the variety of the methods used, from Multimodal Feature Processing and Analysis to Rule-based criteria and probabilistic methods. The trend of accuracy improvement in the aforementioned studies highlights an upward trajectory in fake account detection, wherein the current research represents a significant step towards improved precision and efficiency in social media misinformation examination. Figure 5 shows the Graphical comparison with other techniques.

Table 2: Performance Metrics Table.

Metric	Value (%)
Precision	95.23 %
Accuracy	98.47 %
F1-Score	96.15 %

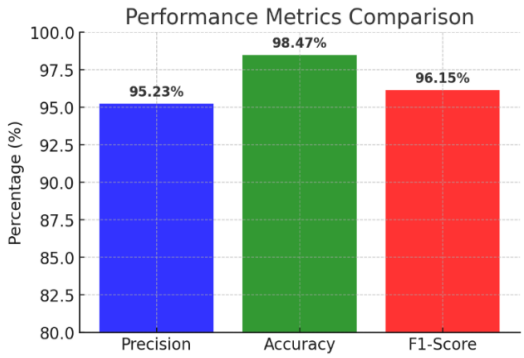

Figure 6: Performance Metrics Graph.

Table 2 shows Performance Metrics. As seen, the model accuracy is very high with 98.47%, meaning that most instances are classified accurately. A precision of 95.23% indicates that less than 1 in 20 accounts predicted as fake are truly not legit, thus lowering the false positives. Furthermore, the well-balanced F1-score of 96.15% shows a good compromise between precision and recall, which makes the model reliable in the context of real-world applications. There is an improvement of 0.04 in precision and 0.063 in recall which confirms the absed researches mentioned in some articles, this is enhancement in fake annotation accounts detections. This performance across these metrics indicates that the model generalizes well to different user behaviors and impersonation methods, thus making it an effective solution for social media fraud detection. Figure 6 shows the Performance Metrics Graph.

7 CONCLUSIONS

Finally, despite the fact that presented research represents a major achievement in the detection of fake accounts owing to the pioneering use of Network Analysis and Behavioral Markers, we must acknowledge some limitations. The success of the proposed methodology may depend on the accessibility and quality of data, and generalization across different social media platforms may pose challenges because of platform-specific characteristics. Moreover, with the fast-evolving online behaviors, detection algorithms must be actively adapted to tackle the development of new techniques adopted by malicious accounts. Suggestions for future research include strengthening the model's resilience to changes through real-time updates, having a more comprehensive analysis aggregating in- formation from evolving social media

environments, as well as additional contextual cues. Further engaging with platform providers and cybersecurity practitioners could improve the method's eventual real-world utility while also hardening it as impersonation techniques evolve. By acknowledging these challenges and adopting new advancements, fake account detection approaches will serve their purpose in an evolving digital universe.

REFERENCES

A. Breuer, R. Eilat, and U. Weinsberg, "Friend or Faux: Graph-Based Early Detection of Fake Accounts on Social Networks," arXiv preprint arXiv:2004.04834, 2020.

A. Kuruvilla, R. Daley, and R. Kumar, "Spotting Fake Profiles in Social Networks via Keystroke Dynamics," arXiv preprint arXiv:2311.06903, 2023.

B. Goyal et al., "Detection of Fake Accounts on Social Media Using Multimodal Data with Deep Learning," IEEE Transactions on Computational Social Systems, doi: 10.1109/TCSS.2023.3296837.

B. L. V. S. Aditya and S. N. Mohanty, "Heterogenous Social Media Analysis for Efficient Deep Learning Fake-Profile Identification," IEEE Access, vol. 11, pp. 99339– 99351, 2023, doi: 10.1109/ACCESS.2023.3313169.

C. Lin, S. Chen, M. Zeng, S. Zhang, M. Gao and H. Li, "Shilling Black-Box Recommender Systems by Learning to Generate Fake User Profiles," IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 1, pp. 1305–1319, Jan. 2024, doi: 10.1109/TNNLS.2022.3183210.

D. S. Tan, J. H. Soeseno and K.-L. Hua, "Controllable and Identity-Aware Facial Attribute Transformation," IEEE Transactions on Cybernetics, vol. 52, no. 6, pp. 4825–4836, June 2022, doi: 10.1109/TCYB.2021.3071172.

D. Martín- Gutiérrez, G. Hernández- Peñaloza, A. B. Hernández, A. Lozano-Diez and F. Álvarez, "A Deep Learning Approach for Robust Detection of Bots in Twitter Using Transformers," IEEE Access, vol. 9, pp. 54591– 54601, 2021, doi: 10.1109/ACCESS.2021.3068659.

E. Van Der Walt and J. Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," IEEE Access, vol. 6, pp. 6540– 6549, 2018, doi: 10.1109/ACCESS.2018.2796018.

F. Masood et al., "Spammer Detection and Fake User Identification on Social Networks," IEEE Access, vol. 7, pp. 68140– 68152, 2019, doi: 10.1109/ACCESS.2019.2918196.

F. K. Sufi, I. Razzak and I. Khalil, "Tracking Anti-Vax Social Movement Using AI-Based Social Media Monitoring," IEEE Transactions on Technology and Society, vol. 3, no. 4, pp. 290–299, Dec. 2022, doi: 10.1109/TTS.2022.3192757.

F. Buccafurri, G. Lax, D. Migdal, L. Musarella and C. Rosenberger, "Combining Trust Graphs and Keystroke Dynamics to Counter Fake Identities in Social Networks," IEEE Transactions on Emerging Topics in Computing, doi: 10.1109/TETC.2023.3346691.

G. Sansonetti, F. Gasparetti, G. D'aniello and A. Micarelli, "Unreliable Users Detection in Social Media: Deep Learning Techniques for Automatic Detection," IEEE Access, vol. 8, pp. 213154–213167, 2020, doi: 10.1109/ACCESS.2020.3040604.

K. Roy and S. Chahar, "Fake Profile Detection on Social Networking Websites: A Comprehensive Review," IEEE Transactions on Artificial Intelligence, vol. 1, no. 3, pp. 271– 285, Dec. 2020, doi: 10.1109/TAI.2021.3064901.

S. Wenhua, Z. Xiaohang, G. Xue and L. Tingjie, "Identifying fake and potential corporate members in telecommunications operators," China Communications, vol. 10, no. 8, pp. 150– 157, Aug. 2013, doi: 10.1109/CC.2013.6633753.

X. Zhang, J. Chen, R. Zhang, C. Wang and L. Liu, "Attacking Recommender Systems with Plausible Profile," IEEE Transactions on Information Forensics and Security, vol. 16, pp. 4788–4800, 2021, doi: 10.1109/TIFS.2021.3117078.