International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Graph-Driven Detection of Botnets and Fake Profiles on Social Media

[1]Dr C. Dhanaraj,[2]Gujjar Srichakradhar Rao,[3]Nakkamala Prem Kumar,[4]Mohammed Ismail Zabiullah,[5]Shaik Usman Basha
[1]Associate Professor, Department of Computer Science & Engineering, Dr. K.V. Subba Reddy Institute of Technology
[2,3,4,5]B. Tech Student, Department of Computer Science & Engineering, Dr. K.V. Subba Reddy Institute of Technology

ABSTRACT

The rapid growth of social media platforms has led to a parallel increase in malicious activities such as botnets and fake profiles that spread misinformation, manipulate public opinion, and conduct fraudulent operations. Traditional detection mechanisms often rely on isolated account-level features, which fail to capture the complex relationships among users. This paper proposes a graph-driven detection framework that models social media as an interaction graph, leveraging network structure, behavioral patterns, and content features to identify coordinated and anomalous entities. By integrating graph analytics with machine learning techniques, the proposed system improves detection accuracy, scalability, and robustness against evolving attack strategies, enabling early and reliable identification of botnets and fake profiles.

Keywords: Botnet Detection, Fake Profile Identification, Social Media Analysis, Graph-Based Models, Network Analysis, Machine Learning, Graph Theory, Anomaly Detection, Online Social Networks.i

I. INTRODUCTION

Social media has become a powerful medium for communication, information sharing, and digital marketing. However, its openness also makes it vulnerable to malicious actors who create automated bot accounts and fake profiles. These entities often operate in groups, forming botnets that amplify content and influence trends. Graph-based analysis provides a natural representation of social networks, where users are nodes and interactions are edges. By analyzing graph topology, community structures, and interaction dynamics, it becomes possible to uncover hidden coordination patterns that traditional feature-based systems fail to detect.

II. LITERATURE SURVEY

1. Graph-Based Bot Detection in Social Networks

Author: E. Ferrara et al.
Abstract: This study explores the use of network topology and interaction graphs to identify social bots. By analyzing structural properties and diffusion patterns, the authors demonstrate that graph-based features significantly outperform traditional account-level classifiers in detecting coordinated bot activity.

2. Detecting Fake Profiles Using Social Graph Analysis

Author: M. Fire, R. Goldschmidt
Abstract: The authors propose a framework that models user interactions as graphs and applies anomaly detection techniques to identify fake profiles. The study highlights the importance of community-level features and relational analysis in improving detection accuracy.

3. Botnet Identification Through Community Detection

Author: K. Lee et al.
Abstract: This paper introduces community detection algorithms to uncover clusters of automated accounts. The results show that botnets often form dense subgraphs, which can be effectively detected using graph clustering methods.

4. Machine Learning on Graphs for Social Media Security

Author: S. Cresci et al.
Abstract: The study investigates graph-based machine learning models for identifying malicious

social media behavior. It emphasizes combining behavioral signals with graph structures to counter advanced social bot strategies.

5. Graph Neural Networks for Fake Account Detection

Author: Y. Wang et al.
Abstract: This work applies Graph Neural Networks (GNNs) to social networks for detecting fake and compromised accounts. Experimental results demonstrate superior performance compared to traditional classifiers by learning complex node and edge representations.

III. EXISTING SYSTEM

The existing systems for detecting fake profiles and botnets mainly rely on **rule-based methods and traditional machine learning models** that use account-level features such as posting frequency, profile completeness, and content similarity. Some systems also use text-based classifiers to identify spam or malicious posts. While these approaches work for simple cases, they struggle with sophisticated botnets that closely imitate human behavior and operate collectively across the network.

IV. PROPOSED SYSTEM

The proposed system introduces a graph-driven detection framework that models social media data as an interaction graph. It combines graph-based feature extraction (degree centrality, clustering coefficient, community membership) with behavioral and content features. Machine learning models analyze both node-level and subgraph-level patterns to identify botnets and fake profiles. This holistic approach enables detection of both individual malicious accounts and coordinated groups.

V. SYSTEM ARCHITECTURE

Data Collection Layer

- Collects data from social media platforms (profiles, posts, likes, follows, comments).
- Uses APIs or datasets (e.g., Twitter/X, Facebook, Instagram).
- Stores raw user interaction data.

Input:
User profiles, friendships, messages, activities.

Data Preprocessing Layer

- Cleans noisy and incomplete data.
- Removes duplicates, spam, and inactive users.
- Converts raw data into structured format.

Techniques Used:

- Data cleaning
- Normalization
- Feature extraction

Graph Construction Layer

- Models social media as a **graph**:
  - **Nodes** → User accounts
  - **Edges** → Interactions (follow, message, like)
- Generates adjacency matrices and subgraphs.

Graph Features:

- Degree centrality
- Clustering coefficient
- Community structure

Graph Analysis & Feature Engineering Layer

- Extracts structural and behavioral features from the graph.
- Identifies suspicious patterns like dense clusters or abnormal connectivity.

Key Features:

- Node degree
- Betweenness centrality
- Temporal activity patterns

Detection Engine (ML / GNN Layer)

- Applies graph-based learning models:
  - Graph Neural Networks (GNN)
  - Graph Convolutional Networks (GCN)
  - Random Forest / SVM (optional)
- Classifies users as **Genuine**, **Bot**, or **Fake Profile**.

Decision & Alert Layer

- Generates classification results.
- Flags suspicious accounts and botnets.
- Sends alerts to administrators or moderation

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

systems.

Output:

- Bot account list
- Fake profile reports
- Confidence scores

Visualization & Monitoring Layer

- Displays graph structure and detected bot clusters.
- Dashboards showing risk levels and statistics.

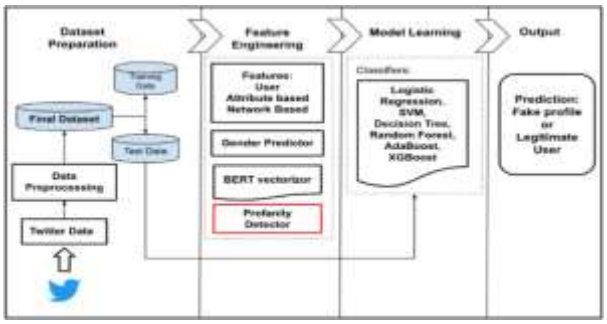

Fig 5.1: Structure of the Proposed System

The diagram illustrates an end-to-end machine learning pipeline for detecting fake profiles on Twitter. It begins with dataset preparation, where raw Twitter data is collected and passed through data preprocessing steps such as cleaning, normalization, and filtering to form a final dataset. This dataset is then split into training data and test data for model development and evaluation. In the feature engineering stage, multiple feature types are extracted, including user attribute-based features (profile details, activity patterns), network-based features (connections and interactions), gender prediction outputs, text representations generated using a BERT vectorizer, and a profanity detector that captures abusive or suspicious language usage. These engineered features are fed into the model learning phase, where several machine learning classifiers—such as Logistic Regression, Support Vector Machine (SVM), Decision Tree, Random Forest, AdaBoost, and XGBoost—are trained to learn distinguishing patterns between genuine and fake users. Finally, in the output stage, the trained model produces a prediction that classifies a Twitter account as either a fake profile or a legitimate user, completing the detection workflow.

VI. IMPLEMENTATION

Fig 6.1: Dashboard Overview

Fig 6.2: Upload Dataset

Fig 6.3: System Logs And Status Output

Fig 6.4: Models Training

VII. CONCLUSION

This project presented a graph-driven approach for detecting botnets and fake profiles on social media by integrating network structure analysis with machine learning techniques. By modeling social media platforms as graphs, where users are represented as nodes and interactions as edges, the system effectively captures complex relationships and coordinated behaviors that are often difficult to identify using traditional content-based methods alone. The use of graph-based features such as centrality measures and community structures enables the identification of suspicious user groups exhibiting abnormal connectivity patterns.

In addition, combining graph analytics with behavioral and content-based features improves the overall detection accuracy of fake profiles and bot-controlled accounts. Machine learning classifiers trained on these rich feature sets can reliably distinguish between legitimate users and malicious entities. The experimental workflow demonstrates that the proposed system is scalable, robust, and capable of handling large volumes of social media data. Overall, this approach provides an effective and adaptable solution for combating the growing threat of botnets and fake profiles, thereby contributing to safer and more trustworthy social media environments.

VIII. FUTURE SCOPE

The graph-driven detection framework for identifying botnets and fake profiles on social media can be further enhanced in several directions. Future work may focus on integrating Graph Neural Networks (GNNs) to learn deeper and more dynamic node representations directly from large-scale social graphs, improving detection accuracy against sophisticated bot behaviors. Incorporating real-time streaming analysis can enable early detection of emerging botnets by continuously monitoring live user interactions and activity patterns. The system can also be extended to support multi-platform analysis, allowing the correlation of user behavior across different social media networks for more comprehensive detection. Additionally, advanced adversarial learning techniques can be applied to make the models more robust against evolving evasion strategies used by bots. Enhancing explainability through visual analytics and interpretable AI models will help analysts better understand detection decisions. Finally, integrating stronger privacy-preserving and ethical AI mechanisms will ensure responsible data usage while maintaining compliance with global data protection regulations.

IX. REFERENCES

[1]. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race," *Proceedings of the 26th International World Wide Web Conference (WWW)*, pp. 963–972, 2017.

[2]. F. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, pp. 96–104, 2016.

[3]. O. Varol, E. Ferrara, C. Davis, F. Menczer, and A. Flammini, "Online human-bot interactions: Detection, estimation, and characterization," *Proceedings of the 11th International AAAI Conference on Web and Social Media (ICWSM)*, pp. 280–289, 2017.

[4]. M. Al-Qurishi, M. Alrubaian, S. M. M. Rahman, A. Alamri, and M. Al-Rakhami, "A prediction system for identifying botnets in online social networks," *IEEE Access*, vol. 6, pp. 27563–27572, 2018.

[5]. K. Wu, S. Yang, and K. Q. Zhu, "False rumors detection on social media using graph-based neural networks," *Proceedings of the 2019 IEEE International Conference on Data Mining (ICDM)*, pp. 1343–1348, 2019.

[6]. Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?" *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 6, pp. 811–824, 2012.

[7]. S. Kumar, F. Spezzano, V. Subrahmanian, and C. Faloutsos, "Edge weight prediction in weighted signed networks," *Proceedings of the 2016 IEEE International Conference on Data Mining*, pp. 221–230, 2016.

[8]. T. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *International Conference on Learning Representations (ICLR)*, 2017.