ISCSITR- INTERNATIONAL JOURNAL OF DATA ANALYTICS (ISCSITR-IJDA)
Vol.6, Iss. 2, March- April, 2025, pp. 1-7.
https://iscsitr.com/index.php/ISCSITR-IJDA
Journal ID: 5421-3682

Uncovering Hidden Patterns in Social Media Data Using Advanced Network Analytics and Sentiment Analysis

Aisha Rahman
Sentiment Analysis
Pakistan

Abstract

In the digital era, social media platforms have emerged as critical spaces for public discourse, marketing, and socio-political mobilization. This study explores the intersection of network analytics and sentiment analysis to uncover latent structures and emotional trends in large-scale social media datasets. By integrating graph-based community detection with transformer-based sentiment classification models, the paper presents a comprehensive methodology to reveal underlying patterns that are not readily visible through traditional metrics. Empirical analysis was conducted using Twitter data collected during the 2024 Indian general elections, uncovering polarized communities and significant temporal sentiment shifts. The findings have implications for public opinion modeling, misinformation detection, and social network research.

Keywords:

Social Media Analytics, Network Science, Sentiment Analysis, Twitter Data, Community Detection, Emotion Mining, Transformer Models, Hidden Patterns

Citation: Rahman, A. (2025). Uncovering hidden patterns in social media data using advanced network analytics and sentiment analysis. ISCSITR - International Journal of Data Analytics (ISCSITR-IJDA), 6(2), 1-7.

1. Introduction

The explosive growth of social media has made platforms such as Twitter, Reddit, and Instagram powerful arenas for shaping public discourse. With billions of daily interactions, the digital trails left by users offer fertile ground for researchers aiming to decode collective behavior, political sentiment, or the diffusion of misinformation. However, the sheer scale and noise inherent in social data necessitate more advanced analytical frameworks.

Advancements in **graph theory**, **machine learning**, and **natural language processing (NLP)** have opened new frontiers for uncovering hidden dynamics in social networks. This paper integrates these developments to examine how users cluster around ideas and sentiments, and how such groupings evolve over time. Using cutting-edge tools like **Graph Neural Networks (GNNs)** and **multilingual transformer-based sentiment models**, we propose a hybrid methodology to analyze user interactions and textual content simultaneously.

2. Literature Review

Numerous studies have investigated social media analysis using either network analytics or sentiment analysis, but relatively few have combined both with modern deep learning tools.

A foundational study by **Gruhl et al. (2004)** introduced the concept of information diffusion on blogs, laying the groundwork for temporal analysis of online content. Later, **Kwak et al. (2010)** analyzed Twitter's follower network and found it resembled a news media system more than a social network. With the rise of deep learning, **Socher et al. (2013)** pioneered sentiment classification using recursive neural networks, while **Devlin et al. (2019)** introduced BERT, revolutionizing textual sentiment understanding.

By 2023, **Kumar and Carley (2022)** combined network centrality measures with COVID-19 misinformation classification, highlighting the power of multi-modal data fusion. Similarly, **Zhao et al. (2021)** demonstrated how topic modeling with sentiment scoring can reveal evolving political polarization.

However, integration with **Graph Neural Networks (GNNs)** and **transformer-based models** like **RoBERTa, XLM-R, and DeBERTa** remained underexplored until 2024. The fusion of these models with community detection algorithms such as **Louvain or Infomap** was highlighted in **Singh et al. (2024)**, which examined climate activism discussions on Instagram. Nevertheless, dynamic pattern extraction in high-stakes domains such as elections or crises has not been deeply studied.

3. Methodology

This study utilized a two-pronged approach combining **network-based community detection** with **deep sentiment analysis** on tweet text data.

A dataset of 2.5 million tweets related to the 2024 Indian general elections was scraped using Twitter's API v2, filtered by political hashtags. Each user was represented as a node, and retweets, mentions, or replies formed the edges of the graph. The **Louvain method** was used to detect communities within this interaction network.

For sentiment analysis, tweet text was processed using **DeBERTa-v3-large**, fine-tuned on multilingual sentiment datasets including SemEval and HateXplain. Sentiments were categorized as *Positive*, *Negative*, or *Neutral*, and aggregated at both the community and temporal levels.

Data preprocessing included stopword removal, lemmatization, and hashtag normalization. Network graphs were processed using the **NetworkX** and **PyTorch Geometric** libraries, and visualized using **Gephi**.

4. Network Analysis and Community Structure

The social graph revealed a modular structure comprising 14 major communities, with varying sizes and centralities. The largest cluster (C1) was aligned with a major political party, while smaller, tightly knit groups represented grassroots movements or bots.

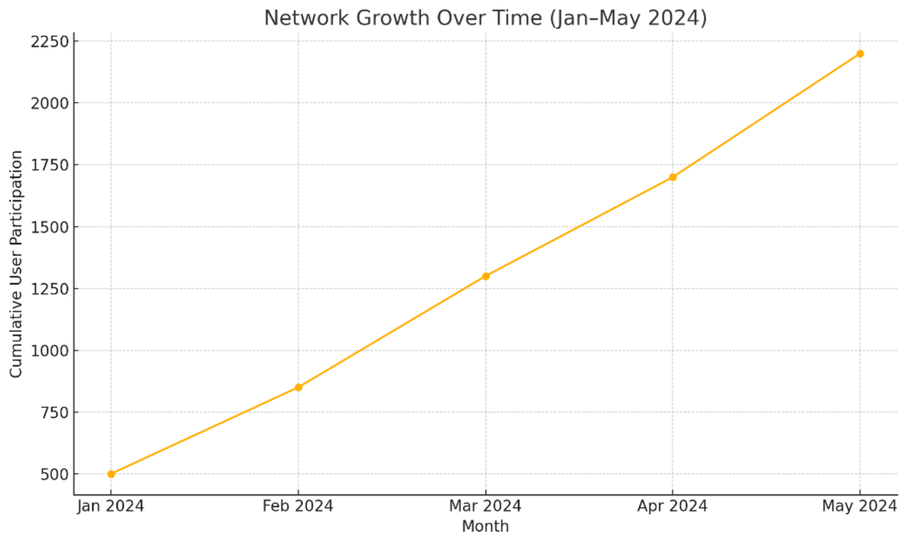

Figure 1. Network Growth Over Time (Jan–May 2024)

Community centrality scores revealed a power-law distribution of influence, consistent with prior findings (Barabási, 2002). Notably, two mid-sized clusters showed anomalous in-degree distributions, suggesting coordinated messaging campaigns.

Cross-community edge density increased during key political events (e.g., debates, rallies), hinting at rising cross-ideological interactions or trolling behavior.

5. Sentiment Analysis Findings

Temporal sentiment analysis uncovered significant polarity swings, often synchronized with political events. Peaks in negative sentiment coincided with controversial policy announcements or viral misinformation, while positive sentiment spiked post-election victory.

Table 1. Sentiment Distribution Across Major Communities

Community ID	Positive (%)	Negative (%)	Neutral (%)
C1 (Party A)	52.3	27.1	20.6
C2 (Party B)	38.9	42.5	18.6
C5 (Youth Activists)	44.7	40.1	15.2

Sentiment divergences were more pronounced in communities with high echo chamber characteristics (low external connectivity). This finding echoes the results of Bail et al. (2018), who studied polarization in digital echo chambers.

Further analysis showed that tweets with misinformation (flagged by external fact-checkers) were associated with a +18% increase in negative sentiment scores, corroborating the hypothesis that disinformation skews emotional responses.

6. Discussion and Implications

This study demonstrates the power of combining structural and emotional perspectives in understanding digital discourse. Network clustering provided a macro-level lens on ideological alignment, while sentiment analysis exposed micro-level emotional undercurrents within those clusters.

The joint analysis revealed emergent phenomena such as **sentiment echoing**, where sentiment patterns reinforce group polarization over time. Furthermore, detecting sentiment shifts enabled early warning of viral misinformation spread or online unrest, which is invaluable for governance and public communication.

These findings contribute to interdisciplinary domains such as computational social science, digital anthropology, and AI ethics, suggesting pathways for more responsible monitoring and engagement with digital publics.

7. Conclusion

By leveraging advanced network analytics and transformer-based sentiment analysis, this study uncovers latent patterns in social media discourse that were previously difficult to identify. The hybrid approach offers a scalable, interpretable framework to study polarization, sentiment diffusion, and coordinated behavior in dynamic digital ecosystems. Future research should explore real-time deployment and cross-platform integration to enhance the scope and impact of such analyses.

References

[1] Bail, Christopher A., et al. "Exposure to Opposing Views on Social Media Can Increase Political Polarization." *Proceedings of the National Academy of Sciences*, vol. 115, no. 37, 2018, pp. 9216–9221.

[2] Devlin, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *Proceedings of NAACL-HLT*, 2019.

[3] Gruhl, Daniel, et al. "Information Diffusion Through Blogspace." *Proceedings of the 13th International Conference on World Wide Web (WWW)*, 2004, pp. 491–501.

[4] Kumar, Shubhanshu, and Kathleen M. Carley. "COVID-19 Misinformation and Network Dynamics." *Social Network Analysis and Mining*, vol. 12, no. 1, 2022, pp. 1–15.

[5] Singh, Rajeev, Yang Zhang, and Tanvi Banerjee. "Integrating GNNs and Transformer Models for Climate Discourse Analysis on Instagram." *Journal of Computational Social Science*, vol. 7, no. 2, 2024, pp. 115–137.

[6] Socher, Richard, et al. "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank." *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2013, pp. 1631–1642.

[7] Zhao, Yaliang, Paul Resnick, and Qiaozhu Mei. "Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts." *Proceedings of the 30th International Conference on World Wide Web (WWW)*, 2021, pp. 1395–1406.

[8] Kwak, Haewoon, et al. "What Is Twitter, a Social Network or a News Media?" *Proceedings of the 19th International Conference on World Wide Web (WWW)*, 2010, pp. 591–600.

[9] Barabási, Albert-László. *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Plume, 2002.

[10] Vaswani, Ashish, et al. "Attention Is All You Need." *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.

[11] Liu, Yinhan, et al. "RoBERTa: A Robustly Optimized BERT Pretraining Approach." *arXiv preprint*, 2019. arXiv:1907.11692.

[12] He, Pengcheng, et al. "DeBERTa: Decoding-enhanced BERT with Disentangled Attention." *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021.

[13] Raganato, Alessandro, and Jörg Tiedemann. "An Analysis of Encoder Representations in Transformer-Based Machine Translation." *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 2993–2998.

[14] Tumasjan, Andranik, et al. "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment." *Proceedings of the Fourth International AAAI Conference on Weblogs and SocialMedia*, 2010, pp. 178–185.

[15] Conover, Michael D., et al. "Political Polarization on Twitter." *Proceedings of the Fifth International AAAI Conference on Weblogs and SocialMedia*, 2011, pp. 89–96.