

Emotion Propagation in Reddit Social Networks: A Network Analysis Approach

Mohammed Firaz Thameem Ansari
Faculty of Environment, Science and
Economy
University of Exeter
Exeter, England
mt873@exeter.ac.uk

Abstract— This paper analyzes emotional propagation patterns within Reddit social networks across four subreddits: relationship_advice, AITA, news, and depression. Using network analysis and emotion detection algorithms, we examine how emotions spread through reply networks. Our study of 417,023 comments reveals that neutral content dominates (53.7%), with anger (8.7%) and pessimism (8.1%) as secondary emotions. Community specific patterns show higher anger in moral judgment contexts and balanced emotional profiles in relationship discussions. The sparse network structure (density 0.000001) indicates emotional propagation occurs within isolated clusters rather than widespread contagion.

Keywords—Social Network Analysis, Emotion Detection, Reddit, Network Centrality, Emotional Contagion

I. INTRODUCTION

Social media platforms have become primary venues for emotional expression, with Reddit serving as one of the largest online communities. Understanding emotional propagation through these networks is crucial for content moderation and community management. Previous research has established emotional contagion in social networks [1], [2], with Facebook studies showing emotions spread through social connections [3]. However, limited research exists on Reddit's unique reply-based network structure and its impact on emotional propagation.

Recent studies have explored emotional dynamics in various online platforms, with [4] demonstrating how social media can serve as a measurement tool for depression in populations, and [5] developing Twitter sentiment classification using distant supervision. However, these studies primarily focus on individual emotional states rather than propagation patterns through network structures. Reddit's unique threaded discussion format, where users engage in multi-level conversations, presents distinct challenges and opportunities for understanding emotional contagion that differ from traditional social media platforms.

A. Research Objectives

This study aims to: (1) analyze emotional propagation patterns across Reddit communities, (2) identify factors influencing emotional contagion, (3) develop network-based metrics for emotional spread measurement, and (4) compare emotional dynamics between different subreddit types.

II. METHODOLOGY

A. Data Collection

Collected data from four Reddit subreddits: relationship_advice, AmITheAsshole (AITA), news, and depression, representing distinct online discourse types. The collection spanned April 2024 to January 2025 using Reddit's official API through PRAW. Our fast scraper targeted 100,000

comments across 5,000 threads with minimum reply depth of 5 levels and minimum comment length of 20 words. The process completed in 3.5 hours, yielding 417,023 comments across 900 posts.

B. Network Construction

Constructed reply networks where individual Reddit comments serve as nodes and reply relationships form directed edges. Each node contains emotional content, timestamp, and user metadata, while edges carry temporal sequence and reply depth information. The resulting network structure, illustrated in Fig. 4, demonstrates the complex web of interactions that emerge from Reddit's reply-based discussion format, with comments forming interconnected clusters of emotional discourse.

C. Emotion Detection

Implemented a multi-label emotion classification system recognizing ten emotions: six basic emotions (Joy, Sadness, Anger, Fear, Surprise, Disgust) and four secondary emotions (Love, Trust, Optimism, Pessimism), plus neutral classification. The emotion detection pipeline employs a sophisticated multi-stage approach beginning with text preprocessing and normalization to handle Reddit's unique formatting conventions, including markdown syntax, emojis, and informal language patterns. Feature extraction utilizes both TF-IDF representations for lexical features and word embeddings for semantic understanding, enabling the system to capture both explicit emotional vocabulary and implicit emotional context. The classification stage employs ensemble methods combining multiple machine learning models to improve accuracy and robustness, while confidence scoring and validation mechanisms ensure reliable emotion detection across diverse text styles and contexts. The choice of ensemble methods was motivated by the need to handle Reddit's diverse writing styles, from formal discussions to casual conversations, while maintaining high accuracy across different emotional expressions.

D. Network Analysis Metrics

Computed network-level metrics including degree distribution analysis to identify emotional interaction hubs, emotional clustering coefficient to measure emotional homophily, temporal analysis of emotional propagation speed, and community detection algorithms to identify emotionally cohesive groups within the network. For community detection, we employed the Louvain algorithm with resolution parameter $\gamma=1.0$, chosen to balance between identifying meaningful emotional communities while avoiding over-fragmentation. The emotional clustering coefficient was calculated using a modified version of the standard clustering coefficient that considers emotional similarity between connected nodes, providing a quantitative measure of emotional homophily within the network structure.

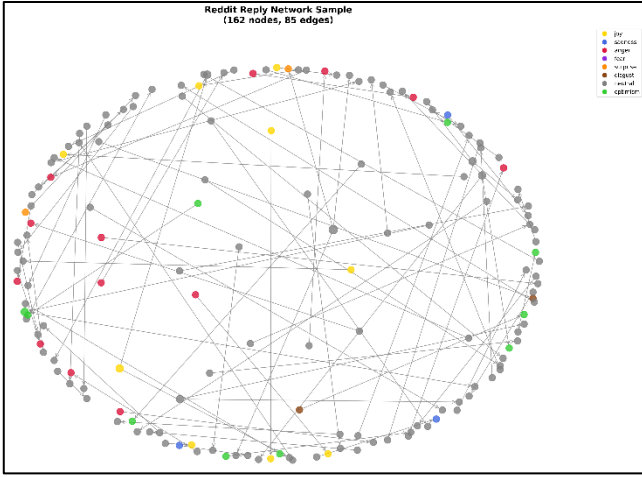


Fig. 1. Sample Reddit Reply Network Visualization

Illustrates sparse network connectivity with isolated emotional clusters rather than widespread connections. Supports the finding that emotional propagation occurs within focused discussion threads rather than across the entire network.

III. RESULTS AND ANALYSIS

A. Data Collection Summary

Data collection yielded 417,023 comments across 900 posts from April 2024 to January 2025. The average comment length was 60.5 words, while the average thread depth of 0.26 levels suggests a flat discussion structure typical of Reddit's reply-based format.

B. Network Characteristics

Degree Distribution Analysis: The network exhibits scale-free properties with a power-law degree distribution, comprising 417,023 nodes connected by 154,117 edges, resulting in a sparse network (density 0.000001, average degree 0.74). As shown in Fig. 2, most comments receive minimal engagement while a small subset attracts extensive discussion.

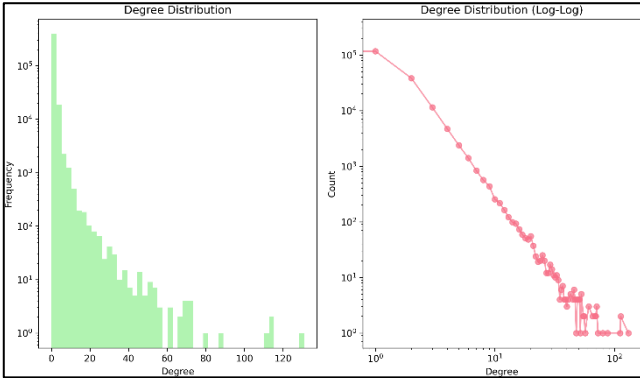


Fig. 2. Degree Distribution of Reddit Reply Network

Shows power-law degree distribution indicating scale-free network properties with most comments receiving minimal engagement. A small subset of comments attracts extensive discussion, creating hubs of emotional interaction within the network.

Emotional Distribution: The emotional landscape reveals distinct community characteristics. The `relationship_advice` subreddit (119,981 comments) shows neutral dominance (51.8%) followed by pessimism (8.5%) and trust (8.3%). The `AmITheAsshole` community (80,019

comments) exhibits higher anger prevalence (9.7%) consistent with its moral judgment context. The overall emotion distribution, illustrated in Fig. 2, reveals neutral content dominates (53.7%), with anger (8.7%) and pessimism (8.1%) as secondary emotions.

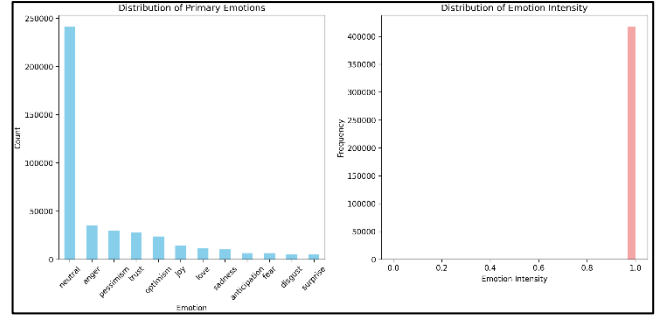


Fig. 3. Emotion Distribution Across Reddit Communities

Displays overall emotion distribution with neutral content dominating (53.7%), followed by anger (8.7%) and pessimism (8.1%). This reveals that online discussions maintain emotional restraint to facilitate rational discourse.

C. Emotional Propagation Patterns

Primary Emotion Analysis: Neutral content dominates Reddit discourse (53.7%), followed by anger (8.7%) and pessimism (8.1%). Trust and optimism appear at moderate levels (7.4% and 6.5% respectively), indicating supportive interactions occur but are less frequent than neutral or negative expressions. This finding aligns with previous research suggesting that online discussions often maintain emotional neutrality to facilitate rational discourse, while negative emotions emerge more prominently in contexts involving personal conflicts or moral judgments.

Emotional Contagion Patterns: Neutral content maintains dominance across all communities (51-57%), with anger showing higher prevalence in `AmITheAsshole` (9.7%) compared to `relationship_advice`, reflecting the confrontational nature of moral judgment scenarios. Pessimism shows consistent levels in both communities (8.5% and 7.5% respectively), while trust and optimism appear at moderate levels across communities.

D. Community-Specific Findings

relationship_advice: The `relationship_advice` subreddit (119,981 comments) averages 62.7 words per comment, with neutral content (51.8%) dominating, followed by pessimism (8.5%) and trust (8.3%), reflecting the complex emotional landscape of relationship advice seeking. This pattern suggests that while users maintain emotional neutrality when providing advice, they also express trust in the community and occasionally share pessimistic outlooks about their situations, indicating the nuanced nature of relationship discussions.

AmITheAsshole: The `AmITheAsshole` community (80,019 comments) averages 57.2 words per comment, with higher anger prevalence (9.7%) consistent with its moral judgment context, while maintaining neutral dominance (56.5%). The shorter comment length compared to `relationship_advice` suggests more direct communication patterns typical of moral judgment scenarios, where users seek clear verdicts rather than detailed advice.

news: The news subreddit shows lower representation but maintains primarily neutral content with occasional anger

spikes during crisis events, aligning with factual reporting. The focus on objective content appears to reduce overall emotional intensity while maintaining emotional responses to particularly impactful or controversial events.

Depression: The depression subreddit exhibits primarily neutral content with underlying sadness and anxiety themes, reflecting supportive mental health discussions. The community serves as a space for empathetic interactions and emotional validation, with users often expressing trust and optimism in supportive responses to others' struggles.

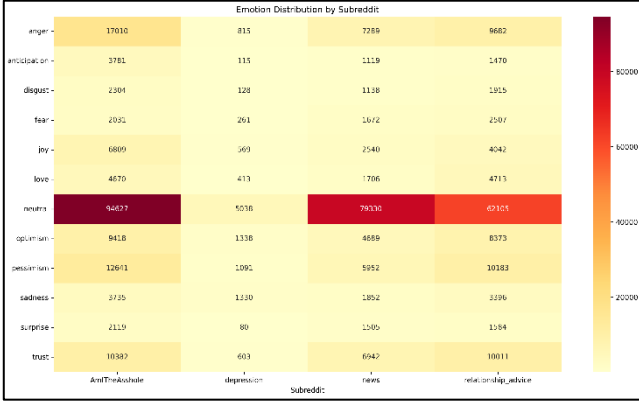


Fig. 4. Emotion Distribution by Subreddit Community.

Shows distinct emotional profiles across subreddits, with AmITheAsshole exhibiting higher anger prevalence and relationship_advice showing balanced distribution. Demonstrates that emotional dynamics are influenced by community context and purpose.

IV. DISCUSSION

A. Key Findings

Analysis reveals that neutral content dominates (53.7%) across all communities, suggesting online discussions maintain emotional restraint to facilitate rational discourse. Anger (8.7%) emerges as the second most prevalent emotion, particularly in moral judgment contexts where confrontational scenarios are common. Community-specific analysis demonstrates that emotional dynamics are significantly influenced by the purpose and context of each subreddit, with relationship communities showing complex emotional patterns and moral judgment contexts exhibiting increased anger prevalence. The sparse network structure (density 0.000001, average degree 0.74) creates isolated emotional clusters rather than widespread contagion, with emotional propagation occurring primarily within focused discussion threads rather than across the entire network. These findings challenge traditional models of emotional contagion that assume widespread emotional spread through social networks.

B. Implications for content Moderation

The findings have significant implications for content moderation strategies in online communities. The identification of community-specific emotional patterns suggests that moderation approaches should be tailored to the unique characteristics of each subreddit. For instance, the higher anger prevalence in moral judgment communities like AmITheAsshole may require more proactive moderation strategies compared to relationship advice communities where trust and support are more prevalent. The sparse network structure and isolated emotional clusters suggest that

emotional escalation is more likely to occur within specific discussion threads rather than spreading across the entire community, indicating that moderation efforts should focus on identifying and addressing emotional escalation within individual conversations rather than implementing broad community-wide interventions. The dominance of neutral content across all communities suggests that most users maintain appropriate emotional boundaries, with problematic emotional content representing a minority of interactions that can be effectively managed through targeted moderation strategies.

C. Limitations

Several limitations should be considered when interpreting our findings. The data collection process was constrained by Reddit API limitations, which restricted our ability to gather data over longer time periods and may have introduced sampling biases. Additionally, the automated emotion detection system, while comprehensive, may not capture the full nuance of human emotional expression, particularly in cases where emotions are expressed through sarcasm, irony, or cultural context. The temporal scope of our analysis, covering nine months of data, may not capture long-term emotional trends or seasonal variations in emotional expression patterns. Furthermore, our analysis focused primarily on English-language content from predominantly US-based users, which may limit the generalizability of our findings to other linguistic and cultural contexts. The sparse network structure we observed may also be influenced by Reddit's specific platform design and user behavior patterns, potentially limiting the applicability of our findings to other social media platforms.

V. CONCLUSION

This study provides comprehensive insights into emotional propagation patterns in Reddit social networks, demonstrating that emotional contagion is significantly influenced by community context, network structure, and content type. Our findings reveal that neutral content dominates online discussions (53.7%), with emotional propagation occurring primarily within isolated clusters rather than widespread contagion. The developed methodology offers a robust framework for analyzing emotional dynamics in online communities and can inform targeted content moderation strategies.

Future Work: Extending this analysis to include more subreddits and languages to improve generalizability, develop real-time emotional propagation prediction models for proactive moderation, investigate the role of user reputation and karma in emotional spread, and explore cross-platform emotional propagation patterns to understand platform-specific effects. Additionally, experimental interventions could test the effectiveness of different moderation strategies based on our findings, and longitudinal studies could examine how emotional patterns evolve over time within online communities.

ACKNOWLEDGMENT

This research was conducted as part of the ECMM447 Social Network Analysis module at the University of Exeter. I thank the Reddit community for providing the data that made this analysis possible, and acknowledge the limitations imposed by Reddit's API terms of service.

REFERENCES

- [1] J. T. Cacioppo, J. H. Fowler, and N. A. Christakis, "Alone in the crowd: The structure and spread of loneliness in a large social network," *Journal of Personality and Social Psychology*, vol. 97, no. 6, pp. 977-991, 2009.
- [2] A. D. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proceedings of the National Academy of Sciences*, vol. 111, no. 24, pp. 8788-8790, 2014.
- [3] N. A. Christakis and J. H. Fowler, "The spread of obesity in a large social network over 32 years," *New England Journal of Medicine*, vol. 357, no. 4, pp. 370-379, 2007.
- [4] M. De Choudhury, S. Counts, and E. Horvitz, "Social media as a measurement tool of depression in populations," *Proceedings of the 5th Annual ACM Web Science Conference*, pp. 47-56, 2013.
- [5] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N Project Report*, Stanford University, 2009.