Probing Information Dissemination Barriers: An Identification Perspective

Group #20: Harsh Victor Challa, Akshay Jayasoorya, Kritshekhar Jha,
Shuchi Talati, Saurabh Zinjad
School of Computing and Augmented Intelligence, Arizona State University
{hvictorc, ajayasoo, kjha9, stalati1, szinjad}@asu.edu

***Abstract*—The rapid proliferation of social media platforms ushered in unprecedented challenges in information dissemination. The objective of this study is the interaction between the three crucial areas of bot identification, comprehension of communities and echo chambers, and influence operation mitigation. While typically treated in isolation, understanding their interconnection offers a more all-encompassing strategy for preserving the integrity of digital communication, promoting discourse, and thwarting the influence of bad actors. In an era where the impact of information is profound, our research helps create a healthier social discourse.**

I. Introduction

The way information is disseminated, digested, and influenced in our society has changed dramatically as a result of the rapid expansion of social media platforms, particularly Twitter. This change, nevertheless, has not been without its share of difficulties. Even the user of today suffers from cognitive decision fatigue due to the presence of other social media platforms like Bluesky, Mastodon, etc. The growth of influential actors who use digital communication for diverse reasons, such as information disinformation and the proliferation of automated bots and echo chambers, are among the most noteworthy.

With its millions of daily users, Twitter serves as a hub for real human contact as well as a breeding ground for automated programs or bots. These bots pose serious dangers to the integrity of public discourse because they are frequently used to spread extreme propaganda and meddle in social affairs [1] [2] [3]. Fake information has primarily drawn recent attention in a political context but it also has been documented in information promulgated about topics such as vaccination, nutrition, and stock values. It is particularly pernicious in that it is parasitic on standard news outlets, simultaneously benefiting from and undermining their credibility [4]. Therefore, effective bot identification systems are urgently needed. Parallel to this, the digital era has given rise to echo chambers—virtual environments where people are primarily exposed to discourse supporting their preexisting opinions while protected from opposing viewpoints [5] [6]. This behavior has significant ramifications since it discourages critical thinking, encourages the spread of false information, and widens socioeconomic gaps. To solve this, current research is focused on effectively measuring the size and influence of echo chambers on social media platforms and investigating ways to lessen their negative consequences.

Additionally, disinformation efforts are being carried out at never-before-seen scales and rates via social media and digital communications [7] [8]. The affordances of social media make it possible to influence operations, which can spread false information quickly and effectively. The proliferation of disinformation on social media has developed from a socio-technical mix of platform design, algorithms, human factors and political and commercial incentives. Actors and technologies involved provide a starting point for targets of governance within an accountability network [9]. These procedures present a significant problem, necessitating the creation of automated systems for their detection, characterization, and mitigation. While each of these issues; detection of social media bots, comprehension of communities or echo chambers, and thwarting online influence operations—has been addressed separately, it is impossible to overlook the underlying connections between them. Automated bots may be essential for sustaining echo chambers and enhancing the effects of influence campaigns. This study aims to close the gap between these three crucial areas of research because it is believed that by tackling them all at once, we can create more thorough and efficient plans for preserving the integrity of digital communication, fostering a diversity of viewpoints, and thwarting the influence of bad actors. By doing this, we support the larger objective of promoting a more resilient and healthy digital environment in a time when information is more powerful and consequential than ever before.

II. Motivation & Problem Statement

The challenges of identifying artificial bots, understanding communities and echo chambers, and countering online influence operations have historically been studied in isolation. However, recognizing their interconnections and devising comprehensive solutions is crucial to enhancing our defenses in the digital information ecosystem. This research aims to explore these areas individually before integrating them into a holistic solution.

- **Communities/Echo Chambers Identification**:

S.No	Task	Task Description	Owner	ETA	Status
1	Project Selection	Explore all the project options shared by the instructor, and select a project.	Team	08/30/2023	Completed
2	Problem Statement Identification	Identify and read different papers on Information dissemination; to understand the current research problem statements and their presence in the various domains.	Team	09/08/2023	Completed
Milestone-0 Identification of the Problem Statement				09/08/2023	Completed
3	Background Study	Dive deeper into the chosen problem statement and explore various approaches to information dissemination and how it leads to communities and echo chambers in social media networks.	Team	10/15/2023	In-Progress
Milestone-1 Project Proposal Submission				09/30/2023	Completed
4	Dataset Exploration	Identification of the dataset	Team	09/30/2023	In-progress
5	Topic Detection	We plan to evaluate the dataset using topic detection. This will help us to identify the key conversation contexts present in the dataset	Team	10/05/2023	In-Progress
6	Community Detection	Objective: To use user embeddings to train a self-supervised GNN; to capture ideological similarities	Kritshekhar, Akshay, Shuchi	10/30/2023	In-Progress
		Code Implementation			-
		Model training & hyperparameter tuning; testing and results.			-
Milestone-2 Phase-1 Implementation					
7	Bot Detection/ Classification	Objective: Using relational GCNN on social graphs encoded with user tweets and user property data	Shuchi, Akshay, Harsh, Saurabh	10/30/2023	In-progress
		Code Implementation			-
		Model training & hyperparameter tuning; testing and results			-
8	Influential actor Classification	Objective: To develop a composite metric combining follower counts and network-based interactions. This method assigns influence scores, gauging each actor's impact within the community.	Saurabh, Harsh, Kritshekhar	10/30/2023	In-progress
		Code Implementation			-
		Model training & hyperparmeter tuning; testing and results.			-
Milestone-3 Phase-II Implementation				10/30/2023	-
9	Breaking the Barrier	An attempt to change the stance of the identified influencial actor; in a zero sum game.	Team	11/27/2023	-
10	Integration	Integration all modules into one comprehensive automated solution.	Team	11/27/2023	-
Milestone-4 Final Test & Demo				11/27/2023	-
10	Final Report	Documentation	Team	12/03/2023	-
Mileostone-5 Project Documentation & Code Submission				12/03/2023	-

TABLE I
PROJECT PLAN, MILESTONES AND TASK DISTRIBUTION

- *Motivation:* Existing methods like analyzing ideological differences between users and their neighbors and studying interaction graphs to understand communities & echo chambers methods have limitations. These graphs [10] do not always indicate communities and challenges with categorizing users or using semi-supervised methods with weak labels [11] [12].
- *Problem Statement:* We plan to study the measures of how similar users are to their own communities and how different they are from users in other communities. Later, use user embeddings to train a self-supervised GNN to capture ideological similarities among users by analyzing their interactions and shared posts.

- **Social Media Bot Identification**:
  - *Motivation:* Deep learning and feature engineering methods can be used to categorize current bot detection techniques [13] [14]. However, the dynamic nature of social media brings two problems: bots impersonating real users and communities & bots operate independently but collaborate as a team to impact individuals negatively.
  - *Problem Statement:* We intend to investigate how the disguise challenge can be addressed by encoding user tweets with pre-trained language models and using diverse user property data. Constructing a diversified graph from the social media network and then using relational graph convolutional networks might help in addressing the challenges.
- **Influential actor Identification**:
  - *Motivation:* Detecting influential actors is vital for understanding community dynamics and network structures. These individuals often serve as key links between groups and impact community development. Identifying them provides valuable insights into community structure and behavior, enhancing our ability to overcome barriers posed by influential actors. [15]
  - *Problem Statement:* We aim to implement the detection of influential actors by formulating a combination of follower and network-based interaction features, such as follows, mentions, retweets, posts, etc., for all users in the dataset. This enables us to generate an impact factor for each actor, & thus to assess their level of influence within the community. [16]
- **Breaking the Barrier**:

– *Motivation:* In order to promote free conversation, lessen polarization, combat radicalization, improve information circulation, and protect society from the effects of online extremism, it is essential to dismantle echo chambers [17] [18].
– *Problem Statement:* We plan to leverage pretrained large language models. We intend to create a zero-sum game between influential actors from opposing communities across echo chambers. This zero-sum game aims to change the stance of other influential actors, thereby contributing to the flow of ideologies and communication.

III. Related Work

Yang et al. [19] applied a random forest algorithm with minimal account metadata to address the problem of bot detection in social media networks. Miller et al. [20] entails extracting 107 elements from user tweets and property information and using anomaly detection algorithms to detect bots. Using the sequence of a user's online actions to find bot groups by analyzing the longest common substrings was studied by Cresci et al. [14]. Kudugunta et al. [13] Suggested a deep learning architecture that incorporates tweet content and metadata. for effective bot detection. Several other kinds of research were focused on using word embeddings and a three-layer BiLSTM [21], using graph convolutional networks [22]. SATAR [23] creates a self-supervised task for learning Twitter user representations and applies it to bot detection with fine-tuning.

Research has explored the echo chamber effect on information dissemination, revealing that certain groups can reach a broader audience on average [24]. By analyzing shared news, An et al. [25] analyzed whether Facebook users were exposed to diverse information. Understanding Echo Chambers (ECs) content, as demonstrated by Calder et el. [26] includes two approaches: measuring stance on a subject and identifying emotions and their intensity. It's worth noting that content-based methods often rely on unsupervised NLP techniques, potentially leading to less accurate labels. The meso-scale approach delves deeper into node divisions using community recognition algorithms to identify echo-chamber-like structures with similar ideological leanings. Villa et al. [27] proposed a hybrid meso-scale and content-based method that was applied to study the presence of ECs in COVID-19-related tweets. By constructing an interaction network and employing the METIS community discovery technique, the network was divided into distinct communities. [28]

Different techniques, such as statistical measures, propagation algorithms, machine learning, topic modeling, similarity measures, and optimization algorithms, are used for Influential actor detection. Panchendrarajan et al. [29] covers topic-based influential emphasizing the importance of labeled data in machine learning methodologies and the effectiveness of topic models in integrating topic and influence modeling. Jain et al. [30] proposed a novel Twitter-based influence measure that combines the relative impact of different features categorized under two varied feature sets and takes into account the contribution of structural centrality measures for computing the feature set. On the other hand, The study by Yang et al. [31] used a three-step methodology to identify influential Twitter users during natural disasters, analyzing the relationship between a user's increase in the number of followers and the content of their tweets, exploring the relative importance of different features for users. Due to the Page rank algorithm's potent ability to automatically rank nodes in a big network, it has been used in the bulk of research studies. Influential user detection approaches use statistical measures, such as the number of reactions to a topic-related post, but fail to capture indirect influence. Chen Avin et al. [32] discuss how to regulate social media to mitigate the problem of echo chambers and break the barriers to the flow of opposing views in social discourse. It argues that the echo chamber problem could be considered a fairness problem and examines the research field of fairness in recommender systems. The paper explores three main types of solutions to the problem of unfair recommender systems: preprocessing solution, in-processing, and post-processing.

IV. Project Plan and Task Allocation

We have carefully divided the project tasks among team members to enable parallel work in order to ensure efficient progress and effective use of our team's resources. We have created a Table I highlighting the organized task distribution; we can maximize our team's skills and ensure that everyone can contribute to the project's success at the same time. Each team member clearly understands their assigned problem statement, which allows for improved collaboration and accountability. By employing this strategy, we hope to accelerate the development of our project and increase overall effectiveness. As the project progress and more granular tasks are identified the table will be updated.

References

[1] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics reports*, vol. 659, pp. 1–44, 2016.
[2] A. Deb, L. Luceri, A. Badaway, and E. Ferrara, "Perils and challenges of social media and election manipulation analysis: The 2018 us midterms," in *Companion proceedings of the 2019 world wide web conference*, 2019, pp. 237–247.
[3] E. Ferrara, "Disinformation and social bot operations in the run up to the 2017 french presidential election," *arXiv preprint arXiv:1707.00086*, 2017.
[4] D. M. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild *et al.*, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
[5] S. Fortunato, "Community detection in graphs," *Physics reports*, vol. 486, no. 3-5, pp. 75–174, 2010.
[6] L. Terren and R. Borge-Bravo, "Echo chambers on social media: A systematic review of the literature," *Review of Communication Research*, vol. 9, pp. 99–118, 2021.
[7] N. Confessore, G. J. Dance, R. Harris, and M. Hansen, "The follower factory," *The New York Times*, vol. 27, 2018.

[8] R. S. Mueller III, "Report on the investigation into russian interference in the 2016 presidential election. volumes i & ii.(redacted version of 4/18/2019)," 2019.

[9] F. Saurwein and C. Spencer-Smith, "Combating disinformation on social media: Multilevel governance and distributed accountability in europe," *Digital journalism*, vol. 8, no. 6, pp. 820–841, 2020.

[10] M. Wolfowicz, D. Weisburd, and B. Hasisi, "Examining the interactive effects of the filter bubble and the echo chamber on radicalization," *Journal of Experimental Criminology*, vol. 19, no. 1, pp. 119–141, 2023.

[11] A. Matakos, E. Terzi, and P. Tsaparas, "Measuring and moderating opinion polarization in social networks," *Data Mining and Knowledge Discovery*, vol. 31, pp. 1480–1505, 2017.

[12] C. Musco, C. Musco, and C. E. Tsourakakis, "Minimizing polarization and disagreement in social networks," in *Proceedings of the 2018 world wide web conference*, 2018, pp. 369–378.

[13] S. Kudugunta and E. Ferrara, "Deep neural networks for bot detection," *Information Sciences*, vol. 467, pp. 312–322, 2018.

[14] S. Cresci, "A decade of social bot detection," *Communications of the ACM*, vol. 63, no. 10, pp. 72–83, 2020.

[15] C. Zheng, Q. Zhang, S. Young, and W. Wang, "On-demand influencer discovery on social media," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. New York, NY, USA: Association for Computing Machinery, 2020.

[16] F. Gammoudi, M. Sendi, and M. N. Omri, "A survey on social media influence environment and influencers identification," *Social Network Analysis and Mining*, vol. 12, no. 1, p. 145, 2022.

[17] A. Tommasel, J. M. Rodriguez, and D. Godoy, "I want to break free! recommending friends from outside the echo chamber," in *Proceedings of the 15th ACM Conference on Recommender Systems*, 2021, pp. 23–33.

[18] F. Masrour, T. Wilson, H. Yan, P.-N. Tan, and A. Esfahanian, "Bursting the filter bubble: Fairness-aware network link prediction," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 01, 2020, pp. 841–848.

[19] K.-C. Yang, O. Varol, P.-M. Hui, and F. Menczer, "Scalable and generalizable social bot detection through data selection," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 01, 2020, pp. 1096–1103.

[20] Z. Miller, B. Dickinson, W. Deitrick, W. Hu, and A. H. Wang, "Twitter spammer detection using data stream clustering," *Information Sciences*, vol. 260, pp. 64–73, 2014.

[21] F. Wei and U. T. Nguyen, "Twitter bot detection using bidirectional long short-term memory neural networks and word embeddings," in *2019 First IEEE International conference on trust, privacy and security in intelligent systems and applications (TPS-ISA)*. IEEE, 2019, pp. 101–109.

[22] S. Ali Alhosseini, R. Bin Tareaf, P. Najafi, and C. Meinel, "Detect me if you can: Spam bot detection using inductive representation learning," in *Companion proceedings of the 2019 world wide web conference*, 2019, pp. 148–153.

[23] S. Feng, H. Wan, N. Wang, J. Li, and M. Luo, "Satar: A self-supervised approach to twitter account representation learning and its application in bot detection," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2021, pp. 3808–3817.

[24] R. Interian, R. G. Marzo, I. Mendoza, and C. C. Ribeiro, "Network polarization, filter bubbles, and echo chambers: an annotated review of measures and reduction methods," *International Transactions in Operational Research*, vol. 30, no. 6, pp. 3122–3158, 2023.

[25] J. An, D. Quercia, and J. Crowcroft, "Partisan sharing: Facebook evidence and societal consequences," in *Proceedings of the second ACM conference on Online social networks*, 2014, pp. 13–24.

[26] F. H. Calderón, L.-K. Cheng, M.-J. Lin, Y.-H. Huang, and Y.-S. Chen, "Content-based echo chamber detection on social media platforms," in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2019, pp. 597–600.

[27] G. Villa, G. Pasi, and M. Viviani, "Echo chamber detection and analysis: a topology-and content-based approach in the covid-19 scenario," *Social Network Analysis and Mining*, vol. 11, no. 1, p. 78, 2021.

[28] E. Cau, V. Morini, and G. Rossetti, "Trends and topics: Characterizing echo chambers' topological stability and in-group attitudes," *arXiv preprint arXiv:2307.15610*, 2023.

[29] R. Panchendrarajan and A. Saxena, "Topic-based influential user detection: a survey," *Applied Intelligence*, vol. 53, 07 2022.

[30] S. Jain and A. Sinha, "Identification of influential users on twitter: A novel weighted correlated influence measure for covid-19," *Chaos, Solitons & Fractals*, vol. 139, p. 110037, 2020.

[31] Y. Yang, C. Zhang, C. Fan, W. Yao, R. Huang, and A. Mostafavi, "Exploring the emergence of influential users on social media during natural disasters," *International Journal of Disaster Risk Reduction*, vol. 38, p. 101204, 2019.

[32] C. Avin, H. Daltrophe, and Z. Lotker, "Breaking the echo chamber effect in social media: An impossibility result," 2023.