Identification and Characterization of Content Traps in YouTube Recommendation Network

Md Monoarul Islam Bhuiyan*, Nitin Agarwal*†
*COSMOS Research Center, University of Arkansas at Little Rock, Arkansas, USA
†International Computer Science Institute, University of California, Berkeley, California, USA
e-mail: {mbhuiyan, nxagarwal}@ualr.edu

***Abstract*—YouTube's recommendation algorithm accounts for a substantial portion of total video views, influencing what users see and engage with. This study investigates how the algorithm may contribute to the formation of content traps, which are clusters of videos that repeatedly expose users to topically similar content. We employ Focal Structure Analysis (FSA), a Social Network Analysis (SNA) approach, to identify structurally cohesive groups of videos within the recommendation network, focusing on the China–Uyghur dataset as a case study. Topic modeling and divergence metrics are used to evaluate the thematic composition of each focal structure, revealing reduced topical diversity in areas where content traps are present. Building on this, we characterize each focal structure by its topical dominance, clustering coefficient, and the relative size of the focal structures, which allows us to distinguish between structurally dense traps and large, loosely connected ones. Our results show that content traps often exhibit strong topical alignment through tightly interconnected nodes. This study contributes a framework for identifying and characterizing content traps and offers insights relevant to understanding algorithmic reinforcement in content recommendation systems.**

Keywords-Content Traps*; *Characterization*; *YouTube Recommendation Network*; *Social Network Analysis.

I. INTRODUCTION

With the increasing influence of social media platforms, content sharing, news consumption, and community interaction have become deeply embedded in everyday digital behavior. YouTube, as the leading video-sharing platform and the second-most visited social media site globally, also plays a central role in this transformation. Operating in over 100 countries and 80 languages [1], YouTube's recommendation algorithm is responsible for 70% of the platform's watch time [2], making it a key driver of user engagement and content exposure. While this algorithm effectively suggests personalized content, it can also lead to the formation of content traps, which are sets of videos that repeatedly promote thematically similar material. This effect is especially concerning in sensitive domains, such as the China–Uyghur, where algorithmic patterns may amplify narrow topical exposure and limit access to diverse perspectives. Understanding how these traps form and persist is essential for evaluating the broader implications of recommendation systems [3].

In this study, we examine the emergence of content traps within YouTube's recommendation network by applying FSA [4], a Social Network Analysis (SNA) technique, to detect cohesive groups of nodes that may reinforce algorithmic exposure. We construct a directed graph based on video recommendation paths and identify focal structures that may act as attractor sets. To evaluate their thematic consistency, we apply topic modeling and measure Jensen–Shannon (JS) [5] and Kullback–Leibler (KL) [6] divergence scores between topic distributions, allowing us to quantify topical uniformity. These measures are particularly well-suited for evaluating topic concentration and distributional shifts in recommendation networks, where small differences in topic probability vectors may indicate the presence of algorithmic bias or thematic redundancy within focal structures. We further characterize each focal structure using structural features, such as average clustering coefficient and the relative size of the focal structures, enabling a multi-dimensional analysis of how content traps differ in form and scale. Thus, combining structural and semantic metrics contributes to a deeper understanding of content reinforcement in recommendation networks and its implications for algorithmic exposure.

To guide our investigation, we address the following research questions:

- **RQ1**. How can content traps be identified through focal structures within YouTube's recommendation network?
- **RQ2**. How can topic modeling and divergence metrics (JS/KL) be used to evaluate the topical consistency of focal structures?
- **RQ3**. How can content traps be characterized based on structural properties (e.g., average clustering coefficient, size) and topical dominance?

The remainder of this paper is structured as follows: Section II summarizes prior work on influential node detection and content traps in social media; Section III details our analytical approach; Section IV presents key findings; and finally Section V outlines implications and future directions.

II. RELATED WORK

This section is divided into two parts. First, we review methods for identifying influential nodes or sets of nodes in social networks. Then, we examine prior work related to content traps and content homogeneity in recommendation systems.

A. Identifying Influential Sets in Social Networks

Identifying structurally significant nodes or sets of nodes is central to Social Network Analysis. Classical methods, such as HITS [7] and PageRank [8] have measured node influence, while community detection techniques have aimed to group similar or densely connected nodes [9]. Moving beyond individual influence, recent work has focused on identifying smaller sets of key players that maximize information flow. FSA, introduced and extended in later studies [10], which identifies compact, relevant subgraphs overlooked by global centrality metrics. Alassad et al. [4] proposed a more comprehensive approach by combining user-level centrality with group-level modularity in a bi-level optimization framework to detect dense and sparse influential structures. Beyond detection, resilience and fragmentation metrics have been used to assess how these structures influence overall network stability [11].

B. Content Traps and Topical Homogeneity

Recommendation algorithms can create content homogeneity, reinforcing user exposure to repetitive themes. This phenomenon, closely related to filter bubbles [12], has been observed in various platforms, such as Facebook [13], and during events like the 2018 Brazilian election [14]. Research has proposed mitigation strategies, including diversification algorithms and fairness-aware link prediction models [15]. In addition, tools have also been developed to raise user awareness of algorithmic bias and promote content diversity [16][17]. Despite these efforts, there remains a gap in systematically identifying and characterizing content traps which are defined here as topically consistent clusters of videos within YouTube's recommendation network. Prior studies have focused primarily on conceptual or behavioral dimensions with limited empirical frameworks. Our study addresses this gap by applying FSA to detect potential traps and by using topic modeling and divergence metrics to evaluate their topical uniformity and diversity. We further assess their structural role in network connectivity, contributing to a clearer understanding of algorithm-driven content reinforcement on large-scale platforms.

To the best of our knowledge, no prior work systematically detects content traps using both structural and topical analyses. In this study, we aim to fill that gap by applying FSA and divergence metrics to YouTube's recommendation network.

III. METHODOLOGY

This section outlines our systematic approach to analyzing the YouTube recommendation network, detecting focal structures, and evaluating the presence of content traps. We start by summarizing our approach in collecting data, dataset background, and building YouTube recommendation networks. After that, we present the network resiliency approach taken to rank key focal structures. In addition, we lay the foundation for the analysis of the topics using the BERTopic model. Lastly, this section explores several metrics to investigate the topical consistency across different topics within the focal structures.

A. Data Collection

The data collection process in this study was designed to systematically capture YouTube's algorithmic behavior through its 'watch-next' recommendations. In this study, we analyzed the China–Uyghur. Below, we provide background details for this context and the motivation for studying them.

1) China–Uyghur Dataset: The situation in Xinjiang centers on the challenges faced by the Uyghur Muslim minority, including cultural repression, ethnic marginalization, and state-driven policies [18]. Scholars have examined the dataset through multiple perspectives, such as identity politics, language regulation, interethnic relations, and movements for greater autonomy [19]. Between 2018 and 2022, the issue gained increased international attention due to growing concerns over human rights violations.

We selected the China–Uyghur dataset for its geopolitical and ideological relevance in examining algorithmic content amplifications and recommendation dynamics within the recommendation network.

2) Keyword Generation and Crawling: We began by organizing workshops with subject matter experts to develop a focused set of keywords associated with the China–Uyghur. These keywords were used as search queries on YouTube to collect an initial set of seed videos. Below is a listing that shows the selected keywords for the collection of our dataset.

- Penindasan/oppression + Uighur/Uyghur
- Kejam/cruel + Uighur/Uyghur
- Saudara muslim/muslim brother + Uighur/Uyghur
- Kalifah/caliph + Uighur/Uyghur
- Khilafah/caliphate + Uighur/Uyghur
- "China is Terrorist"; "Stop Genocide"; "Save Muslim Uyghur"
- "Get Out China"; "I Love Muslim Uyghur"; "Peduli Uyghur" / "Care Uyghur"
- "Bebaskan muslim Uyghur dari penindasan China" / "Free Uyghur Muslims from China's oppression"
- "Do'a kan saudaramu" / Pray for Muslim Uyghur
- Hizbul Tahrir (HTI) + Uighur/Uyghur; Front Pembela Islam (FPI) + Uighur/Uyghur
- Nahdlatul Ulama + Uighur/Uyghur; Muhammadiyah + Uighur/Uyghur
- Hebibulla Tohti + Indonesia; Mohammed Salih Hajim + Indonesia
- Yusuf Martak + Uighur/Uyghur; Slamet Ma'arif + Uighur/Uyghur
- Xiao Qian + Uighur/Uyghur; Pendidikan/education + Uighur/Uyghur

We used a custom crawler to extract YouTube video recommendations up to five levels recursively, balancing data depth with computational feasibility [20][21]. Metadata and engagement statistics were collected via the YouTube Data API, while transcripts were obtained using an external method [22][23].

B. Recommendation Network Construction

The China–Uyghur dataset was constructed by initiating a recursive crawl starting from a curated set of seed videos that were retrieved using targeted keyword queries. YouTube's recommendation system was then used to capture up to four additional hops of recommended videos, resulting in a five-level directed network. This process yielded a graph consisting of 9,748 unique videos and 14,307 directed edges representing recommendation pathways. Our analysis was conducted on the recommendation graph, allowing us to examine how groups of interconnected videos that had contributed to patterns of topical consistency could be identified through focal structure analysis.

C. Focal Structure

Focal Structures (FSs) refer to distinct groups of nodes within a social network that play a central role in shaping influence or coordination. In the context of YouTube, we define focal structures as sets of videos that act as attractor content, potentially reinforcing specific themes and limiting exposure to diverse perspectives, thereby contributing to content traps.

We model the recommendation network as a graph $G = (V, E)$, where V represents videos and E denotes the recommendation links. Focal structures are defined as subgraphs $G' = (V', E')$, with $V' \subseteq V$ and $E' \subseteq E$, and are grouped into a collection $F = \{G'\}$. To ensure distinctiveness, no focal structure in F may fully contain another, i.e., for all $G_i, G_j \in F$, such that $i \neq j$, it holds that G_iG_j and G_jG_i [4]. This constraint guarantees that each focal structure represents a unique, non-overlapping attractor set suitable for analysis.

D. Network Resiliency Assessment

We conducted a network resilience analysis to evaluate the structural importance of focal structures within the recommendation network. Each focal structure was removed from the graph, and the number of resulting clusters measured the resulting network fragmentation. A greater number of disconnected components indicates that the removed structure played a central role in maintaining network cohesion. This method highlights the influence of focal structures in preserving content flow and structural integrity within the recommendation system [24].

E. Topic Modeling with BERTopic

We applied BERTopic [25] to video transcripts to uncover dominant themes associated with each focal structure. BERTopic was selected over traditional models like Latent Dirichlet Allocation (LDA) [26] for its ability to capture semantic and contextual nuances more effectively. Due to BERTopic's input size constraint (512 tokens), transcripts were split into coherent chunks along sentence boundaries. Topics were then mapped back to videos to analyze thematic distribution.

To identify content traps, we used a topic dominance threshold. If a single topic accounted for more than 50% of the videos in a focal structure, it was classified as a content trap. This condition is expressed as:

$$T = \frac{n_{topic}}{n_{total}} > 0.5 \tag{1}$$

where n_{topic} is the number of videos assigned to the dominant topic, and n_{total} is the total number of videos in the focal structure. This approach allowed us to identify clusters exhibiting low content diversity systematically.

F. Divergence Metrics

To further evaluate topical uniformity, we used two statistical measures, namely Kullback-Leibler (KL) Divergence and Jensen-Shannon (JS) Divergence, to compare topic distributions.

1) Kullback-Leibler (KL) Divergence: KL Divergence quantifies how one probability distribution diverges from a reference distribution:

$$D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \tag{2}$$

where P is the topic distribution of a focal structure and Q represents a uniform topical distribution across the topics of the focal structure. Lower values indicate higher similarity and, thus, stronger topical concentration.

2) Jensen-Shannon (JS) Divergence: JS Divergence is a symmetric, bounded variant of KL Divergence, defined as:

$$D_{JS}(P||Q) = \frac{1}{2} \left(D_{KL}(P||M) + D_{KL}(Q||M) \right) \tag{3}$$

$$M = \frac{1}{2}(P + Q) \tag{4}$$

A lower JS Divergence score similarly indicates high topic uniformity. KL and JS Divergence help quantify the degree of thematic consistency within focal structures.

In our analysis, low divergence values indicated content traps, while higher values suggested greater content diversity. These metrics offer a complementary quantitative basis for evaluating the presence of content traps in recommendation networks.

IV. RESULTS

This section presents key factors contributing to content traps in YouTube's recommendation network. We begin by examining focal structures and their impact on network cohesion. We then assess KL and JS divergence metrics to evaluate topical consistency, and conclude by studying the identification and characterization of content traps through topic dominance and structural features.

A. Structural Role of Focal Structures in Network Connectivity

FSA is a network-based method aimed at identifying influential groups of nodes that collectively shape the structure and flow of information. In this study, we applied FSA to the recommendation network built from the China–Uyghur dataset and identified 105 focal structures. We removed each focal structure individually to evaluate its structural significance and analyzed the resulting network fragmentation. An increase in disconnected components following removal indicated a higher structural dependency on that focal structure. This process enabled us to determine the most critical structures supporting network cohesion, with the top five listed in Table I.

 ISBN: 978-1-68558-272-2

TABLE I: KEY METRICS FOR FOCAL STRUCTURES IN THE UYGHUR RECOMMENDATION NETWORK, INCLUDING SIZE, DOMINANT TOPIC, TOPIC UNIFORMITY, AND DIVERGENCE SCORES. STRUCTURES WITH UNIFORMITY ABOVE 50% ARE FLAGGED AS POTENTIAL CONTENT TRAPS.

Datasets	Focal Structure (FS)	No. Videos in FS	No. of Videos in Dominant Topic	No. of Clusters	Topic Uniformity	Content Trap	KL Divergence	JS Divergence
China-Uyghur	3	105	64	185	61%	YES	0.680	0.158
	9	30	17	44	57%	YES	0.004	0.001
	1	25	15	41	60%	YES	0.012	0.003
	102	13	7	31	54%	YES	0.067	0.234
	101	13	5	28	38%	NO	0.154	0.043

B. Topic Uniformity and Content Trap Identification in Focal Structures

To examine thematic concentration within the China–Uyghur recommendation network, we applied BERTopic to extract topics from video transcripts across the identified focal structures. Topic uniformity was measured by calculating the proportion of videos within each structure that shared the most dominant topic. A focal structure was classified as a content trap if over 50% of its videos aligned with a single topic. Focal Structures 3 (FS3) and 9 (FS9) met this criterion with other focal structures, with most of their videos associated with one dominant theme, indicating low topical diversity. This suggests that FS3 and FS9 may contribute to content traps by repeatedly exposing users to a narrow range of content. Table I reports the topic distribution statistics for FS3, FS9, and other key focal structures, while Figures 1 and 2 visualize the concentration of topics across the structure. These findings demonstrate how topic dominance within a focal structure can limit exposure to diverse content and reinforce algorithmically driven content loops, directly addressing **RQ1**.

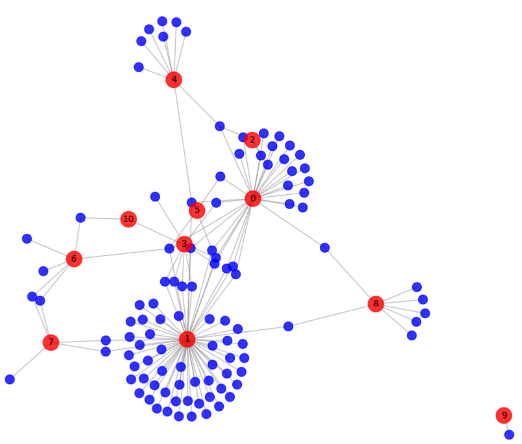

Figure 1: Network visualization of focal structure 3 in the China–Uyghur dataset, with red nodes as topics, blue nodes as video IDs, and edges indicating their associations.

C. Divergence Metrics and Their Role in Identifying Content Traps

In our study, we employed Jensen-Shannon (JS) and Kullback-Leibler (KL) divergence metrics to assess the topical consistency within focal structures. These measures quantify the similarity between topic distributions within a focal structure, which in turn helps us determine the presence of content traps. Low divergence values indicate high topical uniformity, which suggests repetitive content exposure, while higher values reflect greater topic diversity.

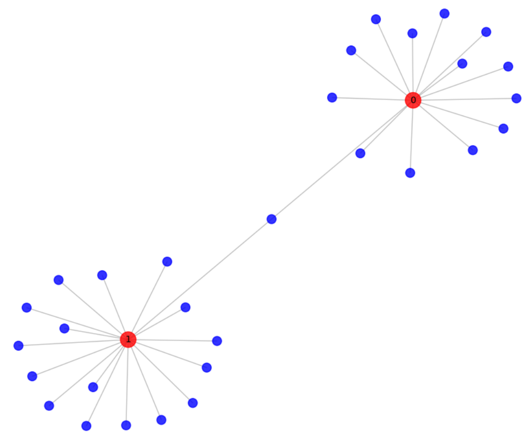

Figure 2: Network visualization of focal structure 9 in the China–Uyghur dataset, with red nodes as topics, blue nodes as video IDs, and edges indicating their associations.

Focal Structures 3 (FS3) and 9 (FS9) in the China–Uyghur dataset exhibited a relatively low JS divergence value alongside a moderately low KL divergence, indicating limited thematic variation across its constituent videos. This combination suggests that, although there is some distributional variability, FS3 and FS9 are still characterized by dominant topics that reduce content diversity. These metrics, reported in Table I, reinforce the classification of FS3 and FS9 as content traps, where algorithmic recommendations predominantly reinforce a narrow thematic scope. This finding contributes to our evaluation of **RQ2**, demonstrating how divergence metrics can reveal the extent of topic concentration within influential structures.

D. Characterizing Content Traps in Focal Structures

To understand the structural and topical properties of content traps within YouTube's recommendation network, we mapped each FS along two axes: topical dominance and either (i) average clustering coefficient or (ii) size, represented by the number of constituent nodes. This enabled a quadrant-based interpretation to characterize focal structures according to their structural cohesion and topical uniformity, both of which indicate their potential to function as content traps.

In the first analysis, as shown in Figure 3, we observe that Q1, representing focal structures with high topical dominance and average clustering coefficient, exhibits a dense aggregation

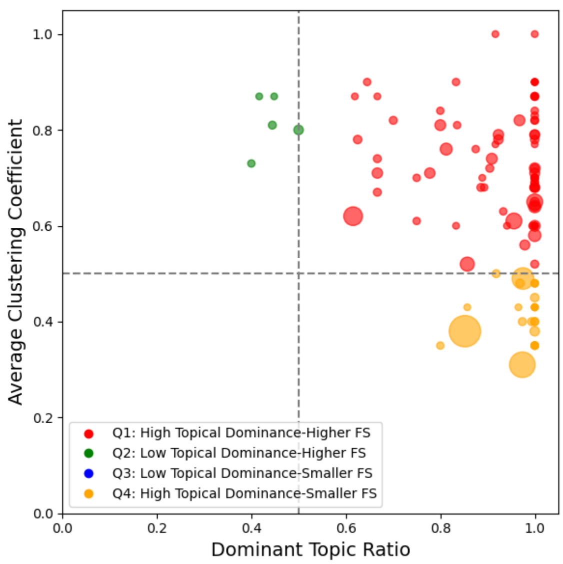

Figure 3: Characterization of content traps by dominant topic ratio and average clustering coefficient, illustrating structural and topical properties of focal structures.

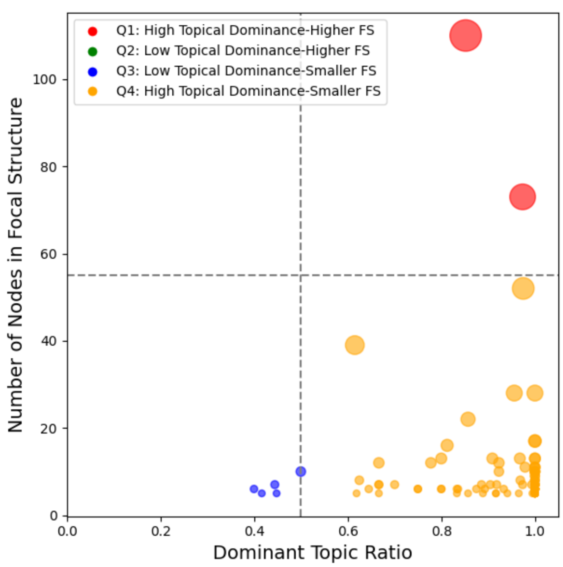

Figure 4: Content trap characterization by topic dominance and focal structure size.

of red-coded nodes. These structures are thematically uniform and structurally cohesive, forming tight-knit recommendation loops. Their high clustering suggests redundancy and limited escape routes for users, reinforcing their potential to function as strong content traps. Conversely, Q2 (low topical dominance, high clustering; green) reflects structurally dense but thematically diverse structures. These may serve as hubs of varied content but are less likely to trap users within a single narrative. Additionally, Q4 (high topical dominance, low clustering; orange) comprises topically consistent but loosely connected structures. Despite weak cohesion, some of the largest focal structures appear here, suggesting that scale and topical uniformity alone can sustain content traps. In other words, even without dense connectivity, large and uniform structures can act as broad-reaching traps. These may still act as traps due to content repetition with less structural reinforcement.

In the second analysis, as shown in Figure 4, we replace the clustering coefficient with the size of the focal structure to assess how node count interacts with topical dominance. Q1 again highlights high-risk traps: large, topically homogeneous structures dominate this quadrant. Their size and thematic alignment indicate both reach and reinforcing potential. In contrast, Q4 reveals numerous small, topically concentrated clusters, which may act as micro-traps that are limited in reach but still repetitive in exposure. Q2 is absent in this plot, reinforcing the rarity of large, thematically diverse structures. Q3 appears minimally, further supporting that small, diverse clusters are less likely to retain user attention. Together, these quadrant analyses suggest that content traps are best characterized by the convergence of structural density and topical uniformity, particularly in large focal structures. These insights support the development of targeted mitigation strategies that disrupt topical alignment (e.g., via content diversification) or structural reinforcement (e.g., reducing internal clustering). Furthermore, this structural–topical characterization lays the groundwork for interpreting how such traps may interact with user behavior, especially in the context of elevated engagement observed in high-uniformity structures, thus contributing to our understanding of **RQ3**.

V. Conclusion and Future Work

This study presented a network-based approach for identifying and characterizing content traps within YouTube's recommendation system, with a focus on the China–Uyghur context. By applying FSA, we extracted cohesive sets of videos that function as attractor content within the recommendation network. Through topic modeling and information divergence metrics (JS and KL), we evaluated topical uniformity across focal structures, revealing clusters with limited thematic variation. Our characterization further incorporated structural properties, such as clustering coefficient and size, enabling a nuanced understanding of how content traps differ in form and intensity. Engagement metrics provided additional support, highlighting user interactions that may reinforce the persistence of these traps. Our findings show that content traps are not solely defined by structural cohesion; even large, loosely connected focal structures can exhibit strong topical alignment and influence user navigation. This underscores the need to consider both network structure and content semantics in assessing algorithmic influence on content exposure.

Future work will focus on extending our analysis to include more network structure dimensions and content dimensions, and more topics/datasets/platforms to evaluate the generalizability of our findings. Additionally, we aim to compare our focal structure-based approach with other SNA methods to effectively identify content traps. We plan to integrate semiotic analysis [27] to examine how symbols impact the formation and reinforcement of content traps, and to explore content infusion strategies as a means of mitigating these effects.

ACKNOWLEDGEMENTS

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920), U.S. Office of the Under Secretary of Defense for Research and Engineering (FA9550-22-1-0332), U.S. Army Research Office (W911NF-23-1-0011, W911NF-24-1-0078), U.S. Office of Naval Research (N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Laboratory, U.S. Defense Advanced Research Projects Agency, the Australian Department of Defense Strategic Policy Grants Program, Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment, and the Donaghey Foundation at the University of Arkansas at Little Rock. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

REFERENCES

[1] C. Agency, *Youtube statistics, https://www.charleagency.com/articles/youtube-statistics/*, Accessed: April 10, 2025.

[2] A. Gallagher, L. Cooper, R. Bhatnagar, and C. Gatewood, *Pulling back the curtain: An exploration of youtube's recommendation algorithm, https://www.isdglobal.org/isd-publications/pulling-back-the-curtain-an-exploration-of-youtubes-recommendation-algorithm/*, Accessed: March 17, 2025.

[3] M. I. Gurung, M. M. I. Bhuiyan, A. Al-Taweel, and N. Agarwal, "Decoding youtube's recommendation system: A comparative study of metadata and gpt-4 extracted narratives," in *Companion Proceedings of the ACM on Web Conference 2024*, 2024, pp. 1468–1472.

[4] M. Alassad, M. N. Hussain, and N. Agarwal, "Comprehensive decomposition optimization method for locating key sets of commenters spreading conspiracy theory in complex social networks," *Central European Journal of Operations Research*, vol. 30, no. 1, pp. 367–394, 2022.

[5] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information theory*, vol. 37, no. 1, pp. 145–151, 1991.

[6] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[7] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM)*, vol. 46, no. 5, pp. 604–632, 1999.

[8] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1-7, pp. 107–117, 1998.

[9] P. Bedi and C. Sharma, "Community detection in social networks," *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, vol. 6, no. 3, pp. 115–135, 2016.

[10] F. Şen, R. Wigand, N. Agarwal, S. Tokdemir, and R. Kasprzyk, "Focal structures analysis: Identifying influential sets of individuals in a social network," *Social Network Analysis and Mining*, vol. 6, pp. 1–22, 2016.

[11] M. J. Alenazi and J. P. Sterbenz, "Comprehensive comparison and accuracy of graph metrics in predicting network resilience," in *2015 11th international conference on the design of reliable communication networks (DRCN)*, IEEE, 2015, pp. 157–164.

[12] E. Pariser, *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

[13] A. Bechmann and K. L. Nielbo, "Are we exposed to the same "news" in the news feed? an empirical analysis of filter bubbles as information similarity for danish facebook users," *Digital journalism*, vol. 6, no. 8, pp. 990–1002, 2018.

[14] G. M. Lunardi, G. M. Machado, V. Maran, and J. P. M. de Oliveira, "A metric for filter bubble measurement in recommender algorithms considering the news domain," *Applied Soft Computing*, vol. 97, p. 106 771, 2020.

[15] 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, 2020, pp. 841–848.

[16] A. Amrollahi, "A conceptual tool to eliminate filter bubbles in social networks," *Australasian journal of information systems*, vol. 25, pp. 1–16, 2021.

[17] L. Burbach, P. Halbach, M. Ziefle, and A. Calero Valdez, "Bubble trouble: Strategies against filter bubbles in online social networks," in *Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Healthcare Applications: 10th International Conference, DHM 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part II 21*, Springer, 2019, pp. 441–456.

[18] A. M. Dwyer, "The xinjiang conflict: Uyghur identity, language policy, and political discourse," *Policy Studies*, vol. 15, 2005.

[19] R. Hasmath, "What explains the rise of majority–minority tensions and conflict in xinjiang?" *Central Asian Survey*, vol. 38, no. 1, pp. 46–60, 2019.

[20] M. C. Cakmak, O. Okeke, U. Onyepunuka, B. Spann, and N. Agarwal, "Investigating bias in youtube recommendations: Emotion, morality, and network dynamics in china-uyghur content," in *International Conference on Complex Networks and Their Applications*, Springer, 2023, pp. 351–362.

[21] M. C. Cakmak, N. Agarwal, and R. Oni, "The bias beneath: Analyzing drift in youtube's algorithmic recommendations," *Social Network Analysis and Mining*, vol. 14, no. 1, p. 171, 2024.

[22] Google Developers, *YouTube Data API*, https://developers.google.com/youtube/v3, Accessed: April. 10, 2025.

[23] M. C. Cakmak and N. Agarwal, "High-speed transcript collection on multimedia platforms: Advancing social media research through parallel processing," in *2024 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2024, pp. 857–860. DOI: 10.1109/IPDPSW63119.2024.00153.

[24] M. M. I. Bhuiyan, S. Shajari, and N. Agarwal, "Resilience and node impact assessment in youtube commenter networks leveraging focal structure analysis," *The Eleventh International Conference on Human and Social Analytics (HUSO 2025)*, 2025.

[25] M. Grootendorst, "Bertopic: Neural topic modeling with a class-based tf-idf procedure," *arXiv preprint arXiv:2203.05794*, 2022.

[26] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[27] M. I. Gurung, N. Agarwal, and M. M. I. Bhuiyan, "How does semiotics influence social media engagement in information campaigns?" In *Proceedings of the 58th Hawaii International Conference on Systems Science (HICSS-58)*, Big Island, Hawaii: Hawaii International Conference on System Sciences, Jan. 2025, pp. 6474–6483.