Characterizing Suspicious Commenter Behaviors

Shadi Shajari, Mustafa Alassad, Nitin Agarwal
COSMOS Research Center, University of Arkansas - Little Rock, USA
Email: sshajari@ualr.edu, mmalassad@ualr.edu, nxagarwal@ualr.edu

***Abstract*—YouTube has revolutionized content consumption and global user interaction. It has become a central hub for video sharing, entertainment, and information dissemination. However, as the user base continues to expand and actively engage with the platform, concerns have arisen regarding the presence of suspicious behavior among commenters. This study presents an approach based on social network analysis to detect suspicious commenter behaviors and identify similarities across various YouTube channels in relation to such behaviors. The analysis involves 20 YouTube channels that disseminated false views about the U.S. Military. The dataset included 7,782 videos, 294,199 commenters, and 596,982 comments. We employ a combination of methods, including Graph2vec, UMAP, K-means, Hierarchical clustering, qualitative and quantitative analyses. The objective is to categorize channels based on the level of suspicious behavior and reveal common patterns exhibited among them. To assess the effectiveness of the proposed methodology, the outcomes revealed the presence of commenter mobs and significant similarities among these channels, providing valuable insights into the prevalence of suspicious commenter behavior.**

***K*eywords–Social Network Analysis, YouTube, Commenter Network Analysis, Graph2vec, UMAP, Suspicious Behaviors.**

I. Introduction

YouTube has revolutionized online content consumption and interaction. It boasts a vast and diverse user base that encourages creators and viewers to connect. However, a notable issue is the presence of commenter mobs in the comment sections. These groups of commenters collaborate to artificially boost engagement on videos, which can harm user interactions and the platform's authenticity.

To address the problem of identification of commenter mobs and to better understand such problematic behavior; this paper presents a social network analysis-based methodology for analyzing 20 YouTube channels that spread false information about the U.S. Military. These channels are selected according to the similar patterns that were observed such as suspicious thumbnails and titles suggesting U.S. military aggression towards Russia and China, pictures of world leaders reacting to these events and some videos having misleading commentary while other videos having titles and thumbnails that are completely unrelated to the video content.

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

ACM ISBN 979-8-4007-0409-3/23/11…$15.00
https://doi.org/10.1145/3625007.3627309

Our goal is to identify similarities between these channels, we employed combinations of advanced techniques, such as Graph2vec [11] a powerful tool for graph representation learning, and Uniform Manifold Approximation and Projection (UMAP), a powerful nonlinear dimension reduction technique that excels at capturing complex patterns in data, especially in graph-related scenarios. To preserve local and global structures, UMAP is ideal for graph-based tasks, including dimensionality reduction [18]. In addition, we used clustering methods such as K-means and Hierarchical clustering [1] to observe similar channels behavior. Furthermore, we conducted both qualitative and quantitative analyses to assess the quality of the identified clusters [2]. The analysis revealed significant similarities among the channels and shedding light on the prevalence of suspicious commenter behavior.

The remaining sections of the paper are structured as follows: In section II, we provide a review of suspicious behavior on YouTube, along with a discussion on Graph2vec, UMAP, and the current understanding of the topic. Section III outlines the data collection methods. It also describes the methodology deployed in the study. Section IV presents the study's findings, including an in-depth analysis of the data, commenter and channel behavior. Finally, section V summarizes the research with directions for future research.

II. Related Studies

We classify the related works into three main studies, namely, Suspicious Behavior on YouTube, Graph2vec and UMAP to cover the presented research topics.

Suspicious Behavior on YouTube. There is not much work on suspicious behavior detection on YouTube. We identified three major studies that come close to computational analysis of suspicious commenter behavior detection on YouTube. We summarize these studies next. Shajari et al. (2023) employed PCA and clustering methods to analyze co-commenter networks on YouTube channels, the current study takes a different approach that distinguishes itself by employing Graph2vec and UMAP [7]. Hussain et al. (2018), proposed an engagement scores approach to identify inorganic behaviors on YouTube networks [5]. Likewise, Kirdemir et al. (2022), proposed an unsupervised methodology that combines multiple layers of analysis to explore coordinated inauthentic behavior assessment on YouTube [20].

Graph2vec. is a powerful tool for graph representation learning that uses the Weisfeiler-Leman Subtree Kernel algorithm [3] to identify suspicious commenter behavior by capturing graph attributes and structure, representing labels as fixed-length vectors. Narayanan et al. (2017) applied Graph2vec in a deep learning framework to analyze social networks, extracting information about centrality values, neighborhoods, and structural details [11]. Yanardag et al. (2015) incorporated comparable techniques using the Wisfeiler-Lehman Subtree Kernel, Grapghlet Kernels, and Shortest-Path graph kernels to capture structural information related to the large-scale graphs and their similarities in the social network domain [9]. In addition, Ribeiro et al. (2017) introduced the struc2Vec framework to employ nodes' structural identity and neighborhood information to enhance node representation in graphs and boost classification performance [14].

Uniform Manifold Approximation and Projection (UMAP) is a potent dimension reduction technique capable of visualizing structural information similarity in high-dimensional datasets [18], performing general nonlinear dimension reduction [17], and an approach to capture the global structure of high-dimensional data accurately as presented in [19]. Furthermore, UMAP is constructed from a theoretical framework based on riemannian geometry and algebraic topology, where UMAP has no computational restrictions on the embedding dimension [16]. It is important to mention that the choice of datasets depends on the research questions, application domain, and availability of the data. In contrast to the existing studies mentioned in this section, our approach employs advanced graph analysis techniques such as Graph2vec to gain a deeper understanding of suspicious commenter behavior and address the problem of identifying commenter mobs on YouTube.

III. METHDOLOGY

In this section, we explore the process of data collection and then introduce our proposed approach.

Data Collection. Subject matter experts helped us select a list of 20 channels that promoted false views of the U.S. military. Python-based multi-threaded script [21] was used to collect data from these channels using YouTube APIs. Several data elements from these channels were collected, including 7,782 videos, 596,982 comments and 294,199 commenters.

Creating Co-commenter Network. This network comprises edges between commenters who comment on the same video in one or more channels [20] where the edges are governed by a threshold [7]. The weight of these edges is determined by the number of shared videos on which they have commented. To utilize this network in the model, we exported it as a gexf file, which is a format based on XML that facilitates the storage of a singular graph structure [10].

Proposed Approach. After creating a co-commenter network, we delve into the concepts of Graph2vec and UMAP. Subsequently, our model employs both K-means and hierarchical clustering methods [1]. Finally, we use the resulting embeddings to cluster the co-commenter networks based on their similarity and evaluate the quality of clusters using various measures presented in section IV.

For example, Figure 1 represents the co-commenter network of the USA Military Channel (left) and the co-commenter network of the US Military System (right). In both co-commenter networks, (A) corresponds to the co-commenter network of the entire channel, and (B, C, and D) represent the three largest commenter mobs. Moreover, the structures of these communities indicate leaders and followers in organizations based on the degree centrality values.

Graph2vec. By encoding structural information into fixed-length feature vectors, Graph2vec effectively learns continuous vector representations of graphs. This enables us to leverage powerful machine learning techniques on the resulting embeddings, facilitating various downstream tasks such as clustering and visualization as proposed in [11]. For this research, Graph2vec utilized the Weisfeiler-Leman Algorithm (WLA) [15], to analyze the co-commenter networks.

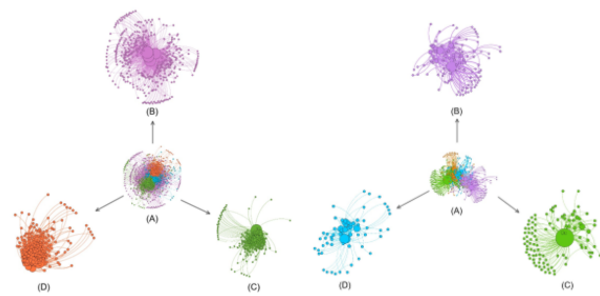

Fig. 1. USA Military Channel's (left) and US Military System (right) co-commenter networks

The WLA is a graph isomorphism test that provides a powerful means to analyze complex networks. WLA iteratively labels nodes and considers the local neighborhood structures. This iterative labeling process allows to capture intricate details of complex patterns within the co-commenter networks, enabling a comprehensive understanding of the relationships and interactions among the commenters. For this purpose, the WLA's key parameters were carefully selected; for example, firstly, to identify structural patterns beyond the immediate neighborhoods, the WLA's iteration was set at 2. Secondly, to capture balanced information from the structures, dimensionality was set at 128. Thirdly, to obtain a stable optimization process, the learning rate was set at 0.025. Finally, a minimum word count of 5 was utilized for semantic understanding.

Using these parameters, the fit method of the Graph2vec object trained the model on multiple co-commenter networks. These parameters were selected based on empirical evaluation and are not arbitrary. While we have strived to find the optimal values, our experimentation and validation procedures confirm that the selected parameter values were in an appropriate balance for our analysis and the anticipated outcomes levels.

Moreover, WLA's procedure creates hashed neighborhood labels [3]. To comprehend these labels in the co-commenter network, the TaggedDocument objects and the Doc2Vec model were employed to generate and learn the continuous vector representations of the co-commenter networks [4]. To obtain a meaningful visualization of the outcomes, UMAP method was utilized, as explained below.

UMAP. We carefully selected all default parameters to ensure optimal performance and meaningful visualizations. The procedure includes the number of neighbors, set to 5, to balance local structure preservation; the minimum distance between points was set to 0.1, for appropriate spread and separation; and the number of components was set to 4, to capture significant structural information and facilitate the visualization. Additionally, the standard Euclidean distance metric was employed to successfully observe the performance of the UMAP algorithm in reducing dimensionality, preserving the global structure, and enabling accurate visualization and clustering of the co-commenter networks.

IV. Result

This section presents the results of the clustering analyses using the K-means and hierarchical clustering methods [1]. To further assess the quality of the outcomes, we conducted both qualitative and quantitative clusters quality analysis [6].

K-Means and Hierarchical Clustering. Two distinct clusters were identified by both techniques, as shown in Figure 2. The optimal number of clusters for K-means was determined using the silhouette score, which measures the similarity among channels within each cluster as described in [13]. Similarly, for hierarchical clustering as presented in [22], the single linkage method was employed to create a dendrogram showing the closeness of each channel to others. The cut-tree method was then used to determine the optimal number of clusters and separate the channels based on their similarity.

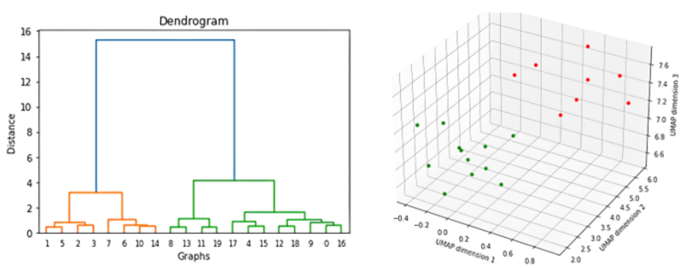

Fig. 2. Channel categorization using k-means (right) and hierarchical clustering (left).

Figure 2 represents the results of K-means clustering (right), the green group corresponds to cluster 0, while the red group corresponds to cluster 1. Visualizations, such as scatter plots, revealed well-separated clusters, indicating effective classification of data points within each cluster. This demonstrates the quality and accuracy of the clustering approaches in grouping similar channels together accurately. On the left side, the plot presents the results obtained from hierarchical clustering, the channel index refers to the unique channel ID assigned to each channel (graph) in our dataset.

Qualitative and Quantitative Cluster Quality Analysis. The qualitative analysis required visually examining the scatter plots to observe the clear distinction and unity among the clusters, as shown in Figure 2. It also involved assessing the similarity of videos shared among the channels within each cluster. Moreover, the quantitative analysis concerned calculating relevant metrics such as the cophenetic correlation coefficient [12], and the Davies-Bouldin index [8]. These measures provide objective indicators of how well the clustering algorithm successfully grouped similar channels together and separated dissimilar ones.

In the quantitative cluster quality analysis, we assessed the performance of the clustering algorithm using two measures. The first measure, the cophenetic correlation coefficient, yielded a score of 0.82. A higher score indicates better separation between the clusters, indicating that the algorithm successfully grouped similar channels together. The second measure, the Davies-Bouldin index, resulted in a score of 0.44. A lower score indicates better separation between the clusters, suggesting that the algorithm effectively differentiated dissimilar channels. Furthermore, to analyze the distinctive features of the identified clusters, we investigated the presence of cliques within the co-commenter network and specifically focused on cliques with at least five members. Cliques raised suspicions because they displayed strong connections and coordinated actions among commenters of a channel. By examining these suspicious commenter cliques, we could investigate individual channels to determine if they are involved in suspicious or coordinated commenters' mob-like activities.

In our analysis, we compared cluster 0, and cluster 1 to understand the presence of suspicious channels within each cluster. Cluster 0 emerged as the cluster with channels comprising of commenters with the highest number of cliques containing at least five members, indicating a significant concentration of suspicious channels. Within cluster 0, we identified channels such as, "Armageddon TV" with 12,241, "USA Military Channel" with 9,246, and "USA Military Channel 2" with 782 number of cliques as the top three most suspicious channels.

These findings highlight the potential for coordinated activities and interconnectedness within cluster 0. In contrast, cluster 1 exhibited a smaller number of cliques, with Channels "US Military System", "Brittany Lewis", and "U.S. Military Technology" being the top three most suspicious channels, possessing 26, 24, and 20 cliques respectively. Although the number of cliques in cluster 1 is relatively lower than in cluster 0, it still raises concerns regarding potential suspicious activities within this cluster.

In the context of qualitative analysis, we observed distinctive traits in the identified clusters by closely examining YouTube channels and their content. Cluster 0 includes larger channels claiming official news, focusing more on US military news compared to news from other countries. These insights provide valuable information about the composition and unique characteristics of channels in each cluster, helping us understand their differences and similarities. Figure 3 presents example videos representing the channels within cluster 0.

Fig. 3. Thumbnail images of a video from USA Military channel (left) and USA Military channel 2 (right).

Moreover, cluster 1 includes smaller channels associated with non-official news sources and personal accounts. These channels mainly featured videos about the US military and covered topics related to countries like Ukraine, and Russia. This highlights the diversity of content and perspectives available on YouTube, ranging from individual viewpoints to alternative sources of news. Figure 4 illustrates the sample videos that depict the channels belonging to cluster 1.

Fig. 4. Thumbnail images of a video from ACTUAL NEWS TODAY (left) and NICKY MGTV (right).

V. Conclusion and Future Works

In this study, we propose a methodology for identifying and detecting suspicious commenter behaviors of 20 YouTube channels that disseminate false information about the U.S military. Our approach involves utilizing Graph2vec to learn embeddings for the networks, enabling accurate clustering based on similarity. Additionally, we incorporated dimensionality reduction using UMAP to enhance the clustering process. Through qualitative and quantitative analysis, we comprehensively evaluated the clustering algorithm, validating the quality of the resulting clusters. These findings underscore the significance of leveraging advanced techniques and fine-tuning parameters for effective network analysis in the context of YouTube co-commenter networks.

In future research, we aim to expand our model's predictive capabilities to larger datasets, study the link between channel characteristics and suspicious activities (especially clique numbers), and thoroughly investigate content creators in suspicious channels to enhance our approach to countering suspicious behaviors on YouTube

Acknowledgments

This research is funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920, IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of the Under Secretary of Defense for Research and Engineering (FA9550-22-1-0332), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2675, N00014-17-1-2605, N68335-19-C-0359, N00014-19-1-2336, N68335-20-C-0540, N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Laboratory, U.S. Army Research Office (W911NF-20-1-0262, W911NF-16-1-0189, W911NF-23-1-0011), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059), Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094).

References

[1] Aastha Gupta, Himanshu Sharma, and Anas Akhtar, "A COMPARATIVE ANALYSIS OF K-MEANS AND HIERARCHICAL CLUSTERING," EPRA International Journal of Multidisciplinary Research (IJMR), pp. 412–418, Sep. 2021, doi: 10.36713/epra8308.
[2] V. Osinska, "A qualitative–quantitative study of science mapping by different algorithms: The Polish journals landscape," Journal of Information Science, vol. 47, no. 3, pp. 359–372, Jun. 2021, doi: 10.1177/0165551520902738.
[3] N. Huang and S. Villar, "A Short Tutorial on The Weisfeiler-Lehman Test And Its Variants," in ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Jun. 2021, pp. 8533–8537. doi: 10.1109/ICASSP39728.2021.9413523.
[4] J. H. Lau and T. Baldwin, "An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation." arXiv, Jul. 18, 2016. Accessed: May 04, 2023. [Online]. Available: http://arxiv.org/abs/1607.05368.
[5] M. N. Hussain, S. Tokdemir, N. Agarwal, and S. Al-Khateeb, "Analyzing Disinformation and Crowd Manipulation Tactics on YouTube," in 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona: IEEE, Aug. 2018, pp. 1092–1095. doi: 10.1109/ASONAM.2018.8508766.
[6] D. Henry, A. B. Dymnicki, N. Mohatt, J. Allen, and J. G. Kelly, "Clustering Methods with Qualitative Data: a Mixed-Methods Approach for Prevention Research with Small Samples," Prev Sci, vol. 16, no. 7, pp. 1007–1016, Oct. 2015, doi: 10.1007/s11121-015-0561-z.
[7] S. Shajari, N. Agarwal, and M. Alassad, "Commenter Behavior Characterization on YouTube Channels," 2023, doi: 10.48550/ARXIV.2304.07681.
[8] J. Xiao, J. Lu, and X. Li, "Davies Bouldin Index based hierarchical initialization K-means," IDA, vol. 21, no. 6, pp. 1327–1338, Nov. 2017, doi: 10.3233/IDA-163129.
[9] P. Yanardag and S. V. N. Vishwanathan, "Deep Graph Kernels," in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney NSW Australia: ACM, Aug. 2015, pp. 1365–1374. doi: 10.1145/2783258.2783417.
[10] GEXF File Format. [retrieved: July, 2023]. [Online]. Available: https://gexf.net
[11] A. Narayanan, M. Chandramohan, R. Venkatesan, L. Chen, Y. Liu, and S. Jaiswal, "graph2vec: Learning Distributed Representations of Graphs." arXiv, Jul. 17, 2017. Accessed: Jun. 05, 2023. [Online]. Available: http://arxiv.org/abs/1707.05005
[12] J. S. Farris, "On the Cophenetic Correlation Coefficient," Systematic Zoology, vol. 18, no. 3, p. 279, Sep. 1969, doi: 10.2307/2412324.
[13] L. Lovmar, A. Ahlford, M. Jonsson, and A.-C. Syvänen, "Silhouette scores for assessment of SNP genotype clusters," BMC Genomics, vol. 6, no. 1, p. 35, Dec. 2005, doi: 10.1186/1471-2164-6-35.
[14] L. F. R. Ribeiro, P. H. P. Saverese, and D. R. Figueiredo, "struc2vec: Learning Node Representations from Structural Identity," in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax NS Canada: ACM, Aug. 2017, pp. 385–394. doi: 10.1145/3097983.3098061.
[15] M. Grohe, M. Lichter, and D. Neuen, "The Iteration Number of the Weisfeiler-Leman Algorithm," 2023, doi: 10.48550/ARXIV.2301.13317.
[16] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform Manifold Approximation and Projection," JOSS, vol. 3, no. 29, p. 861, Sep. 2018, doi: 10.21105/joss.00861.
[17] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," 2018, doi: 10.48550/ARXIV.1802.03426.

[18] B. Ghojogh, A. Ghodsi, F. Karray, and M. Crowley, “Uniform Manifold Approximation and Projection (UMAP) and its Variants: Tutorial and Survey.” arXiv, Aug. 24, 2021. Accessed: Jun. 05, 2023. [Online]. Available: http://arxiv.org/abs/2109.02508.

[19] H. Jeon, H.-K. Ko, S. Lee, J. Jo, and J. Seo, “Uniform Manifold Approximation with Two-phase Optimization,” 2022, doi: 10.48550/ARXIV.2205.00420.

[20] B. Kirdemir, O. Adeliyi, and N. Agarwal, “Towards Characterizing Coordinated Inauthentic Behaviors on YouTube,” Apr. 2022.

[21] J. Kready, S. A. Shimray, M. N. Hussain, and N. Agarwal, “YouTube Data Collection Using Parallel Processing,” in 2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), New Orleans, LA, USA: IEEE, May 2020, pp. 1119–1122. doi: 10.1109/IPDPSW50202.2020.00185

[22] F. Nielsen, “Hierarchical Clustering,” in Introduction to HPC with MPI for Data Science, Cham: Springer International Publishing, 2016, pp. 195–211. doi: 10.1007/978-3-319-21903-5_8.