Identifying and Analyzing Communities in YouTube Comments

Abstract

YouTube is the largest video streaming platform on the web, attracting billions of users daily who watch and engage with content through comments. A significant portion of these users consist of children and teenagers, who frequently interact with one another in the comment sections, forming active communities. However, these communities can also experience negative interactions between users, including instances of online bullying. To explore these communities, we used the YouTube Data API v3 to collect comments from videos produced by creators targeting this demographic. We used these comments to construct co-commenter networks. We then applied topic modeling and sentiment analysis to gain deeper insights into the content and dynamics within these communities. This study identified distinct community patterns. Smaller communities demonstrated stronger internal connections, while larger networks showed weaker ties. The analysis uncovered two significant findings: one community showing coordinated negative behavior and another revealing signs of targeted youth advertising practices.

CCS Concepts

• **Information systems → Social networking sites**.

Keywords

YouTube, Community Detection, Topic Modeling, Sentiment Analysis

ACM Reference Format:

. 2025. Identifying and Analyzing Communities in YouTube Comments. In *Proceedings of 17th ACM Web Science Conference 2025 (WebSci'2025).* ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

YouTube has grown exponentially since its launch in 2005. With billions of users, it is now the largest video streaming platform on the web and a major source of online entertainment content. As highlighted by [1], many children and teenagers use the platform as an alternative to traditional entertainment sources, such as television, with the majority of their browsing time spent on YouTube.

One way children and teenagers interact on the platform is through the comment section. This feature allows users to share their thoughts on the video with the content creator and form online communities. These interactions can significantly impact their social development, behavior, and attitudes. Additionally, the contrasting one-way relationship between users and content creators further amplifies this impact by creating emotional connections that shape how they experience social media [8]. Furthermore, children's YouTube usage patterns are influenced by factors such as self-regulation abilities. These patterns not only affect their engagement with content but are also linked to emotional and behavioral problems [6].

Analyzing the communities formed in YouTube comments can provide valuable insights into the dynamics of user interactions, providing a deeper understanding of the platform's social ecosystem. This process involves identifying groups of users that share similar behaviors, typically modeled as a graph, where users and edges represent nodes represent interactions, and unsupervised techniques as well as deep learning techniques are frequently employed to detect them [9]. For instance, a study utilized co-commenter networks to analyze suspicious commenter behaviors, leveraging network structural features and clustering methods to identify coordinated activities and behavioral similarities among channels [10].

The study by [1] demonstrates that many current developments focus on identifying inappropriate content targeting young children. Advanced machine learning and deep learning techniques are being used to improve the detection of harmful videos and enhance children's safety on the platform. However, there is limited research focusing on analyzing the social network communities formed in the comment section of YouTube creators who produce content aimed at children and teenagers.

This study aims to analyze the communities formed within the comment sections of YouTube creators who produce content aimed at children and teenagers, as well as the content of the comments themselves. To achieve this, we selected 10 Brazilian YouTube creators based on their subscriber count and total views. The analysis begins with the construction of a commenter dataset for each creator by collecting all comments from 50 videos. Subsequently, we generate co-commenter networks by linking users who commented on the same video. Then, we compute graph metrics and apply the Louvain Algorithm to detect community structures within these networks. Additionally, we analyze the content of the comments using Topic Modeling and Sentiment Analysis to gain further insights into their thematic and emotional characteristics. The results reveal notable community structures within the networks. Larger networks tend to exhibit weaker internal connections, while smaller communities show higher densities and clustering coefficients. Sentiment analysis indicates that positive comments typically express fan appreciation, whereas negative comments often reflect dissatisfaction with the content. Notably, we identified a specific community that exhibits negative sentiments due to possible organized activities, as well as signs of targeted advertising aimed at children and teenagers in another community.

The remainder of this study is structured as follows. Section 2 reviews related work, highlighting prior research on community analysis, co-commenter networks, and sentiment analysis in the context of YouTube and children's content. Section 3 details the

WebSci'2025, May 20–24, 2025, New Brunswick, NJ, USA

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

methodology, including data collection, co-commenter network construction, community detection using the Louvain Algorithm, and the application of topic modeling and sentiment analysis. Section 4 presents the analysis results, focusing on the community structures identified, their characteristics, and insights from sentiment and topic modeling. Finally, Section 5 summarizes key findings, discusses limitations, and proposes directions for future research.

2 Related Work

A social network community can be defined as a set of members of that specific network that interact with each other, similar to real-world communities but in an online environment. These online communities are shaped based on how users interact on each platform, and the analysis of such communities can uncover hidden relationships between users and how they engage [9].

Recent research by [7] uses community analysis to investigate coordinated inauthentic campaigns on YouTube, focusing on characterizing suspicious behaviors through a multi-step analysis of engagement trends and co-commenter networks. These co-commenter networks are built by connecting two commenters if they comment on the same video. By analyzing these co-commenter networks and channel engagement trends, they were able to identify patterns indicative of manipulation, such as increasing views paired with decreasing subscribers, and highlight channels exhibiting coordinated commenting behaviors. In [10], the authors follow this methodology to explore the problematic behaviors of commenters on YouTube, specifically focusing on "commenter mobs." that manipulate engagement metrics to distort public perception. By analyzing 20 targeted channels through social network analysis, the research fills gaps in understanding how suspicious commenter activities boost engagement, which has been insufficiently addressed in prior literature. Employing a co-commenter network model and clustering techniques, the study identifies distinct groups with varying levels of suspiciousness and highlights collusion among commenters across channels. Key findings reveal central figures driving discussions and coordinated efforts to amplify specific narratives, suggesting these channels significantly contribute to spreading misinformation.

With a similar methodology, the study by [5] investigates disinformation tactics used on a specific conspiracy theory channel, distinguishing its approach by focusing on user engagement patterns rather than just spam detection. Data was collected using the YouTube Data API monitoring metrics like views, likes, dislikes, and comments, revealing patterns indicative of potential manipulation. The analysis identified two commenter groups, peripheral and core, with the second group exhibiting higher instances of inorganic behaviors. Co-commenter network analysis revealed clusters around specific conspiracy topics and identified bot-like and spam actions.

In [11], the authors investigate the prevalence and nature of abusive comments on YouTube, particularly focusing on the impact of hate speech on users, especially teenagers, by utilizing exploratory data analysis and topic modeling. The study employed Latent Dirichlet Allocation (LDA) for topic modeling and sentiment analysis using the TextBlob library to assess the emotional tone of comments. Findings revealed that certain YouTubers received significantly harsher comments, with varying sentiment levels indicating the need for interventions to protect young creators from severe online harassment. The study concludes by recommending measures such as disabling comments on particularly abused videos and emphasizes the importance of understanding the dynamics of cyberbullying.

In the context of children's content on YouTube [1], machine learning and deep learning techniques are used to analyze video, audio, and user comments to detect inappropriate material. These methods rely on natural language processing for text analysis and computer vision for visual content detection. Despite recent advancements, the challenge of accurately detecting inappropriate content targeted at This demographic, which uses machine learning and deep learning approaches, still persists.

This study differs from the previously mentioned works by focusing specifically on the co-commenter networks formed in the comment sections of YouTube creators who produce content aimed at children and teenagers. While prior research has explored co-commenter networks to detect suspicious behaviors, coordinated activities, and abusive language, these studies have primarily targeted general or adult-oriented content. Additionally, research on children's content has largely concentrated on detecting inappropriate material within videos rather than analyzing user interactions. By combining community detection through the Louvain Algorithm with Topic Modeling and Sentiment Analysis, this work provides a perspective on the thematic and emotional dynamics within these communities, addressing a gap in understanding the social interactions surrounding child-focused content on YouTube.

3 Methodology

This section outlines the methodology employed, which is divided into three main phases: data collection, community analysis, and comment analysis, as illustrated in Figure 1. The data collection phase describes the process of gathering comments from selected YouTube channels using the YouTube Data API v3. The community analysis phase explains the construction of co-commenter networks, the application of the Louvain Algorithm for community detection, and the computation of graph metrics to characterize these networks. Finally, the comment analysis phase details the use of BERTopic for topic modeling and XLM-RoBERTa for sentiment analysis to uncover thematic and emotional patterns within the comments. We describe each phase in detail to provide a comprehensive overview of the methodological steps taken to analyze user interactions in YouTube comment section.

3.1 Data Collection

Initially, we used the YouTube Data API v3 to gather comprehensive channel metadata for a group of YouTube creators. Through Viewstats[1], an analytics platform that facilitates ranking the most viewed channels by country and category, we selected ten creators based on the size of their subscriber base, total view count, and content type, as summarized in Table 1.

Subsequently, we retrieved video IDs and titles for 50 videos from each channel. Finally, we construct a dataset of user-generated comments and replies by aggregating all comments and their replies associated with the selected videos. We also collected additional contextual information, such as the video titles and the number of

[1]https://www.viewstats.com/

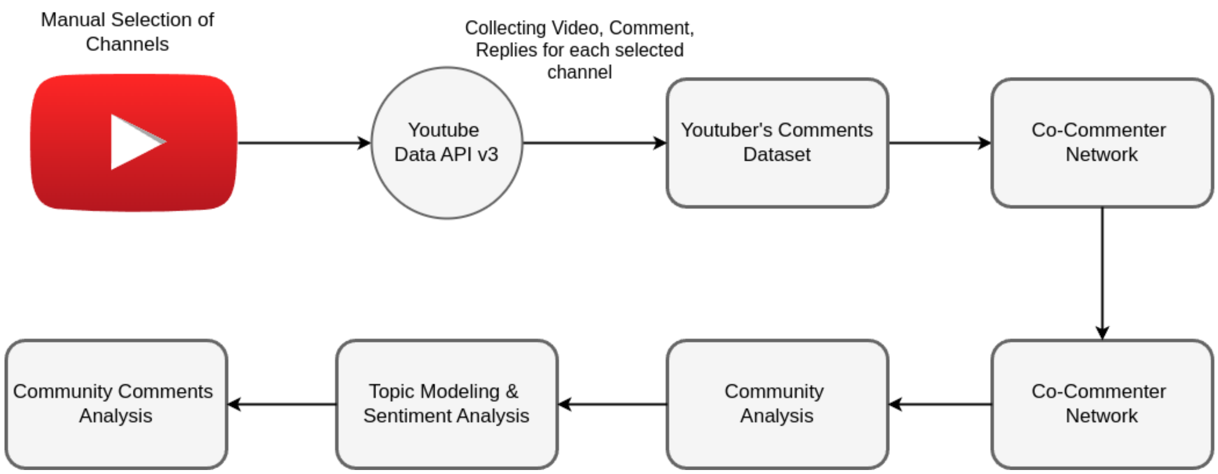

Figure 1: Methodology

Table 1: YouTubers with Subscribers and Views

YouTuber	Subscribers	Total Views	Content Type
Felipe Neto	46,800,000	18,171,376,895	Entertainment, vlogs
Enaldinho	40,000,000	18,009,391,826	Entertainment, vlogs
Natan por Aí	24,800,000	16,409,135,985	Entertainment, vlogs
Rezendeevil	33,800,000	13,686,976,806	Entertainment, vlogs
Robin Hood Gamer	21,500,000	11,309,149,204	Gaming (Minecraft, Roblox)
Brancoala	13,100,000	9,897,577,900	Family vlogs, Games
AuthenticGames	20,100,000	8,967,486,579	Gaming (Minecraft)
Camila Loures	15,600,000	5,509,570,500	Lifestyle vlogs
Geleia	10,200,000	3,207,064,265	Gaming (Minecraft)
CadresPlayer	8,640,000	2,640,042,941	Gaming (Minecraft)

likes each comment received. The number of comments collected per channel ranges approximately 11,000 to 43,000.

3.2 Community Analysis

To analyze the communities that emerge within each YouTuber's comment section, we used the datasets constructed in the previous phase to construct the co-commenter networks. We connect each commenter to another if they commented on the same video, increasing the weight of the edge if they commented on more than one video together. To maintain only the stronger connections, we filtered out co-commenters who commented on less than 10 videos together, as done by [10] and [7].

To identify the communities of each co-commenter network we used the Louvain Algorithm. We chose this algorithm because of its high efficiency and performance on real-world networks without ground truth [12].

Following the method developed by [7], we computed a set of metrics computed for each co-commenter network, including average degree, number of nodes and edges, average coefficient, modularity, coverage, the number of communities identified with the Louvain Algorithm. We also computed graph clique metrics, including the number of maximal cliques that have at least five members, average clique size, median clique size, average degree of clique members, average clustering coefficient of clique members. To discover the similarities between the channels, we used these metrics with KMeans and Hierarchical Clustering, along with Principal Component Analysis to reduce the complexity while maintaining important features.

3.3 Comments Analysis

To further understand user behavior, we implement topic modeling and sentiment analysis. For the topic modeling phase, we employ the BERTopic model [4] and for the sentiment analysis we use the multilingual XLM-roBERTa-base model [2]. Initially, we processed and cleaned these comments, masking user handles to preserve anonymity, removing emojis, links and stop words. Then, we use topic modeling to extract themes associated with the communities identified in the previous phase. Finally, the sentiment analysis model evaluates the polarity of the content, offering deeper insights into the emotional tone of the discussions and facilitating a more comprehensive analysis of the communities.

4 Results

In this section, the results are shown and analyzed. Considering that a modularity value above 0.3 is a good indicator of significant community structures in a network [3], the Louvain algorithm successfully identified significant community structures in all the YouTube creator's co-commenter networks. We applied the KMeans and Hierarchical Clustering methods with the community metrics described, with the reduced complexity of the Principal Component Analysis (PCA). Both methods identified 5 distinct clusters, as shown in Figure 2.

The most distinguishable channels are clusters 1 (Rezendeevil) and 3 (Felipe Neto). Both creators had very large co-commenter networks, with Rezendeevil's network having 13,808 nodes and 149,033 edges, and Felipe Neto has 24,987 nodes and 125,165 edges. The creator Rezendeevil had the lowest modularity (0.52) but the highest average degree (21.5), along with the most maximal cliques and the largest average clique size. This indicates that both creators have highly interconnected and engaged communities. The higher

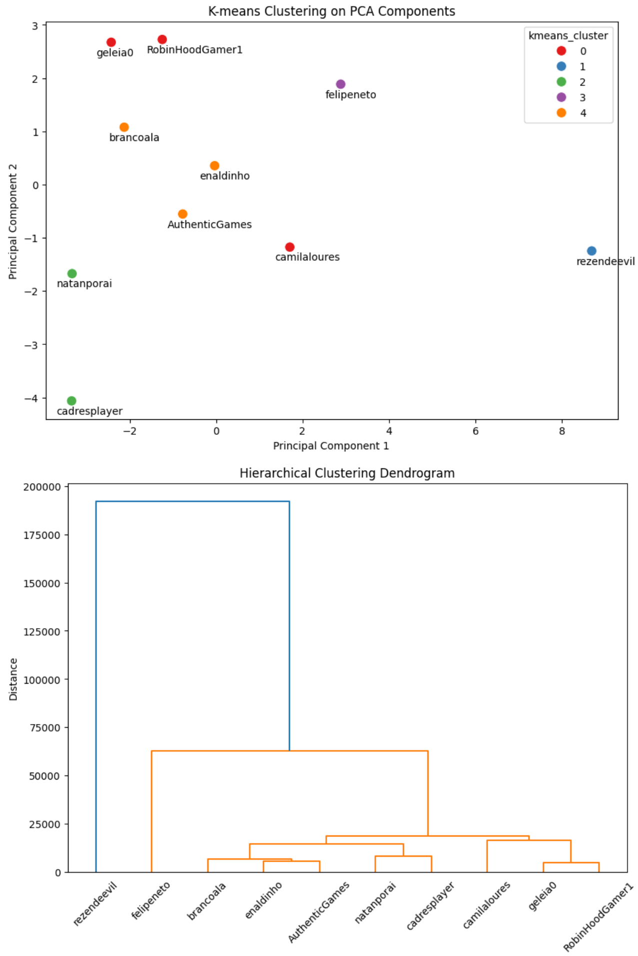

Figure 2: KMeans and Hierarchical Clustering

average degree indicates that Rezendeevil's followers are more engaged than Felipe Neto's, while the higher modularity score shows that Felipe Neto's community structure is more distinct. This can be seen through Figure 4 and Figure 5, with each community colored.

A community sentiment analysis showed that most comments on Rezendeevil's videos were labeled as positive by the XLM-roBERTa-base model, consistent with a highly engaged community of fans. Most positive comments focus on the video content, specific individuals featured, or compliments directed toward the creator. Through topic analysis, a recurring theme labeled "quero" was identified, with comments such as "Eu quero" and "Eu quero muito uma." (translated as "I want" or "I really want one"). These comments predominantly come from members of communities 5 and 6, characterized by higher-density values (0.15) and lower conductance metrics (0.07) relative to other communities Additionally, these communities exhibit lower clustering coefficients, a pattern observed across all communities. A more detailed investigation of the videos receiving these comments could reveal potential signs of direct or implicit advertising aimed at children and teenagers.

In contrast to Rezendeevil, Felipe Neto had more negative than positive comments. By analyzing the polarity per community, it was possible to identify that community 12, that has the highest percentage of negative comments (55.3%), is also the community with the highest clustering coefficient of approximately 0.23, lowest conductance of approximately 0.06, and highest density of 0.21. This may indicate a close-knit, organized community that frequently comments on the same content spreading negative and hateful comments. Topic Modeling also shows recurring themes labeled "Brasil", "Hipocrisia" ("Hypocrisy"), "Dinheiro, Pobres" ("Money, Poor"), with comments such as "é tanta burrice por metro quadrado é um absurdo como tem gente que acredita nestas hipocrisias" ("it's such stupidity per square meter; it's absurd how many people believe in these hypocrisies."), "Esse Felipe merda é uma das desgracas que so existe aqui no Brasil, ..." ("This Felipe piece of trash is one of the disasters that only exists here in Brazil, ..."). Figure 3 provides a word cloud visualization of these comments. The term 'user' is used as a placeholder to mask instances where one commenter mentions another user's name. The other communities had lower densities, lower average clustering coefficients, and higher conductance values. These communities displayed a wider variety of topics and a higher number of positive comments compared to community 12.

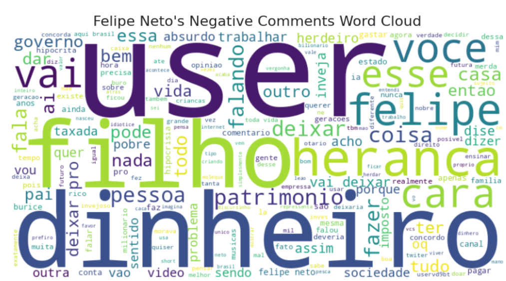

Figure 3: Felipe Neto's negative comments word cloud

Cluster 2 represents creators with smaller, tightly-knit networks with stronger internal connections, which suggests that users within their communities frequently interact with each other but not much across different subgroups. Topic Modeling shows that topics related to fan appreciation dominate across all communities, with comments like "Te amo Cadres" ("I love you Cadres"), "Oi Natan eu sou teu fan" ("Hi Natan, I'm your fan"). The analysis of both creators' communities with the highest proportion of negative comments showed that while their communities can be critical of their content, they are not entirely hostile. Most negative feedback arises from unmet expectations or dissatisfaction with specific content, with comments like "Não faz isso Natan voce pode morrer sem ar" ("Don't do this, Natan, you could suffocate") or "Natan faz algo melhor que isso" ("Natan, do something better than this"). For Cadresplayer, many criticisms focused on the creator's voice, with topics frequently labeled by the keywords "voice","strange","AI","different". This was reflected in comments such as "Cadres sua voz mudou, está parecendo um robô" ("Cadres, your voice has changed, it sounds like

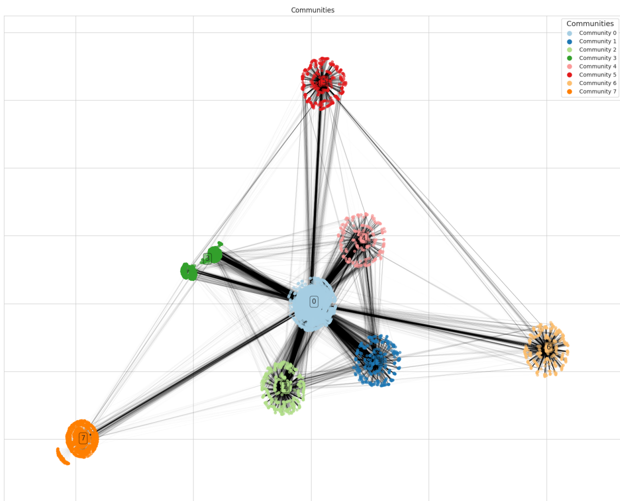

Figure 4: Rezendeevil's Communities

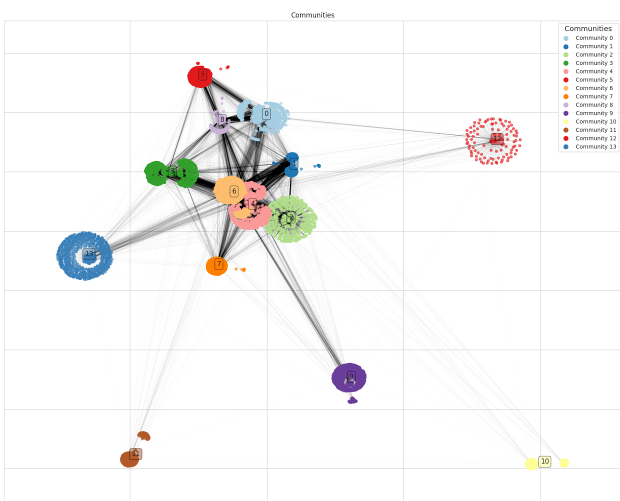

Figure 5: Felipe Neto's Communities

a robot"). For both creators, the communities with the highest proportion of negative comments have varying numbers of clustering coefficient, density, and conductance.

The networks of Camila Loures, Robin Hood Gamer, and Geleia0 in Cluster 0 exhibit moderately engaged networks, with significant differences in their community dynamics. All three creators have networks characterized by high modularity values (Geleia0: 0.90, Robin Hood Gamer: 0.84, Camila Loures: 0.64), indicating well-defined subgroups, though Camila Loures's network is more interconnected with fewer communities (7) compared to the more fragmented networks of Robin Hood Gamer (24 communities) and Geleia0 (31 communities). The average degree is highest for Camila Loures (12.44), suggesting a higher interaction levels among her commenters, while the clustering coefficients are low across all creators, indicating sparse local connectivity within communities.

In terms of topics and sentiment polarity, Camila Loures's the audience engages positively with lifestyle content, though criticism arises around repetitive video formats. The analysis of the creator's less positive communities revealed negative feedback on video styles or lack of originality ("Essa temporada ta ruim demais muito chata", "This season is really bad, very boring"), as well as neutral comments with content requests ("Você poderia fazer um vídeo mais interativo com os fãs.", "You could make a more interactive video with fans"). These communities demonstrated higher density values, lower conductance metrics, and a significantly elevated average degree (approximately 25, compared to less than 10 in all other communities). However, they exhibited a low average clustering coefficient.

Robin Hood Gamer follows a similar pattern of feedback. His fans engage positively around Minecraft videos, with many supporting comments from fans. The analysis of the less positive communities also revealed negative feedback related to content, often focused on content quality, as well as suggestions for future content improvements. However, these communities did not show significant differences compared to the rest of the creator's communities, and they also showed varying values of clustering coefficient, density, and conductance. Geleia0 shows a similar pattern to Robin Hood Gamer, but with emphasis on the communities with the highest proportion of negative comments (44.8% and 38.3%), with comments often expressing frustration, insults, or dissatisfaction with the content or creators. These comments frequently include abusive language, criticism of creators' personalities or actions, and expressions of anger across different themes. Examples include "O dream é um monstro matou o próprio pet que raiva" ("Dream is a monster; he killed his own pet, how infuriating."), "Aaaaaaa o geleia é muito burroooooooo que ódio" ("Aaaaaaa Geleia is so dumb, how hateful."), "Nao foi justo ele tirou o messi porque ele nao gosta da argentina" ("It wasn't fair, he took out Messi because he doesn't like Argentina"). Figure 6 provides a word cloud visualization of these comments.

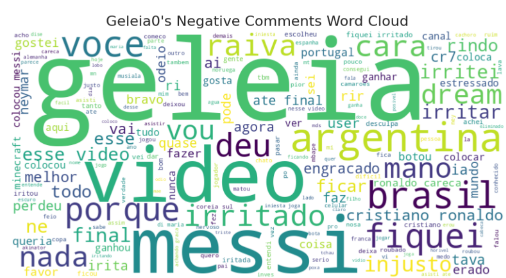

Figure 6: Geleia0's negative comments word cloud

Cluster 4 represents creators with moderately sized networks that show similar community dynamics and thematic engagement, exhibiting network structures with node counts ranging from 11,000 to 12,000, and edge counts varying between 26,000 and 37,000. The creators in this cluster (Enaldinho, AuthenticGames, and Brancoala) have networks characterized by high modularity values (ranging from 0.75 to 0.86), though their clustering coefficients are low, reflecting sparse local connectivity within communities. Enaldinho has 14 communities, AuthenticGames has 20, and Brancoala has 19, which shows varying levels of audience fragmentation. Topic Modeling revealed the same pattern of content requests, negative feedback over repetitive content, and fan appreciation. For all creators of this cluster, the communities with the highest proportion of negative comments showed varying numbers of average clustering coefficient, density, and conductance.

Overall, the analysis of YouTube co-commenter networks revealed distinct patterns of community structure, engagement, and thematic focus across the identified clusters. Each cluster exhibits unique characteristics based on graph metrics, sentiment polarity, and topic modeling. Across all co-commenter networks, the communities identified with the Louvain algorithm showed varying clustering coefficient, density, and average degree. Table 2 shows the main metrics used per cluster.

Table 2: Network Characteristics by Cluster

Cluster	Nodes	Edges	Avg Degree	Modularity
0	8,800-19,900	54,800-72,000	7.2-12.4	0.64-0.90
1	13,808	149,033	21.58	0.52
2	2,700-5,300	5,200-13,000	3.7-4.8	0.78-0.83
3	24,987	125,165	10.01	0.65
4	11,000-12,800	26,000-37,500	4.5-5.9	0.72-0.87

Across all clusters, neutral comments dominate, reflecting passive engagement or constructive feedback. While positive sentiments typically reflect fan appreciation, negative responses generally stem from unmet expectations or frustration with repetitive content, with occasional hateful and targeted comments. Larger communities serve as broader hubs with weaker internal connections, possibly with users that don't often interact with others.

Within the networks of clusters 0, 2, and 4, smaller communities generally show higher densities and clustering coefficients compared to larger ones. Overall, polarity within these communities is generally non-hostile, with exceptions sporadically distributed across communities. Analysis of the most negative communities reveals diverse characteristics: while some display more elevated density and clustering coefficient values, others exhibit lower values. Complementing this sentiment analysis, topic modeling provides additional insights, for example, fans of Cadresplayer expressed dissatisfaction with the creator's voice, describing it as robotic, while Geleia0's fans engaged in discussions on sports-related themes.

In contrast, the analysis of Cluster 3's network identified a distinct group disseminating negative comments directed at the creator and other users, with topic modeling revealing recurring themes within these interactions. Meanwhile, the analysis of Cluster 1's network with topic modeling highlighted potential direct or implicit advertising indicators.

5 Conclusion

This study analyzed the communities formed in the comment sections of YouTube creators who produce content aimed at children and teenagers. The research uncovered significant insights into the structure, engagement, and thematic dynamics of these online communities by leveraging co-commenter network analysis, community detection with the Louvain Algorithm, topic modeling, and sentiment analysis.

Our findings highlighted the existence of distinct community structures within YouTube comment sections. By grouping channels into clusters, the analysis allowed for a detailed comparison of the similarities and differences across these networks. Topic modeling and sentiment analysis revealed consistent patterns across all creators: positive comments were primarily associated with fan appreciation, neutral comments indicated passive engagement and negative comments often reflected unmet expectations or dissatisfaction.

However, we noted exceptions. For instance, in Cluster 3 the network analysis revealed a distinct community spreading negative comments with higher clustering coefficient values and density and low conductance. This pattern was unique to Cluster 3. In other networks, communities with similar network metrics did not necessarily correspond to higher levels of negativity. Communities with high negativity were found across clusters 0, 2, and 4, both in groups with similar network characteristics of Cluster 3's community (high density and clustering coefficient) and those with lower network metric values.

This research presents the value of examining social interactions in YouTube comments by combining community analysis, topic modeling, and sentiment analysis. Future studies would benefit from exploring alternative community detection algorithms and analyzing larger datasets spanning multiple videos and channels, investigating overlapping community structures, and extending the analysis to different social media platforms. This expanded scope would provide deeper insights into user behavior and community dynamics, ultimately contributing to the development of a safer online environment for children and teenagers.

Acknowledgments

To my family and friends, for their love and support, and to the professors, for their guidance and teachings throughout this journey.

References

[1] Saeed Ibrahim Alqahtani, Wael M. S. Yafooz, Abdullah Alsaeedi, Liyakathunisa Syed, and Reyadh Alluhaibi. 2023. Children's Safety on YouTube: A Systematic Review. *Applied Sciences* 13, 6 (2023). https://doi.org/10.3390/app13064044

[2] Francesco Barbieri, Luis Espinosa Anke, and Jose Camacho-Collados. 2022. XLM-T: Multilingual Language Models in Twitter for Sentiment Analysis and Beyond. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. European Language Resources Association, Marseille, France, 258–266. https://aclanthology.org/2022.lrec-1.27

[3] Aaron Clauset, M. E. J. Newman, and Cristopher Moore. 2004. Finding community structure in very large networks. *Phys. Rev. E* 70 (Dec 2004), 066111. Issue 6. https://doi.org/10.1103/PhysRevE.70.066111

[4] Maarten Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* (2022).

[5] Muhammad Nihal Hussain, Serpil Tokdemir, Nitin Agarwal, and Samer Al-Khateeb. 2018. Analyzing disinformation and crowd manipulation tactics on YouTube. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 1092–1095.

[6] Donghee Kim, Sangha Lee, Hyojin Kim, and Yunmi Shin. 2024. From temperament to YouTube: exploring the link between childhood temperament, YouTube usage patterns, and emotional/behavioral problems among children. *BMC Public Health* 24 (2024).

[7] Baris Kirdemir and Oluwaseyi Adeliyi. 2023. Towards Characterizing Coordinated Inauthentic Behaviors on YouTube.. In *The 2nd Workshop on Reducing Online*

Misinformation through Credible Information Retrieval (ROMCIR 2022) held with the 44th European Conference on Information Retrieval (ECIR 2022).

[8] Raquel Lozano-Blasco, Marta Mira-Aladrén, and Mercedes Gil-Lamata. 2023. Social Media Influence on Young People and Children: Analysis on Instagram, Twitter and YouTube. *Comunicar: Media Education Research Journal* 31, 74 (2023), 117–128.

[9] Mahsa Nooribakhsh, Marta Fernández-Diego, Fernando Gonzá lez Ladrón-De-Guevara, and Mahdi Mollamotalebi. 2024. Community detection in social networks using machine learning: a systematic mapping study. *Knowledge and Information Systems* (2024), 1–55.

[10] Shadi Shajari, Nitin Agarwal, and Mustafa Alassad. 2023. Commenter Behavior Characterization on YouTube Channels. arXiv:2304.07681 [cs.SI] https://arxiv.org/abs/2304.07681

[11] Shubhanshu Shekhar, Akanksha, and Aman Saini. 2021. Utilizing Topic Modelling To Identify Abusive Comments On YouTube. In *2021 International Conference on Intelligent Technologies (CONIT)*. 1–5. https://doi.org/10.1109/CONIT51480.2021.9498368

[12] Xuemei You, Yinghong Ma, and Zhiyuan Liu. 2020. A three-stage algorithm on community detection in social networks. *Knowledge-Based Systems* 187 (2020), 104822. https://doi.org/10.1016/j.knosys.2019.06.030