

Unveiling Digital Distraction: An In-depth Analysis of YouTube Channels on Khalistani Movements Across Educational, Influencer, and News Domains

Student Name: Saurav

Roll Number: 2020243

Independent Project report submitted in partial fulfillment of the requirements for the Degree of B.Tech. in Electronics and Communication engineering.

Abstract

The escalating concern regarding the role of digital media in fostering distraction has prompted an investigative analysis of 30 YouTube channels distributed across three distinct domains. These domains include educational content creators, influencer material content creators, and news channels. We have chosen 30 videos revolving around Khalistani movements based on YouTube recommendations. We collected data about these channels and videos and represented the data in tabular and graphical representations, enhancing the visual comprehension of the content disseminated by the chosen media. We focused on the comment section to observe user engagement and collected all the comments. All comments are subjected to a comprehensive analysis of user emotions, patterns, and insights into audience interactions. The main objective of this study is to identify which domain intentionally creates content that distracts the audience from the main topics. The investigation seeks to delineate the markets these domains target and ascertain whether they cater to specific audience demographics. We have used YouTube APIs to collect data and perform network analysis on this data. Under the network analysis, we are utilizing concepts such as centrality. This methodology allows for a nuanced examination of the relationships and interactions within the network of YouTube channels, revealing potential strategies employed to captivate and distract audiences. The research concludes by drawing comparisons between the digital environment and the allegorically described "Weapons of Mass Distraction," illuminating the possible intentional attempts made by some domains to grab users' attention. The results of this study lay the groundwork for future discussions on the ethical production and consumption of online material while offering insightful information about the mechanics of digital distraction.

Contents

1	Introduction	2
1.1	Introduction	2
1.2	Motivation	2
2	Methodology	3
2.1	Data Collection and Visualization	3
2.2	Network Analysis	5
3	Conclusion	9

Chapter 1

Introduction

1.1 Introduction

In today's time, where digital platforms have become a part of daily life, the issue of digital distraction has become a crucial topic to show concern. This study helps to understand the role of YouTube in shaping the user's mind's attention and contributing to digital space. They are motivated by "Weapons of Math Destruction," by Cathy O'Neil, which emphasizes the frequently invisible impact of algorithms in our daily lives. YouTube is a platform for seeking attention from users. Our study categorizes YouTube channels into educational content creators, influential channels, and news channels. This classification is done to find out their respective impacts on user distraction. Our main objective is to determine which domain most effectively caters to a significant audience and how YouTube recommendations play an important role.

1.2 Motivation

YouTube provides a platform to share content and allows people to engage with the platform by showing their emotions through likes, dislikes, and comments. According to one report, the user base of YouTube is around 758 million across the globe monthly. In 2017, a time came when YouTube surprised everyone, as most people were using it as a news source. YouTube has a two billion user base, which allows us to analyze the pattern of distraction and user engagement. Recently Khalistani movement became a global issue, which gives us a space to explore the way of attention and distraction caused by these domains. As this topic is complex and has socio-political nuances, it offers the opportunity to examine how these domains, including educational content creators and influencer-driven news channels, navigate sensitive issues, what impact it has on user attention, and how the YouTube recommendation system contributes to this.

Chapter 2

Methodology

2.1 Data Collection and Visualization

For data collection, we used YouTube APIs for fetching the data about YouTube channels and respective videos; for choosing the videos related to the topic Khalistani movement, I randomly picked 30 channels that mainly target socio-political issues and chose random 30 videos from these channels based on the YouTube recommendation system. I started with plotting several plots, such as channel name vs. subscriber count and medium vs. number of views; similarly, for the videos, we have calculated the numbers of likes vs number of views ratio and also plotted the video vs number of comments, which somewhere shows user engagement to some extent. To identify which domain intentionally creates content that distracts the audience from the main topics, I used network analysis—concepts such as centrality. To show the interaction between the users, I plotted intra-domain and inter-domain network graphs. I also plan the word clouds for better visuals and representation of user comments.

Various plots for data representation and visuals:

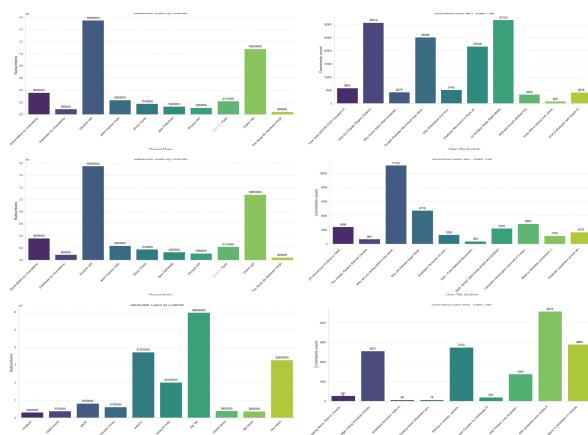


Figure 2.1: Plots for data representation



Figure 2.2: Word Cloud of educational channel video comments



Figure 2.3: Word Cloud of Study channels Video Comments

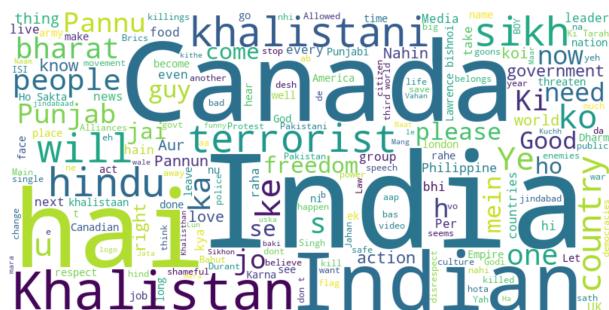


Figure 2.4: Word Cloud of News channels Video Comments

2.2 Network Analysis

The primary objective of this study is to identify which domain is responsible for distracting the audience. As narrated earlier, we have chosen thirty channels, which we categorize into three categories. Each domain consists of ten videos that talk about the Khalistani incident. To analyze the engagement trend, we first studied each part and tried to plot the intra-domain graph. We took each video as a node, and if they had any everyday users, we increased the count of edges by one; the edge shows the simple user between the two videos. The graph for influencers has the most everyday users within the same domain compared to study content creators and news channels.

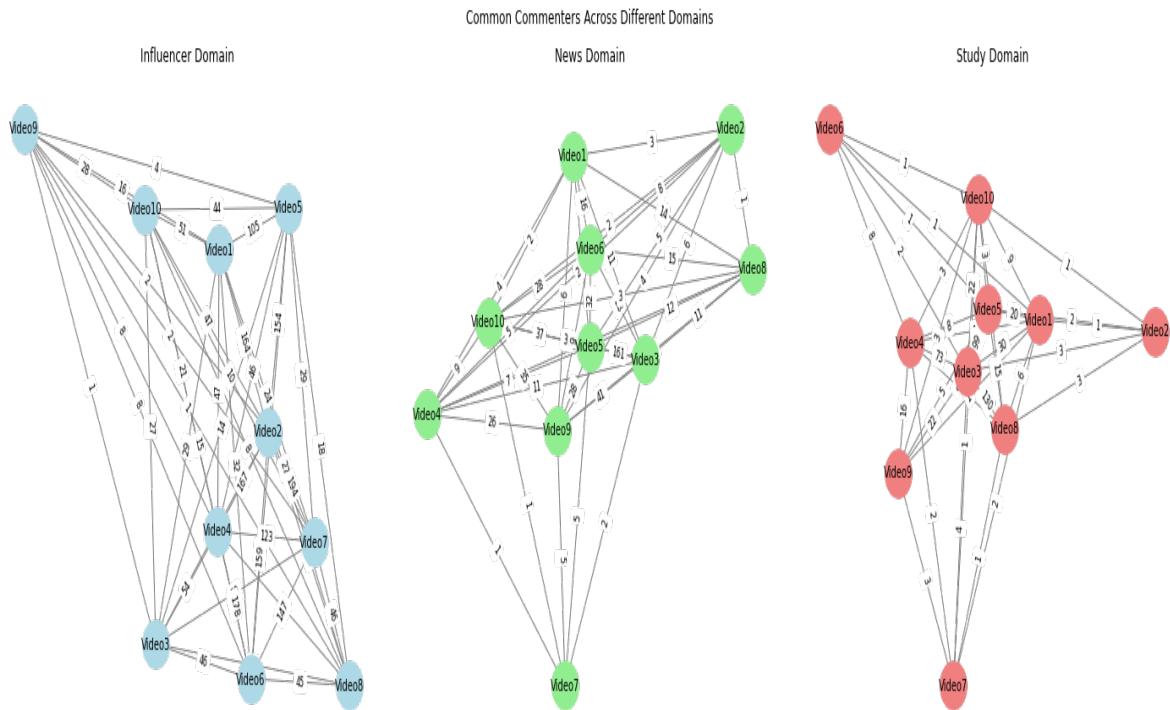


Figure 2.5: User engagement through various domains

Now, to analyze more about the intra-domain engagement of users, we used the concept of centrality and plotted the network graphs; those videos which have more engagement among the users occupied the central space of the graph, and those videos that had less engagement seemed to be far from the major nodes. The central node in this graph has to be enlarged in size as compared to another node that is far from the center node. The central node has thicker edges because they have more user interaction than the nodes that are far from the centre. We can conclude that those nodes that occupy central space are enlarged and the edges also thick, and as we move away from the center, the node size decreases and the edges' thickness also decreases. Here, we can also observe that the Influencers have more central nodes than domains and thicker edges than other domains.

Now, talking about the cross-domain network graphs, for cross-domain graphs, we have studied the common user between each domain; for better visuals, we chose the common user between two parts by the single edge with edge count, which represents the common user between two

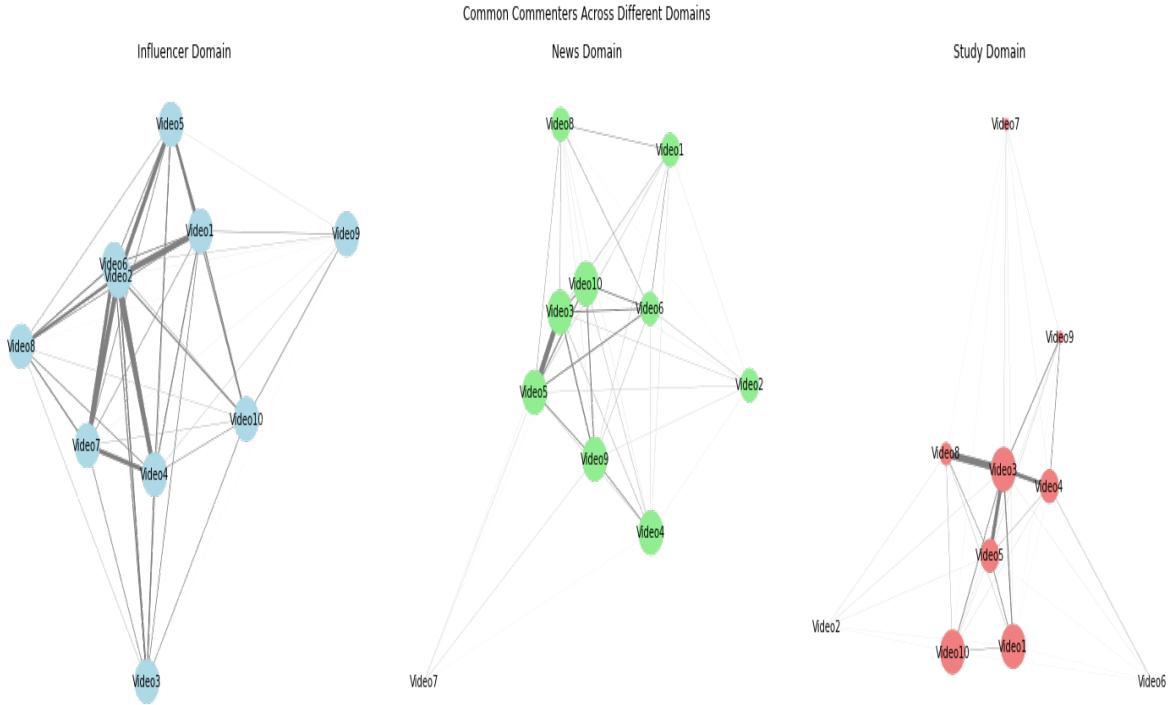
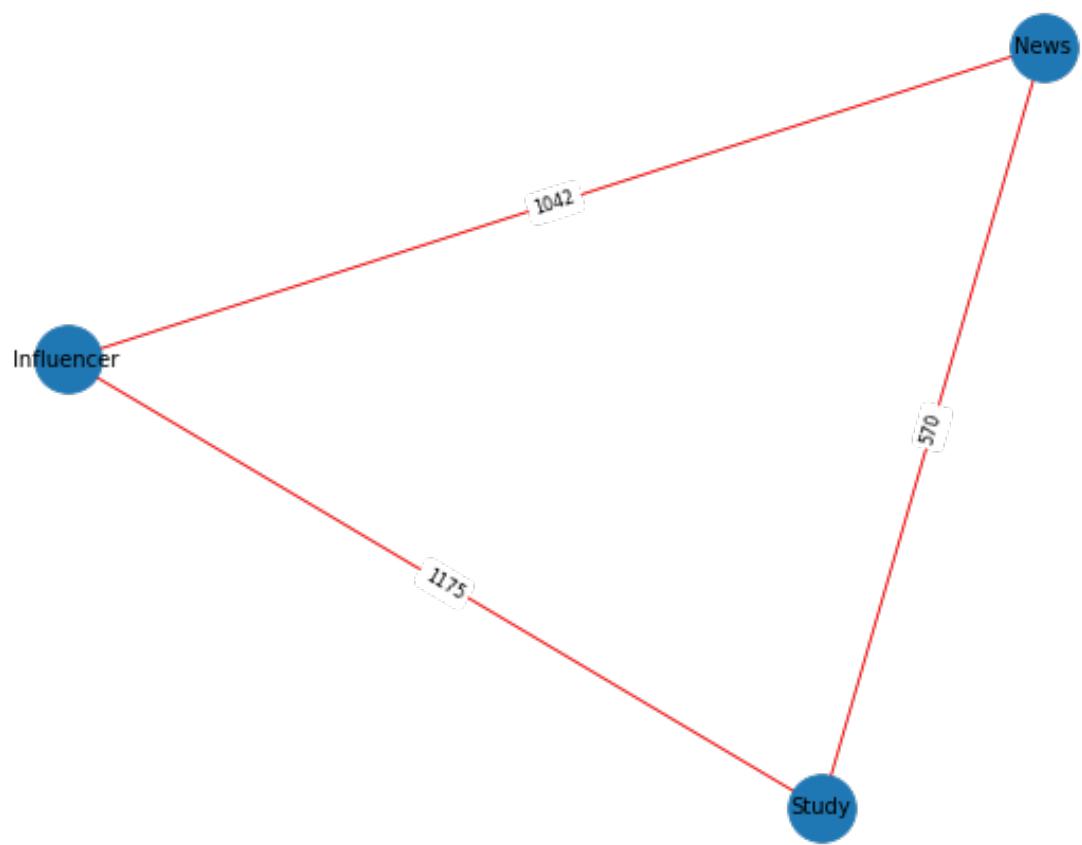


Figure 2.6: Network graphs of each domain based on centrality

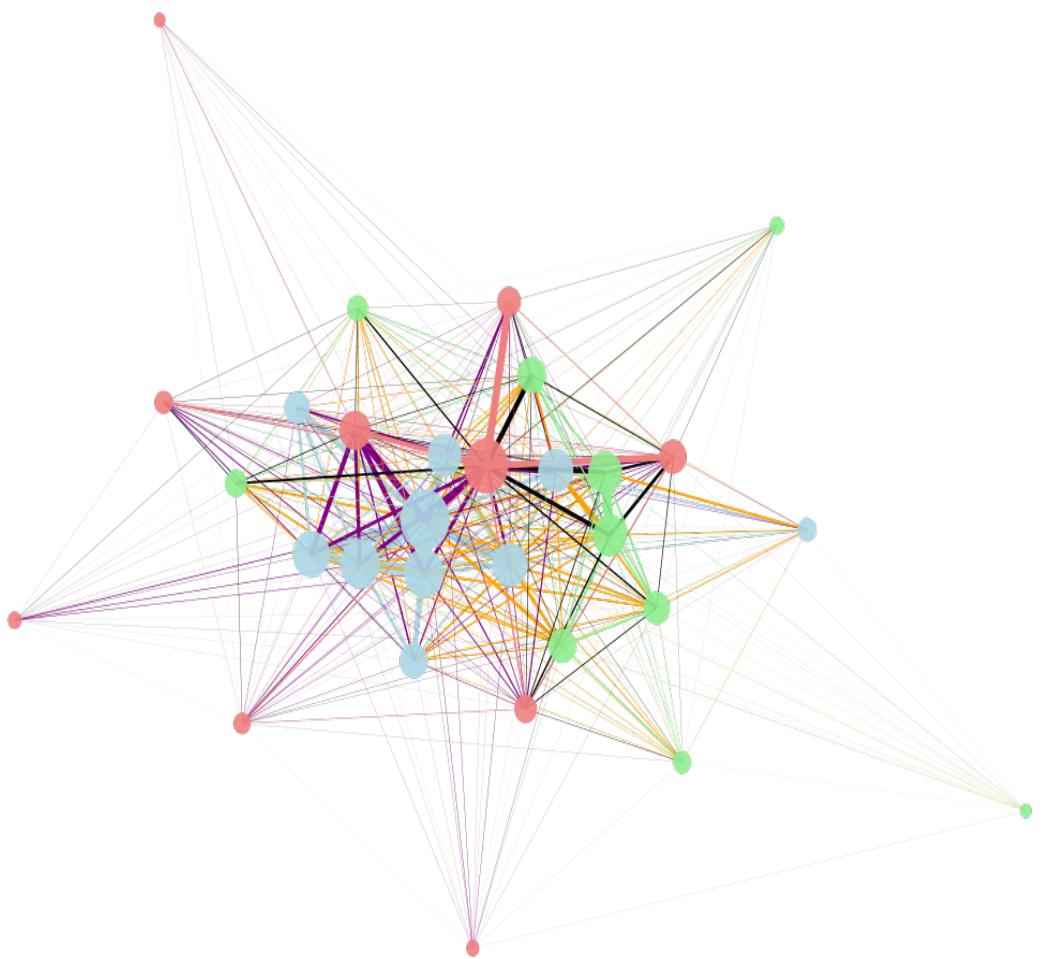
domains. Here, we figure out an interesting thing: the interaction between the influencer and the study domain is higher, and the interaction between the news and the study domain is less.

Now, we will merge all the domains and try to plot them in a common space. This graph plots on the basis of centrality; that node which has a higher number of common users must come in the center, and the thickness of edges is also more in the case of central nodes. Those nodes with a smaller number of common users are far away from the center, and the thickness of the edges is also not too much compared to central node edges, which show less user engagement. In this graph, the main space is also occupied by the influencer domain and a little bit by the study domain.

Cross-Domain Connections Based on Common Commenters



Common Commenters (Based on Channel URL) Within and Between Videos of Different Domains



Chapter 3

Conclusion

In conclusion, our data representation reveals YouTube to be an online social platform enabling user engagement primarily through comments. This study highlights how content creators distract the users from their purpose of opening YouTube. The first graph that we get from the network analysis shows us that in the case of the influencer-driven channel, the user engagement is highest as compared to study and news channel, which mean they have catered the fixed audience from the platform, people are watching their content on a daily basis and followed them. Talking to the second graph, we have seen that the edges between the intra-domain in the case of the Influencer are too thick as compared to the other two domains, which also shows user engagement towards the Influencer, whereas in study one users prefer to stick to their favorite channels they do not interact too much with another same domain channel. From the third graph, where we had shown the common user cross-domain, we have found that the influencer and study domain have the most number of common user followed by Influencers and news, and the news and study domain have least number of common user. This graphs clearly point out Influencer dominate the platform and seems to be a primary player in distraction. According to the theoretical aspect, it should be like the news and study domain together should have the most number of common users all study channels are civil service preparation channels, so from the perspective of students, those who are preparing for the examination must follow study channels and also need to engage in news channels but that contradicting with the graphs data. Here, Influencer-driven content occupies a particular market and people are blindly following them. From the last, which is plotted on the concept of centrality, this graph is plotted across domains in common space here influencers also occupy the central space of the frame and have very enlarged nodes as compared to other domains, which shows that this domain has most user within the domain and also cross-domain, somewhere study domain also some central space but none of the news channels nodes occupied the central space, the width of the edge seem to be thick within the Influencer domain and also between Influencer and study domain. We can conclude that the influencer domain is the distracting element in aspect to the other two domains as we had from the study also that there is Influencer and study channel interaction, Influencer and News channels interaction, but when it comes to Study and News channels interaction, this interaction seems to be almost half as compared to other interaction.