Complex Networks as a Methodology for Social Studies: The Case of 'Son de Amores'

Carlos Redroban[1,*], Luis Rivas-Tovar[2], Augusto Pino[3] and Lissete Reyes[4]

[1]University Ecotec, Samborondon, Ecuador
[2]National Politechnic Institute, Ciudad de Mexico, Mexico
[3]Catholic University of Argentina, Buenos Aires, Argentina
[4]University of Guayaquil, Guayaquil, Ecuador

Abstract

This work was based on the analysis of a social phenomenon that was generated by the parody of a song of the famous Spanish group Andy and Lucas with one of their songs "Son de Amores", for its analysis we used the methodology of complex networks and information search with web scraper to understand the reasons for its success considering that its creator Lis Padilla was not a well known or influential person in social networks. The results showed that in the structure of the network created with the information obtained, applying the centrality metrics, it was determined that with the centrality of Eigenvector the influencer "Hey Santana" could be a determinant for the success of the parody in the youtube network, despite the fact that the origin of the parody was in the TikTok social network.

Keywords

Complex network, Eigenvector,, Centrality, Son de Amores

1. Introduction

In this decade, the increasing interconnection and complexity of human interactions have driven the development of new methodological tools to understand social phenomena. Complex networks, an interdisciplinary field that merges concepts from graph theory, statistical physics, and data science, have emerged as a powerful approach for analyzing the structure and dynamics of social systems. This approach allows for modeling relationships and interactions between individuals, groups, and organizations as nodes and links in a network, revealing underlying patterns that may not be evident through traditional methods.

The study of social phenomena through complex networks offers a holistic view, capturing both the microdynamics of individual interactions and the emergent properties that arise at the macro level. Phenomena such as the spread of information, opinion formation, power dynamics, and social cohesion can be quantitatively explored, enabling researchers to identify the structural features that influence collective behavior. Moreover, the ability of networks to model nonlinear and adaptive systems opens new possibilities for understanding how small perturbations can lead to significant changes in society.

This article explores how complex networks can serve as a robust and versatile methodology for studying social phenomena, highlighting their ability to unravel the intricate interactions underlying contemporary social dynamics. Specifically, a detailed analysis will be conducted on the case of "SON DE AMORES," a parody created by Lis Padilla in Peru, which went viral and propelled her to fame. This case provides a unique opportunity to examine how social networks and the structure of digital interactions can amplify and transform local content into a global phenomenon, revealing the mechanisms behind virality and the construction of fame in the digital age.

ICAIW 2024: Workshops at the 7th International Conference on Applied Informatics 2024, October 24–26, 2024, Viña del Mar, Chile
*Corresponding author.
credrobano@ecotec.edu.ec (C. Redroban); larivas33@hotmail.com (L. Rivas-Tovar); agupinov@gmail.com (A. Pino); liss2706@gmail.com (L. Reyes)
0000-0002-0672-5323 (C. Redroban); 0000-0002-5186-9895 (L. Rivas-Tovar); 0009-0005-1600-617X (A. Pino); 0000-0002-1363-3443 (L. Reyes)

2. Materials and methods

The analysis of the article was based on the methodology of complex networks, and the obtaining of the data by means of the use of a web scraper, the complex networks in their minimum expression are more than structures conformed by nodes and their relations [1], the success of the analysis of the application of the methodology of complex networks is in knowing how to identify the nodes and to define the interactions between them in the context of an investigation[1, 2, 3], that nodes will be obtained with scraping.

2.1. Complex Network

Few disciplines can identify their point of origin among them the science of networks, to understand it, we must go back to 1736, when the mathematician Euler, tried to solve the Königsberg bridges problem, which gave rise to the graph theory, although the graph theory is part of mathematics, this served as the basis for the study of networks [4].

Paul Erdös and his collaborator Alfred Rényi, invented the formal theory of random graphs, in which a new node is linked with equal probability with the rest of the network; currently, this premise is used as a theoretical basis in the generation of other networks under the framework of the theory of random networks. In 1967, the psychologist Stanley Milgram, conducted an experiment under the hypothesis that the "world was small" for the effect, he considered the world as a huge network of social relationships in which you could contact anyone in the world through a network of friends in just a few steps; the results of the experiment was precisely that with only 6 steps or contacts you could contact everyone, this result led to coin the phrase "six degrees of separation" [1].

Watts and Strogatz, demonstrated that small-world networks are not only present in social structures, -from which the idea arose-, but are present in all kinds of interconnected networked systems; generically it is a model of generation of random graphs with small average distances and high values of the grouping coefficient which is known as real-world networks [5].

On the other hand, Barabasi and Albert, starting from an analysis of Erdös random graph model, showed that many real-world networks have degree distributions whose behavior obeys what is called a power law; they concluded that many real networks such as the www, internet, actor-network, metabolic networks, etc., were scale-free networks [6].

A network is a catalog of the components of a system, called nodes or vertices, and the interactions between them, called links or edges, this representation provides a common language for studying diverse systems that have different greatly in nature, appearance, scope, etc.

Due to the exponential increase and availability of digital information, a major movement in network research has emerged in recent years, shifting the focus from the analysis of small individual graphs to the study of statistical properties in large-scale graphs. This shift in focus has been driven primarily by improved computing power and communication networks that allow us to collect and analyze large-scale data [7].

In network science networks are often distinguished by some elementary property of the underlying graph, Laslo Barabási and Réka Albert present a broader classification of network types based on their elementary characteristics (Figure 1), for example:

- **Undirected Network:** A network whose links don't have a defined direction. For instance: Internet, power grid, and science collaboration networks.
- **Self-loops:** In many networks nodes don't interact with themselves, so the diagonal elements of the adjacency matrix are zero, $A_{ii} = 0\ i = 1, ...N$. In some systems self-interactions are allowed; in such networks, self-loops represent the fact that node i interacts with itself. Examples: WWW, protein interactions.
- **Multigraph/Simple Graphs:**In a multigraph nodes are permitted to have multiple links (or parallel links) between them. Hence A_{ii} can be any positive integer. Networks that do not allow multiple links are called simple. Multigraph Examples: Social networks, where we distinguish friendship, family, and professional ties.

- **Directed Network:** A network whose links have selected directions. Examples: WWW, mobile phone calls, citation network.
- **Weighted Network:** A network whose links have a defined weight, strength or flow parameter. The elements of the adjacency matrix are $A_{ij} = W_{ij}$ if there is a link with weight W_{ij} between them. For unweighted (binary) networks, the adjacency matrix only indicates the presence $(A_{ij} = 1)$ or the absence $(A_{ij} = 0)$ of a link. Examples: Mobile phone calls, email network.
- **Complete Graph (Clique):** In a complete graph, or a clique, all nodes are connected to each other. Examples: Actors in the cast of the same movie, as they are all linked to each other in the actor-network." [6]

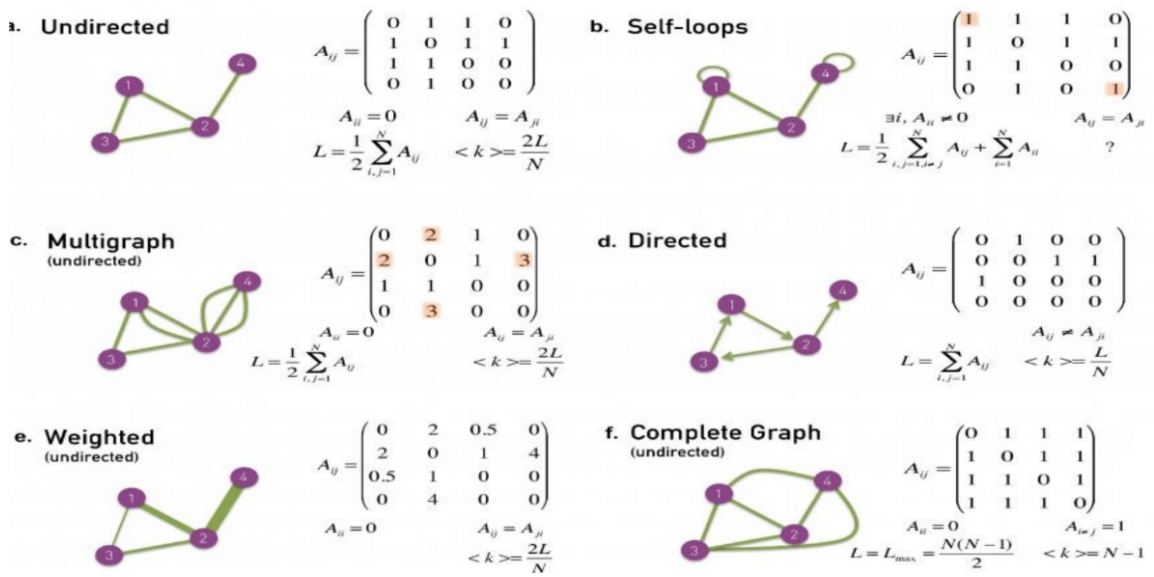

Figure 1: Topology of Complex Network. Barabási y Albert.

2.1.1. Centralities

There are several metrics to determine the behavior and relevance of both nodes and their relationships in a network; however, in this study, we used the most important metrics regarding centrality, these help to identify the most relevant or central nodes within a network; the most common measures used to represent different attributes of centrality are: degree, betweenness, closeness and eigenvector [3, 5, 6, 8, 9, 10, 11, 12].

Degree centrality captures the basic connectivity of the nodes, it determines how connected the nodes in the network are, for directed networks, must define in-degree and out-degree centrality separately.

$$EC_D(i) = \frac{deg(i)}{n-1} \tag{1}$$

Betweenness centrality determines the importance of a node as an intermediary in the network.

$$C_B(i) = \frac{1}{(n-1)(N-2)} \sum_{j \neq i, k \neq i, j \neq k} \frac{N_{sp}(ji \rightarrow k)}{N_{sp}(j \rightarrow k)} \tag{2}$$

where $N_{sp}(j \rightarrow k)$ is the shortest number from node *j* to node *k*, and $N_{sp}(ji \rightarrow k)$ is the shortest number from node *j* to node *k* go through node *i*. This metric can also be defined for edges in a similar way, which is called edge betweenness [3].

Closeness centralitymeasures the shortest path between two nodes *i* and *j*, it helps to identify nodes that are closer to other nodes in the network.

$$C_C(i) = \left(\frac{\sum_j d(i \to j)}{n-1}\right)^{-1} \tag{3}$$

This is an inverse of the average distance from node *i* to all other nodes. If $C_C(i) = 1$, that means must research any other node from node *i* in just one step [3, 6].

Eigenvector centrality determines the importance of being connected to other important nodes in the network, this centrality is proportional to the sum of the centrality of the neighboring nodes.

$$EC(i) = x_i = c\sum_{j=1}^{n} a_{ij}x_j \tag{4}$$

where *c* is a constant of proportionality to $x = [x_1, x_2, x_3 ... x_n]^T$

2.2. Web Scraper

To building the scraper, must be used a Web Scraper that takes advantage of the Chrome browser from Google [13]. The scraping structure was designed based on two variables:

Table 1
Structure of database

	Featured Videos		Recommended video	
Youtube Channel	Name of video	Url of video	Name of video	Url of video

- The **featured videos** on a YouTube channel.
- **recommended videos** within each of these featured videos.

2.2.1. Featured Videos

To start with the scraper process must use videos featured by Lis Padilla YouTube channel, the creator of the parody "Son de Amores"[1]

Scrapers are thus a general means for searching the title and URL of videos on this page (see Figure 2).

Although it is true that Lis Padilla's parody was created in the social network TikTok, it became famous in the youtube network when rumors began to circulate that she had won a house for the success obtained with the video "son de amores", however in her youtube repository, Lis Padilla has not uploaded this video, but a video called "me compre una casa", with more than 161 k., views.

To build the scraper you must click on the scraper icon and then press F12 or Ctrl+Shift+I, to open developer tools liks figure 3:

Create a sitemap The new sitemap is created in the following way: Clicking the "Create new sitemap" tab (under the Web Scraper tab) and selecting "Create Sitemap".

Once we have created the sitemap, we'll get a new empty directory named _root.

Create list of videos Now we need to tell the scraper what to select. We add a new selector for our scraper by clicking Add new selector button (see Figure 5).

We'll call our first selector, "featured" since they're the featured video of the channel. Switch the Type to be "Link" (we don't only want the Text, but the title text and the URL), then click on the "Select" button under "Save Selector" (see Figure 6 bottom).

[1] https://www.youtube.com/@lispadilla5512/videos

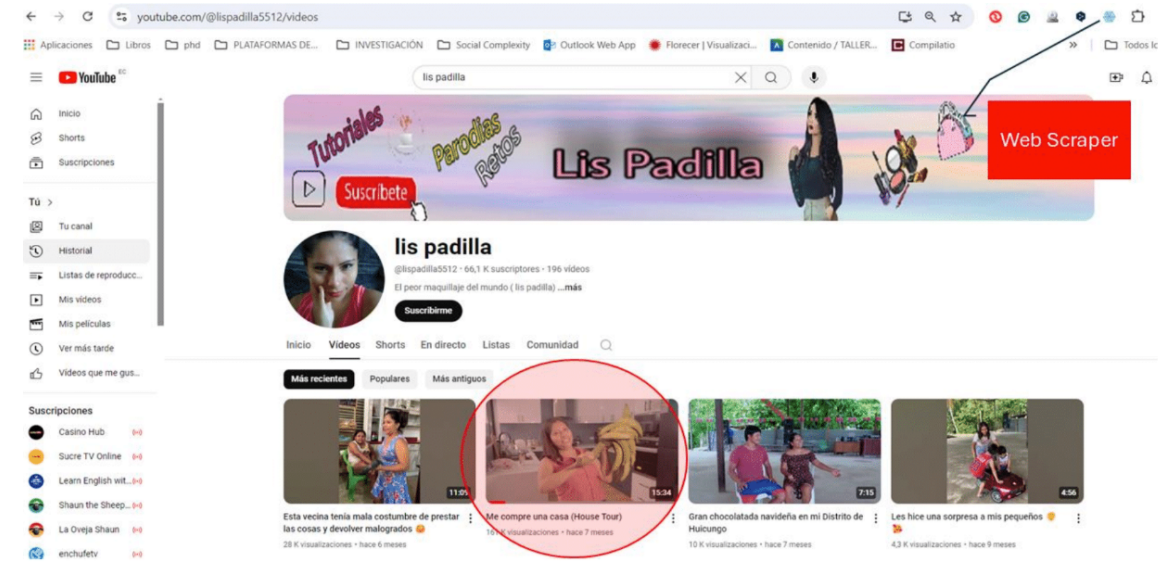

Figure 2: 2024 Lis Padilla Youtube Channel. [Public domain]

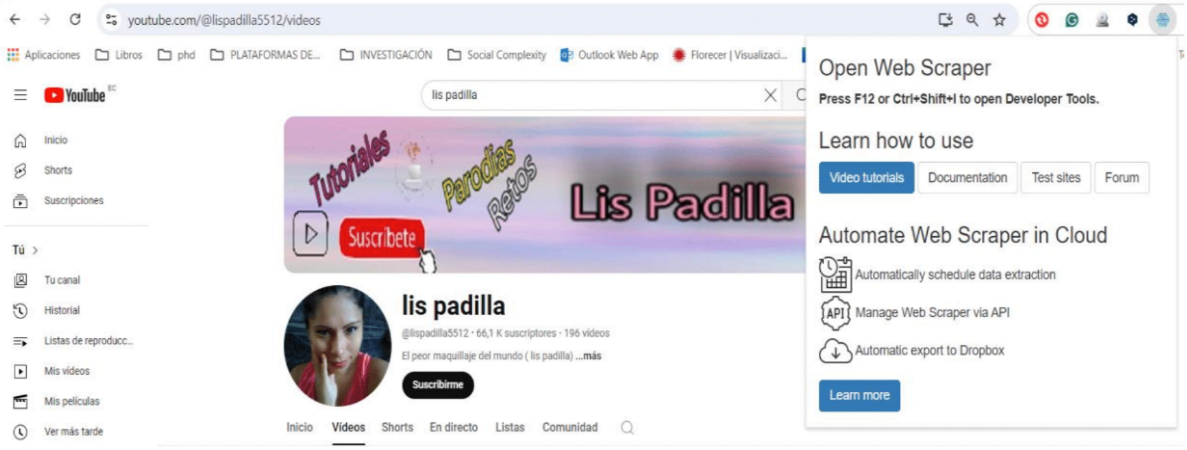

Figure 3: 2024 Lis Padilla Youtube Channel. [Public domain]

Figure 4: Create Sitemap

Now start the scraper process, and move the mouse cursor over the first video until you select the video title, only the title. Then click on the “green colored title” (see Figure 6). It should turn from green to red. Now you have to save this selection by clicking the “Done selecting” button.

2.2.2. Recommended Videos

We also need the title and URL of videos for all the "recommended videos" from within each of the featured videos but in this case only for “me compre una casa”

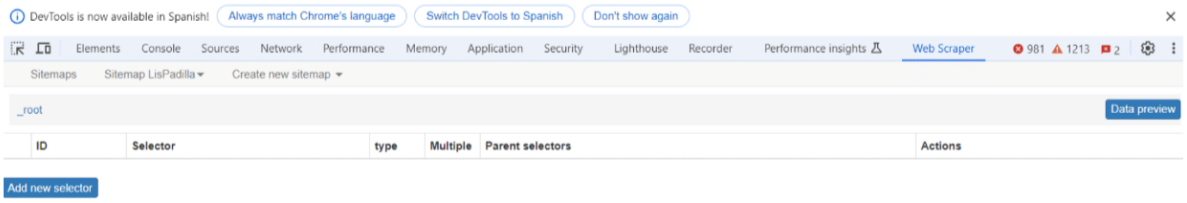

Figure 5: Directory

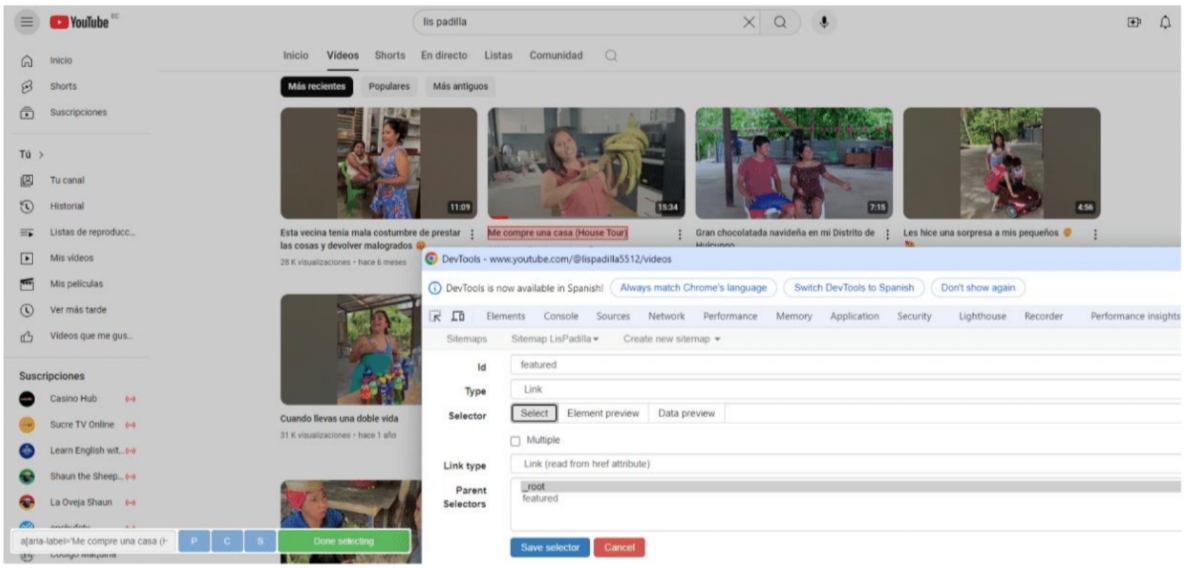

Figure 6: Select featured videos to scraper

Must now create a selector similar to the featured selector we did earlier, but for the recommended videos (see Figure 7).

- Add new selector.
- Call Id: “recommended”.
- Type: Link.
- Selector: Select.
- Hover over the first recommended video to mark it, then go to the second.
- Save with Done selecting.
- Remember the Multiple check box.
- Press: Save selector.

2.2.3. Scrape and export data

Start the scraper going under “Sitemap LisPadilla” and clicking the >Scrape< button. Then click >Start scraping< (see Figure 8).

After the scraping is done, we can output the data to a CSV file that we can open in MS Excel (see Figure 9).

2.3. Gephi

The software Gephi is a complex network visualization and analysis tool that allows users to explore and understand the complexity of relationships between elements in a dataset. Once the data related to the case study was obtained, we proceeded to structure an adjacency matrix to load it into Gephi.

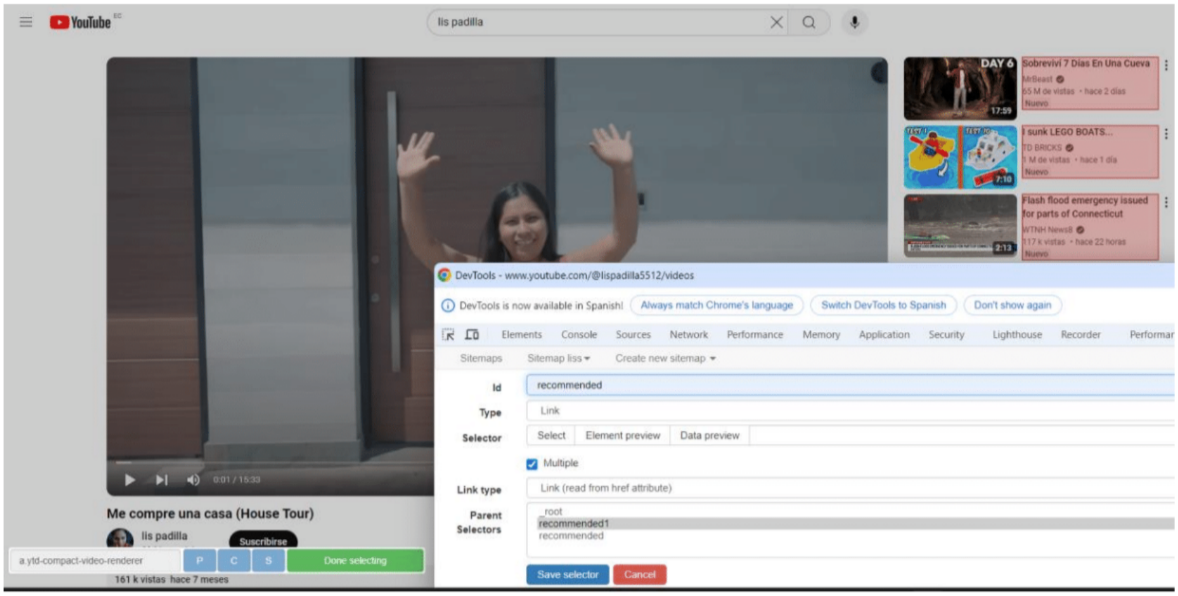

Figure 7: Select recommended videos to scraper

Figure 8: Web Scraping

3. Results

The analysis of this article was based on the parody of the song “Son de Amores”, created by Lis Padilla on June 16, 2024, in the social network TikTok, which to date has 2. 1 M likes, 38 k comments, and has been shared 701 k times; however, after circulating rumors on the YouTube social network that its creator bought a house due to the success of this parody, the search for information began to corroborate technically through the methodology of complex networks. How was this success structured in the YouTube social network?

In the YouTube social network, Lis Padilla (@lispadilla5512) now has a channel with 66k subscribers, and 196 videos uploaded with an average of 20k views; however, there are no videos related to the parody created in the TikTok social network, on the contrary, there are two videos related to the topic of buying a house with more than 100k views, the last one created at the beginning of the year 2024 that has 161k visualizations called “Me compre una casa” was selected to see if there was any relationship between the success of the parody “Son de Amores” between the social network TikTok and YouTube.

Figure 9: Browse and export data

When searching on the YouTube social network for the video "Son de Amores", the results showed the user Jesica Paucar (@JesicaPaucar-hy2uh) with the greatest viewing impact, who has 5k subscribers and has only 3 videos on her channel, one of them called "Son de Amores" uploaded two months ago; When playing the video, it was in fact the parody created by Lis Padilla on the TikTok social network and which to date had 759k views, the other two videos did not even reach 1k views.

Another user who identified herself was Isadora Baima (@isadorabaima), with 231 subscribers and a low viewing impact on her videos. However, a month ago she uploaded a video emulating what Lis Padilla did in her TikTok video, which she called "son amores meme", this video had 26k views.

The most relevant finding was from the user Hey Santana (@HeySantanaYT), a YouTuber with 233k subscribers, 206 high-impact videos with over a million views, and whose repository has 2 videos uploaded a month ago, explaining the parody made by Lis Padilla in TikTok which have 14M “Son de Amores Meme” and 7.3M “Son de Amores TikTok” views.

The scraping was applied to these four users of the YouTube social network, to structure the network that allows the data to be analyzed under the complex network methodology, the results were the following (see Figure 10):

The structured network analysis presented the results of the centrality metrics in the most important nodes (see Table 2, and Figure 11):

The results show that the user @HeySantanaYT, plays a leading role in this network, on the one hand, he has a high rating of views in each video he publishes and on the other hand, in the case analyzed, he caused the name of Lis Padilla to be positioned when explaining what he did. with his video on the social network TikTok.

In order to connect the success of the video "Son de Amores" with its owner, the betweenness centrality plays a fundamental role in this objective, as well as identifying low-impact users who took advantage of said success to raise their ratings in YouTube.

The centrality metrics (see Figure 12): allowed in this case to confirm that the YouTuber @HeySantanaYT with one of his videos related to Lis Padilla was the trigger for success on YouTube since the recommended videos are already related to other entertainment and news channels. in which they also echoed the success of Lis Padilla.

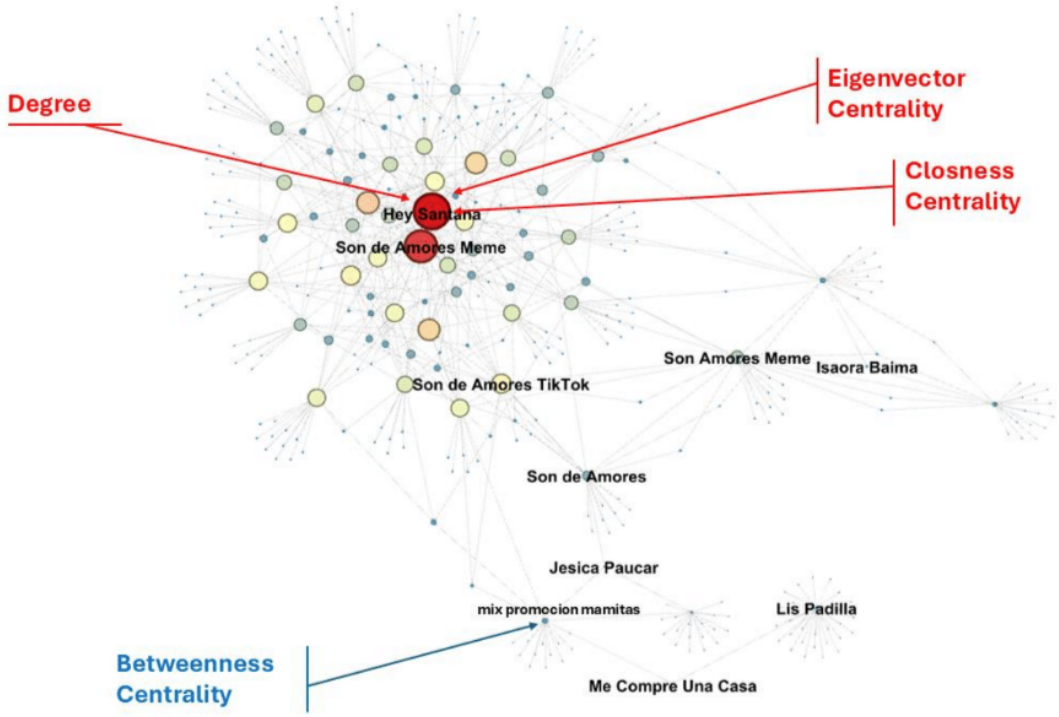

Figure 10: Complex Network for "Son de Amores"

Table 2
Centrality measures

Id	Label	Eigenvector	Closeness	Betweenness	Degree
https://www.youtube.com/@HeySantanaYT/videos	Hey Santana	1	0.37071651	0.17642455	28
https://www.youtube.com/watch?v=zPGBO6dYVoQ&t=2s	Son de Amores Meme	0.89662656	0.35172414	0.10768745	24
https://www.youtube.com/watch?v=RaIcjpqoJIU	Son de Amores TikTok	0.47517875	0.34261036	0.08049197	17
https://www.youtube.com/watch?v=zPGBO6dYVoQ	Son Amores Meme	0.27771696	0.30617496	0.17664522	23
https://www.youtube.com/watch?v=ZKZ5hcY7__c	Son de Amores	0.23772498	0.33055556	0.10950416	19
https://www.youtube.com/watch?v=S2RvzSBY1QA	Mix promocion mamitas	0.14050569	0.30177515	0.27709772	19
https://www.youtube.com/@lispadilla5512/videos	Lis Padilla	0.08896554	0.2	0.14573695	20
https://www.youtube.com/@isadorabaima	Isaora Baima	0.06405003	0.24536082	0.00205995	3
https://www.youtube.com/@JesicaPaucar-hy2uh	Jesica Paucar	0.05580324	0.27503852	0.03807058	3
https://www.youtube.com/watch?v=VPNinNoCDtc	Me Compre Una Casa	0.03369973	0.24072825	0.24496585	

4. Discussion

The data analyzed from the social network YouTube, show that the success obtained by Lis Padilla with her parody "Son de Amores", is not due to the fact that this video was in her video repository. is not because this video was in her video repository, but because a YouTuber @HeySantanaYT published two videos to analyze her parody that managed to capture millions of views, a situation that did not happen even in the original video created by Lis Padilla in the social network TikTok.

Likewise, the results show a strategy to increase the impact or rating on the YouTube social network,

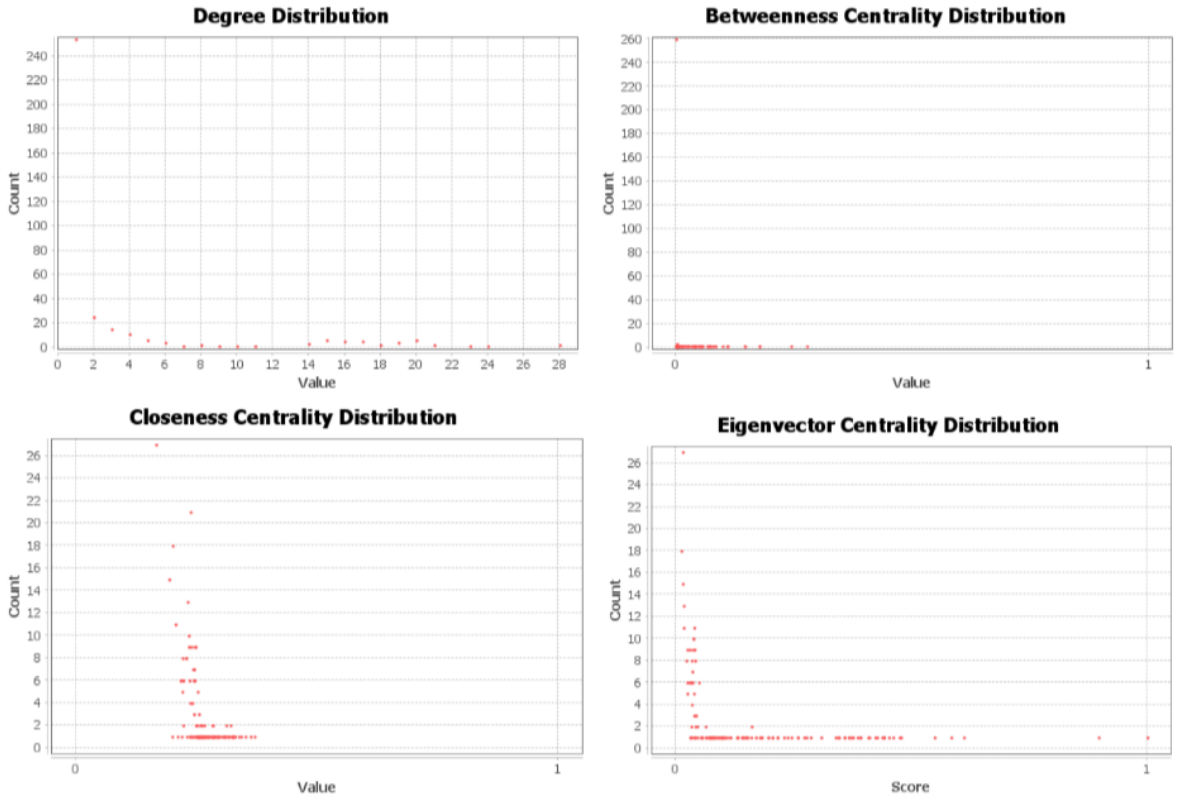

Figure 11: Centrality Distribution

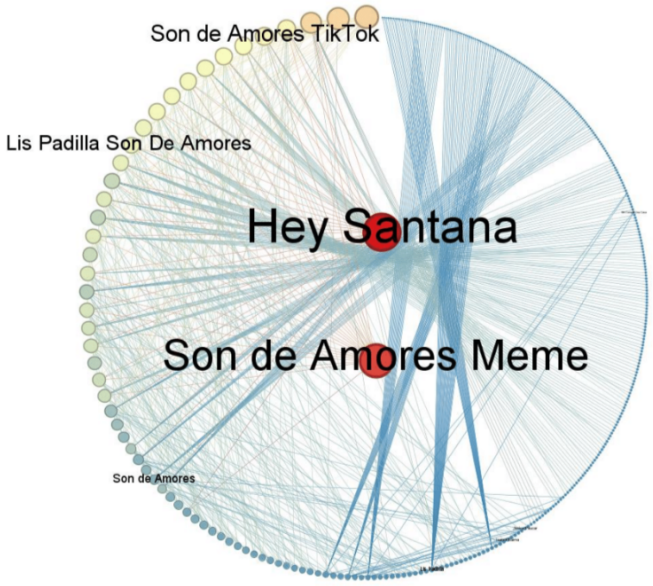

Figure 12: Eigenvector Centrality in Dual Circle Layout

publishing videos with the highest trend at the moment in your repository can position you in good places or at least increase your number of views or followers, as what happened to @JesicaPaucar-hy2uh and @isadorabaima.

Being lucky enough that an influencer or YouTuber can spread your video or explain its content, relieving its author, can catapult the content creator to success, hence the importance of YouTubers or influencers on social networks.

To identify the ideal YouTuber or influencer to achieve success, for example, of an advertising campaign or a video that you want to position in a short time among the most viewed, the complex network methodology is an ideal tool for this objective. within which, centrality metrics can technically identify the best option.

References

[1] D. Watts, Seis grados de separación, La ciencia de las re (2006).

[2] A. Barabási, Network science." philosophical transactions of the royal society a: Mathematical, Physical and Engineering Sciences 371 (2013).

[3] H. Sayama, Introduction to the modeling and analysis of complex systems, Open SUNY Textbooks, 2015.

[4] A.-L. Barabási, Ciencia de redes (Signa_Lab), ITESO, 2022.

[5] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world'networks, nature 393 (1998) 440–442.

[6] M. Pósfai, A.-L. Barabási, Network science, Citeseer, 2016.

[7] G. Budel, Y. Jin, P. Van Mieghem, M. Kitsak, Topological properties and organizing principles of semantic networks, Scientific Reports 13 (2023) 11728.

[8] L. Rivas-Tovar, Topological properties and organizing principles of semantic networks, 2024.

[9] M. E. Newman, The structure of scientific collaboration networks, Proceedings of the national academy of sciences 98 (2001) 404–409.

[10] M. E. Newman, Assortative mixing in networks, Physical review letters 89 (2002) 208701.

[11] L. C. Freeman, et al., Centrality in social networks: Conceptual clarification, Social network: critical concepts in sociology. Londres: Routledge 1 (2002) 238–263.

[12] L. C. Freeman, Visualizing social networks, Journal of social structure 1 (2000) 4.

[13] M. Hilbert, Uccss lab scrapingcomments, 2024.