The social network of the Reddit homeless community

Analysis of the r/Homeless subreddit

Fatima, Liaqat

University of Calgary, fatima.liaqat1@ucalgary.ca

Emma, K, Towlson

University of Calgary, emma.towlson@ucalgary.ca

This study employs social network analysis to investigate the structure of the r/homeless subreddit on Reddit. The study focuses on the social network nature of the subreddit, examining the likelihood of a scale-free distribution, high clustering, and low path length. The results indicate that the network follows a power-law distribution, consistent with a scale-free network. The network is dominated by a few highly connected nodes that form hubs. The hubs in the network are crucial for information flow, and the results of centrality measures identified key individuals who act as hubs. Community detection and topic identification were also employed, providing a better understanding of the social dynamics of this subreddit. Ultimately, this study presents a unique opportunity to examine the complex social structure of the homeless population and create a sense of empathy to reduce the gap between homeless people and society.

CCS CONCEPTS • Mathematics of computing~Discrete mathematics~Graph theory

Additional Keywords and Phrases: Homelessness, Social Science, Network Analysis, Social Media

1 INTRODUCTION

Homelessness is a complex, multi-faceted problem that affects millions of people worldwide and can have significant impacts on mental and physical health, education, employment, and overall well-being [1]. With the rapid growth of online social networks, people experiencing homelessness now have new opportunities to connect, share their experiences, and access resources and support [2]. One such platform is Reddit, a popular social news and discussion website that hosts a subreddit dedicated to discussing homelessness, r/homeless. While previous research has explored how homeless people use social media, there is limited understanding of the structure and dynamics of these networks [3]. Social network analysis is a powerful tool that can be used to gain insights into the structure, functioning, and dynamics of social networks. As Brandes, Ulrik, quoted in his paper:

> "Networks are important because if we don't understand networks, we can't understand how markets function, organizations solve problems, or how societies change." [4]

Many governments and not-for-profit organizations have put in place strategies to research this area and billions of taxpayers' dollars are spent every year in the hope to understand, reduce, and ultimately eradicate this problem. Unfortunately, governments are not close to reaching this goal [5]. This study employs network analysis techniques to investigate the structure of the r/homeless subreddit on Reddit, with a view to understanding the social dynamics of this platform and its potential impact on the lives of those experiencing homelessness. Specifically, the social network nature of the subreddit is examined, with known features of human social networks, such as the likelihood of a scale-free distribution, high clustering, and low path length, serving as critical benchmarks for assessment and comparison.

Further analysis is conducted on the interaction behavior of users, as well as the key individuals and groups that shape the network, providing a better understanding of the social dynamics of this subreddit. Community detection and topic identification are also employed, with a focus on the formation of meaningful communities that can be interpreted coherently and logically. The presence of coherent themes and groupings would validate the meaningfulness and reliability of the identified communities. Ultimately, this study presents a unique opportunity to examine the complex social structure of the homeless population using a social network analysis approach and a non-traditional data collection methodology. The goal of this study is to learn something meaningful about this population, create a sense of empathy, and reduce the gap between homeless people and society. This, in turn, may motivate others to develop targeted interventions and resources to support this population.

2 RELATED WORK

Mago et al. [2013] paper focuses on analyzing the factors of homelessness. In this article, they discuss that there is no single cause of homelessness, as well as it is very difficult to define and measure homelessness because this problem is dynamic in nature, which makes it harder for the government and related organizations to effectively address it. They proposed a new approach to analyzing these factors and highlighted the need for targeted policies and interventions. Our report also emphasizes the importance of addressing the issues raised by the r/homeless subreddit member and suggests that network analysis can be used to identify key actors and connections within the community that can inform policy and intervention strategies.

Bandari et al. [2021] used a qualitative method (thematic content analysis) to analyze 360 unique posts from the r/homeless subreddit from the year 2019. They individually analyze all 360 posts and sub-categorize them. They found 4 main themes in the dataset: (a) social issues (like a shortcoming of social services, violence against LGBQT+ people, etc.) (b) communication of needs and concerns, (c) offering care and support, and (d) online community management and engagement. Their research supported that r/homeless subreddit was a valuable resource to access the hidden homeless groups (e.g., cars, dwellers, rural homeless) who often get neglected in the literature. This supports our use of data from the subreddit to gain insights into the behavior and interactions of homeless individuals online. This study only focused on the content of the network, which opens a door for our study to analyze the structure of their interaction behavior.

The paper by Buntain et al. [2014] aims to understand user interaction and identify answer-person roles on Reddit. The authors collected data from events where users participate in "IAMI" events on Reddit and constructed a network to identify the answer-person role. They used a machine learning algorithm to design a model to identify these roles across their dataset with 80% accuracy. The study shows the role of the answer-person user in information spreading and bridging information gaps in the community. The paper supports the use of network analysis techniques in answering questions on users' interaction behavior and promotes future researchers to investigate such behaviors across borders.

This study is by Rice et al. [2012] evaluated the acceptability of a hybrid face-to-face and online social networking HIV prevention program for homeless youth. Peer-led intervention successfully recruited peers and online participants, proving that the new intervention model was acceptable and feasible for providing intervention prevention to vulnerable populations. The study promotes collaboration between multiple stakeholders like the homeless population, service providers, and policymakers. However, limitations include relying on self-reported data and a small dataset. Our report also highlights the importance of collaboration between different stakeholders in addressing issues related

to homelessness and suggests that network analysis can provide a way to identify potential collaborators and key connections within the r/homeless subreddit.

3 METHODOLOGY

This section will discuss the dataset and methods adopted by this paper to perform network and content analysis. To answer our main question, we sub-divided our research into two main categories; 1) finding structure and hub nodes of our network and 2) how communities are formed and what are their mutual topics of discussion.

3.1 Data Collection and Cleaning

In this study, we collected data from Reddit using the Python Reddit API Wrapper (PRAW) library with an authenticated account and a single PRAW instance. Reddit's strict data crawling limitations required us to extract posts using different categories that included: "new," "hot," "controversial," and "Top." This resulted in 2512 posts and their associated data, including the author's username, post score, title, content, and timestamp. We retrieved a total of 36379 comments and replies made between January 2020 and February 2023 using this information.

The data was divided into eight different files and stored in two main categories: network creation files and content analysis files. We merged the network creation files as a Pandas data frame for cleaning and wrangling. Duplicate and null values were removed from the post id, comment ids, and reply's ids columns. The comments and replies related details were initially stored in separate columns, but we merged these columns as comments that required columns renaming, merging, and dropping. Afterward, we stored our nodes and edges in CSV files.

In the content analysis files, we cleaned the posts file by removing any duplicates and null values, merged the post title and post content parts, and filtered our post content to keep only the content created by the users in our weakly connected subgraph. We then merged our comment content into Pandas Dataframe which was initially stored in three separate files due to the large file sizes. Like what we did to network files, we had to merge the comment and replies to related columns and remove duplicates and null values. We filtered the content of our comments based on our weakly connected component users (nodes). Lastly, we merged the post and comment Dataframes into one content Dataframe.

3.2 Network Creation and Measures

For our project, we utilized Pandas library for easy data manipulation and used the networkx library to construct a weighted undirected network graph. The graph aimed to capture the flow of communication within the network and to identify distinct communities within it. The graph nodes corresponded to the authors of the posts, comments, and replies, with the edges denoting any communication between the author of the post and that of the comment or reply, and the weights captured the number of interactions between them. Our initial graph contained 7150 nodes and 36379 edges but after removing self-loops, and parallel edges and keeping only the largest weakly connected component of the original graph we got 7126 nodes and 19743 edges.

We calculated network metrics using network methods. To answer our first question, we plotted the degree distribution of the graph on a log-log scale. To check if our graph is scale free[1] (a network with a degree distribution that follows a power law distribution, that is, a straight line on a logarithmic plot with a negative slope. This property of scale-free networks is characterized by the presence of a few highly connected nodes, known as "hubs," which account for a disproportionate number of edges in the network. [6]), we fitted power law distribution on graph data

[1] A network that has a few nodes of very high degree (hubs), while most nodes have relatively low degrees.[6]

using the power-law library from Pipy and calculated the clustering coefficient using networkx. Clustering coefficient measures the degree to which nodes in a network tend to cluster together, i.e., how likely it is for two nodes that share a neighbor to also be connected to each other. [6]. We also calculated the average shortest path length of the graph to see the small-world effect in the network. Small-world effect is the phenomenon where many real-world networks have both a high degree of clustering and a short average path length between nodes [6]. Scale-free and small-world properties are common characteristics of many social networks, reflecting the underlying mechanisms of network formation and growth in these systems. We further calculated centrality measures, a family of metrics that aim to quantify the relative importance or influence of nodes in a network [6]. We calculated degree[2], betweenness[3], and closeness[4] centrality to identify hub nodes in the network. We created an ensemble of null model by running 1000 iterations of a connected double-edge swap model, it's a model for generating random graphs that preserves the degree distribution of the original network and ensures the resulting graph is connected while introducing randomness through edge swaps. We compare its result with the original graph to rule out any randomness in the original graph.

To detect communities[5] in our network we used networkx's Louvain Community Detection algorithm. It is a greedy, hierarchical clustering algorithm that maximizes modularity (a value between -1 and 1 that measures the density of links inside communities as compared to links between communities) [11]. We ran 400 iterations of the algorithm and found 0.9500 as the optimum value for the resolution[6] parameter, as it gave the highest modularity. To find the communities in the network we ran 1000 iterations of the algorithm with a resolution value of 0.9500 to find the best partitions/communities of the network that had the highest modularity.

3.3 Data Preparation for Content Analysis

We used pandas, nltk and re libraries, and N_grams model (a statistical language model that is used to predict the probability of the next word in a sequence of words given the previous words in the sequence. [12]) to create a function that processes the content to make it ready for analysis. The function checked for Nan values, replaced any URLs, usernames, or alphanumerical values, and tokenize text. It removed all stopwords and punctuation from the text as they do not contribute much to the topic. We filter the words from each post such that no duplicates appear in the processed set of each post or comment. To find the themes in each community we grouped the content based on their community numbers and considered the top 15 most common bigrams[7] from each community. To obtain a concise comprehension of the hubs' nodes' significance within the network, based on the content, we analyzed the highest-ranked content attributable to the top four hub nodes that were identified in centrality measures.

4 RESULTS

This section will discuss the results of our network and content analysis.

---

[2] The degree of a node is just its number of links or connections to other nodes, so the degree centrality of a node is just its degree. [6]

[3] Betweenness centrality is a measure of the extent to which a node lies on paths between other nodes in the network. [6]

[4] Closeness centrality is a measure of the average distance of a node to all other nodes in the network. [6]

[5] A community is a set of nodes that are more densely connected to each other than to the rest of the network.

[6] Resolution parameter is a tuning parameter, which determines the level of granularity in the community structure by affecting the balance between the preference for smaller versus larger communities. A higher resolution parameter tends to produce smaller communities, while a lower resolution parameter tends to produce larger communities. [11]

[7] a 2-gram (which we'll call bigram) is a two-word sequence of words [12]

4.1 Network Analysis

We started our analysis with the observation of the degree distribution of the graph, these results showed a heavy-tailed distribution, and the graph showed a power-law distribution with a scaling exponent of approximately $\alpha = 3.01$. The graph is an excellent fit to the Barabási-Albert (BA) model[8], which predicts an exponent of exactly three. The model is rooted in preferential attachment mechanism, which gives an advantage to nodes that are already well-connected. As a result, the network evolves in a way that the rich get richer, and nodes with high degree tend to attract more links, resulting in the power-law degree distribution, see Figure 1, and a scale-free network. The value of 0.0277 for the average clustering coefficient suggested that the network has a relatively low level of clustering, which is consistent with what we would expect from a scale-free network. The results of power law and clustering coefficient verify that it is indeed a scale-free network, which means that the network is dominated by a few highly connected nodes that form hubs, while most nodes have relatively low degrees and are not as well connected. The graph shows an average Shortest Path Length of 4.0152. These combinations of properties mean that the nodes in the network tend to form communities, while still allowing for efficient communication and information flow across the network.

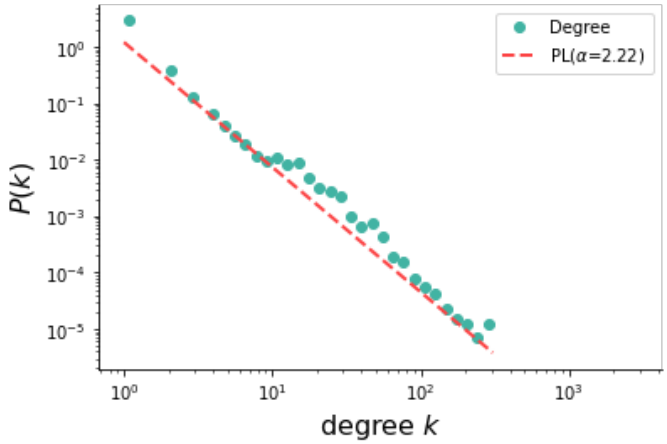


Figure 1: The power-law distribution and the degree distribution of the r/homeless subreddit. The graph is plotted on logarithmic scales.

To evaluate the significance of our results we compare it with the null model. The results of the null model showed a clustering mean of 0.0191 with a standard deviation of 0.0010. The null model results also showed an average shortest path length of 4.0062 with a standard deviation of 0.0078. This indicates that the graph has a small-world property, which is a characteristic of many real-world networks. These results suggested that our original graph has a higher clustering coefficient and a slightly longer average shortest path length than the null model see Table 1 for comparison. Surprisingly, the difference in average shortest path length of the original graph and null model is not as significant as we had expected. Nonetheless, these results indicate that the structure of the original graph is not random and has some underlying organization and pattern.

8 The Barabási-Albert (BA) model is a widely used model for generating scale-free networks.[6]

Table 1: Comparing values of original graph with null model.

Measure	Original Graph	CDES null Model	Standard Deviation
Clustering Coefficient	0.0277	0.0191	+ / - 0.0010
Shortest Path Length	4.0152	4.0062	+ / - 0.0078

The network's scale-free property with low clustering suggests that the nodes are less likely to form clusters around the hubs which is due to a small-world structure of the network where nodes are connected to distant parts of the network through a few intermediary nodes.

As mentioned above, the hubs exert a significant influence over the network's structure and dynamics. The hubs in the network are crucial for information flow, as they have the potential to reach many nodes with a few connections. The results of centrality measures identified such hubs in our graph, see Figure 2. We found that users "Grassyhobo" and " veryberryblue " are the hub nodes with the highest degree in our graph, respectively, see Table 2. Indicating that these users are a central point of information flow in the network, as they receive a lot of information from others, and they are also the originators of information. The users "Grassyhobo" also has the highest betweenness centrality with "MrArmenian" having the second highest score in this centrality measure and can be seen as a bridging node between different group of nodes within the network. That indicates that these nodes are important for maintaining the network's overall information flow, connectivity, and structure. Lastly, we identify the nodes' closeness centrality measure and found that the user "Liquidmemer" has the highest rank with "Veryberryblue" being in the second position in this centrality measure. These users are those that can reach other nodes in the network quickly and are therefore important for information dissemination and efficient communication within the network. Their high closeness centrality makes them more accessible and influential within the network, allowing them to play a key role in decision-making processes or other important functions. We will discuss hubs more in the content analysis section.

Table 2: Top ten hub nodes based on degree centrality measure.

Node (User)	Degree	Node (User)	Degree
Grassyhobo	306	DJ44x	235
veryberryblue	290	survivalmany	206
MrArmenian	268	HomelessOnReddit	203
Liquidmemer	265	Hannahpenns	196
Iamshamtheman	246	periwinkletweet	183

Using Louvain community detection algorithm we identified 13 different communities in the network, see Figure 2. This partition of the network has the highest modularity value of 0.5111 out of all the 1000 iterations. Each community represents a subset of nodes in the network that are more strongly connected to each other than to nodes in other communities. This is useful information for understanding the structure of the network and identifying groups of nodes that are functionally related, serve a similar purpose or look for common topic of discussion. We will discuss the results of these findings more in the content analysis section.

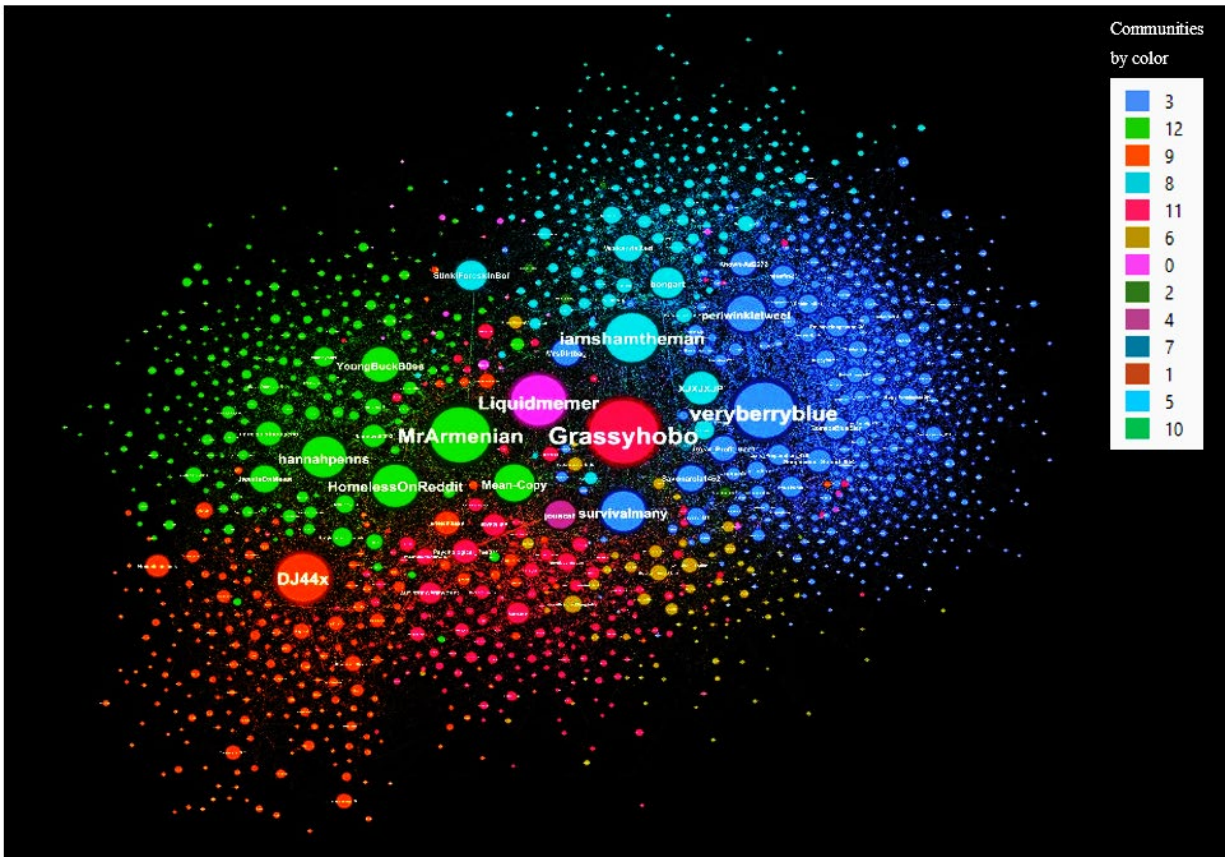


Figure 2. r/homeless network colored by communities and node size by degree centrality (Layout used – ForceAtlas [is a force-directed layout algorithm used in Gephi, a popular network visualization and analysis tool.])

4.2 Content Analysis

The top 15 most frequently occurring bigrams[9] in the text corpus were identified, including phrases related to job searching, financial difficulties, and the need for shelter and social support. These findings suggest that these topics are the most pressing concerns for the homeless population on the r/homeless subreddit, see Figure 2. Furthermore, we conducted content analysis for each community and identified community wise top 15 most common bigrams. These bigrams revealed the primary areas of interest and concerns of each community. For community 0, the phrases like "Merry Christmas", "missing Christmas", and "Covid Watford" suggest users (nodes) are discussing Christmas and the pandemic, see Figure 3. While they are celebrating Christmas with cozy feelings, they are also concerned about dangerous sleeping conditions due to the extreme temperature and COVID-19 pandemic in Watford, UK. Community 1 talks about legal and social issues in Canada and Community 2 talks about finding the nearest toilets, rescue situations, health costs, and camps in Texas. Community 3 seems to be concerned about living situations such as the need for shelter, work, and money to pay rent and buy food. They also express hope and ask for help from friends and

9 Word pairs

family, see Figure 3. Community 4 discussed the need for work in England and Community 5 discussed raising awareness via different modes. Community 6 discusses the course of life and various other concerns, including emergency shelters, park living, and money to pay for shelter. Community 7 talks about family and the struggle to buy things, especially during winter, and their religious beliefs. Community 8 discusses issues such as staying social and friends' help, working to afford rent, and staying in a place. They also express concerns about drugs, health, reading places, and phone issues. Community 9 talks about facing financial challenges, and a need for assistance with basic needs such as food, shelter, and healthcare. They are also seeking support for their loved ones and a desire for a sense of normalcy and stability in their lives. However, despite these difficulties, there are some positive expressions of gratitude and hope, such as "bless god", see Figure 3. Community 10 is talking about housing and queers and their health. Community 11 appears to be struggling with basic needs and seeking support for finding stable housing, employment, and resources for their families. Lastly, Community 12 expresses a desire for a stable living situation, help with rent, assistance with finding jobs, and support from friends and family, see Figure 3. The community also values cleanliness and a safe environment. Overall, this analysis provides insight into the varied interests and concerns of different communities within the larger population of interest, which can inform the development of targeted interventions and resources.

According to nodal level content analysis on our top four hub nodes, Grassyobo shared the most (93) predominantly comments, focusing on sharing homeless camp life experience and social support. The highest-scored content by the user reflected the experience of losing their camp to a Bulldozer, which received 175 upvotes. Veryberryblue shared 44 primarily posts, reflecting on experiences of homelessness and seeking advice. The user's highest-scored content reflected the experience of feeling unsupported during times of adversity, despite previous assurances of social network support, which received 169 upvotes. MrArmenian shared 29 mostly positive and supportive posts on available support services and celebrated their 1-month sobriety and received 614 upvotes. This user's communication is notable for receiving a high number of upvotes, indicating a desire within the subreddit community for positive and supportive content. Liquidmemer shared 28 mixed posts and comments focused on camaraderie among homeless individuals, gratitude, and the challenges of living on the streets. The user's highest-scored content focused on being thankful for finding wifi and power. Again, this user received high upvotes for their positive communication, indicating a desire within the subreddit community for such content.

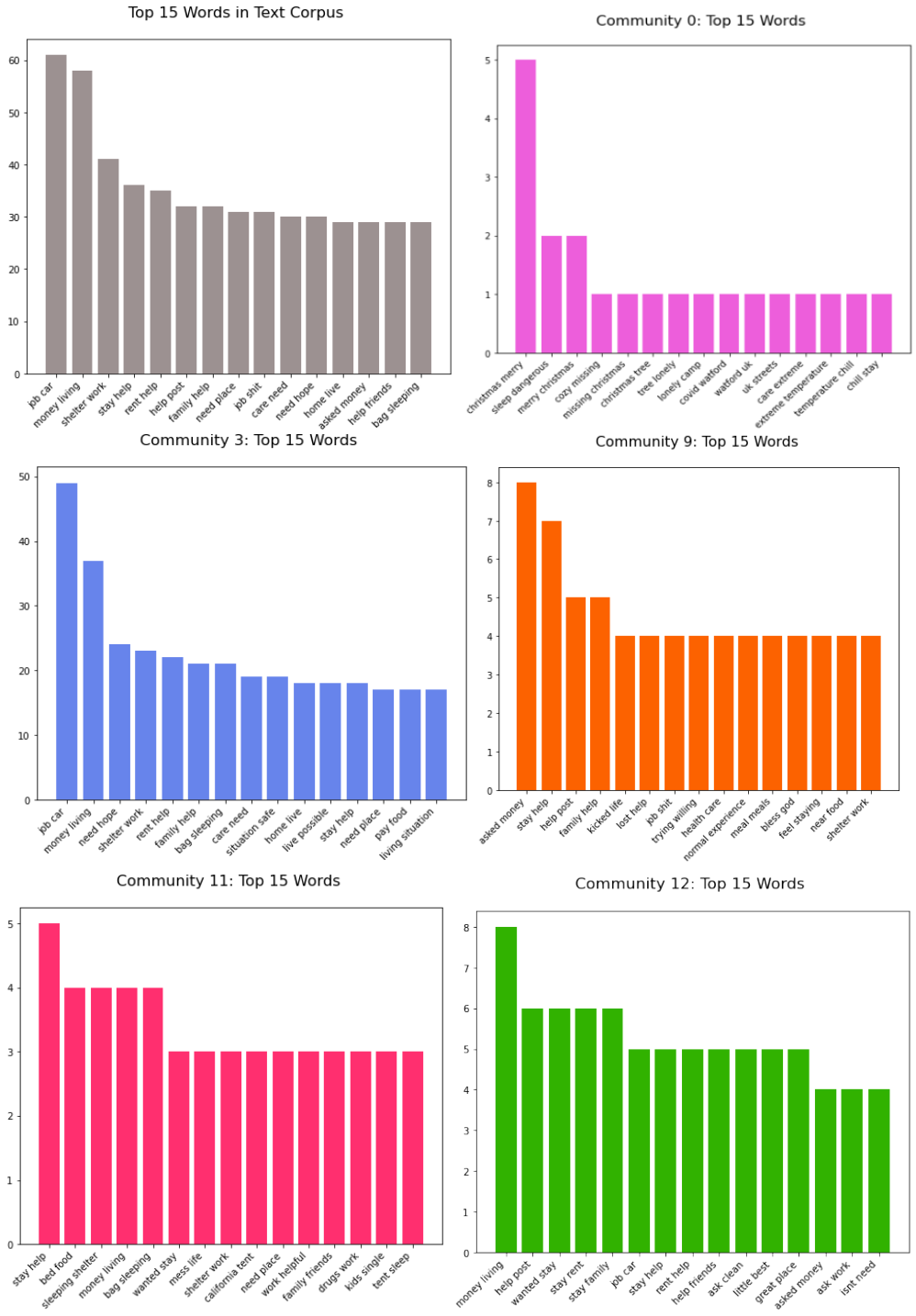


Figure 3. The bar charts show the top 15 bigrams in both the network and individual community. The color variations of bar charts correspond to the distinct hues assigned to each community in Figure 2.

5 DISCUSSION

The results show that this network has many similarities to other social networks that have been studied in the literature, thus verifying that this network is like any other social network that has scale-free behavior and small world property. Though the results that we found are limited to a subset of this subreddit. This network is the representative of the two years (January 2020 – February 2023) and we may gain additional findings if we consider other years. Especially the results of the centrality measure might change. Nonetheless, we can say that the network would still be centralized and hierarchical in nature, where some nodes will be more dominating or influential than others. Governments and not-for-profit organizations can effectively monitor or access hub nodes to identify valuable properties or organizations. The development of strategies using subreddits and hubs to spread critical information about resources and support programs could prove to be a cost-effective and time-efficient approach to reaching a larger and underprivileged homeless population. Therefore, this study presents a holistic report on the potential benefits of utilizing such strategies, which could serve as a crucial contribution to addressing the complex issue of homelessness.

It is necessary to delve deeper into the nature of hub nodes, examining not only their centrality in the network but also the type of nodes they represent and how they interact - within one community or across many. Furthermore, investigating whether the scale-free property in the graph can be attributed to homophily, i.e., the tendency of individuals to form connections with others who share similar characteristics, is a crucial next step. In future work, further investigation is needed to understand the factors driving the formation of communities around specific topics. It is also vital to examine deeply the nature of the interactions among hub nodes, which hub nodes provide support, which seek support, and which activities garner the most attention from other users in the network. These findings could provide valuable insights for policymakers, non-profit organizations, and other stakeholders in developing targeted support programs and interventions for people experiencing homelessness.

6 CONCLUSION

In conclusion, the network analysis of the r/homeless subreddit provided insights into the structure, dynamics, and concerns of these online communities. The result of power law distribution proved that the Barabási-Albert model provided a good fit to the data, indicating that the network follows a scale-free model with highly connected hub nodes. The presence of highly connected hub nodes with high degrees, betweenness centrality, and closeness centrality indicates their significance in maintaining the network's structure and dynamics. The identification of communities within the network highlights the existence of functionally related groups of nodes. The content analysis of the r/homeless subreddit revealed the most pressing concerns of the homeless population and the significant role of individuals with lived experience of homelessness in providing support and advice. These findings have important implications for the development of targeted interventions and resources to support the homeless population. Future research could explore the dynamics of these networks over time and examine how external factors or interventions affect their structure and dynamics. Overall, this study demonstrates the value of network analysis and content analysis in understanding the complex social dynamics of online homeless population.

ACKNOWLEDGMENTS

We would like to acknowledge the usage of the Text Processing and N-Grams functions code from the CPSC 572 tutorial example file on supervised machine learning. This code has been instrumental in providing us with the necessary tools to analyze and process textual data effectively.

REFERENCES

[1] Vijay K Mago, Hilary K Morden, Charles Fritz, Tiankuang Wu, Sara Namazi, Parastoo Geranmayeh, Rakhi Chattopadhyay & Vahid Dabbaghian 2013. Analyzing the impact of social factors on homelessness: a fuzzy cognitive map approach. Journal of BMC medical informatics and decision making, 13, 94 (2013). Retrieved January 20 from DOI: https://doi.org/10.1186/1472-6947-13-94.

[2] Mary Yost 2012. The Invisible Become Visible: An Analysis of How People Experiencing Homelessness Use Social. Elon Journal of Undergraduate Research in Communications 3 (2012), 3 pages. Retrieved January 21, 2023, DOI: http://www.inquiriesjournal.com/articles/830/the-invisible-become-visible-an-analysis-of-how-people-experiencing-homelessness-use-social-media.

[3] Aparajita Bhandari, & Billie Sun 2021. An online home for the homeless: A content analysis of the subreddit r/homeless. Journal of New Media & Society 0, 0 (Oct. 2021). Retrieved January 18, 2023, from DOI: https://journals.sagepub.com/doi/10.1177/14614448211048615.

[4] Ulrik Brandes, Garry Robins, Ann McCranieAnn, Stan Wasserman 2013. Journal of Network Science 1, 1 (Apr. 2013), 1-15 pages. Retrieved January 27, 2023, from doi:10.1017/nws.2013.2.

[5] Casey Thomas. 2022. Chronic Homelessness. Independent Auditor's Report. Office of the Auditor General of Canada, Ottawa, Canada. Retrieved December 15, 2022, from DOI: https://www.oag-bvg.gc.ca/internet/English/parl_oag_202211_05_e_44151.html.

[6] Mark. E. J. Newman. 2010. *Networks: An Introduction* (1st. ed.). Oxford, New York, NY. Retrieved on March 20 from DOI: https://math.bme.hu/~gabor/oktatas/SztoM/Newman_Networks.pdf

[7] Reddit 2023. Reddit r/homeless. Retrieved January 17, 2023, from DOI: https://www.reddit.com/r/homeless/.

[8] Praw 2023. PRAW: The Python Reddit API Wrapper. Retrieved January 18, 2023, from https://praw.readthedocs.io/en/stable/.

[9] Cody Buntain and Jennifer Golbeck. 2014. Identifying social roles in reddit using network structure. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion)*. ACM Inc, New York, NY, USA, 615–620. https://doi.org/10.1145/2567948.2579231.

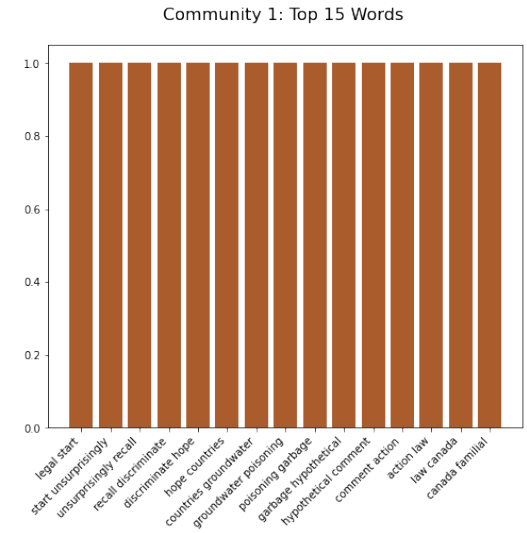
[10] Eric Rice, Eve Tulbert, Julie Cederbaum, Anamika Barman Adhikari, Norweeta G. Milburn 2012. Mobilizing homeless youth for HIV prevention: a social network analysis of the acceptability of a face-to-face and online social networking intervention. Journal of Health Education Research 27, 2 (Apr. 2012), Pages 226–236. Retrieved on February 10 from DOI: https://doi.org/10.1093/her/cyr113.

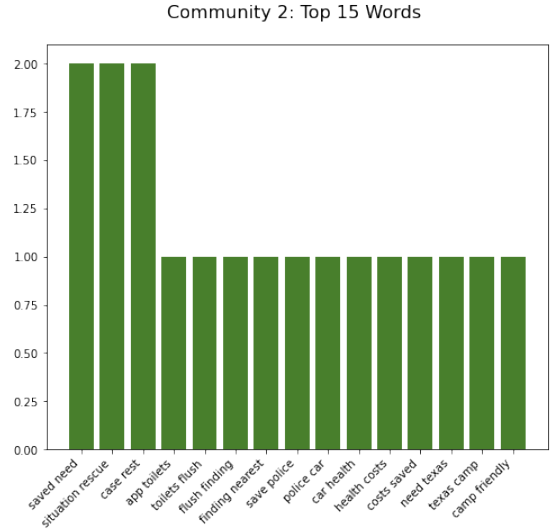
[11] Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, & Etienne Lefebvre 2008. Journal of Statistical Mechanics Theory and Experiment 2008. Retrieved on April 10 from DOI: 10.1088/1742-5468/2008/10/P10008

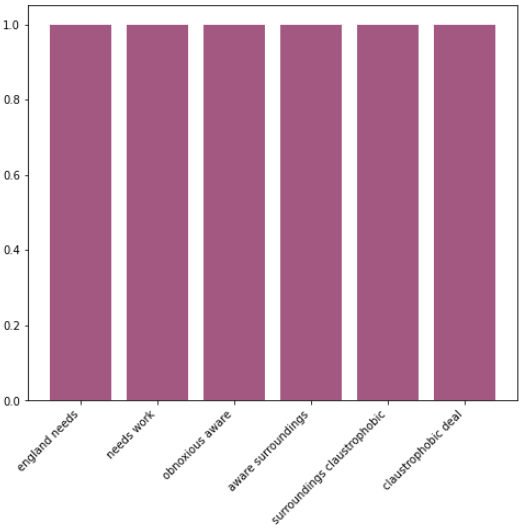
[12] Jurafsky, D., & Martin, J. H. (2022). Speech and Language Processing (3rd ed. draft). Retrieved on April 11 from DOI: https://web.stanford.edu/~jurafsky/slp3/3.pdf

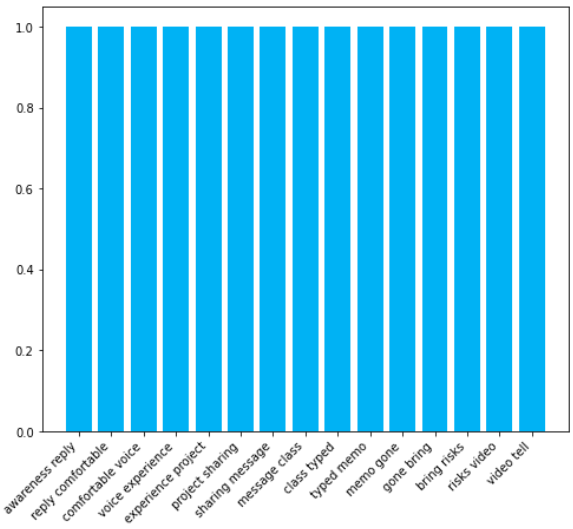
APPENDIX

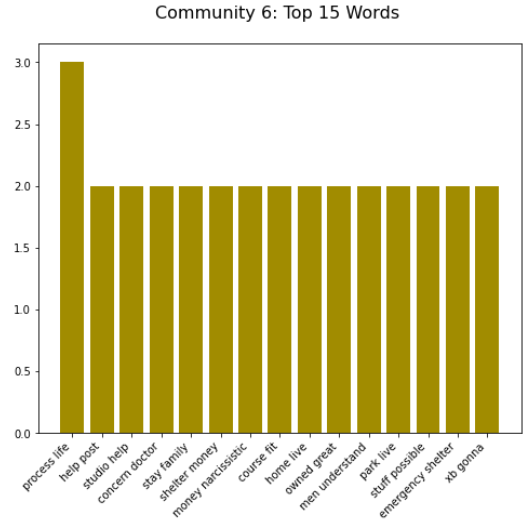
This section provides the visuals of the remaining communities' content analysis.

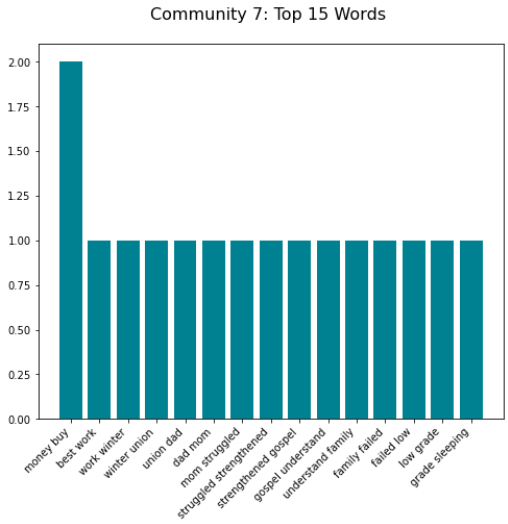


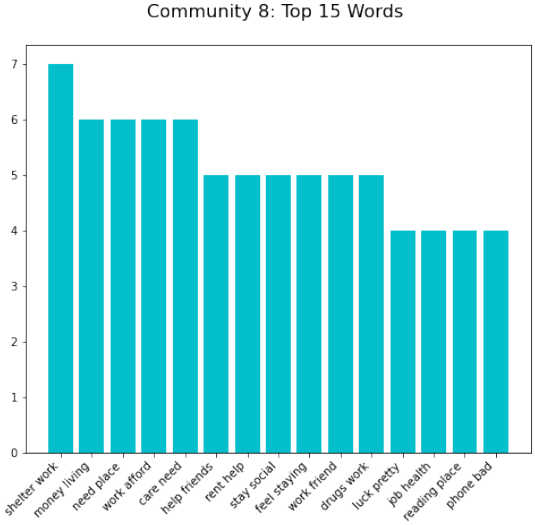


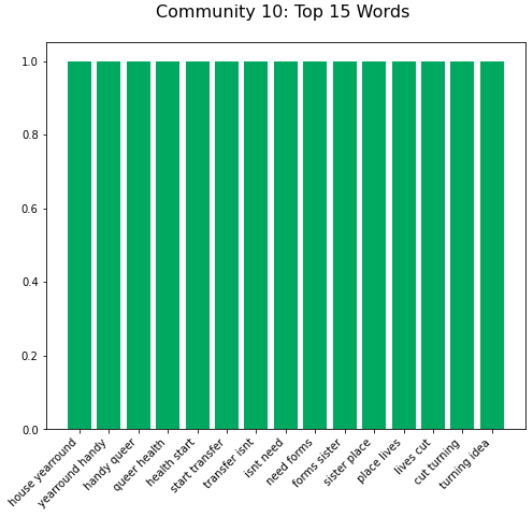





Community 6: Top 15 Words
process life
help post
studio help
concern doctor
stay family
shelter money
money narcissistic
course fit
home live
owned great
men understand
part live
stuff possible
emergency shelter
xb gonna


Community 7: Top 15 Words
money boy
best work
work winter
winter union
union dad
dad mom
mom struggled
struggled strengthened
strengthened gospel
gospel understand
understand family
family failed
failed low
low grade
grade sleeping


Community 8: Top 15 Words
shelter work
money living
need place
work afford
care need
help friends
rent help
stay social
feel staying
work friend
drugs work
luck pretty
job health
reading place
phone bad


Community 10: Top 15 Words
house yearround
yearround handy
handy queer
queer health
health start
start transfer
transfer isnt
isnt need
need forms
forms sister
sister place
place lives
lives cut
cut turning
turning idea