Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2024) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

5th International Conference on Innovative Data Communication Technologies and Application (ICIDCA 2024)

Echoes of Truth: Unraveling Homophily in Attributed Networks for Rumor Detection

Rithish S.V, Prabu C.R, Anuush M.B, Deepthi L.R

[a]*Department of Computer Science and Engineering,*
Amrita School of Computing,
Amrita Vishwa Vidyapeetham,
Amritapuri, India

Abstract

The surge in information and misinformation during the COVID-19 pandemic, particularly regarding 5G-related rumors, was prominent on various social media platforms, notably Twitter. This study addresses the pervasive issue of rumor propagation on Twitter by employing diverse machine learning models, clustering techniques, and node detection algorithms to gain insights into the dynamics of rumor dissemination. The primary focus is identifying influential individuals who significantly impact the spread of these rumors and had an accuracy of 0.86 while classifying various labels. The study reveals network structures and key influencers within each cluster by categorizing users into clusters. While it is commonly observed that users with a more extensive follower base have a broader impact on rumor dissemination, our innovative approach integrates advanced community detection algorithms and methods, the Overlapnodedetect algorithm and Rumorsourcetrace algorithm for identifying influential characters to uncover latent homophily structures within these clusters. The objective is identifying crucial users responsible for initiating and perpetuating these rumors.

Keywords: homophily, community, rumor, complex networks, social networks, attributed networks

1. Introduction

The present social media usage by around 4.9 billion individuals demonstrates these platforms' substantial influence and extent. Nevertheless, besides the expanding number of users, the statement acknowledges the potential

* Rithish S.V
E-mail address: rithish675@gmail.com,crvprabu@gmail.com,mbanuush880@gmail.com, lrdeepthi2002@gmail.com

negative consequence of this expansion—rumors. An estimated increase in social media users to around 5.85 billion by 2027 indicates that a larger population will be exposed to information shared on these platforms. Although advantageous in several ways, the enhanced interconnectedness and dissemination of information also pose issues such as the proliferation of rumors and disinformation. A rumor is a disclosure of information not independently verified throughout the communication. Social media platforms have probably contributed to the spread of rumors and their potential to reach a larger audience than before. Numerous machine learning models, artificial intelligence (AI) models, natural language processing (NLP), and deep learning (DL) algorithms are put forth by the researchers to detect different forms of rumors, however, limited research is focused on rumor detection on attributed networks.

In the world of attributed networks, where people are characterized by an array of features encompassing their social interactions, biological connections, and virtual footprints, forming cohesive communities plays a vital role. This paper delves into the compelling context of records diffusion at some point in the COVID-19 pandemic, particularly inside the sphere of Twitter. This work aims to identify the driving forces behind the propagation of rumors by deploying an overall and comprehensive technique. The proposed method performs the classification of categorized data sets using a spectrum of algorithms, followed by identifying the source of the rumor.[2] The merits of this method yield distinct communities, everything with its particular narrative. These groups provide helpful insights into their members' collective behavior and underlying motivations. This insight, in turn, unravels the critical effects steering the variety of rumors.

Attributed networks offer an effective framework for analyzing the complex nature of Human interests, regardless of whether they occur in the real world or online. Within these networks, people are classified by various characteristics, including their social connections, online behaviours, and relationships. These characteristics can be behavioural patterns and content choices, creating a complicated web of elements that help us understand their places in the network. Due to their ability to provide a more comprehensive examination of controlling the flow of information, influence, and actions through these complex systems, attributed networks have become increasingly popular in identifying an individual interest.

They are especially useful in interpreting the way connected communities develop in these networks, giving knowledge of the mechanisms behind the behaviour and formation of these communities. Attributed networks effectively analyze the complex web of interpersonal relationships, motives, and group behaviors. It additionally assists in understanding how information and rumors spread, as in the COVID-19 pandemic on social media sites like Twitter. Identifying these sources of the rumor spread and revealing the relevant details about them will help ride the unfolding of misinformation. At the same time, popularity-based detection can effectively identify patterns and anomalies in large datasets. Additionally, popularity-based detection may fail to identify legitimate but less popular content or behavior. Currently, the rumor is predicted based on the number of friends and followers and making them into the group. Still, the algorithm cannot find the source person responsible for this rumor spread to break out. This study complements our understanding of the mechanisms underlying rumor dissemination and presents precious insights into effective techniques for countering it.

The major contributions are:

- Classification of rumors based on machine learning algorithms.
- Overlapnodedetect algorithm, which detects overlapping nodes based on attributes.
- Rumorsourcetrace algorithm, which detects rumor sources by employing the attributes.

The paper is organized as follows: Section 2 covers relevant work, Section 3 discusses methodology, Section 4 presents rumor categorization, Section 5 addresses rumor detection, and Section 6 covers the conclusion.

2. Related works

Recent research has shown that combining data preprocessing, hyper-parameter tuning, and multi-model classification can effectively detect rumors. Data preprocessing is essential for rumor detection, as it can help remove noisy and missing data, improving the performance of classification models. The labeling of the data [1] technique is used to assign labels to unlabeled data points based on the labels of their neighbor. Balancing[4] strategies are used to

address the class imbalance problem, which occurs when one class is much less frequent than the other. Hyperparameter tuning is the process of finding the optimal values for the hyperparameters of a machine-learning model. Multi-model classification involves using multiple machine-learning models to classify data points. The predictions of the individual models are then combined to produce a final prediction. This approach can improve the performance of rumor detection by leveraging the strengths of different machine learning algorithms[5]. One specific approach to rumor detection using multi-model classification is to use a K-nearest neighbors model. KNN is a simple but effective machine learning algorithm that classifies data points based on the labels of their K nearest neighbors[6]. Research on the Twitter dataset has been conducted in rumor detection, and the retweet ratio detected rumor. Either way, this is more efficient, but it cannot find how many people retweeted and spread that information [15]. In some research, multiple datasets about rumor were considered and classified using various machine learning algorithms. Classifying the dataset makes the detection algorithm more efficient [16]. Despite the importance of the propagation of rumors and the opportunities for their study presented by online social networks, little is known about rumor propagation on these networks. While information diffusion in online social networks has recently been the subject of considerable attention for everyone, this work has generally focused not only on the distinction between true and false information but also on the source of the rumor. Rumor detection on social media is challenging due to the large volume of data, the fast spread of rumors, and the difficulty of distinguishing rumors from factual information. This paper uses the data set in the paper[2]. Rumor detection seeks to identify the originator of a rumor inside a social network, whether it be an individual or a specific area. Detection can be determined by analyzing many factors, such as network structure, diffusion models, centrality measures, and evaluation metrics. [17]. Single source rumor detection based on the SEIR model is done in [18]. The authors have effectively developed a method to estimate the origin of rumors in online networks. They were able to show that their estimate is equal to Jordan's infection centre. Dayani et al. [19] conducted a retrospective analysis of a tweet dataset gathered in 2009 and discovered that content-based features, specifically word frequencies, are crucial in detecting rumors, in contrast to user-based features. Ma et al. [20] created a rumor detection approach that utilizes recursive neural networks to establish a connection between content linguistics and dissemination evidence. The model utilized a propagation tree that originated from a specific post, with each node representing a post that generated a response. Rumor detection using Recurrent Neural Networks (RNN) with a variation of Autoencoders (AE) was proposed in [21] to effectively learn the typical behaviours exhibited by individual users. The discrepancies between the model's outputs and inputs are utilized to quantify the extent of divergence and then contrasted with the thresholds to ascertain whether it qualifies as a rumor. Although there is a lot of research on rumor classification and detection, very limited research is conducted on detecting rumor sources based on attributes.

3. Methodology

3.1. Dataset

The dataset used in this paper were the tweets from Twitter on January 1, 2020, and July 15, 2020, by using and trying to find keywords related to 5G and COVID-19, inclusive of "5G" and COVID-19, in the text of tweets[2]. The dataset supports machine learning algorithms, graph processing, and related fields in studying the spread of misinformation. Furthermore, it provides a series of baseline experiments using both Graph Neural Networks and other established classifiers that use simple graph metrics as features.

The subgraphs of 3,000 manually classified Tweets from Twitter's follower network were extracted and distinguished into three categories: 5G-Corona Conspiracy: This class consists of tweets explicitly declaring or suggesting a more profound connection between COVID-19 and 5G technology. Other Conspiracy: In this class, the tweets that propagate conspiracy theories unrelated to the 5G-COVID-19 connection are mentioned. Non-Conspiracy: This class encompasses tweets that don't fit into the preceding two labels.

The dataset contains proper segregation of different tweets, but they were not labeled accordingly, so the dataset is to be preprocessed to achieve well-balanced and stable data.

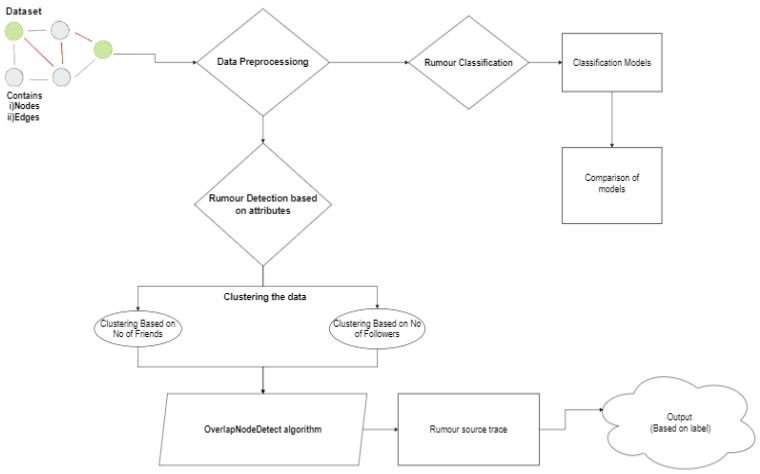


Fig. 1: Block Schematic Diagram

3.2. Data preprocessing

This data set consisted of unlabeled and grouped tweets related to 5G-related rumors. Addressing the task of rumor propagation inherently includes dealing with imbalanced data sets. To overcome this, an oversampling strategy was implemented to balance the data set[1][4]. The oversampling balances the data set by adding value to the minority data based on the other similar data available in this data set. The consciousness of oversampling based on the minority data, wherein additional samples had been synthetically generated to suit their preexisting data. This strategic technique effectively mitigated the magnificence imbalance and made the data set fit to train without noisy data values.

Furthermore, diverse data preprocessing strategies were explored to make the data set fit accurately for the model while encompassing a bagging strategy to improve version robustness and generalization. Hyperparameter tuning before training a model will improve the accuracy. Two different cross-validation strategies were employed to achieve the perfect test fit for the machine: Grid Search CV and Random Search CV. [4]. Grid Search CV systematically traverses more than a few hyperparameters, facilitating the identification of the foremost mixture that maximizes overall performance. In contrast, Random Search CV employs random sampling from the hyperparameter area, calculating ratings for each hyperparameter set, and it chooses the appropriate score from it.

Cross-validation and data-balancing techniques were used to enhance the machine-learning model's accuracy. Cross-validation enabled the evaluation of the model's performance on multiple subsets of the data. Data balancing mitigated the impact of class imbalance, preventing the model from being biased towards the majority class. These rigorous methods were used to improve the performance of the model's predictions, leading to improved accuracy and reliability.

For the statistics preprocessing approach, the Grid Search CV yielded the most favourable results and was appropriate for the model for classification[3]. This technique, together with the alternative implemented strategies, collectively sets up a stable and finely-tuned data set that is nicely prepared for training the model, and the data set is well-balanced and preprocessed for future usage. As the data was preprocessed, the next step of classification of labels using the various machine learning models was done. Figure 1 shows our approach to detecting and identifying the source of the rumor spread.

4. Rumor Classification

The classification process was achieved through various preprocessing steps, including dataset balancing and cross-validation for model robustness. Hyperparameter tuning was performed using grid search. In the study, different

classification methods were explored to address the unique challenges presented by the data set, which included the presence of multiple labels such as 5G, Non-conspiracy, and others. The classification models are methods that were broadly categorized into Binary Classification[5][7], Multi-Class Classification, and Multi-Label Classification. The Multi-Label Classification approach proved to be the most suitable for our case, given that each data point had the potential to be associated with any combination of the three labels[4]. Instead of assigning a single label, multi-label classification allowed for the simultaneous prediction of multiple labels. A range of metrics was employed to assess the models' performance, encompassing accuracy, precision, recall, and F1 score. The models' interpretability was also considered, as it was essential to understand the basis for their predictions.

SVM is renowned for its robustness and effectiveness in various machine-learning tasks, especially classification. However, it comes at a substantial computational cost; this became evident as the training will take a lot of computer resources, which cannot be achieved by the normal or usual devices that were in use[5]. Even though SVM is renowned for numerical data classification, with this dataset, the accuracy score of each label is minimal. After careful analysis and applying a stratified cross-validation method, the overall accuracy is 0.67, which is a very low accuracy score, and it requires a high computational resource for training. The other is the Random Forest classifier, an ensemble learning method where the decision tree plays a significant role. The random forest performance with this dataset is not as the expected performance. The accuracy and f1 score are not very low after training the model.

The next model considered for implementation is the bagging technique, often called bootstrapping, utilizing random sampling with a substitute. The type version underwent training through bootstrapping, resulting in an accuracy of 0.73. The ongoing method of improving the version entailed running with numerous subsamples of the records and combining their results. This increased the version's accuracy and gave it extra flexibility to adapt to exclusive data sets. As it was a moderate dataset and not big data, the bagging might provide biased results based only on the training dataset.

The KNN classification model was observed to be the most suitable option for this data set, with a preprocessing rate of an impressive 0.87 percent. The model became particularly powerful in categorizing the facts into three unique labels with varying tiers of accuracy. The KNN model has a specific algorithm to identify that point's k-nearest neighbors from the training set. "Nearest" is often determined by Euclidean distance or other similar metrics for better accuracy. In this case, the KNN furnished a 0.87 accuracy for the label '5G', 0.95 for non-conspiracy, and 0.78 for 'Other.' The overall accuracy for the test data is 0.86, which indicates its robustness and dependability, hence justifying its use for classifying the data set on these labels. The comparative analysis of all the accuracy scores is in Table 1

Table 1: Accuracy Scores

Model Used	***5G***	***Other***	***Non Conspiracy***	***Overall Accuracy***
SVM	0.68	0.61	0.65	0.67
Random Forest	0.12	0.19	0.84	0.71
Bagging	0.31	0.46	0.07	0.73
Logistic Regression	0.06	0.1	0.85	0.74
KNN	0.93	0.86	0.75	0.86

5. Rumor Detection

People with more followers and friends on social media are more likely to be influential. This is because they have a larger audience to share their content with. If these followers and friends also share the same content, then the person with the larger audience is even more influential. So, to identify the most influential people tweeting about 5G, we can look for the people with the most followers and friends who are also tweeting about 5G. These people will likely significantly impact the conversation about 5G on social media.

Here is an example:

Suppose we have two people tweeting about 5G:

- Person A has 100 followers and ten friends.
- Person B has 1000 followers and 100 friends.

Person B is more likely to be influential than Person A because they have a larger audience. If Person B's followers and friends share their content, then Person B is even more influential.

To identify the most influential people tweeting about 5G, we can use a variety of methods, such as:

- Looking at the number of followers and friends that people have.
- Looking at the number of people who share and retweet people's content.
- Looking at the engagement rate of people's content (e.g., the number of likes and comments).

A three-step approach has been used for rumor detection:

- Clustering- Grouping nodes that share similar attributes.
- Overlapnodedetect-Detecting the set of overlapping nodes from the clusters formed.
- Rumorsourcetrace-Detecting the source of rumors based on the overlapping nodes.

5.1. Cluster Formation

The next step involves the identification of overlapping nodes within the well-classified dataset[9]. These overlapping nodes are characterized by their membership in multiple clusters [6][7]. Specifically, the focus is on identifying overlapping nodes with more friends and followers than others, as these nodes are believed to wield more influence. A novel algorithm named OverlapNodeDetect is proposed to determine the overlapping nodes.

Initially, K-means clustering is performed, which performs the clustering based on the attributes of friends and followers. K-means clustering is an unsupervised machine learning algorithm that groups data points into a predefined number of clusters. The algorithm works by iteratively assigning data points to clusters based on their similarity to the cluster's center, which is predefined.

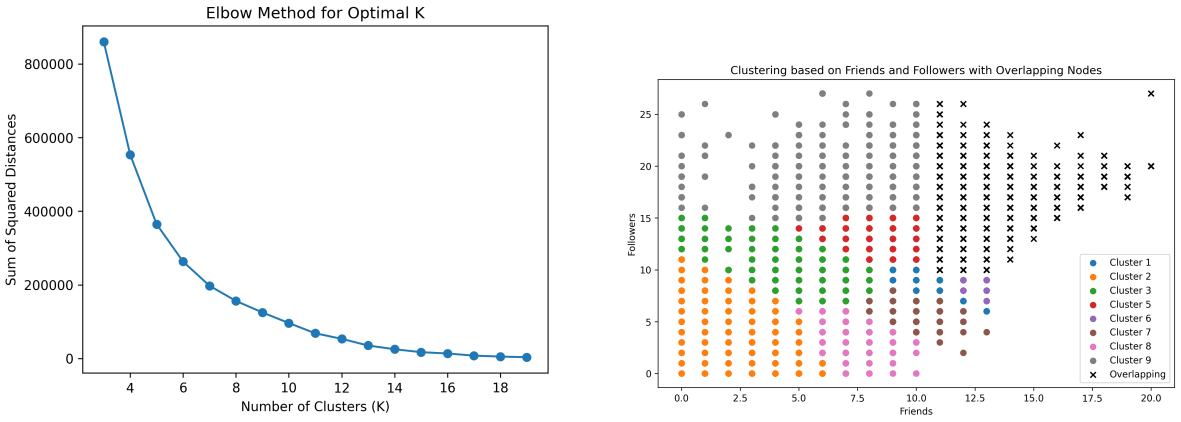
There were many approaches for assigning a K value for the K-means clustering algorithm. The elbow approach is used here, and the elbow method is a graphical approach that illustrates the sum of squared distances (SSD) between facts points and their cluster centroids concerning the price of K. The optimal k value is identified when the SSD curve exhibits a horizontal bend. To achieve an optimal K value for the clustering. We concluded that the optimal value of k is 10. Then, K-means clustering to group the data points into 10 clusters. Fig 2(a) shows the result of the elbow method plotted with the plot function from the pyplot library.

5.2. Overlapnodedetect Algorithm

To pinpoint the overlapping nodes, a specific range threshold for the number of friends and followers was defined[8]. Any node surpassing this threshold regarding friends and followers was categorized as an overlapping node. Subsequently, all the overlapping nodes were exported to another file as specified in Algorithm 1. The Fig 3 gives the approach of the overlapnodedetect Algorithm.

This process yielded a compilation of overlapping nodes that may be used for in-intensity records evaluation. For instance, this list helps identify the maximum influential nodes within the network and the invention of carefully related companies of overlapping nodes, finally finding the overlapping nodes using these clusters. The final overlapping nodes after using the algorithm were visually plotted in Fig 2(b).

This approach involved clustering the id of users into distinct communities based on the number of followers and friends each individual has. Instead of adhering to a one-length-fits-all technique, the type became custom-designed to guarantee that the recognized communities were distinguished using significant criteria. This meticulous methodology



(a) Elbow method (b) Overlapping nodes

Fig. 2: Elbow method and Overlapping nodes

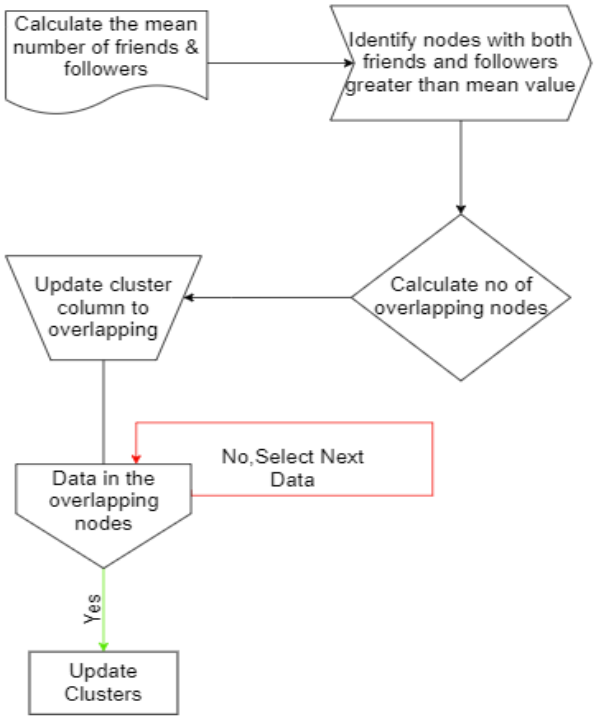


Fig. 3: Overlapnodedetect: Overlapping algorithm

played a crucial role in presenting a comprehensive perspective of our social network, enriching the comprehension of rumor propagation within this intricate web of interconnections.

Algorithm 1 depicts the approach for identifying the overlapping nodes from the cluster by combining the followers and friends cluster.

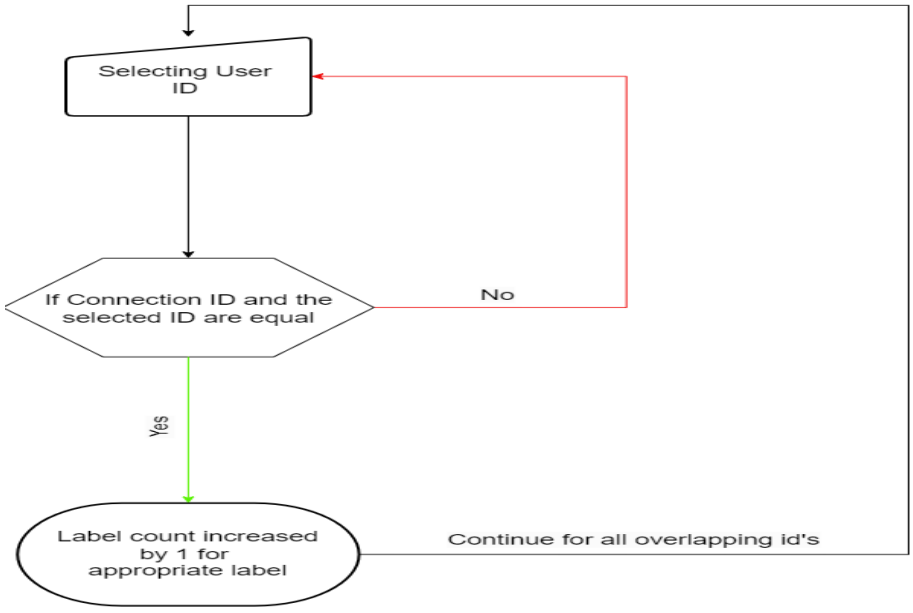


Fig. 4: Sourcetrace Algorithm

Algorithm 1 Overlapnodedetect Algorithm

Input: Clustered Dataset
Output: Overlapping nodes based on attributes

1. Calculate the mean number of friends and followers in the data set.
2. Identify nodes with friends and followers greater than the respective means.
3. Calculate the number of overlapping nodes for each label.
4. Update the 'cluster' column of data to overlapping data.
5. Repeat the step 4 for all the nodes.
6. Visualize the clusters using a scatter plot.
7. Update Cluster Column if data is in overlapping nodes.

5.3. RumorsourceTrace: Detection of Rumor Sources

After establishing community clusters, the subsequent pivotal phase involves the identification of the source individual responsible for the initial spread of these rumors. These connections are thoroughly examined to enhance the utility of edge connections given in the data set, and labels are added to each connected ID within the edge connection file. Following this enhancement, a comparative analysis is conducted to the outcomes of the clustering manner with the recently received edge connection records. A meticulous inspection and alignment process is carried out to identify matches, ensuring that each ID in the current overlapping data file is correctly associated with the user ID column in the edge connection file. As in algorithm 2, if any of the IDs in overlapping data is found in the user ID column of the edges file, the label of the connection ID, which is parallel to the user ID, is noted and the connection count is

increased accordingly for the appropriate tables. The count of all the connections id for all the appropriate labels '5g' and 'Other is added as shown in Fig 4.

Upon coming across a connection, the research proceeds by reviewing the label related to the corresponding ID. This comprehensive approach helps the compilation of a depend for every label, presenting deeper insights into influential individuals within the community based on the rumors they propagate. The output file incorporates the ID, the number of connections linked to various labels, and the total count of connections associated with each ID (Table II). This output record presents vast info regarding the number of connections related to every ID. In specific, it delves into how connections are disbursed amongst one-of-a-kind labels. After analyzing all the nodes, the source rumour propagator finally concluded as was in 2. Notably, the individual with the most substantial connections within the "5G" label emerges as the focal point of our research and is recognized as the most influential figure within the community. This significant finding from our study underscores the pivotal role played by the individual with the highest connection count in the context of the "5G" label, offering valuable insights into the dynamics of rumor propagation within the community. The algorithm 2 depicts the rumor source identification.

Algorithm 2 Rumorsourcetrace Algorithm

Input: Overlapping Nodes
Output: Total connections counts on different labels based on connection IDs.

1. Find all the user IDs present in the overlapping nodes:
   - Extract a list of all user IDs from the dataset. These user IDs represent the nodes in the overlapping nodes.
2. Within each cluster, find the people who are connected with those user IDs (connected IDs):
   - For each user ID, identify their connections within the data set.
3. Identify the label of each user from the clustered data and increase the count on that label by
   - For each user ID, determine the label associated with that user.
   - Increase the count of the respective label by 1.

The classification of the different labels is achieved using kNN, and accuracy is achieved as 0.86. Based on the output in Table 2, The node in the first row with a total connection count of 390 is supposed to be a source of rumor spread, and the table orderly shows the top 5 rumor spread nodes. From this, it can be concluded that the source of a rumor is likely to be a person with more friends and followers than the threshold specified. This conclusion is important because it can help us to identify potential sources of rumors and misinformation on social media. By identifying these potential sources, we can take steps to mitigate the spread of misinformation and protect the public. The detection of this will prevent a disaster before it happens by identifying the sources.

Table 2: Top five nodes which are rumor sources

id	friends	followers	5G	other	Total Connection Count
23518497	16	16	212	138	390
10075640	16	16	68	74	144
58381306	13	15	46	17	68
16045379	18	18	43	0	99
32224445	19	19	39	52	118

6. Conclusion

The WiCo data set was analyzed in this study, primarily focusing on friends and followers as attributes. Through this research, we were able to locate and track down the sources of rumors inside the network. By closely examining the characteristics, the origins of the rumor were found and tracked. By highlighting the possibility of using social connections to identify the sources of rumors, our work sheds light on the dynamics of information dissemination in the digital age. This research has demonstrated the effectiveness of homophily rumor detection, which is one of the main conclusions. This method showed how useful it is for identifying groups of people with similar traits and interests. Through the identification of these homophilous communities, we were able to obtain essential insights about the network's structure and information flow. This information can be used to create more focused information diffusion and control tactics, ultimately improving our understanding of online communication dynamics.

References

[1] Zhu, Xiaojin, and Zoubin Ghahramani. "Learning from labelled and unlabeled data with label propagation" ProQuest Number: INFORMATION TO ALL USERS (2002).
[2] Pogorelov, Konstantin, et al. "Wico text: a labeled dataset of conspiracy theory and 5g-corona misinformation tweets." Proceedings of the 2021 Workshop on Open Challenges in Online Social Networks. 2021.
[3] Santhosh, Nikita Mariam, Jo Cheriyan, and Lekshmi S. Nair. "A multi-model intelligent approach for rumor detection in social networks." 2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS). IEEE, 2022.
[4] Werner de Vargas, Vitor, et al. "Imbalanced data preprocessing techniques for machine learning: a systematic mapping study." Knowledge and Information Systems 65.1 (2023): 31-57.
[5] Schaffer, Cullen. "Selecting a classification method by cross-validation." Machine learning 13 (1993): 135-143.
[6] Guo, Gongde, et al. "KNN model-based approach in classification." On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings. Springer Berlin Heidelberg, 2003.
[7] Ganesh, Parvathy, Lekshmi Priya, and R. Nandakumar. "Fake news detection-a comparative study of advanced ensemble approaches." 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2021.
[8] Priyanga, V. T., et al. "Exploring fake news identification using word and sentence embeddings." Journal of Intelligent Fuzzy Systems 41.5 (2021): 5441-5448.
[9] Zhu, Tieying, et al. "Bidirectional label propagation for community detection." 2017 4Th International conference on information science and control engineering (ICISCE). IEEE, 2017.
[10] Dileep, P. Radhika, and L. R. Deepthi. "Analysis of Link Prediction Methods in Weighted and Unweighted Citation Network." 2022 International Conference on Connected Systems and Intelligence (CSI). IEEE, 2022.
[11] Huang, Zhihao, et al. "Detecting community in attributed networks by dynamically exploring node attributes and topological structure." Knowledge-Based Systems 196 (2020): 105760.
[12] Nair, Athira, et al. "Classification of Trust in Social Networks using Machine Learning Algorithms." 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT). IEEE, 2022.
[13] Ge, Jinhuan, et al. "LPX: Overlapping community detection based on X-means and label propagation algorithm in attributed networks." Computational Intelligence 37.1 (2021): 484-510.
[14] Pillai, Reshma S., and L. R. Deepthi. "Citation Recommendation Using Deep Learning Approach." ICT Systems and Sustainability: Proceedings of ICT4SD 2022. Singapore: Springer Nature Singapore, 2022. 359-369.
[15] Tetsuro Takahashi, et al. "Rumor detection on twitter" The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems 2012.
[16] Pathak, Ajeet Ram, et al. "Analysis of techniques for rumor detection in social media." Procedia Computer Science 167 (2020): 2286-2296.
[17] Shelke, Sushila, and Vahida Attar. "Source detection of rumor in social network–a review." Online Social Networks and Media 9 (2019): 30-42.
[18] Zhou, Yousheng, et al. "Rumor source detection in networks based on the SEIR model." IEEE access 7 (2019): 45240-45258.
[19] Dayani, Raveena, et al. "Rumor detection in twitter: An analysis in retrospect." 2015 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS). IEEE, 2015.
[20] Ma, Jing, Wei Gao, and Kam-Fai Wong. "Rumor detection on twitter with tree-structured recursive neural networks." Association for Computational Linguistics, 2018.
[21] Chen, Weiling, et al. "Unsupervised rumor detection based on users' behaviors using neural networks." Pattern Recognition Letters 105 (2018): 226-233.