Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 10, October 2021: 1470-1482

RUMOR SOURCE IDENTIFICATION IN SOCIAL NETWORK WITH LOWEST SEARCH SPACE

Sushila Shelke[1], Vahida Attar[2]

[1]Dept. of Computer Engineering & IT, College of Engineering Pune (COEP), Pune, Savitribai Phule Pune University, Pune, India

[2]Dept. of Computer Engineering & IT, College of Engineering Pune (COEP), Pune, Savitribai Phule Pune University, Pune, India
sushila.shelke@cumminscollege.in, vahida.comp@coep.ac.in

ABSTRACT

The rise in social network data leads to extensive information propagation where the information shared without authenticity results in massive diffusion of rumor. Rumors during the present pandemic situation of COVID-19 create fear, anxiety and a negative impact on individuals. Identification of rumor sources helps control the undesirable effects of rumor diffusion in a social network. This research targets discovering the starting point of a rumor in the social network with improved accuracy and reduced network space. The proposed algorithm of TPSD achieved this by applying the search space reduction method and reverse propagation technique together. The network is examined using a monitor-based approach and rumor is diffused by a discrete-time susceptible-infected model and using incremental propagation delay. The incremental delay helps to detect the nominee partition precisely with the help of a partition-connected graph. In this nominee partition, rumors from the monitor nodes are reproduced in the reverse direction to identify the source. The experiment is performed in synthetic and real-world data collected from a semantic web of Twitter. The previous work shows the accuracy concerning distance error 0-4 hop distance. The experimental result illustrates that the actual source is identified within 0-1 hops in a real-world social network such as Twitter and Facebook. The experimental results reveal that our approach outperforms the current methods.

Keywords: Diffusion Model, Rumor, Source Estimation, Social Network

I. INTRODUCTION

Individuals are currently connected to many microblogging sites due to the more comfortable communication method, sharing various data formats in a single platform and connecting the public worldwide. Facebook, Twitter and Reddit are popular microblogging websites also known as social networks. There are 4.2 billion energetic social media users all over the world (Tankovska, 2021). The rise in the trend of social networking websites is confirmed as really supportive in catastrophic conditions such as ordinary disasters (Flood, Storms, earthquakes) (Lifang et al., 2020), human-made disasters (Terrorist attacks, Shootouts) and crises (Luna & Pennock, 2015). These conditions speedily lead to extensive data propagation. During the COVID-19 situation, the spread of information helps people be aware of health precautions such as masks, handwash, and physical distancing (Tasnim et al., 2020). Despite the supportive feature of a social network, there is difficulty verifying the information propagated by these vast crowds. The information that circulates rapidly without verifying its reliability (Zannettou et al., 2019) and later proved false information is called a rumor.

Today, the whole world is under the anxiety of COVID-19. There were many stories circulated related to this disease. Examples of few rumors spread around us in this pandemic situation are Holding breath is a credible way to test for coronavirus (O'Rourke, 2020), Drinking garlic water cures coronavirus (Mikkelson, 2020), One of the first nurses to receive the vaccine in AL is now dead (Funke, 2020). Such news spreads fear and anxiety among society, which needs to be reduced or stopped using effective strategies. There are many specialty-based fact-checking websites such as Politifact (*Politifact Website*, 2021), FactCheck (*Fact-Checking Website*, 2021a), Snopes (*Fact-Checking Website*, 2021b), which work for debunking rumors or fake news. Due to manual efforts, news verification is time-consuming, which may not control the impact of rumor diffusion at the early stage and does not detect all possible source(s). Resisting the wave of a rumor is demanding and vital for organizations, personalities, election commission, government agencies, etc., wherever there is a requirement to discover the source. Overcomplicated broadcasting, and continuous enhancements in the network, differentiating the quick and accurate starting point of rumors in a semantic microblogging website.

In recognition of the rumor source, several factors such as the construction of the network, methods for diffusion, centrality metrics, and evaluation measures need to be considered are studied (Jiang et al., 2017) (Shelke & Attar, 2019). This research focuses on identifying the source of rumor with more considerable accuracy. The current methods detected the origin with 0-4 hops distance. In source detection of rumor, accuracy is more important; therefore, this research focused on improving the accuracy of rumor source identification and decreasing the source estimation time by reducing the search space.

In the previous work of (Shelke & Attar, 2020), a method for identifying the origin of rumor in a social network described the correctness of the path of 0 - 4 hops in a real-world social network. The research in this paper focused on improving the accuracy of previous work and also demonstrating the improvement over existing benchmarked methods. A rumor source identification model is proposed and contributions of this work are:

1. Put forth a methodology for data collection from Twitter
2. Proposed a method for identifying nominee partition to diminish the search space for source estimation.
3. A source estimation method that finds the estimated sources from various observers using the reverse propagation method.

The experimental results evaluated with the existing method (Pinto et al., 2012), (Louni & Subbalakshmi, 2018) on real-world twitter datasets, which express that the planned method outperforms distance error (DE). There is minimal work on improved accuracy of source detection (Paluch et al., 2018) focused on finding the source precisely and faster. This work is mainly inspired by (Louni & Subbalakshm, 2018), (Paluch et al., 2018), where the proposed method increases accurateness. Section 2 discusses related work. Section 3 covers the methodology. The experimental outcomes are presented in section 4 and finally, the conclusion and future scope are explained in section 5. The main objective is to present the research of rumor source detection on semantic Twitter networks with improved accuracy.

II. BACKGROUND OF STUDY:

The dissemination of rumors in a network produces several hazards like inappropriate conclusions in terrible circumstances and targeting the reputation of an association or individuals. The dispersal of stories in a network can be forced by detecting rumors and a starting point of rumor at the primary stage. The rumor source is nothing but the origin or the first user who propagates the rumors message in the network. This section explores the literature on source detection approaches. This section explores the incredible advancements in source identification methods. Few examples are finding the source of widespread disease in a temporal network (Choi, 2020), discovering leakage of gas in a wireless sensor network (Shu et al., 2016), the root of a rumors story in a social network which are by inference related to source identification of a rumor.

III. REVIEW OF LITERATURE

In the method of finding the source, networks can be observed by a snapshot-based and monitor-based approach. Various diffusion models can diffuse the rumors as Independent Cascade (IC), Susceptible-Infected (SI) and Susceptible-Infected-Recovered (SIR). In a snapshot-based approach, multiple snapshots of the network are taken at different times(Rácz & Richey, 2020). In the monitor-based approach, various users in the networks are considered observers to collect rumor data. For processing, numerous snapshots required more computation time as compared to monitor nodes. Majorly used diffusion models in monitor-based approaches are SI (Louni & Subbalakshm, 2018), (Paluch et al., 2018), SIR (Jiang et al., 2016) and IC (Xu & Chen, 2016).

The approach, in which observer nodes keep a record of receiving time of posts, assumes that the rumor disperses through the BFS tree are proposed by (Pinto et al., 2012). Considering data collection from all the monitor nodes requires a long time. (Paluch et al., 2018) consider only the nearest monitor nodes with the least infection time and discrete-time SI model with the Gaussian dispersal for the suspension in rumor. They improvise a method of source detection as a distinction to (Pinto et al., 2012). (Xu & Chen, 2016) applied rumor quantifier measure and an active IC model to identify a rumor's source where the correctness of source detection depends on monitor nodes' size. A time-based network method by applying the SIR diffusion model and proposed an innovative maximum likelihood estimation (MLE) to guarantee the active growths in the network by (Jiang et al., 2016). They conclude that monitor based inspection shows good precision for source recognition of rumor.

(Louni & Subbalakshmi, 2018) make use of a weighted graph along with a normal distribution for the arbitrary delay in propagation. They follow the continuous-time SI model for rumor diffusion and Louvain's method (Blondel et al., 2008) to divide the graph into various partitions. The algorithm is proposed in two stages. They use several network instances, then divide the network and determine the nominee partition where the origin node fits. They used approximately alike MLE to discover the nominee partition and assessed the source on various graph instances.

In the previous work of origin identification (Shelke & Attar, 2020), a diffusion tree which was built with the help of monitor nodes and MLE, was utilized to identify the origin of rumor. The accuracy of origin identification was identified with DE and presented 0-2 hops on a synthetic dataset and 0-4 hops on real-world datasets. The DE was large in number for a huge network as the diffusion tree was built with the help of monitor nodes that were selected approximately. Therefore, to improvise the accuracy reverse propagation method is utilized.

The literature study shows that the technique for identifying nominee partition helps reduce search network and to increase the accuracy, the reverse propagation method is helpful. Therefore, these methods are being used with modification in this research to design a proposed strategy. However, the accuracy in terms of DE shows 0-4 hops distance in existing work, which needs improvement. In distinction with existing research, this research work used a discrete-time SI model for rumor dissemination. Also, recognize the nominee partition using the node that passes rumors to the vertex with the least receiving time in the connected partition graph and using a snap of the graph to decrease the calculations of many sample graphs.

IV. METHODOLOGY

The literature study concludes that the rumor source identification process involves many factors such as network topology, network observation, rumor propagation model and source estimation approach. This research considers network topology as an undirected graph and network observation using a monitor-based approach. Consistently weights are allotted in the graph from 0 and 1. Further, this section explores the details of data collection from real-world network, diffusion model and source estimation method.

- **Dataset**

The experiments are performed on synthetic and real-world dataset. The synthetic dataset of Erds-Rnyi (ER) random graphs (Erds & Rényi, 1960) is used to generate the graph. The graphs are produced with a probability of 0.5 and the weights are consistently allotted between 0 and 1. The real-world dataset of Facebook and Twitter (Erds & Rényi, 1960) openly available on (Leskovec & Krevl, 2014) are used as benchmarked datasets.

The collection of real-world data from Twitter is one of the major contributions of this research. Initially rumors news identified from debunking website snopes.com and tweets for that particular news are collected using tweepy API using different keywords in search query to get maximum number of tweets for that news. For the collected tweets, a user network is designed for level 1 by identifying the followers of each user; similarly, followers are identified recursively for 3 levels. For simplicity and to get the dense network, users with less than 5000 followers are considered in data collection. Also, from the list of followers, only users who were active in the last 30 days from the date of data extraction are measured in data curation. Finally, after collecting followers in each level the dataset has total users as 56479 and their interconnections 75805. The diameter of this newly created real-world network is 7. The particulars of the benchmarked and collected real-world datasets are shown in Table 1.

Table 1. Real-World Data Sets Details

Network	Facebook	Twitter	Twitter (Collected)
# Nodes	4039	81306	56479
# Edges	88234	1768149	75805
Diameter	8	7	7

- **Diffusion Model**

Information diffusion models are used to describe and recreate the information in the network. There are only two statuses in discrete-time SI diffusion model: one is susceptible, i.e., the node where the minimum single vertex among its neighbors has received the rumor and the other one is infected i.e., that vertex acquires the rumor. Figure 1 shows the rumor diffusion model utilized in this work. The rumor's starting point is arbitrarily chosen at time 0 from all the vertices and continues the diffusion with contagion rate α. The node which receives the rumor can contaminate all its susceptible neighbors by the equivalent contagion rate α and immutable delay for every neighbor. The rumor circulation delay for every neighbor is equal because in a real-time network such as Twitter, it is reachable to all the user's followers when a user posts a message. At each time period, the infected node will diffuse its susceptible neighbors with the rate α and increase the time by one. Consequently, the vertex which receives rumors late can be differentiated by initially infected nodes.

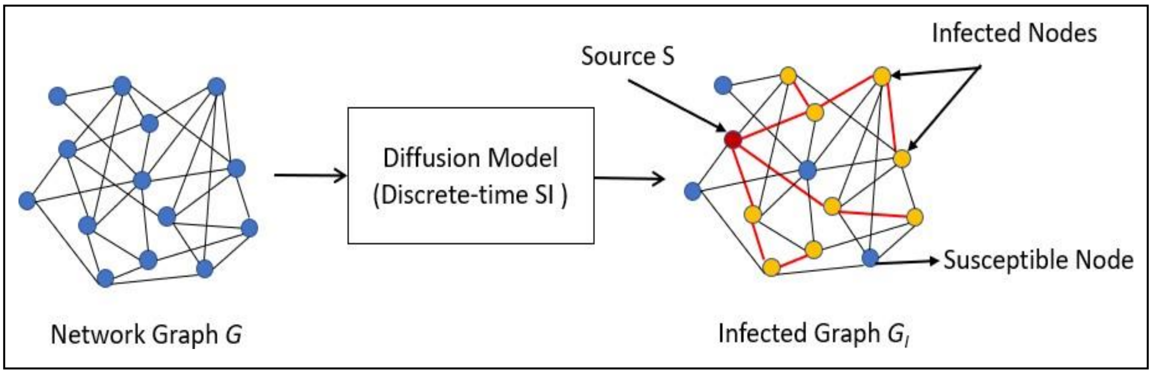

Figure 1. Rumor Diffusion with Discrete-time SI model

- **Source Estimation Method**

This research mainly focused on minimizing the search space to reduce computation. Therefore, the projected source detection method works in two phases; in phase-I, it recognizes the nominee partition where the source mainly presents and in phase -II estimated source is detected with the help of observers. Architecture for source identification is shown in Figure 2 where the green blocks indicates that there is change in techniques utilized from previous work. The impression for selecting nominee partition is taken from Louni & Subbalakshmi, 2018, where they choose the candidate partition based on the MLE, which helps to reduce the network size for source detection. However, in this research, candidate or nominee partition is identified with the help of discrete-time Susceptible Infected (SI) as diffusion model and minimum propagation delay, which helps in reducing the search space for source detection. Also, the method of reverse propagation is referred from Jiang et al., 2016, where they applied reverse propagation on the entire network. In this research work, reverse propagation is performed on the network designed from the nominee partition. Although the reverse propagation approach helps improvise the accuracy, it is not suitable for the entire network as it is time-consuming. Dotted blocks are used to recognize the phase-wise blocks in the architecture.

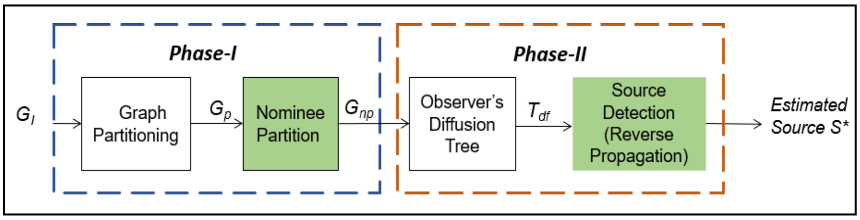

Figure 2. Architecture of Rumor Source Detection

Nominee Partition

The rumors get viral very speedily in the early phase, i.e., within hours (Friggeri et al., 2014). There is an assumption that the vertex with the least receiving time is the single that become infected at the initial phase is examined for recognizing nominee partition. The original network is partitioned into P partitions using Lovains's partitioning method (Blondel et. al. 2008). The vertices connecting different partitions are used to form the Gp connected partition graph and then find the vertex b from Gp with the least arrival time. Recognize the node a who contaminates b and determine the partition P1 where vertex a exist, then partition P1 is targeted as nominee partition. Figure 3 explains the flow of identifying nominee partition.

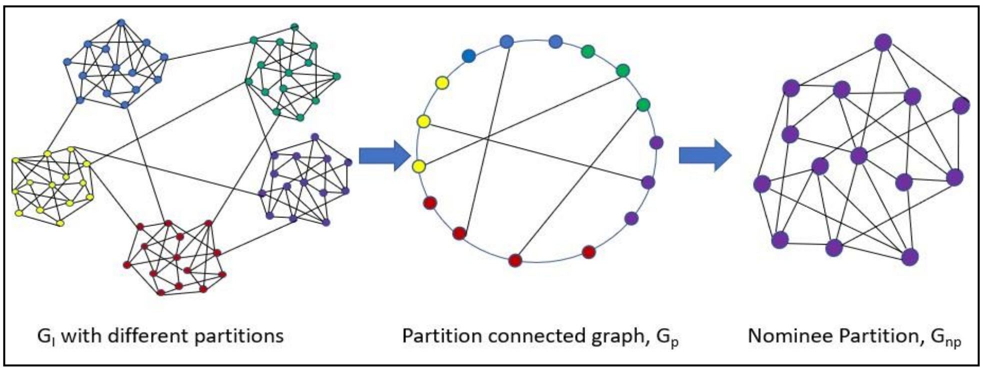

Figure 3. Nominee Partition Selection

- **Source Identification**

The rumor source identification needs a network where rumor is disseminated. Due to the unavailability of ground truth rumor sources, diffusion models are used to propagate rumors in the network. Hence, SI diffusion model is utilized to propagate the rumor in the network to get infected network. The proposed method is designed in two segments. The original graph disseminated by the SI model and diffusion time is specified as input to phase-I, which will recognize the nominee partition. In Phase-II the source is identified on nominee partition. Algorithm 1 shown in Figure 4, Two Phase Source Detection (TPSD) explains the complete procedure of source detection, where it shows that the proposed method works in two phases.

In phase-I, the original rumor diffusion graph is divided into different partitions, a nominee node is recognized by determining the node with smallest infection time from the connected partition graph which gives the nominee partition as an output. Algorithm 2 shown in Figure 5, explains about the selection of nominee partition and the detailed description for selection of nominee partition is explained in previous section.

Algorithm 1: Two-Phase Source Detection (TPSD)

Input: Infected Graph G_I (V, E, W, I_t)
Output: Estimated Source, es
1. **procedure** TPSD (V, E, W, I_t)
2. V^p , E^p = FNP (V, E, W, I_t) // First Phase
3. es = FES (V^p , E^p) // Second Phase
4. return es
5. **end procedure**

Figure 4. Algorithm 1 Two-Phase Source Detection Algorithm

Algorithm 2: Find Nominee partition (Phase 1)

Input: Infected Graph G_I (V, E, W, I_t)
Output: Nominee Partition Graph, G_{np} (V^p , E^p)
1. **procedure** FNP(V, E, W, I_t)
2. P_d= BP (V, E, W) // Louvain's Best partition
3. G_{pc} = Build partition connected Graph from P_d
4. Sort $V^{pc} \varepsilon G_{pc}$ for infection time
5. V_{mint}= min $(G_{pc} (V^{pc} , E^{pc} , I_t))$ // Select vertex with minimum Infection Time
6. Determine vertex Y which infects V_{mint}
7. cp= Find P_1 from P_d where $Y \varepsilon P_d$
8. return V^p, E^p
9. **end procedure**

Figure 5. Algorithm 2 Find Nominee Partition

Algorithm 3: Find Estimated Source (Phase 2)

Input: Nominee Partition Graph, G_{cp} (V^p , E^p)
Output: Estimated Source, *es*
1. **procedure** FES (V^p , E^p)
2. Select top K observers, k=Top(BC(G_{cp}))
3. Sort observers, O_K with infection time, $\{O_1, O_2, \ldots O_k\}$
4. Build Diffusion tree from neighbors of O_1 to O_K
5. Find srcl from each o_i in O_K using reverse propagation
6. *es = max(srcl, key=count) // Max frequency of node*
7. return *es*
8. **end procedure**

Figure 6. Algorithm 3 Find Estimated Source

In Phase-II, the betweenness centrality (BC) metric is used to choose the observer nodes from which only top k observers with higher BC are selected. The observers are sorted according to their infection time and the first observer is selected. Then the diffusion tree is built from neighbors of the first observer to the remaining k-1 observers. Find the provisional sources from each observer in the diffusion tree using the reverse propagation method. Finally, select the estimated source as the vertex having a higher frequency as a provisional source. Algorithm 3 shown in Figure 6, gives a brief of the source estimation method in nominee partition. Figure 7 explains about diffusion tree and process of reverse propagation applied on nominee partition to detect the source of the rumor. The correctness of source detection is confirmed by identifying the shortest distance among the real and estimated root node.

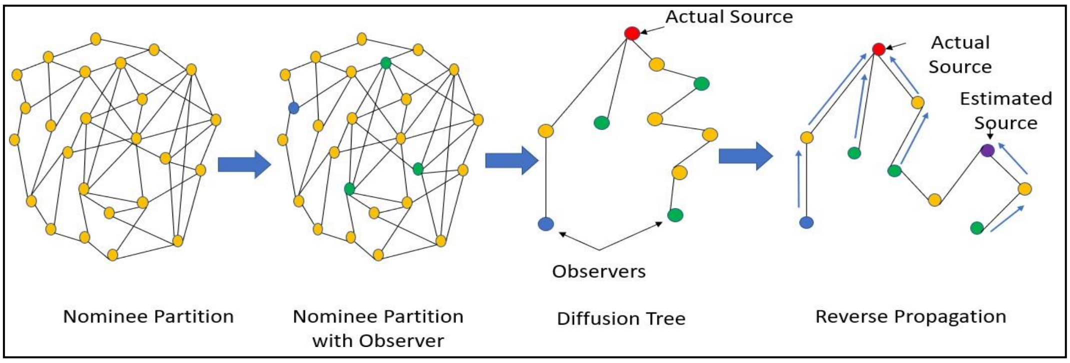

Figure 7. Process of Source Identification

Algorithm 2 is used for network reduction and identifying nominee partition. Algorithm 3 will work on these partitions to recognizes the source using reverse propagation method.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed work is experimented on Spyder-anaconda a scientific python development environment. The details of the evaluation metrics, baseline algorithms and evaluation of the proposed method are discussed in this section.

- **Evaluation Metric:**

There are various evaluation metrics used by researchers such as distance error (DE), Average Distance Error (ADE) and rank (Wang et al., 2017). This research adopted metrics as DE the lowest path of hops between the accurate root and the source assessed by the algorithm. The real source is the node, which initiates the rumor using the diffusion model, and the estimated source is the node identified using the proposed method of rumor source identification. ADE shows the average of DE where the source is identified for each run of rumor diffusion and source estimation. DE 0 indicates that the real source identified accurately 100%, and 1 indicates a one-hop distance between the actual origin and the estimated source. ADE shows the average of DE where the source is identified for each run of rumor diffusion and source estimation. Lower the value of ADE directs higher accuracy of source estimation. To determine the ADE the rumor is diffused every time to get an estimated source for that diffusion. Subsequently, 10 to 100 iterations of diffusion and source estimation models are performed.

- **Baseline Algorithms:**

    - **PTVA:** Pinto et al., 2012 proposed a method based on a monitor-based approach and the SI diffusion model presents accuracy of 0-6 hops distance in a real-world Twitter network.

    - **Louni:** Louni & Subbalakshmi, 2018 proposed a two-stage algorithm that follows the SI diffusion model and shows the accuracy of 0-4 hops in a real-world Twitter network.

- **Experimental Results:**

The proposed TPSD algorithm is evaluated on a synthetic and real-world network of Facebook and Twitter.

- **Synthetic Network**

The synthetic network of ER graph has been constructed for four graphs with different sizes of nodes such as 200, 500, 1000 and 2000. The ER model designs arbitrary graphs with N nodes that have an identical probability of edge formation. The experiment has been performed on ER graph to show the reduction of network for source identification. Table 2 shows the details of the actual size and reduced size of the network. Table 2 concludes that the search space gets minimized roughly by 80%.

An analysis of execution time in the ER network is shown in Figure 8, which showcases the execution time needed for rumor diffusion and source detection. The execution time is determined separately for rumor diffusion and source estimation time, where source estimation time involves the time of identifying nominee partition and source identification in the network built from nominee partition. The experiment is performed 10 times for varying network sizes such as 1000, 2000, 3000 and 4000 nodes. Figure 8 shows the average execution time for each network size. It can be revealed that the source estimation time does not rise along with diffusion time when the network size increases because the source is identified on a reduced network size and rumor diffused in the entire network.

Table 2. Details of Actual Vs Reduced Network (ER)

Graph No.	Actual N/W size	Reduced N/W size
1	200	35
2	500	115
3	1000	219
4	2000	517

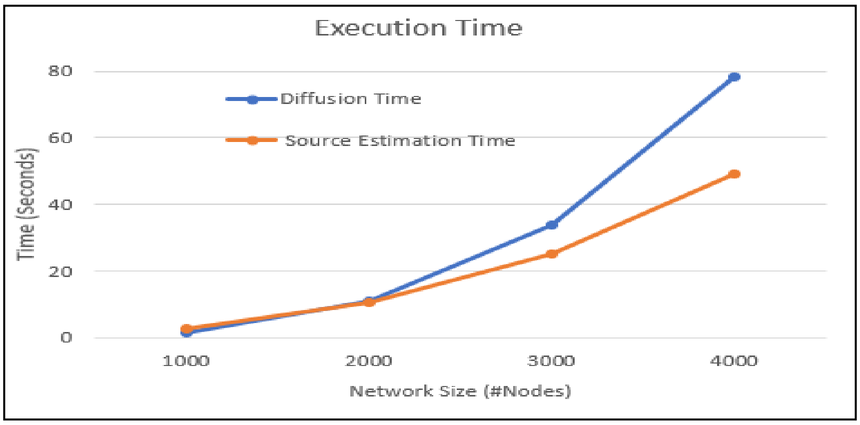

Figure 8. Execution Time for Rumor Diffusion and Source Estimation

- **Real-world Network**

The proposed model is evaluated on a real-world dataset of Facebook and Twitter network. The statistics of these benchmarked datasets and the real-world data collected from Twitter is presented in Table 1. The proposed TPSD method was tested on a Facebook dataset for 10 independent executions with different densities of observers and evaluated for ADE, shown in Figure 9. Observer density indicates the percentage of nodes selected as observer nodes having the highest betweenness centrality. The ADE gets increased when there are a smaller number of observers. Observing a density of 15% shows good accuracy; therefore, observer density is assumed as 15% only for all the remaining experiments.

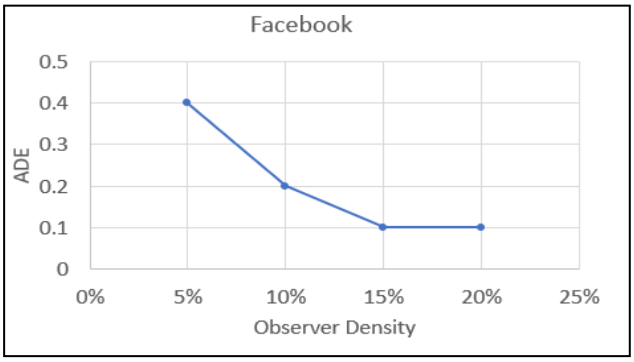

Figure 9. Average Distance Error (ADE) on Facebook

Figure 10 demonstrates the frequency of DE for PTVA, Louni and the Proposed method of TPSD, where experiments were performed independently for 100 runs on the Twitter dataset. The results of baseline algorithms are taken from the values mentioned in Louni & Subbalakshmi, 2018. This work can conclude that the results are improved using the proposed algorithm. Compared to baseline methods, TPSD shows the DE in the range of 0-1 hops, whereas DE presented by PTVA is 0-6 and Louni in the range of 0-4 hops. Overall, on real-world benchmarked datasets of Facebook and Twitter, DE has been improved from 0 - 4 hops to 0 - 1 hop by the proposed method. Figure 8 compares the proposed TPSD algorithm with baseline approaches in terms of ADE, where ADE is determined by considering the average of all DE's when executed 100 times. The minimum value of ADE specifies that the method shows good accuracy towards source detection. Figure 11 presents ADE for TPSD as 0.3, which shows that the proposed method outperforms the existing methods.

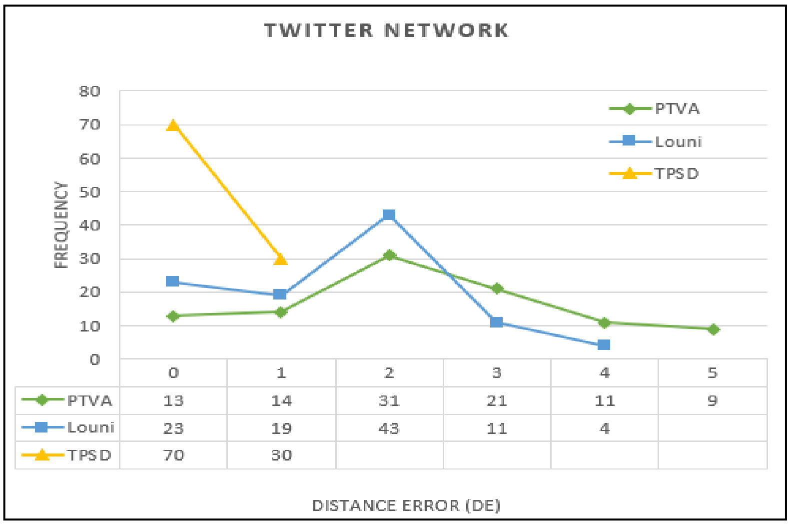

Figure 10. Distance Error Vs Frequency for PTVA, Louni and Proposed method

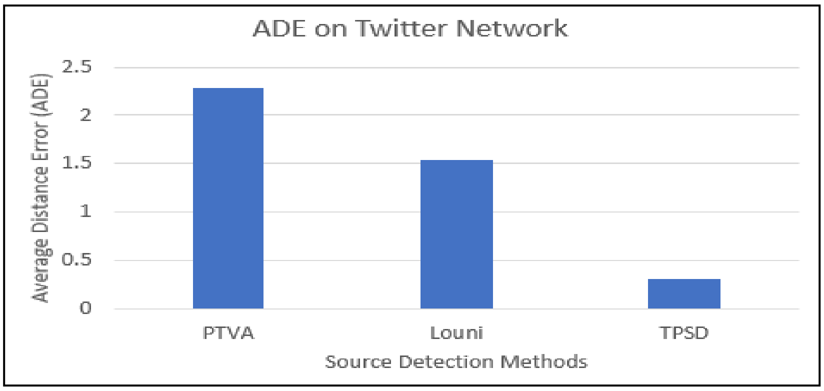

Figure 11. Comparison of ADE with Source Detection Methods

Figure 12 shows the comparison of DE and various real-world datasets, where Real-Twitter indicates the real-world data collected from Twitter. The experiment is performed for 10 runs and frequency depicts the occurrence of DE by proposed TPSD for particular dataset. It can be observed that the proposed method shows the accuracy of 0-1 hop on Facebook, Twitter as benchmarked dataset and Real-Twitter as a collected dataset.

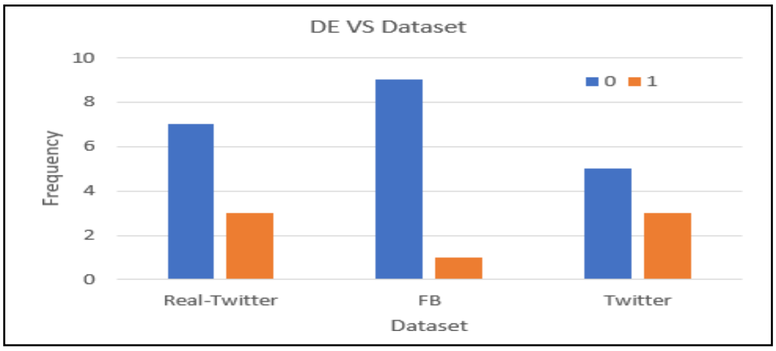

Figure 12. Distance Error Vs Real-world Dataset

The limitations of the research work include that the proposed method will not work for the continuous-time SI diffusion model and the projected algorithm is applicable to identify only a single source of the rumor.

VI. CONCLUSION

This research targeted improving the accuracy of source identification by reducing the exploration space for discovering the source of a rumor in the network. To minimize the search sector and improve the accuracy, we have proposed a model that divides the graph into the different partition and then finds out the nominee partition with the least arrival time in the connected partition graph. The reverse propagation method is applied in nominee partition, which improves the accuracy of source detection. An experiment can reveal that nominee partitions decrease the search space and minimize the computations. An experiment has been performed on real-world data collected from Twitter using follower network approach. The research work in this paper shown that the proposed model outperforms well on synthetic and real-world networks. In a real-world network, it demonstrates 0-1 hops distance as distance between actual and estimated source of rumor. The researchers are planning to extend the real-world data collected from social networks and plan to design a generalized model for identifying single and several sources In the future.

REFERENCES

1. Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.
2. Choi, J. (2020). Epidemic Source Detection over Dynamic Networks. *Electronics*, *9*(6), 1018.
3. Erds, P., & Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, *5*, 17–61.
4. Friggeri, A., Adamic, L., Eckles, D., & Cheng, J. (2014). Rumor Cascades. In *Icwsm* (pp. 101–110). http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/download/8122/8110
5. Funke, D. (2020). *One of the first nurses to receive the vaccine in AL is now dead.* https://www.politifact.com/factchecks/2020/dec/17/facebook-posts/alabama-nurse-did-not-die-after-taking-coronavirus/
6. Jiang, J., WEN, S., Yu, S., Xiang, Y., & Zhou, W. (2016). Rumor Source Identification in Social Networks with Time-varying Topology. *IEEE Transactions on Dependable and Secure Computing*, *15*(1), 1. https://doi.org/10.1109/TDSC.2016.2522436
7. Jiang, J., Wen, S., Yu, S., Xiang, Y., & Zhou, W. (2017). Identifying Propagation Sources in Networks: State-of-the-Art and Comparative Studies. *IEEE Communications Surveys and Tutorials*, *19*(1), 465–481. https://doi.org/10.1109/COMST.2016.2615098
8. Leskovec, J., & Mcauley, J. J. (2012). Learning to discover social circles in ego networks. *Advances in Neural Information Processing Systems*, 539–547.
9. Leskovec, J., & Krevl, A. (2014). *SNAP Datasets: Stanford Large Network Dataset Collection*.
10. Lifang, L. I., Zhiqiang, W., Zhang, Q., & Hong, W. E. N. (2020). Effect of anger, anxiety, and sadness on the propagation scale of social media posts after natural disasters. *Information Processing & Management*, *57*(6), 102313.
11. Luna, S., & Pennock, M. (2015). Social media in emergency management advances, challenges and future directions. *Systems Conference (SysCon), 2015 9th Annual IEEE International*, 792–797.
12. Louni, A., & Subbalakshmi, K. P. (2018). Who Spread That Rumor: Finding the Source of Information in Large Online Social Networks With Probabilistically Varying Internode Relationship Strengths. *IEEE Transactions on Computational Social Systems*, *5*(2), 335–343. https://doi.org/10.1109/TCSS.2018.2801310

13. Mikkelson, D. (2020). *Will Garlic Water Cure Coronavirus?* https://www.snopes.com/fact-check/garlic-cure-coronavirus/
14. O'Rourke, C. (2020). *Monitor yourself for COVID-19 with a breath test.* https://www.politifact.com/factchecks/2020/mar/13/viral-image/dont-hold-your-breath-isnt-credible-way-test-coron/
15. Paluch, R., Lu, X., Suchecki, K., Szymański, B. K., & Hołyst, J. A. (2018). Fast and accurate detection of spread source in large complex networks. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-20546-3
16. Pinto, P. C., Thiran, P., & Vetterli, M. (2012). Locating the source of diffusion in large-scale networks. *Physical Review Letters*, *109*(6), 68702.
17. Rácz, M. Z., & Richey, J. (2020). Rumor source detection with multiple observations under adaptive diffusions. *IEEE Transactions on Network Science and Engineering*, *8*(1), 2–12.
18. Shelke, S., & Attar, V. (2019). Source detection of rumor in social network--A review. *Online Social Networks and Media*, *9*, 30–42.
19. Shelke, S., & Attar, V. (2020). Origin Identification of a Rumor in Social Network. In *Cybernetics, Cognition and Machine Learning Applications* (pp. 89–96). Springer.
20. Shu, L., Mukherjee, M., Xu, X., Wang, K., & Wu, X. (2016). A survey on gas leakage source detection and boundary tracking with wireless sensor networks. *IEEE Access*, *4*, 1700–1715.
21. Stanford Large Network Dataset Collection. [Online]. Available: https://snap.stanford.edu/data/.
22. Tankovska, H. (2021). *Social media - Statistics & Facts*. https://www.statista.com/topics/1164/social-networks/
23. Tasnim, S., Hossain, M. M., & Mazumder, H. (2020). Impact of rumors and misinformation on COVID-19 in social media. *Journal of Preventive Medicine and Public Health*, *53*(3), 171–174.
24. Wang, D., Wen, S., Xiang, Y., Zhou, W., Zhang, J., & Nepal, S. (2017). Catch Me if You Can: Detecting Compromised Users Through Partial Observation on Networks. *Proceedings - International Conference on Distributed Computing Systems*, 2417–2422. https://doi.org/10.1109/ICDCS.2017.47
25. Xu, W., & Chen, H. (2016). Scalable Rumor Source Detection under Independent Cascade Model in Online Social Networks. *Proceedings - 11th International Conference on Mobile Ad-Hoc and Sensor Networks, MSN 2015*, 236–242. https://doi.org/10.1109/MSN.2015.36
26. Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality*. https://doi.org/10.1145/3309699
27. Zhou, Y., Wu, C., Zhu, Q., Xiang, Y., & Loke, S. W. (2019). Rumor source detection in networks based on the SEIR model. *IEEE Access*, *7*, 45240–45258.
28. Politifact website. http://www.politifact.com/
29. Fact-Checking website. https://www.factcheck.org/
30. Fact-Checking website. https://www.snopes.com/