International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 02 | Feb -2017 www.irjet.net p-ISSN: 2395-0072

# DETECTING ROOT OF THE RUMOR IN SOCIAL NETWORK USING GSSS

**S.NIVETHA[1],R.PRIYADHARSHINI[2],P.BALAKUMAR[3],R.K.KAPILAVANI[4]**

*[1]B.E,Department of Computer Science and Engineering, Prince Shri Venkateshwara Padmavathy Engineering College,Tamilnadu,India.*
*[2]B.E,Department of Computer Science and Engineering, Prince Shri Venkateshwara Padmavathy Engineering College,Tamilnadu,India.*
*3 Professor ,Department of Computer Science and Engineering, Prince Dr.K.Vasudevan College of Engineering and Technology,Tamilnadu,India.*
*4 Assistant Professor ,Department of Computer Science and Engineering, Prince Dr.K.Vasudevan College of Engineering and Technology,Tamilnadu,India*.

----------***----------

## ABSTRACT

*Detecting source of the rumor in social network plays a role in limiting the damage caused by them. However rumor spreading in social network to a shorter distance only can be identified by using some of the methodologies. In this paper, we introduce a concept to detect root of the rumor that spread in the social network in wider range by using two concepts. First, we make use of monitor nodes in order to record the data and report it to the server. Second, Greedy Source Set Size (GSSS) Algorithm to find the exact solution and also improve the efficiency for the problem. The root of the rumour is identified by three methodologies and they are Identification method, Reverse dissemination method and microscopic rumor spreading model. The identification method reduces time varying network into series of static network and reverse dissemination method resolve the set of suspect and finally microscopic method establish the real root of rumor by calculating maximum likelihood(ML) value for each suspect. The experiment result shows that it can reduce 80-95% of the root of the rumor in social networks in dynamic time varying network topology.*

**Keyword:-** *Monitor nodes, rumor spreading, GSSS, source identification, and maximum likelihood .*

## 1.INTRODUCTION

In today's world, internet has become the most important medium to circulate information. Social networks are an interesting class of graphs likely to become an increasing importance in the future times [1]. Rumor spreading in social networks plays a critical role in our society and is one of the basic mechanisms for the information dissemination in the networks [2] .For instance, in October 2011, a rumor message in social network that "Apple CEO had heart attack". When the word hit the internet, in first hour of trading the stock lost 10% of its values, spurred by panicked investor who believe that entire job is done by Steve Jobs. The ubiquity and speed access not only improve efficiency of social media but also main reason for rapid spreading of rumor about different communities [3].The solution to this problem are applied in many applications such as identifying the source of infectious disease and finding the source of leaked confidential information.2

### 1.1 PREDICTION OF RUMOR

The identification of rumor message in social network is the most preliminary basis to detect the root of rumor. Existing works mainly detected rumor by analyzing only shallow features of messages. In many scenarios, it is not satisfactory in differentiating rumor message from normal message. But then several methods used a combination of shallow features and implicit features of messages in order to identify the rumor message with efficiency[9].Three methods are mainly consider in detecting the rumor that comprise of profile based, information based and traverse based.

### 1.2 RELATED WORKS

Nowadays, social networks has been incorporated with several communities in sharing malicious information such as computer virus and rumors cause damage to our society[4],[5]. Development of mobile devices had created a great effect in spreading of dynamic information in social

network [6]. Several methods are based on identifying spanning tree in network. The rumor source is considered to be root of the spanning tree [7], [8] ·The main issues are involved in identifying the root of the rumor and emergence of **diffusion [5].**

## 2. SIR MODEL:

Besides the SI or SIS model, information source detection under SIR model is more efficient SIR model refer to susceptible infected recovered. Initially, all nodes are in susceptible stages once it accepts the rumor it is in infected stage but if the node decline or reject the rumor it is in recovered stage. We adopt SIR model because it can represent the state transition of user when they hear a rumor, from being susceptible to being recovered. Generally, people will not believe the rumor again after they know the truth. Therefore, recovered user will not transit their states any more. In order to identify the source we implement three methodologies and are given as novel identification method, reverse dissemination method, novel microscopic method.

### 2.1: IDENTIFICATION METHOD:

This method will convert the time varying social network, we reduce it as a sequence of static network by introducing a time integrating window. Each integrating window aggregates all edges and nodes present in the specified time period. For example the user may move from one place to another and also interaction between the individual appears and disappear in online social networks (i.e., online /offline).

### 2.2: REVERSE DISSEMINATION METHOD:

This method is mainly to identify the set of suspects by sending the copies of rumor from observed nodes in reverse manner. This method is similar to that of Jordan's method whereas Jordan method deals with static network but our method deals with time varying networks. The user which receives the copies simultaneously is considered to be the real root of the rumor. The traditional method is only used for static networks whereas this method is used in time varying networks. This method starts disseminating from the observed nodes with respect to the infection time of the nodes until it reaches the sensor that is infected first.

**Algorithm 1: Reverse dissemination**

**Input**: A set of observed nodes O= {o1, o2, on}, a set of infection times of the observed nodes {t1, t2,…tn}, a threshold α, and a threshold tmax.

**Initialize:** A set of suspect U=ϕ, and t1=..=tn=T if O is a snapshot/wavefront, otherwise T=max{t1,t2,///tn}.

for(t starts from 1to a tmax)do

for(oi:I starts from 1 to n)do

if(oi has not started to disseminate the rumor)then

start to propagate the rumor from use

end2

end

for(u:any node in whole network)do

if(user u received n separate rumor from o)then

compute the maximum likelihood

add user u into set u.

end

end

if(u≥αN)then

keep the first α suspects with large maximum likelihood

stop

end

end

**Output**: A set of suspect.

### 2.3: MICROSCOPIC RUMOR SPREADING MODEL:

This method is mainly to identify the real root of the rumor. This model analytically estimates the probability of each suspect. Maximum likelihood (ML) is calculated for each suspect and the suspect that has maximum ML is considered to be the real root of the rumor. This method is mainly to expose the real source from the set of suspects. Traditional methods used BFS trees instead of original networks and mainly based on node centralization. But this method is mainly based on complex networks.

## 3. PROPOSED SYSTEM

In this paper, we implement two methods to overcome the challenges. First, a number of monitor nodes is injected into the network whose job is to report data they receive. The algorithm identifies rumors and their sources by observing monitor nodes that receives the given piece of information and which do not. We show that, with a sufficient number of monitor nodes, it is possible to recognize most rumors and their sources with high accuracy. Second, Greedy Source Set Size algorithm is mainly used to optimize the output result that is produced.

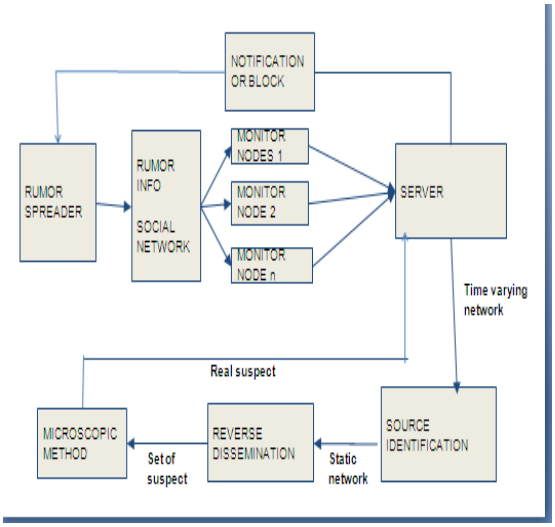



**Fig-1:** System architecture

### 3.1GREEDY SOURCE SET SIZE:

Greedy is an algorithmic paradigm that produce a solution of bit by bit, always choosing the next content that offers the most obvious and immediate benefits. Greedy algorithms are used for optimization problems. At every step, we can make a choice that looks best as a best choice, and can get the optimal solution of the complete problem. Many types of algorithm are under greedy are

**1. kruskal's Minimum spanning Tree :**

In this algorithm, create a MST by picking edges next to next. The greedy choice is to pick the smallest weight edges that doesn't cause cycle in the MST constructed so far.

**2.Prim's Minimum spanning tree:**

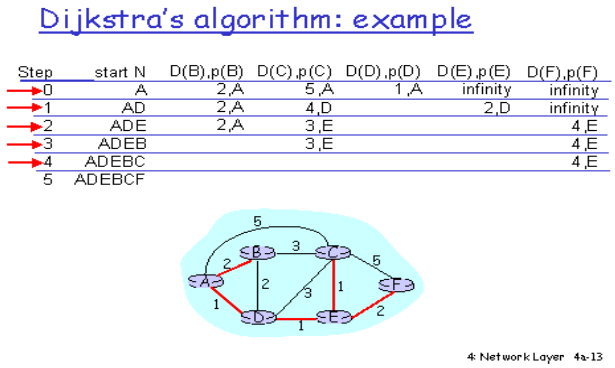
In this algorithm create a MST by picking edges one by one. we maintain two sets: a set of the vertices already included in MST and the set of the vertices that are not included. The greedy choice is to pick the lowest weight edge that connects the two sets.

**3.Dijkstra's shortest path:**

In this, shortest path is build up, edge by edge. we should maintain two sets: set of vertices already included in tree and the set of the vertices not yet included.

### 3.1.1Dijkstra's algorithm:

Dijkstra's algorithm is an algorithm for finding the shortest path between nodes in a graph, for example, traffic networks. It was given by computer scientist Edsger W. Dijkstra in 1956 and published. For a given source node in the graph, the algorithm identifies the shortest distance path between that node and every other.




### ALGORITHM:

Let the node that are starting be called as initial node. Let the respective distance of node Y be the distance from initial node to y.

1. Assign value to every node a tentative distance value: set it to zero for our initial node and to infinity for all other nodes.

2. Set the starting node as current node. Mark all other nodes unvisited nodes. Create a set of all unvisited nodes called as the unvisited set.

3. For the current node, consider all of its unvisited set of neighbor and calculate their respective distances. Compare the newly calculated distance to current assigned value and select the smaller one.

4. When we are considering all of the neighbors of the current node, mark the current node as visited node and remove them from the unvisited set.

5. If the destination node has been marked visited or if the smallest tentative distance among the nodes in the visited node and set it is as the new current node and go back to step 3.

**PSEUDOCODE:**

Function DijkstraAlgm(G,S)

Create vertex set A

For each vertex V in a G;

Dist[V]<-INFINITY

Prev[V]<-UNDEFINED

ADD V to A

Dist[s]<-0

While A is not empty:u<-vertex in?A with min dist[u]

Remove u from A

For each neighbor v of u:

Alt<-dist[u]+length(u,v)

If alt<dist[v]

Dist[v]<-alt

Prev[v]<-u

Return prev[],dist[]

This algorithm is used to find the shortest path and is used to identify the nearby source that are available in the network. This algorithm gives the optimal solution for the given problem.

**EFFICIENCY:**

The complexity/efficiency can be expressed in terms of Big-O Notation. This gives another way of talking about the way input affects the algorithm's running time. And it gives an upper bound of the running time.

In this algorithm, the efficiency varies depending on V=n DeleteMins and E updates for priority queues that were used.

If a fibonacci heap was used then the complexity is O(E+V logV),which is the best bound. The DeleteMins operation takes O(logV).

## 4.CONCLUSION

In this paper, we explore the problem of rumor source identification in time-varying social networks that can be reduced to a series of static networks by introducing a time integrating window. In order to address the challenges posted by time-varying social Networks, we adopted two innovative methods. Based on this model, the number of monitor nodes are injected into the network whose job is to report data they receive. Our algorithm GSSS is used to identifies the sources by which of the monitors received the given piece of information.Further types of observations can be considered in future, such as multiple observations explored in static networks may investigate identifying the sources of rumors spreading across various network platforms.

## REFERENCE:

[1] A.Agaskar and Y. M. Lu, "A fast monte carlo algorithm for source localization on graphs," in SPIE Optical Engineering and Applications. International Society for Optics and Photonics, 2013.

[2] Jiaojiao Jiang , Sheng Wen, Shui Yu,Yang Xiang and Wanlei Zhou,"Rumor Source Identification in Social Networks with Time-varying Topology" ,2016.

[3] M.P.Viana, D.R.Amancio,and L.d.F. Costa,"Ontimevaryingcollaborationnetworks,"Journal of Infometrics, vol.7,no.2 ,pp.371-378,2013.

[4] W.Zang ,P.Zang, C.Zhou,and L.Guou ,"Discovering multiple diffusion source nodes in social networks,"Procedia Computer Science,vol.29,pp.443-452,2014.

[5] J.Jiang, S.Wen, S.Yu, Y.Xiang and W.Zhou "k-center :An Approach on the multi-source identification of information diffusion",2015.

[6] B.Ribeiro, N.Perra , and A.Baronchelli , "Quantifying the temporal resolution on time varying networks," scientific reports,vol.3,2013.

[7] K.Zhu and L.Yiang,"information source detection in SIR model::A sample path based approach,"in Information Theory and Applications Workshop(ITA),2013,pp.1-9.

[8] P.C. Pinto,P.Thiran, and M.Vetterli,"Locating the source

of diffusion in large-scale networks,"phys.Rev.Lett.,vol.109,Aug 2012.

[9] Qiao Zhang,Shuiyuan Zhang,Jian Dong,Jinhua Xiong and Xueqi Cheng,"Automatic Detection of Rumor on Social Network,"2015.

## BIOGRAPHIES

Nivetha.S is currently pursuing her B.E degree in Prince shri venkateshwara padmavathy engineering college,Tamilnadu.
She is interested in Cryptography and network security.
She has authored or coauthored.

Priyadharshini .R is currently pursuing her B.E degree in Prince shri venkateshwara padmavathy engineering college,Tamilnadu.
She is interested in Computer graphics.
She has authored or coauthored.

Balakumar.P received the B.Sc, M.Sc, M.Tech and Ph.D. in Ponnaiyah Ramajayam College, St.Joseph's college and Bharath University in 2003, 2005, 2007, 2011 respectively. He has authored or coauthored 19 papers in journals and conferences. He is interested in Computer networks
He is a professor from 2014 till now in Prince Dr.K.vasudevan college of Engineering and Technology. He received several paper awards from prestigious multimedia journals and conferences

Kapila Vani R.K received the B.E degree in Computer Science and Engineering and M.E degree in Computer Science and Engineering from Dhanalakshmi college of Engineering and alpha college of Engineering, India in 2006 and 2014 respectively. Her interests include Data Mining Algorithms and Software Engineering. She is an assistant professor from 2014 till now in Prince Dr.K.vasudevan college of Engineering and Technology.