IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

Rumour Source Identification in Social Networks with Time-Varying Topology

Ms.M.Narmatha[2], Ms.S.Aruna Devi[2]

Assistant Professor, Department of Computer Science, Sri Krishna Arts and Science College, Coimbatore[1]

M.Sc., Department of Computer Science, Sri Krishna Arts and Science College, Coimbatore[2]

Abstract: I explore the problem of rumour source identification in time-varying social networks that can be reduced to a series of static networks by introducing a time-integrating window. I borrow an idea from criminology and propose a novel method to overcome the challenges. First, I reduce the time-varying networks to a series of static networks by introducing a time-integrating window. Second, instead of inspecting every individual in traditional techniques, I adopt a reverse dissemination strategy to specify a set of suspects of the real rumour source. This process addresses the scalability issue of source identification problems, and therefore dramatically promotes the efficiency of rumour source identification. Third is to determine the real source from the suspects. Information that propagates through social networks can carry a lot of false claims. For example, rumours on certain topics can propagate rapidly leading to a large number of nodes reporting the same (incorrect) observations. In this paper, I describe an approach for finding the rumour source and assessing the likelihood that a piece of information is in fact a rumour, in the absence of data provenance information. I model the social network as a directed graph, where vertices represent individuals and directed edges represent information flow. A number of monitor nodes are injected into the network whose job is to report data they receive.

Keywords: Source identification, Time Varying, Static Networks, Social Networks.

I. INTRODUCTION

Spreading rumours in the social network had been a censorious damage to the society. Social networks are popular media for sharing information. Online social networks enable large-scale information dissemination in a very short time, often not matched by traditional media. Missing information and false claims can also propagate rapidly through social networks. This is exacerbated by the fact that (i) anyone can publish (incorrect) information and (ii) it is hard to tell who the original source of the information is.The access of social network encourages not only the effectiveness of information sharing but also spreading rumours.

In this paper, I focus on two problems related to mitigation of false claims in social networks. First, I study the question of identifying sources of rumours in the absence of complete provenance information about rumour propagation. Second, I study how rumours (false claims) and non-rumours (true information) can be differentiated. Our method is based on an assumption that rumours are initiated from only a small number of sources, whereas truthful information can be observed and originated by a large number of unrelated individuals concurrently. Our approach relies on utilizing network monitors; individuals who agree to let us know whether or not they heard a particular piece of information (from their social neighbourhood), although do not agree to let us know who told them this information or when they learned it. Hence, all I know is which of the monitors heard a particular piece of information. This, in some sense, is the most challenging scenario that offers a worst-case bound on accuracy of rumour detection and source identification. Additional information can only simplify the problem. I show, that even in the aforementioned worst case, promising results can be achieved.

II. RELATED WORKS

I inventory the possible attacks against the integrity of the OLSR network routing infrastructure, and present a technique for securing the network. In particular, assuming that a mechanism for routing message authentication has been deployed [1]. I concentrate on the problem where otherwise “trusted” nodes have been compromised by attackers, which could then inject false (however correctly signed) routing messages. Our main approach is based on authentication checks of information injected into the network, and reuse of this information by a node to prove its link state at a later time. I finally synthesize the overhead and the remaining vulnerabilities of the proposed solution [1].

A rouge node can, indeed, manipulate this assumption and mount attacks against the concerned routing protocol to disrupt routing operations. In addition, a malicious node may also launch Denial of Service (DoS) attacks to deprive legitimate nodes from being serviced. In this chapter, I provide an insight into the various routing attacks available in

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

literature, namely, flooding/resource consumption, wormhole, black hole, link withholding, link spoofing, and replay attacks [2].

The Optimized Link State Routing (OLSR) protocol is a proactive Mobile Ad hoc Network (MANET) routing protocol. Security aspects have not been designed into the OLSR protocol and therefore make it vulnerable to various kinds of attacks. Recent research efforts have focused on providing authentication and encryption techniques to secure the OLSR protocol against attacks from outside intruders. A second line of defence is required to provide intrusion detection and response techniques in protecting the OLSR protocol against attacks from inside intruders [3].

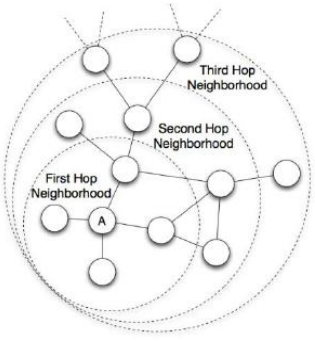
In proactive routing protocols, such as the Optimized Link State Routing (OLSR) protocol, nodes obtain routes by periodic exchange of topology information. Most of these routing protocols rely on cooperation between nodes due to the lack of a centralized administration and assume that all nodes are trustworthy and Ill-behaved. However, in a hostile environment, a malicious node can launch routing attacks to disrupt routing operations or denial-of-service (DoS) attacks to deny services to legitimate nodes [4].

In this paper I investigate security issues related to the Optimized Link State Routing Protocol – one example of a proactive routing protocol for MANETs. I inventory the possible attacks against the integrity of the OLSR network routing infrastructure, and present a technique for securing the network. In particular, assuming that a mechanism for routing message authentication (digital signatures) has been deployed, I concentrate on the problem where otherwise "trusted" nodes have been compromised by attackers, which could then inject false (however correctly signed) routing messages [5]. In this paper I review a specific DOS attack called node isolation attack and propose a new mitigation method. [6] Our solution called Denial Contradictions with Fictitious Node Mechanism (DCFM) relies on the internal knowledge acquired by each node during routine routing, and augmentation of virtual (fictitious) nodes. Moreover, DCFM utilizes the same techniques used by the attack in order to prevent it. The overhead of the additional virtual nodes diminishes as network size increases, which is consistent with general claim that OLSR functions best on large networks.

III.PROPOSED SYSTEM

A. RANK OF TRUE SOURCE

Using the method presented in Section 2.1, all nodes are sorted in the likelihood that they are the actual rumour source. Figure 3 shows the average rank of the actual source in the output. In the ideal case, the rank should be one which means that the top suspect is actually the rumour source. Note that, regardless of the monitor selection method, the rank of the true source generally decreases (i.e., improves by becoming closer to 1) as the number of monitors increases. Dist and NI+Dist generally show a bad accuracy. Random also performs poorly when the number of monitors is small, but it improves as more monitors are added. NI, BC and BC+Dist show better performance than the others. When the number of monitors is very large, the choice of monitor selection does not matter that much anymore, and all algorithms converge. One of the important factors that affect the accuracy of rumour source identification is the number of positive monitors. Figure shows the ratio of experiments in which no monitor received the rumour.



In all monitor selection methods, the ratio decreases as the number of monitor increases. Among the four methods compared, the Dist. selection method has the highest ratio. Dist. basically maximizes inter-monitor distance, so it tends to choose nodes on the boundary of the graph. Therefore, monitors selected by Dist. have low probability of hearing rumours. The Random selection method also has a high ratio of negative monitors when the number of monitors is small. The other methods (NI, NI+ Dist., BC, BC+ Dist.) have small ratio compared to Dist. and Random. When no monitor hears the rumour, it is very hard to find the source accurately as shown in Figure 3 (Random and Dist. when the number of monitors is 20, for example). however, a larger number of positive monitors does not always lead to a more accurate result. Figure 5 shows the average number of positive monitors when the number of monitors is 160. Figure 3 shows that BC has best accuracy, NI has second best, and the others are worse. Note that, the number of

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

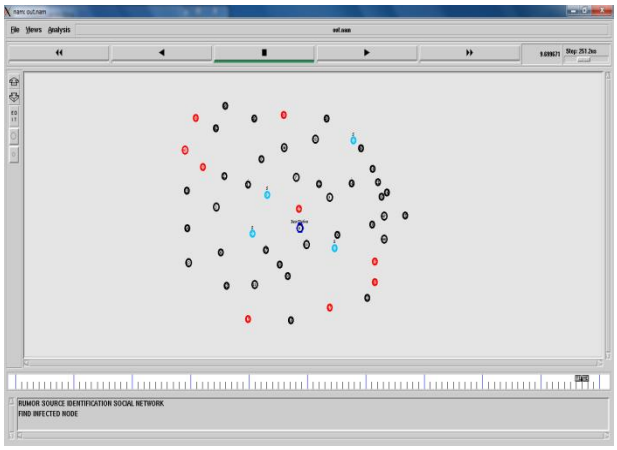
positive monitors in NI is almost double of that in BC, but BC is more accurate. This means that it is not always helpful to have more positive monitors.

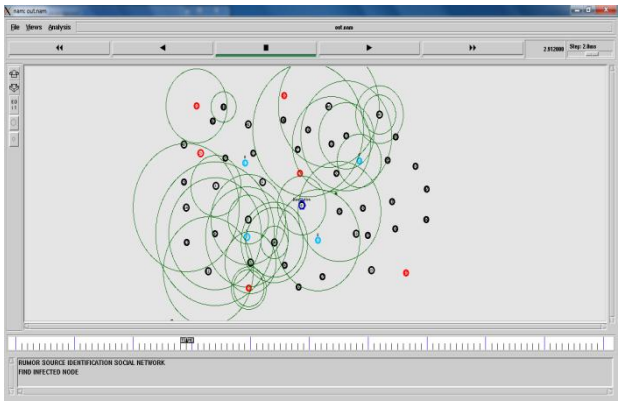
IV. IDENTIFYING RUMORS

A social network is modelled as a directed graph G = (V, E) where V is the set of all people and E is the set of edges where each edge represents information flow between two individuals. I assume that a set of k pre-selected nodes M (M ⊆ V) are our monitors. For rumour investigation purposes, given a specific piece of information, a monitor reports whether they received it or not. I denote the set of monitor nodes who received the rumour by M+, and the set of monitor nodes who have not received it by M− (where M+, M− ⊆ M). I call the former set positive monitors and the latter negative monitors. To identify the source of a rumour, I use the intuition that the source must be close to the positive monitors but far from the negative monitors. Hence, for each node x, our algorithm calculates the following four metrics:

(1) Reachability to all positive monitors I first calculate how many positive monitors are reachable from each node. For a node x to be the rumour source, x must have paths to all monitors in M+. If those paths do not exist, x cannot be a rumour source.
(2) Distance to positive monitors among those nodes that can reach all positive monitors, nodes that are closer, on average, are preferred. In other words, for each node x, I calculate the total distance X m∈ M+ and m is reachable from x d(x, m). Where d(x, m) is the distance from x to m and sort the suspected sources by increasing total distance from positive monitors.

If a rumour is initiated by some person intentionally, it is not independently corroborated by others. Hence, in the absence of collusion, there is only one source of the rumour in the network. If a rumour is initiated by a small colluding group of people, the number of independent sources is just the size of the group. Conversely, if a piece of information is not a rumour, there may be many independent sources of the information. Therefore, it is important to estimate the number of independent sources correctly.

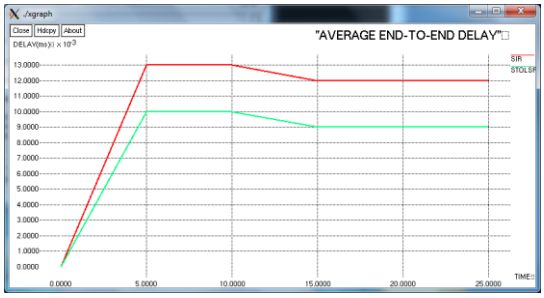


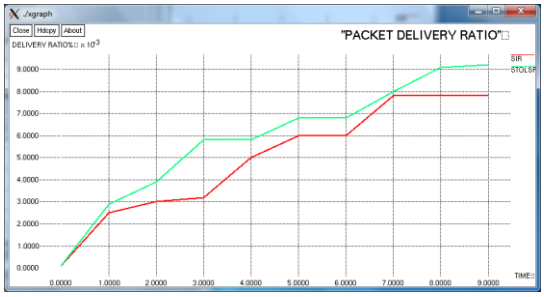


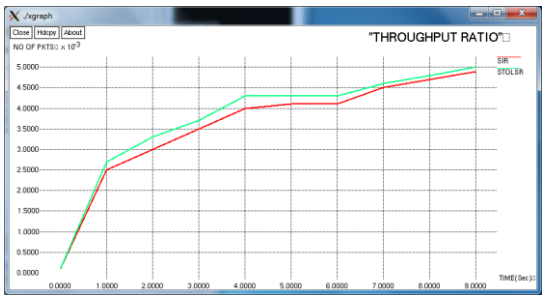
IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

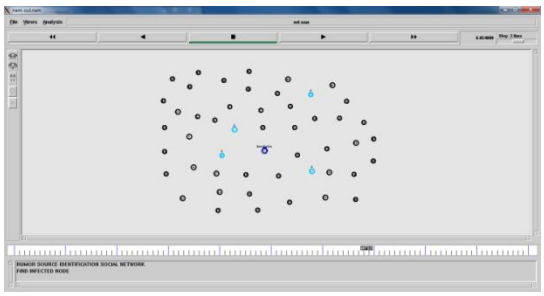
V. RESULTS AND DISCUSSIONS

From the result, I can observe that, if rumours and non-rumours have very large difference in the number of sources, rumour classification can be done with very high accuracy. As the difference in the number of sources of rumours and non-rumours decreases, it gets harder to classify rumours and non-rumours accurately. Another observation from the table is that the algorithms which show good results in rumour source identification (BC and BC+ Dist.) do not always work Ill in rumour classification. This is because the two tasks have different conditions for best performance. In rumour source identification, it is best to have monitors near the rumour source, so that they receive the rumour and estimate the rumour source based on their locations. In rumour classification, it is best to have monitors in various places in the rumour propagation trees so that GSSS and MDGIP can be estimated accurately. I leave finding a monitor selection algorithm that is good for both tasks as future work.









VI. CONCLUSION

In this paper, I proposed an approach for (i) determining whether a piece of information is a rumour or not, and (ii) finding the source of the rumour. Our approach uses a very small amount of provenance information; namely, which of a set of monitors heard the piece of information at hand. To find the rumour source, our algorithm evaluates the likelihood of each node to be the source, calculated from node connectivity and shortest path distances. For rumour classification, I proposed two metrics – Greedy Source Set Size (GSSS) and Maximal Distance of Greedy Information

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 7, Issue 2, February 2018

Propagation (MDGIP) – and used logistic regression. To evaluate the proposed approach, I performed a case study involving a real social network crawled from Twitter. The algorithm shows good potential to help users in identifying rumours and their sources.

REFERENCES

[1] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," IEEE Commune. Survey. Tut. vol. 16, no. 3, pp. 1658–1686, Third 2014.
[2] D. Niyato, E. Hossain, and S. Camorlinga, "Remote patient monitoring service using heterogeneous wireless access networks: architecture and optimization," IEEE J. Sel. Areas Commun., vol. 27, no. 4, pp. 412–423, May 2009.
[3] "Cisco visual networking index: Global mobile data traffic forecast update, 2012-2017," 2013.
[4] H. Lee, K.-J. Park, Y.-B. Ko, and C.-H. Choi, "Wireless LAN with medical-grade QoS for e-healthcare," J. Commun. and Netw. vol. 13, no. 2, pp. 149–159, Apr. 2011.
[5] S. Misra and S. Sarkar, "Priority-based time-slot allocation in wireless body area networks during medical emergency situations: An evolutionary game-theoretic perspective," IEEE J. Biomed. Health Inform. vol. 19, no. 2, pp. 541–548, Mar. 2015.
[6] S. Moulik, S. Misra et al., "Cost-effective mapping betIen wireless body area networks and cloud service providers based on multi-stage bargaining," IEEE Trans. Mobile Compute., to appear.
[7] F. Peter. (2013, April 23) 'bogus' ap tweet about explosion at the white house wipes billions off us markets. The Telegraph,Finance/Market. Washington.
[8] B. Ribeiro, N. Perra, and A. Baronchelli, "Quantifying the effect of temporal resolution on time-varying networks," Scientific reports, vol. 3, 2013.
[9] M. P. Viana, D. R. Amancio, and L. d. F. Costa, "On time varying collaboration networks," Journal of Informetrics, vol. 7,no. 2, pp. 371–378, 2013.
[10] M. Karsai, N. Perra, and A. Vespignani, "Time varying networks and the weakness of strong ties," Scientific reports, vol. 4,2014.