An efficient Technique for the detection of Malicious Users over Large-Scale Social Networks

V Kiran Kumar, N Sai Pranith, L Laxmi Priya, V Kiran Kumar Reddy
Under guidance of **Ankita Sharma**
B. Tech Department of Computer Science and Engineering
St. Martin's Engineering College, Hyderabad, Telangana, India

ABSTRACT

The past few years have witnessed the dramatic popularity of large-scale social networks where malicious nodes detection is one of the fundamental problems. Most existing works focus on actively detecting malicious nodes by verifying signal correlation or behavior consistency. It may not work well in large-scale social networks since the number of users is extremely large and the difference between normal users and malicious users is inconspicuous. In this paper, we propose a novel approach that leverages the power of users to perform the detection task. We design incentive mechanisms to encourage the participation of users under two scenarios: full information and partial information. In full information scenario, we design a specific incentive scheme for users according to their preferences, which can provide the desirable detection result and minimize overall cost. In partial information scenario, assuming that we only have statistical information about users, we first transform the incentive mechanism design to an optimization problem, and then design the optimal incentive scheme under different system parameters by solving the optimization problem. We perform extensive simulations to validate the analysis and demonstrate the impact of system factors on the overall cost.

1. INTRODUCTION

The past few years have witnessed the dramatic popularity of large-scale social networks. They greatly facilitate our daily lives and connect us with a world-wide virtual society [1]–[18]. Meanwhile, security issues in these networks are attracting more and more research attention, one of which is the malicious nodes (users) detection [19], [20]. For example, in [21] the authors collected one month sample of Twitter data, examined 25 million unique URLs and found that over two million (roughly 8%) URLs are scams, malware, and phishing. It is also shown in [22] that 3.6 million U.S. adults lost a total of 3.2 billion dollars due to phishing attacks in 2007. Therefore, the malicious users in social networks have a terrible impact on the network, in terms of degrading the network's performance, reducing the network's efficiency, increasing the cost or even disabling the whole network. It is pressing to detect malicious users and isolate them efficiently.

2. LITERATURE SURVEY

Dynamic channel assignment for wireless sensor networks: a regret matching based approach

Multiple channels in Wireless Sensor Networks (WSNs) are often exploited to support parallel transmission and reduce interference. However, there are many challenges, such as extra communication overhead, posed to the energy constraint of WSNs by the multi-channel usage coordination. In this paper, we propose a Regret Matching based Channel Assignment algorithm (RMCA) to address those challenges.

Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks

In rechargeable sensor networks (RSNs), energy harvested by sensors should be carefully allocated for data sensing and data transmission to optimize data gathering due to time-varying renewable energy arrival and limited battery capacity. Moreover, the dynamic feature of network topology should be taken into account, since it can affect the data transmission. In this paper, we strive to optimize data gathering in terms of

network utility by jointly considering data sensing and data transmission. To this end, we design a data gathering optimization algorithm for dynamic sensing and routing (DoSR), which consists of two parts. In the first part, we design a balanced energy allocation scheme (BEAS) for each sensor to manage its energy use, which is proven to meet four requirements raised by practical scenarios.

Privacy preserving compressive sensing for crowdsensing based trajectory recovery

Location based services have experienced an explosive growth and evolved from utilizing a single location to the whole trajectory. Due to the hardware and energy constraints, there are usually many missing data within a trajectory. In order to accurately recover the complete trajectory, crowdsensing provides a promising method.

Device-to-device communication in lte-advanced networks: a survey

Among the LTE-A communication techniques, Device-to-Device (D2D) communication which is defined to directly route data traffic between spatially closely located mobile user equipments (UEs), holds great promise in improving energy efficiency, throughput, delay, as well as spectrum efficiency. As a combination of ad-hoc and centralized communication mechanisms, D2D communication enables researchers to merge together the long-term development achievements in previously disjoint domains of ad-hoc networking and centralized networking.

Performance modeling for relay cooperation in delay tolerant networks

In this paper, we focus on such relay cooperation and analytically explore its impact on the delivery performance in DTNs. Specifically; we first develop a continuous time Markov chain-based theoretical framework to model the complicated message delivery process in delay tolerant networks adopting the two-hop relay algorithm. We then derive closed-form expressions for both the expected delivery delay and the corresponding expected delivery cost, where the important relay behaviors of forwarding traffic for itself or for other nodes are carefully incorporated into the analysis.

3. OVERVIEW OF THE SYSTEM

ARCHITECHTURE

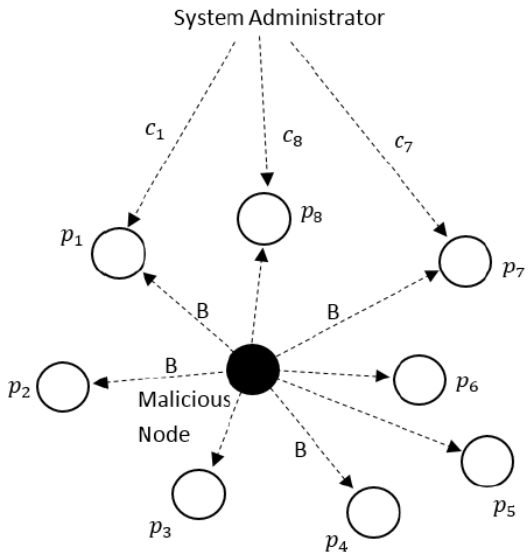

Fig 3.1 Architecture Diagram

EXISTING SYSTEM:

Authors proposed a scheme named Watchdog which is consisted of two parts, watchdog and path rater to detect the malicious nodes and improve the throughput of the network when transmitting data packet. The first part is to detect whether there are malicious misbehaviors. It will count how many times a node unsuccessfully forward a packet and if the time is too much, i.e., exceeds a predefined threshold it thinks the node is malicious. The second part is to find a shortest path without the reported node. However, when the node itself which is reporting other nodes is malicious, it will cause a serious impact on the network.

Authors proposed a scheme named Enhanced AACK (EAACK) with digital signature. An alternative which is based on signal correlation an alternative which is based on signal correlation as sensors have exactly location correlation is used in many works. Used the idea of classification and devise a multivariate classification algorithm to detect the malicious nodes. In, a voting scheme in which the closer nodes have greater weight was used to get a better result.

DISADVANTAGES:

When the node itself which is reporting other nodes is malicious, it will cause a serious impact on the network. Its solution may not be quite practical, as in many applications, the malicious nodes will tamper the packet being transmitted.

Malicious node would contaminate the voting results of its nearest sensors in this scheme.

PROPOSED SYSTEM:

In this paper, we propose an approach to detect malicious users in large-scale social networks from a radical new perspective. The system administrator is not directly participated in the detection process. Instead, it leverages the power of normal users in the social networks to accomplish such a difficult goal, i.e., crowdsourcing the detection tasks to the users.

The system administrator has to decide an incentive scheme to encourage the report of malicious node from users.

We design incentive mechanism under two scenarios: full information and partial information.

ADVANTAGES:

It will provide a common payoff to every user. Considering another fact that incentive scheme to give is strongly dependent on these N users' preferences.

It is hard for the malicious user to make the best strategy to maximize its own interest.

4. MODULES

- OSN System Construction Module.
- Normal Full Profile/Partial Profile User Module
- Malicious User
- Crowd Sourcing Module
- Calculating Malicious User
- Incentive Mechanism Design

Admin:

Calculating Malicious User:

Based on the links shared by different types of users in social network we will calculate malicious user. Firstly full profile users are given preference and if shared links of specific users is more than threshold value use is considered is malicious user.

Incentive Mechanism Design:

Admin will block malicious user and give incentives to users who has reported malicious users.

User:

OSN System Construction Module

In the first module, we develop the Online Social Networking (OSN) system module. We build up the system with the feature of Online Social Networking. Where, this module is used for new user registrations and after registrations the users can login with their authentication.

Where after the existing users can send messages to privately and publicly, options are built. Users can also share post with others. The user can able to search the other user profiles and public posts. In this module users can also accept and send friend requests.

With all the basic feature of Online Social Networking System modules is build up in the initial module, to prove and evaluate our system features.

Normal Full Profile/Partial Profile User Module:

Full Information User:

We consider the case where the system administrator knows the preference of each user in U. If the system administrator has accumulated enough users' information such as age, personality and hobby for a long time or users in the network are willing to share their preference with the system administrator, the system administrator will have the access to

the full information. Without loss of generality, we order users' preferences in the ascending order, i.e., p1 _ p2 _ _ _ _ _ pN0 _ _ _ _ _ pN. For n _ N0, if B + pi _ 0, which means the user himself/herself has motivation to report the malicious user.

Partial Information User:

If the system administrator has not accumulated enough users' information and users in the network are not willing to share their preferences with the system administrator, the system administrator will not have the full information.

Malicious User:

Malicious users perform abnormal activities such as cyber attack or advertisement injection. malicious users in social networks have a terrible impact on the network, in terms of degrading the network's performance, reducing the network's efficiency, increasing the cost or even disabling the whole network.

Crowd Sourcing Module:

Crowdsourcing the detection tasks to the users. When malicious users perform abnormal activities such as cyber-attack or advertisement injection, the users who are the victims of these activities can report them to the system administrator.

6. CONCLUSION

In this paper, we investigated the malicious user detection in the large-scale social networks using crowdsourcing, considering that the malicious user may avoid being reported normal users through providing some incentives and users have different preferences for the malicious user. From the perspective of normal users' preferences, we consider two scenarios: full information and partial information. For full information, we devised the incentive scheme by order users' preferences. For partial information, we focused on two cases where users' preferences follow a uniform distribution and gaussian distribution, respectively. Corresponding incentive schemes were also devised. We have also conducted simulations to illustrate the impact of different factors on the total cost of the system.

In the future work, we will consider the collective impact of multiple malicious users and the incentive mechanism design for scenarios where different users may have different distribution of its preference. Also, we will consider that the malicious user may optimize the constant incentive *B*. In such case, the malicious user may want to maximize its own payoff and the system may want to minimize its cost. The problem can be transformed as a game.

7. REFERENCES

[1] J. Chen, Q. Yu, B. Chai, Y. Sun, Y. Fan, and X. Shen, "Dynamic channelassignment for wireless sensor networks: a regret matching basedapproach," *IEEE Transactions on Parallel and Distributed Systems*,vol. 26, no. 1, pp. 95–106, 2015.

[2] J. Chen, J. Li, S. He, T. He, Y. Gu, and Y. Sun, "On energy-efficient trapcoverage in wireless sensor networks," *ACM Trans. Sen. Netw.*, vol. 10,no. 1, pp. 2:1–2:29, 2013.

[3] G. Han, C. Zhang, L. Shu, and J. J. Rodrigues, "Impacts of deploymentstrategies on localization performance in underwater acoustic sensornetworks," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3,pp. 1725–1733, 2015.

[4] Y. Zhang, S. He, and J. Chen, "Data gathering optimization by dynamicsensing and routing in rechargeable sensor networks," *IEEE/ACM Transactionson Newworking*, 2015. DOI: 10.1109/TNET.2015.2425146, toappear.

[5] C. Zhou, Z. Shi, Y. Gu, and N. A. Goodman, "DOA estimation bycovariance matrix sparse reconstruction of coprime array," in *Proceedingsof IEEE International Conference on Acoustics, Speech, and SignalProcessing (ICASSP)*, pp. 2369–2373, 2015.

[6] M. Dong, X. Liu, Z. Qian, A. Liu, and T. Wang, "Qoe-ensured pricecompetition model for emerging mobile

networks," *IEEE WirelessCommunications*, vol. 22, no. 4, pp. 50–57, 2015.

[7] L. Kong, L. He, X.-Y. Liu, Y. Gu, M.-Y. Wu, and X. Liu, "Privacypreservingcompressive sensing for crowdsensing based trajectory recovery,"in *Proceedings of IEEE International Conference on DistributedComputing Systems (ICDCS)*, pp. 31–40, 2015.

[8] K. Wei, M. Dong, K. Ota, and K. Xu, "CAMF: Context-aware messageforwarding in mobile social networks," *IEEE Transactions on Paralleland Distributed Systems*, vol. 26, no. 8, pp. 2178–2187, 2015.

[9] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communicationin lte-advanced networks: a survey," *IEEE Communications Surveys& Tutorials*, vol. 17, no. 4, pp. 1923–1940, 2015.

[10] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, "Performance modeling forrelay cooperation in delay tolerant networks," *Springer Mobile Networksand Applications*, vol. 18, no. 2, pp. 186–194, 2013.

[11] K. Zheng, Z. Yang, K. Zhang, P. Chatzimisios, K. Yang, and W. Xiang,"Big data-driven optimization for mobile networks toward 5g," *IEEENetwork*, vol. 30, no. 1, pp. 44–51, 2016.

[12] C. Huang, R. Zhang, and S. Cui, "Optimal power allocation for outageprobability minimization in fading channels with energy harvestingconstraints," *IEEE Transactions on Wireless Communications*, vol. 13,no. 2, pp. 1074–1087, 2014.

[13] C. Huang, R. Zhang, and S. Cui, "Throughput maximization for thegaussian relay channel with energy harvesting constraints," *IEEE Journalon Selected Areas in Communications*, vol. 31, no. 8, pp. 1469–1479,2013.

[14] S. Liu, H. Zhu, R. Du, C. Chen, and X. Guan, "Location privacypreserving dynamic spectrum auction in cognitive radio network," in*Proceedings of IEEE 33rd International Conference on DistributedComputing Systems (ICDCS)*, pp. 256–265, 2013.

[15] M. Dong, H. Li, K. Ota, L. T. Yang, and H. Zhu, "Multicloudbasedevacuation services for emergency management," *IEEE Cloud Computing*, vol. 1, no. 4, pp. 50–59, 2014.

[16] L. Wei, H. Zhu, Z. Cao, X. Dong, W. Jia, Y. Chen, and A. V. Vasilakos,"Security and privacy for storage and computation in cloud computing,"*Elsevier Information Sciences*, vol. 258, no. 1, pp. 371–386, 2014.

[17] Q. Yang, S. He, J. Li, J. Chen, and Y. Sun, "Energy-efficient probabilisticarea coverage in wireless sensor networks," *IEEE Transactions onVehicular Technology*, vol. 64, no. 1, pp. 367–377, 2015.