RESEARCH ARTICLE OPEN ACCESS

Adaptive Diffusion of Sensitive Information in Online Social Networks

Mr D.Purushothaman MCA.,M.E.,[1] , K.Giri vardhan raju [2]

[1] Asst. Professor, Department of Computer Applications

[2] Student, Department of Computer Applications

[1], [2] Chadalawada Ramanamma Engineering College (Autonomous)

ABSTRACT

The cascading of sensitive information such as private contents and rumors is a severe issue in online social networks. One approach for limiting the cascading of sensitive information is constraining the diffusion among social network users. However, the diffusion constraining measures limit the diffusion of non-sensitive information diffusion as well, resulting in the bad user experiences. To tackle this issue, in this project, study the problem of how to minimize the sensitive information diffusion while preserve the diffusion of non-sensitive information, and formulate it as a constrained minimization problem where characterize the intention of preserving non-sensitive information diffusion as the constraint and study the problem of interest over the fully-known network with known diffusion abilities of all users and the semi-known network where diffusion abilities of partial users remain unknown in advance. By modeling the sensitive information diffusion size as the reward of a bandit, utilize the bandit framework to jointly design the solutions with polynomial complexity in the both scenarios. Moreover, the unknown diffusion abilities over the semi-known network induce it difficult to quantify the information diffusion size in algorithm design. For this issue, propose to learn the unknown diffusion abilities from the diffusion process in real time and then adaptively conduct the diffusion constraining measures based on the learned diffusion abilities, relying on the bandit framework.

Keywords: - Sensitive information, Complexity, minimization Problem, Diffusion.

I. INTRODUCTION

On the contrast, the semi known network here refers to the case that diffusion abilities of partial users remain unknown in advance. For example, the data of Facebook was reported to be utilized to influence the 2016 election in the US, which then led to a severe trust crisis for Facebook. Thus, due to the privacy concern and potential side effect, even of network managers, it is difficult to obtain the full topology of some global large scale social networks like Facebook, WeChat. Unless the full network topology is known, we cannot evaluate the diffusion abilities of all users. In the IC model, each user has a single chance to successfully diffuse the information to his neighbors with a given probability after this user having received the information. While in the LT model, a user would get the information if a certain fraction of his neighbors have received the information. Since then, a great deal of works study the Influence Maximization (IM) problem, which focuses on efficiently selecting the optimal seed users to trigger a diffusion process in hope of maximizing the final information diffusion size. The system is less effective due to lack of Constraining sensitive Information diffusion. The system doesn't effective due to lack of Mapping Adaptive Diffusion in Fully-known Network into Bandit.

II. RELATEDWORKS

Yuchen Li, Ju Fan∗ , Yanhao Wang, and Kian-Lee Tan proposed Influence Maximization on Social Graphs: A Survey. Influence Maximization (IM), which selects a set of k users (called seed set) from a social network to maximize the expected number of influenced users (called influence spread), is a key algorithmic problem in social influence analysis. Due to its immense application potential and enormous technical challenges, IM has been extensively studied in the past decade. In this Project, survey and synthesize a wide spectrum of existing studies on IM from an algorithmic perspective

Yishi Lin, Wei Chen proposed Boosting Information Spread: An Algorithmic Approach. The majority of influence maximization (IM) studies focus on targeting influential seeders to trigger substantial information spread in social networks. In this project, consider a new and complementary problem of how to further increase the influence spread of given seeders. Our study is motivated by the observation that direct incentives could "boost" users so that they are more likely to be influenced by friends. study the k-boosting problem which aims to find k users to boost so that the final "boosted" influence spread is maximized.

Kempe first propose two classic diffusion models: Independent Cascading (IC) model and linear threshold (LT) model. In the IC model, each user has a single chance to successfully diffuse the information to his neighbors with a given probability after this user having received the information. While in the LT model, a user would get the information if a certain fraction of his neighbors have received the information. Since then, a great deal of works study the Influence Maximization (IM) problem, which focuses on efficiently selecting the optimal seed users to trigger a diffusion process in hope of maximizing the final information diffusion size. Recently, due to the high cost of seeding influential users, Shi propose to let influential users repost the required information while seed the ordinary users for lowering the cost of IM campaign. Similar to the multi-round setting in this paper, the seed selection for maximizing the information diffusion in multiple time rounds is considered in . Moreover, considering the widespread interactions between the cyber (online) and physical (offline) worlds, offline events are utilized in to further improve the performance of IM. On the contrast of the IM problem, there are also abundant researches focusing on minimizing the influence of rumors. One strategy for rumor influence minimization is diffusing the truths over network to counteract rumors. Specifically, the competitive linear threshold (CLT) model that characterizes the competing diffusion of truth and rumor is introduced. Then He and Chen propose to select a set of seed users to maximize the diffusion of truths under the CLT model. Chen extends the IC model to describe the diffusion of positive informations under the effect of negative information, and studies how to maximize the positive information diffusion. However, such clarifying measure cannot be used to constrain the diffusion of private sensitive informations such as personal informations, trade secrets. Another class of rumor blocking measures focuses on blocking a certain number of influential users or social links. On one hand, Song propose to temporarily block a number of users with high diffusion abilities to reduce the diffusion of rumors before a deadline. With the consideration of user experiences, Wang study the online rumor blocking problem that periodically blocking a fraction of users during the rumor diffusion, and set a threshold to controls the blocking time of each user. Further, for coping with the unforeseen events in rumor diffusion, the adaptive blocking strategy is proposed. On the other hand, considering that straightforwardly blocking users is not desirable,propose to block a given number of social links for minimizing the diffusion of rumors. However, as we illustrated before, this kind of measures may incur much information diffusion loss, if being adopted to constrain the diffusion of the sensitive informations considered in this paper. In addition, taking measures to constrain or promote information diffusion is also related to the studies about the effect of human behaviors on diffusion.

Lichao Sun, Weiran Huang, Philip S. Yu proposed Multi-Round Influence Maximization**.** In this project, study the Multi-Round Influence Maximization (MRIM) problem, where influence propagates in multiple rounds independently from possibly different seed sets, and the goal is to select seeds for each round to maximize the expected number of nodes that are activated in at least one round. MRIM problem models the viral marketing scenarios in which advertisers conduct multiple rounds of viral marketing to promote one product. consider two different settings: 1) the non-adaptive MRIM, where the advertiser needs to determine the seed. sets for all rounds at the very beginning, and 2) the adaptive MRIM, where the advertiser can select seed sets adaptively based on the propagation results in the previous rounds.

III. PROPOSED SYSTEM ARCHITECTURE

The system takes the first look into minimizing the diffusion size of sensitive informations while preserving the diffusion of non-sensitive ones. We formulate the problem of interest into a constrained minimization problem where we characterize the intention of preserving non-sensitive information diffusions as the constraint. The system proposes an efficient bandit based framework to jointly explore the solutions over the fully-known and semiknown networks within polynomial running time. Moreover, we design the distributed implementation scheme of our solutions for the further improvement of time efficiency. The system further extend our bandit based solution into a “learning- determining” manner for addressing the challenge of unknown diffusion abilities in semi-known networks. We theoretically prove that the regret bound of our solution is sub-linear to the diffusion time The system is more effective due to ADAPTIVE DIFFUSION IN FULLY-KNOWN NETWORKS. The system is more effective due to Constraining sensitive Information diffusion. The system takes the first look into minimizing the diffusion size of sensitive informations while preserving the diffusion of non-sensitive ones. formulate the problem of interest into a constrained minimization problem where characterize the intention of preserving non-sensitive information diffusions as the constraint. The system proposes an efficient bandit based framework to jointly explore the solutions over the fully-known and semiknown networks within polynomial running time. Moreover, design the distributed implementation scheme of our solutions for the further improvement of time efficiency. The system further extend our bandit based solution into a “learning- determining” manner for addressing the challenge of unknown diffusion abilities in semi-known networks. theoretically prove that the regret bound of our solution is sub-linear to the diffusion time, indicating that the probability variations returned by our solution approximates to the optimal one with the increase of diffusion time. The system performs extensive experiments on both real and synthetic social network datasets. The results demonstrate that the proposed algorithms can effectively constrain the diffusion of sensitive informations, and more importantly, enjoy a superiority over four baselines in terms of 40% less information diffusion loss.

Tweet Server: The following are the functionalities provided by the Tweet Server:

- Login
- View all users & authorize
- View all friends req & response
- Add Tweet category
- Add Filters
- View all tweets
- View diffusion of tweet sensitive information
- View sensitive information re tweet
- View positive information re tweet
- View negative information re tweet
- View all search history
- View all results
- Logout

User: The following are the functionalities provided by the User:

- Register and login
- View profile
- Search friends & friends request
- View all friends create tweet
- View all my created tweets
- Friends tweets & re tweets
- Logout

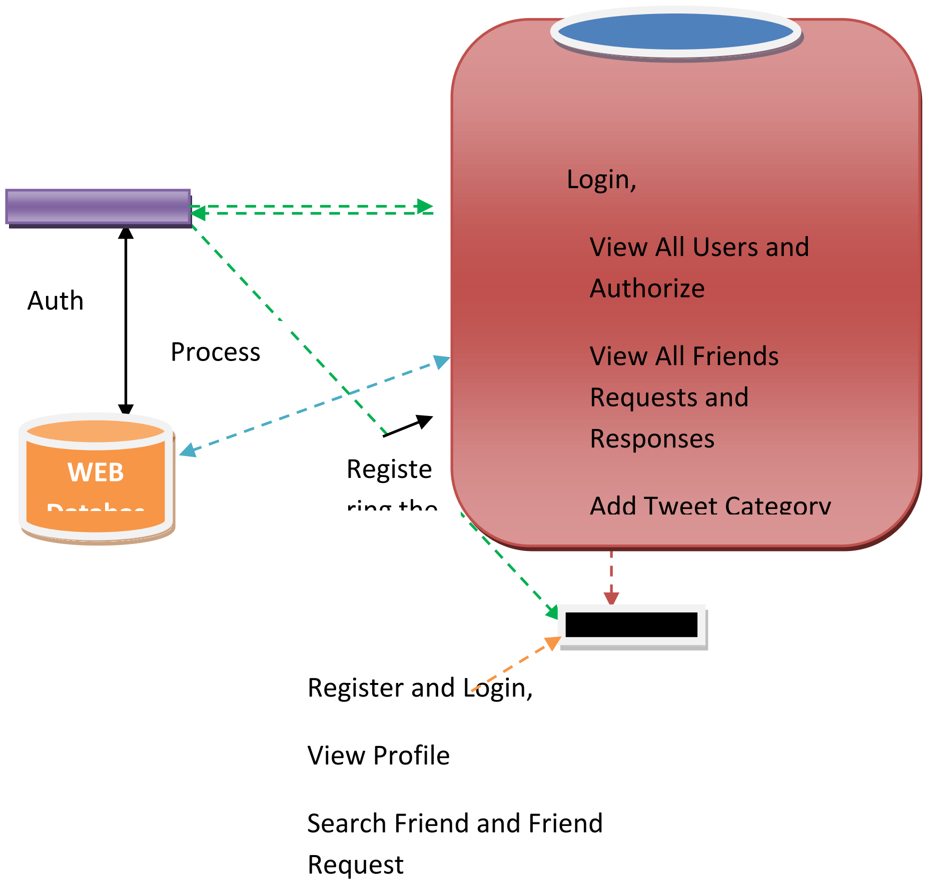

Fig.1 Proposed system Architecture

IV. RESULTS AND DISCUSSION

The output screens obtained after running and executing the system are shown from Fig.2 to Fig.4

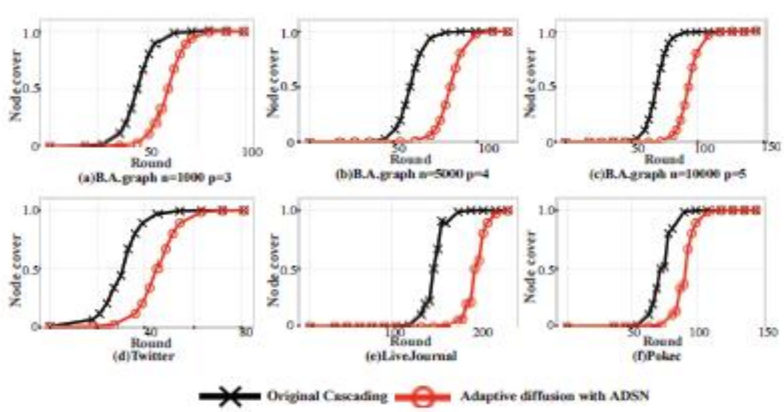

Fig.2 Diffusion size of sensitive information

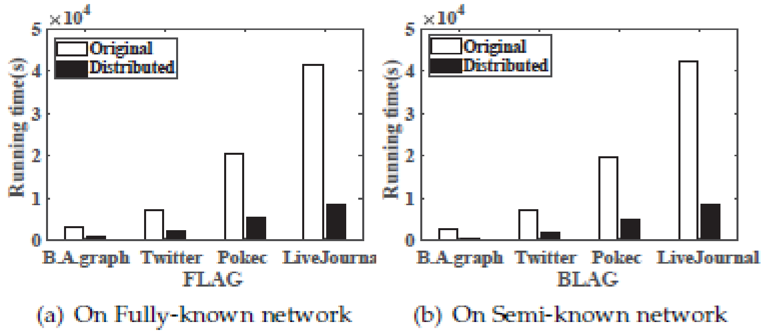

Fig.3 Running time of the model

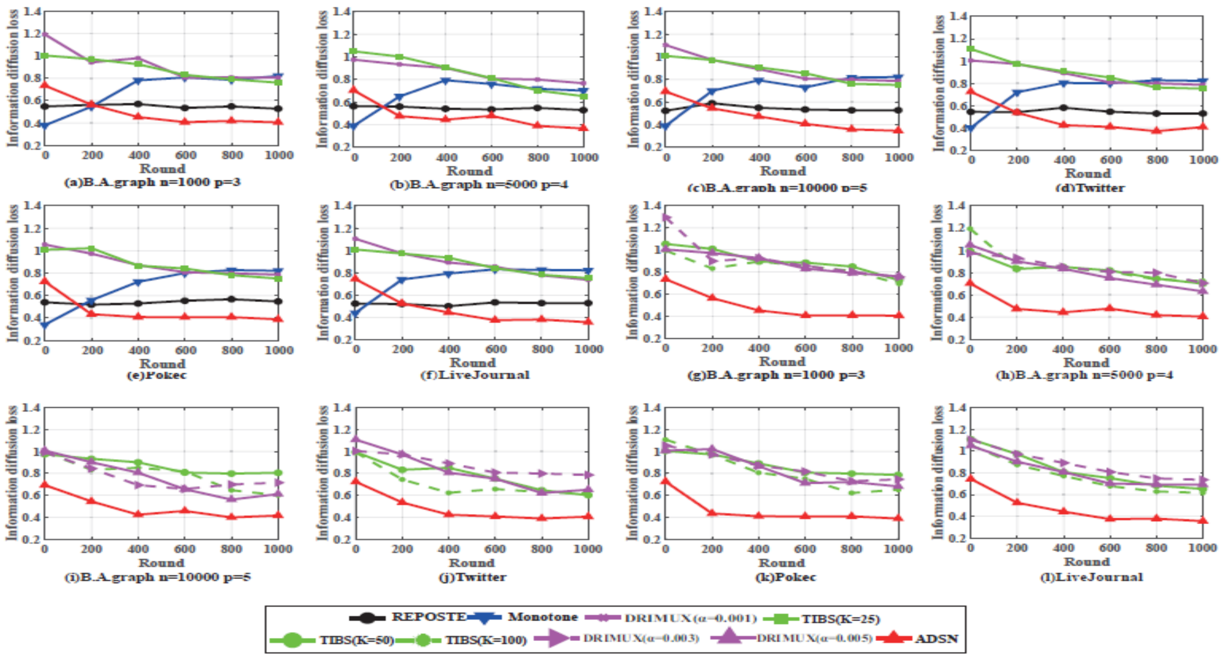

Fig.4 Information loss incurred

V. FUTURE SCOPE AND CONCLUSION

In this paper, we study the problem of constraining the diffusion of sensitive informations in social networks while preserving the diffusion of non-sensitive informations. We model the diffusion constraining measures as the variations of diffusion probabilities via social links, and model the problem of interest as adaptively determining the probability variations through a constrained minimization problem in multiple rounds. We utilize the CCMAB framework to jointly design our solutions in the fully-known and semiknown networks. Over the fully-known network, we propose the CCMAB based algorithm **ADFN** to efficiently determine the probability variations via social links. Over the semi-known network, for tackling the challenge of unknown diffusion abilities of partial users, we propose the algorithm **ADSN** to iteratively learn the unknown diffusion abilities and determine the probability variations based on the learned diffusion abilities in each round. The analysis of regret bound and extensive experiments have been conducted to justify the superiority of our solutions. In addition, in the current work, we define the constraint of maintaining the sum of diffusion probabilities via edges in the objective problem, for the aim of preserving the global diffusion ability of the whole network on diffusing nonsensitive informations. In the future work, we will explore other relevant solutions such as simultaneously minimizing the sensitive information diffusion and maximizing the nonsensitive information diffusion.

REFERENCES

[1] Y. Li, J. Fan, Y.Wang, and K. L. Tan, "Influence maximization on social graphs: A survey", in IEEE Transactions on Knowledge and Data Engineering (TKDE), vol. 30, no. 10, pp. 1852-1872, 2018.

[2] L. Sun, W. Huang, P. S. Yu, and W. Chen, " Multi-round influence maximization", in Proc. ACM SIGKDD, 2018.

[3] Q. Shi, C. Wang, J. Chen, Y. Feng, and C. Chen, "Post and repost: A holistic view of budgeted influence maximization", in Neurocomputing, vol. 338, pp. 92-100, 2019.

[4] X.Wu, L. Fu, Y. Yao, X. Fu, X.Wang, and G. Chen, "GLP: a novel framework for group-level location promotion in Geo-social networks", in IEEE/ACM Transactions on Networking (TON), vol. 26, no. 6, pp. 1-14, 2018.

[5] Y. Lin, W. Chen, and J. C. Lui, "Boosting information spread: An algorithmic approach", in Proc. IEEE ICDE, 2017.

[6] Y. Zhang, and B. A Prakash, "Data-aware vaccine allocation over large networks", in ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 10, no. 2, article 20, 2015.

[7] Q. Shi, C. Wang, J. Che n, Y. Feng, and C. Chen, "Location driven influence maximization: Online spread via offline deployment", in Knowledge-Based Systems, vol. 166, pp. 30-41, 2019.

[8] H. T. Nguyen, T. P. Nguyen, T. N. Vu, and T. N. Dinh, "Outward influence and cascade size estimation in billion-scale networks, in Proc. ACM SIGMETRICS, 2017.

[9] B. Wang, G. Chen, L. Fu, L. Song, and X. Wang, "Drimux: Dynamic rumor influence minimization with user experience in social networks", in IEEE Transactions on Knowledge and Data Engineering (TKDE), vol. 29, no. 10, pp. 2168-2181, 2017.

[10] Q. Shi, C. Wang, D. Ye, J. Chen, Y. Feng, and C. Chen, "Adaptive Influence Blocking: Minimizing the Negative Spread by Observation-based Policies", in Proc. IEEE ICDE, 2019.

[11] S. Wen, J. Jiang, Y. Xiang, S. Yu, W. Zhou, and W. Jia, "To shut them up or to clarify: Restraining the spread of rumors in online social networks", in IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 12, pp. 3306-3316, 2014.