Volume 4 || Special Issue 8 || ICCEME 2019-2020 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

INFLUENTIAL NODE TRACKING ON SOCIAL MEDIA: USING GREEDY APPROACH

Patil Tanmay Rameshwar
Dept. of Computer Engineering SKNSITS Lonavala,India
tp702832@gmail.com

Patil Onkar Nandkumar
Dept. of Computer Engineering SKNSITS Lonavala,India
opatil345@gmail.com

Tehare Sanket Sureshrao
Dept. of Computer Engineering SKNSITS Lonavala,India
tehare.sanket@gmail.com

Desale Pawan Somnath
Dept. of Computer Engineering SKNSITS Lonavala,India
pawan.s.desale@gmail.com

Abstract- The processes and dynamics by which information and behaviors spread through social networks have long interested scientists within many areas. Understanding such processes have the potential to shed light on the human social structure, and to impact the strategies used to promote behaviors or products. While the interest in the subject is long-standing, recent increased availability of social network and information diffusion data (through sites such as Face book, Twitter, and LinkedIn) has raised the prospect of applying social network analysis at a large scale to positive effect. One particular application that has been receiving interest in enterprises is to use word-of-mouth effects as a tool for viral marketing. Motivated by the marketing goal, mathematical formalizations of influence maximization have been proposed and extensively studied by many researchers. Influence maximization is the problem of selecting a small set of seed nodes in a social network, such that their overall influence on other nodes in the network, defined according to particular models of diffusion, is maximized

Keywords: *Social network, Influence maximization, Greedy algorithm, Malicious Attack, etc.*

I INTRODUCTION

As both social network structure and strength of influence between individuals evolve constantly, it requires to track the influential nodes under a dynamic setting. To address this problem, System explore the Influential Node Tracking (INT) problem as an extension to the traditional Influence Maximization problem (IM) under dynamic social networks. While Influence Maximization problem aims at identifying a set of k nodes to maximize the joint influence under one node or user. INT problem focuses on tracking a set of influential nodes that keeps maximizing the influence as the network evolves by means of posts or ads. Utilizing the smoothness of the evolution of the network structure, System propose an efficient algorithm, Upper Bound Interchange Greedy (UBI) and a variant, UBI+. Instead of constructing the seed set from the ground, System start from the influential seed set system find previously and implement node replacement to improve the influence coverage. Furthermore, the system also focuses on detection of posts for their positive or negative views by analyzing the usage history of that posts or ads or product Social network sites like Facebook, Twitter, and Google+ are experiencing incredible growth in users. There are more than a million users as of now. Besides just creating a profile and linking with friends, the social networks are now building platforms to run their website. These platforms are built based on the user profile details. These social applications are soon becoming an example of online communication which makes use of the users private information and activities in social links for various services. The Social networks are popular means of communication among the internet users. Online Social Networks (OSNs) witness a rise in user activity whenever an event takes place. Malicious entities exploit this spur in user-engagement levels to spread malicious content that compromises system reputation and degrades user experience and has recently been reported to face much abuse through scams and other type of malicious content, especially during news making events. People are heavily relaying on online interactions. The internet is giving different options to create and maintain contacts and relations for the user. With the introduction of social media network these options have become even easier to be used. Due to this heavy use of social media network a certain group of internet users called cybercriminal make use of this opportunity for threads.

Cybercriminals use different means to create spams fraud and other attacks on the users. Another means of attack by cybercriminals is the misuse of videos, images and links showed by the user. Attackers upload malicious posts in the season of special events and disasters. They will upload malicious posts which are related to these events and misguide users to click those links. Users who click the links by mistake act as an adversary to the attacker because the malicious posts would automatically re- posts the malicious contents such as links, images or videos on the user profile. Another popular version of this attack results in user profiles to ”like” a Facebook page without their knowledge. In some cases the, spammed posts will lead the users to survey sites which will result in cyber criminals getting profit. Finally, a real world solution in the form of a REST based API and a browser plug-in to identify malicious Facebook application and posts in real time i.e. FraAppe and MyPageKeepe. Once a Facebook user installs My-PageKeeper, it periodically crawls posts from the users wall and news feed. MyPageKeeper then applies URL blacklists as well as custom classification techniques to identify malicious posts. In other words, for every post that it crawls from the wall or news feed of a subscribed user, post does not take into account the application responsible for the post.From these, going to develop desktop application which directly blocks malicious post by analyzing predefined library and graph analysis concept. System going to add one feature of blocking of malicious links also, which was not provided by MyPageKeepe and FraApee. In this work a system of efficient categorization technique for identifying whether a post generated by a third party application is malicious or not. Detecting malicious URLs is now an essential task in network security intelligence. To maintain efficiency of web security, these malicious URLs have to be detected, identified as well as their corresponding links should be found out. Hence users get protected from it and effectiveness of network security gets increased. The malicious users can upload a content he wants to spread. The content that contains malicious data is posted to other users wall under a different form. The user mistakes the posts for a real content and clicks the post, which will take him to another page. Thus the malicious user can benefit from this process. In order to get the attention of the user, the malicious user will include keywords or description of pages that will be of interest to the user. These can be adult content or free downloading sites

II RELATED WORK

This section presents the prior works of the dynamic sensor networks. The author in [6] studied a tendency to advocate a recommendation support for active friending, wherever a user actively specifies a friending target. To the most effective of our data, a recommendation designed to supply steerage for a user to consistently approach his friending target has not been explored for existing on - line social networking services. to maximize the likelihood that the friending target would settle for a call for participation from the user, we have a tendency to formulate a replacement optimization downside, namely, Acceptance likelihood Maximization (APM), and develop a polynomial time rule, known as Selective invite with Tree and In - Node Aggregation (SITINA), to seek out the best resolution.

The author in [7] developed four malicious applications, and evaluated Andromaly ability to detect new malware based on samples of known malware. They evaluated several combinations of anomaly detection algorithms, feature selection method and the number of top features in order to find the combination that yields the best performance in detecting new malware on Android. Empirical results suggested that the proposed framework is effective in detecting malware on mobile devices in general and on Android in particular. The author in [8] tendency to study the economical influence maximization from 2 complementary directions. One is to enhance the first greedy formula and its improvement to more scale back its period of time, and also the second is to propose new degree discount heuristics that improves influence unfold.

The author in [9] discussed completely unique analysis work a couple of new economical approximation algorithmic program for influence maximization that was introduced to maximize the good thing about infectious agent promoting. For potency, we tend to devise 2 {ways| ways that ways in that} of exploiting the 2 - hop influence unfold which is that the influence unfold on nodes inside 2 - hops removed from nodes in a very seed set. Firstly, we tend to propose a brand new greedy methodology for the influence maximization draw back exploitation the 2 - hop influence unfold. Secondly, to hurry up the new greedy methodology, we tend to devise a good manner of removing uncalled - for nodes for influence maximization Based on optimum seed's native influence heuristics.

The author in [10] studied on minimizing the propagation of undesirable things, like pc viruses or malicious rumors, by block a restricted range of links in an exceedingly network, that is converse to the influence maximization downside during which the foremost potent nodes for data diffusion is searched in an exceedingly social network. This minimization downside is a lot of basic than the matter of preventing the unfold of contamination by removing nodes in an exceedingly network. In [11], study temporal patterns related to on - line content and the way the content's quality grows and fades over time. the eye that content receives on t he net varies betting on several factors and happens on terribly totally {different completely different} time scales and at different resolutions. so as to uncover the temporal dynamics of on - line content we tend to formulate a statistic bunch downside employing a similarity metric that's invariant to scaling and shifting. we tend to develop the K - Spectral Centroids (K - SC) bunch algorithmic program that effectively finds cluster Centroids with our similarity live. By applying associate adaptive wavelet - based progressive approach to bunch, we tend to scale K - SC to massive knowledge sets.

The author in [12] suggested info Propagation Game (IPG), a framework that may collect an outsized range of seed choosing ways for analysis. Through the IPG framework, human players aren't solely having fun however additionally serving to contributory the seed choosing ways. Preliminary experiment suggests that spatial relation primarily based heuristics square measure too straightforward for seed choice in a very multiple player surroundings. In [13], explored a unique downside, particularly cogent Node chase downside, as AN extension of Influence Maximization downside to dynamic networks, that aims at chase a group of cogent nodes dynamically such the influence unfold is maximized at any moment. we tend to propose AN economical formula UBI to resolve the INT downside based mostly plan of the Interchange Greedy methodology.

III PROPOSED WORK

This section presents the working model of our proposed model. The main objectives of the study are:

• To propose an efficient algorithm, Upper Bound Interchange Greedy (UBI), to tackle Influence Maximization problem under dynamic social network, which we term as Influential Node Tracking (INT) problem. • To track a set of influential nodes which maximize the influence under the social network at anytime. • To start from the seed set maximizing the influence under previous social network. Then we change the nodes in the existing set one by one in order to increase the influence under the current social network.

The proposed model composes of four phases, namely,

A) User

A user is a person who uses a computer or network service. Users generally use a system or a software product without the technical expertise required to fully understand it. Power users use advanced features of programs, though they are not necessarily capable of computer programming and system administration.

B) Admin

Administrator has the responsibility of ensuring that the administrative activities within an organization run efficiently, by providing structure to other employees throughout the organization. These activities can range from being responsible for the management of human resources, budgets and records, to undertaking the role of supervising other customer. These responsibilities can vary depending on the customer and level of education.

C) Influence maximization

The traditional Influence Maximization problem aims at finding influential nodes for only one static social network. However, real-world social networks are seldom static. Both the structure and also the influence strength associated with the edges change constantly. As a result, the seed set that maximizes the influence coverage should be constantly updated according to the evolution of the network structure and the influence strength.

D) Greedy Approach

Greedy Approach is designed to achieve optimum solution for a given problem. In greedy approach, decisions are made from the given solution domain. As being greedy, the closest solution that seems to provide an optimum solution is chosen. Greedy approach tries to find a localized optimum solution, which may eventually lead to globally optimized solutions. However, generally greedy algorithms do not provide globally optimized solutions.

The following are the merits achieved:
• Our algorithm achieves comparable results as hillclimbing greedy algorithm approximation is guaranteed. The algorithm has the time complexity of O (k n), and the space complexity of O(n), where n is the number of nodes and k is the size of the seed set. • To improves the computation of node replacement upper bound. • To evaluate the performance on large-scale real social network.

IV ALGORITHM

Greedy Algorithm The greedy approach[8] uses heuristic knowledge by selecting local optimum with the goal of achieving global maximum. This approach finds the solution for sub problems with a local maximum as a solution. The final solution can be obtained by combining all sub - solutions into overall solution called optimal solution. 1.A candidate set, from which a solution is created// Subproblem 2.A selection function, which chooses the best candidate to be added to the solution// finding local maximum 3.A feasibility function, that is used to determine if a candidate can be used to contribute to a solution 4.An objective function, which assigns a value to a solution, or a partial solution, and//finding global maximum 5.A solution function, which will indicate when we have discovered a complete solution//optimal solution Authors and Affiliations

V CONCLUSION

Influence maximization (IM), which selects a set of k users (called seeds) to maximize the influence spread over a social network, is a fundamental problem in a wide range of applications such as viral marketing and network monitoring. Existing IM solutions fail to consider the highly dynamic nature of social influence, which results in either poor seed qualities or long processing time when the network evolves. To address this problem, we define a novel IM query named Stream Influence Maximization (SIM) on social streams. Community based Greedy algorithm is used for mining top-K influential nodes. It has two components: dividing the mobile social network into several communities by taking into account information diffusion and selecting communities to find influential nodes by a dynamic programming. Location Based community Greedy algorithm is used to find the influence node based on Location and consider the influence propagation within Particular area. Experimental results have shown the effectiveness of the proposed model

REFERENCES

[1] Guojie Song et al, “Influential Node Tracking on Dynamic Social Network: An Interchange Greedy Approach”, EEE Transactions on Knowledge and Data Engineering, 2017.

[2] M. Kimura, K. Saito, and H. Motoda, “Blocking links to minimize contamination spread in a social network,” ACM Trans. Knowl.Discov. Data, vol. 3, no. 2, pp. 9:1–9:23, Apr. 2009.

[3] L. Fan, Z. Lu, W. Wu, B. Thuraisingham, H. Ma, and Y. Bi, “Least cost rumor blocking in social networks,” in Proc. IEEE ICDCS’13,Philadelphia, PA, Jul. 2013, pp. 540–549.

[4] J. Yang and J. Leskovec, “Patterns of temporal variation in online media,” in Proc. ACM Int. Conf. Web Search Data Minig, 2011, pp. 177–186.

[5] R. Crane and D. Sornette, “Robust dynamic classes revealed by measuring the response function of a social system,” in Proc. of the Natl. Acad. Sci. USA, vol. 105, no. 41, Apr. 2008, pp. 15 649–15 653.

[6] S. Han, F. Zhuang, Q. He, Z. Shi, and X. Ao, “Energy model for rumor propagation on social networks,” Physica A: Statistical Mechanics and its Applications, vol. 394, pp. 99–109, Jan. 2014.

[7] D. Chelkak and S. Smirnov, “Universality in the 2d ising model and conformal invariance of fermionic observables,” Inventiones Mathmaticae, vol. 189, pp. 515– 580, Sep. 2012.

[8] W. Chen, C.Wang, and Y.Wang, “Scalable influence maximization for prevalent viral marketing in large-scale social networks,” in Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2010, pp. 1029–1038.

[9] S. Shirazipourazad, B. Bogard, H. Vachhani, and A. Sen, “Influence propagation in adversarial setting: How to defeat competition with least amount of investment,” in Proc. 21st ACM Int. Conf. Inf. Knowl. Manag., 2012, pp. 585–594.

[10]E. Serrano, C. A. Iglesias, and M. Garijo, “A novel agent based rumor spreading model in twitter,” in Proc. 24th Int. Conf. World Wide Web, 2015, pp. 811–814.

[11] A. Badanidiyuru, B. Mirzasoleiman, A. Karbasi, and A. Krause. Streaming submodular maximization: Massive data summarization on the fly. In KDD, pages 671–680, 2014.

[12] N. Barbieri, F. Bonchi, and G. Manco. Topic-aware social influence propagation models. In ICDM, pages 81– 90, 2012.

[13] S. Bharathi, D. Kempe, and M. Salek. Competitive influence maximization in social networks. In WINE, pages 306–311, 2007