IJCSC 0973-7391
International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 9 • Issue 2 pp. 73-80 March 2018 - Sept 2018 www.csjournals.com

Effectiveness in significant Node Tracking on Dynamic Social Network: An Interchange Greedy Approach

Md Shahid Raza[1], Md Ateeq Ur Rahman [2]
[1]Research Scholar, Dept. of Computer Science & Engineering, SCET, Hyderabad
[2]Professor and Head, Dept. of Computer Science & Engineering, SCET, Hyderabad

Abstract: With the success of on-line social networks and microblogs such as Facebook, Flickr and Twitter, the development of influence exerted by users of such platforms on alternative users, and how it propagates within the network, has recently attracted the interest of pc scientists, info technologists, and marketing specialists. one amongst the key issues during this space is the identification of powerful users, by targeting whom sure desirable selling outcomes are often achieved. As each social network structure and strength of influence between people evolve perpetually, it needs to trace the cogent nodes beneath a dynamic setting. to handle this drawback, we tend to explore the cogent Node trailing (INT) drawback as associate degree extension to the normal Influence Maximization drawback (IM) beneath dynamic social networks. whereas Influence Maximization drawback aims at distinguishing a collection of k nodes to maximise the joint influence beneath one static network, INT drawback focuses on trailing a collection of cogent nodes that keeps increasing the influence because the network evolves. Utilizing the smoothness of the evolution of the network structure, we tend to propose associate degree economical algorithmic program, edge Interchange Greedy (UBI) and a variant, UBI+. rather than constructing the seed set from the bottom, we tend to begin from the cogent seed set we discover antecedently and implement node replacement to enhance the influence coverage. what is more, by employing a quick update methodology by calculative the marginal gain of nodes, our algorithmic program will scale to dynamic social networks with uncountable nodes. Empirical experiments on 3 real large-scale dynamic social networks show that our UBI and its variants, UBI+ achieves higher performance in terms of each influence coverage and period of time. during this article we tend to take a knowledge mining perspective and that we discuss what (and how) can be learned from the out there traces of past propagations. While doing this we offer a quick summary of some recent progresses during this space and discuss some open issues. By no suggests that this text should be supposed as associate complete survey: it's instead (admittedly) a rather biased and private perspective of the author on the subject of influence propagation in social networks.

Index Terms—Influence maximization, trajectory databases, location-aware advertising. Social Networks, Social Influence, Viral Marketing.

I. INTRODUCTION

Traditionally, social network models are descriptive, instead of predictive: they're designed at a really coarse level, generally with solely a number of world parameters, and don't seem to be helpful for creating actual predictions of the longer term behavior of the network. within the past, this was mostly because of lack of data: the networks on the market for experimental study were tiny and few, and contained solely smallest data concerning every node. as luck would have it, the increase of the net has modified this dramatically. large quantities of information on terribly massive social networks ar currently on the market from blogs, knowledge-sharing sites, cooperative filtering systems, on-line play, social networking sites, newsgroups, chat rooms, etc. These networks generally range within the tens of thousands to several nodes, and sometimes contain substantial quantities of data at the extent of individual nodes, spare to create models of these people. collecting these models into models of the larger network they're a part of provides North American nation associate unprecedented level of detail in social network analysis, with the corresponding potential for brand new understanding, helpful predictions, and their productive use in decision-making. we've begun to create social network models at this scale, exploitation information from the Epinions knowledge-sharing web site, the EachMovie cooperative filtering system, and others. The processes and dynamics by that data and behaviors unfold through social networks have long interested scientists inside several areas. Understanding such processes have the potential to shed light-weight on the human social system, and to impact the ways accustomed promote behaviors or merchandise. whereas the interest within the subject is long-standing, recent accumulated handiness of social network and knowledge} diffusion data (through sites like Facebook, Twitter, and LinkedIn) has raised the prospect of applying social network analysis at an oversized scale to positive result. One explicit application that has been receiving interest in enterprises is to use spoken effects as a tool for infective agent promoting. motivated by the promoting goal, mathematical formalizations of influence maximization are planned and extensively studied by several researchers . Influence maximization is that the downside of choosing atiny low set of seed nodes during a social network, specified their overall influence on different nodes within the network, outlined consistent with explicit models of diffusion, is maximized. promoting campaign is typically not a one-time deal, instead enterprises perform a sustaining campaign to market their merchandise by seeding influential nodes incessantly. Often, a promoting campaign could last for months or years, wherever the corporate sporadically allocates budgetsto the chosen influential users to utilize the ability of the spoken result. underneath this example, it's natural and necessary to appreciate that social or data networks ar invariably dynamics, and their topology evolves perpetually over time . for instance, links seem and disappear once users follow/unfollow others in Twitter or friend/unfriend others in Facebook. Moreover, the strength of influence additionally keeps ever-changing, as you're additional influenced by your friends WHO you contact ofttimes, whereas the influence from a follower typically dies down as time elapses if you are doing not contact with one another. As a result, a collection of nodes influential at just once could result in poor influence coverage when the evolution of social network, that suggests that victimisation one static set as seeds across time may lead to disappointing performance. It seems that targeting at completely different|completely different} nodes at different time becomes essential for the success of infective agent promoting. we have a tendency to proceed associate degree example|as an instance|parenthetically|let's say|maybe} the thought of considering the dynamic perspect in influence maximization victimisation an example in Figure one. during this example, users ar connected by edges at completely different time, every of that indicates a user could influence over another user. Numbers over every edge provide the corresponding influencing possibilities. for instance, there's a position between v1 and v3 at t = zero and also the edge is deleted at t = one. And user v1 can influence v2 with a likelihood of zero:7 at t = 0, and also the influencing likelihood is 0:2 at t = one. this implies that user v1 would not influence v3 at t = one and v2 can't be activated by v1 by likelihood zero.7 at t = 1. Suppose we have a tendency to ar asked to search out one seed user to maximise the expected variety of influenced users. with none dynamic constraint, that's all the snapshots ar aggregative into one weighted static graph, user v1 are came because the result. Intuitively, it's expected to influence the largest variety of users among all users. However, if we have a tendency to aim to search out one seed user that influences the largest variety of users at completely different time, user v2 can become the new result at time t = one. Intuitively, this {can be} as a result of v1 can at the most influence v4 at t = one whereas v2 influences v1, v3 and v4 with the next likelihood. However, traditional algorithms for Influence Maximization become inefficient underneath this example as they fail to contemplate the association between social networks at completely different time and have to be compelled to solve several Influence Maximization issues severally for social network at every time. during this paper, we have a tendency to propose associate economical formula, edge Interchange Greedy (UBI), to tackle Influence Maximization downside underneath dynamic social network, that we have a tendency to term as influential Node chase (INT) downside. That is to trace a collection of important nodes that maximize the influence underneath the social network at any time. the most plan of our UBI algorithmic rule is to leverage the similarity of social networks close to in time associate degreed directly discover the important nodes supported the seed set found for previous social network rather than constructing the answer from an empty set. As similarity in network structure results in similar set of nodes that maximize the influence. In our UBI algorithmic rule, we tend to begin from the seed set increasing the influence underneath previous social network. Then we modify the nodes within the existing set one by one so as to extend the influence underneath the present social network. because the optimum seed set differs solely in a very little range of nodes, some rounds of node exchanges ar enough to get a seed set with massive joint influence underneath current social network. Moreover, it are often shown that the on top of node exchange procedure results in a relentless approximation guarantee of 1=2, once bound stopping criteria is applied to node exchanges. Our technique needs an outsized range of computations in evaluating the node substitution gain, that takes unaffordable long-standing if ancient Monte-Carlo simulations ar applied. so as to scale our algorithmic rule up to massive networks, we tend to utilize the boundary primarily based Approach planned by Chou et al. to scale back the calls of Monte-Carlo simulations [13]. we tend to initial tighten their sure by excluding all the influence ways with edges into the seed set associate degreed most vital we tend to style an economical technique to update the boundary because the underlying network structure changes rather than concluding overpriced mathematical process for every individual network, because the result, we tend to propose UBI and its variant, UBI+. intensive experiments ar conducted on 3 real dynamic networks of various sorts and scales. The comparison of our technique to many state-of-arts Influence Maximization algorithms for static network shows that our strategies results in each larger influence coverage and fewer period. we tend to show that our UBI algorithmic rule achieves comparable influence coverage as Greedy algorithmic rule inside solely seconds for networks with lots of nodes across multiple snapshots. Also, the variant algorithmic rule, UBI+ ar conducted on constant networks and show higher performance than UBI. Our contributions are often summarized as follows: nine we tend to explore the important Node trailing (INT) downside as associate degree extension to the normal Influence Maximization downside to maximise the influence coverage underneath a dynamic social network. nine we tend to propose associate degree economical algorithmic rule, boundary Interchange (UBI) to unravel the INT downside. Our algorithmic rule achieves comparable results as hill-climbing greedy algorithmic rule wherever the one 1=e approximation is warranted. The algorithmic rule has the time quality of O(kn), and also the area quality of O(n), wherever n is that the range of nodes and k is that the size of the seed set. nine we tend to propose associate degree algorithmic rule UBI+, supported UBI, that improves the computation of node replacement boundary. nine we tend to judge the performance on large-scale real social network. The experiment results make sure our theoretical findings and show that our UBI and UBI+ algorithmic rule bring home the bacon higher performance of each influence coverage and period. we tend to summarize the connected literatures in section two. In section three, we tend to formally formulate our important Node trailing downside when introducing the diffusion model and also the Influence Maximization downside. we tend to then gift our economical UBI algorithmic rule and its variant, UBI+ algorithmic rule for the INT downside in section four. In section five, we tend to gift our experiment results on 3 real-world large scale dynamic social networks. These models permit North American nation to style "viral marketing" plans that maximize positive viva-voce among customers. In our experiments, this makes it doable to realize abundant higher profits than if we have a tendency to ignore interactions among customers and therefore the corresponding network effects, as ancient selling will. The Network price clients of consumers of shoppers} client price is sometimes outlined because the expected benefit from sales thereto customer, over the life of the link between the client and therefore the company. client price is of crucial interest to corporations, as a result of it determines what quantity it's price outlay to accumulate a specific client. However, ancient measures of client price ignore the very fact that, additionally to purchasing merchandise himself, a client might influence others to shop for them. for instance, if, additionally to seeing a specific flick myself, I persuade 3 friends to examine it with Pine Tree State, my client price with regard to that flick has effectively quadrupled, and therefore the flick studio is so even in outlay a lot of on selling the flick to Pine Tree State than it otherwise would. Conversely, if I tend to form choices on what movies to examine strictly supported what my friends tell Pine Tree State, selling to Pine Tree State could also be a waste of resources, which might be higher spent selling to my friends. we have a tendency to decision the network price of a client the expected increase in sales to others that results from selling thereto client. Clearly, ignoring the network price of shoppers, as is finished in ancient marketing, might cause terribly suboptimal selling choices. But, whereas the existence of network effects has been acknowledged within the selling literature, they need usually been thought-about to be unquantifiable, one notably at the extent of individual customers. this can be what's modified by the information sources currently on the market. Our models change North American nation to live the network price of a client. for every client, we have a tendency to model however probable that client is to shop for some product, as a perform of each the intrinsic properties of the client and therefore the product, and of the influence of the customer's neighbors within the network. By activity probabilistic logical thinking over the joint model of all the purchasers, we will answer queries like "If we have a tendency to market to the current specific set of shoppers, what's the expected benefit from the total network, once the influence of these customers has propagated throughout?" exploitation this capability, we will currently rummage around for the best set of shoppers to promote to, within the sense that selling to the current set can yield the very best come back on investment.

II. Related Works

2.1 Existing System

The processes and dynamics by that data and behaviors unfold through social networks have long interested scientists inside several areas. Understanding such processes have the potential to shed lightweight on the human social organization, and to impact the methods wont to promote behaviors or merchandise. whereas the interest within the subject is long-standing, recent increased accessibility of social network information} diffusion data (through sites like Facebook, Twitter, and LinkedIn) has raised the prospect of applying social network analysis at an oversized scale to positive impact. One explicit application that has been receiving interest in enterprises is to use spoken effects as a tool for microorganism

International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 9 • Issue 2 pp. 73-80 March 2018 - Sept 2018 www.csjournals.com

selling. actuated by the selling goal, mathematical formalizations of influence maximization are planned and extensively studied by several researchers. Influence maximization is that the drawback of choosing alittle set of seed nodes in an exceedingly social network, such their overall influence on different nodes within the network, outlined consistent with explicit models of diffusion, is maximized.

III. PROPOSED SYSTEM

For real dynamic social network, it's unlikely to possess abrupt and forceful changes in graph structure during a short amount of your time. As a result, the similarity in structure of graphs from 2 consecutive snapshots may lead to similar seed sets that maximize the influence underneath every graph. supported the on top of plan, we tend to propose UBI formula for the INT drawback, during which we discover the seed set that maximizes the influence underneath Gt+1 supported the seed set St we've already found for graph Gt. rather than constructing the seed set for graph Gt+1 from the bottom, we tend to begin with St and regularly update by substitution the nodes in St to boost the influence coverage. Our formula 1st uses Associate in Nursing initial set and a number of other rounds of interchange heuristic to maximise the influence, as mentioned within the paper. therefore the interchange heuristic clearly works on a photograph graph. once extended to the dynamic graph, our formula solely must interchange for many a lot of rounds once whenever window and may attain a quicker update. a lot of elaborated descriptions regarding however our technique works on the photograph graphs and dynamic networks are going to be given within the next 2 subsections.

IV. System Architecture

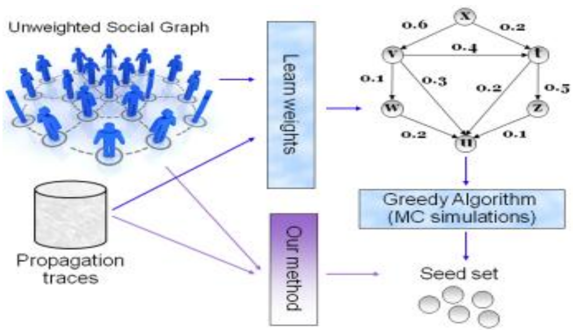

Figure 1: Algorithm Implementation

A greedy algorithmic program, because the name suggests, forever makes the selection that looks to be the simplest at that moment. this implies that it makes a locally-optimal alternative within the hope that this alternative can cause a globally-optimal resolution.

Assume that you simply have Associate in Nursing objective operate that has to be optimized (either maximized or minimized) at a given purpose. A Greedy algorithmic program makes greedy decisions at every step to make sure that the target operate is optimized. The Greedy algorithmic program has only 1 shot to cypher the optimum resolution in order that it ne'er goes back and reverses the choice.

Greedy algorithms have some benefits and disadvantages:

1. it's quite simple to return up with a greedy algorithmic program (or even multiple greedy algorithms) for a haul.
2. Analyzing the run time for greedy algorithms can typically be abundant easier than for alternative techniques (like Divide and conquer). For the Divide and conquer technique,

it's not clear whether or not the technique is quick or slow. this is often as a result of at every level of rule the scale of gets smaller and therefore the variety of sub-problems will increase.

3. The troublesome half is that for greedy algorithms you have got to figure abundant tougher to grasp correctness problems. Even with the proper algorithmic program, it's laborious to prove why it's correct. Proving that a greedy algorithmic program is correct is additional of Associate in Nursing art than a science. It involves plenty of creative thinking.

3.1 Module Description:

In this project, the Modules are:

1.Influence Maximization Module
2. Influential Node Tracking Module
3.Upper bounds comparison Module
4. Upper Bound of Node Replacement Gain Module

Influence Maximization Module :
Marketing campaign is usually not a one-time deal, instead enterprises carry out a sustaining campaign to promote their products by seeding influential nodes continuously. Often, a marketing campaign may last for months or years, where the company periodically allocates budgets to the selected influential users to utilize the power of the word-of-mouth effect. Under this situation, it is natural and important to realize that social or information networks are always dynamics, and their topology evolves constantly over time. For example, links appear and disappear when users follow/unfollow others in Twitter or friend/unfriend others in Facebook. Moreover, the strength of influence also keeps changing, as you are more influenced by your friends who you contact frequently, while the influence from a friend usually dies down as time elapses if you do not contact with each other. As a result, a set of nodes influential at one time may lead to poor influence coverage after the evolution of social network, which suggests that using one static set as seeds across time could lead to unsatisfactory performance.

Influential Node Tracking Module:

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. In this work, we model the dynamic social network as a series of snapshot graphs, G1,. . . , GT . We assume that the nodes remain the same while the edges in each snapshot graph change across different time intervals. Each snapshot graph is modeled as a directed network, Gt = (V;Et), which includes edges appearing during the periods under consideration. Moreover, a set of propagation probabilities Pt u;v is associated with each snapshot graph Gt. Our goal is to track a series of seed sets, denoted as St; t = 1; : : : ; T, that maximizes the influence function t() at each of the snapshot Gt.

Upper bounds comparison Module :
Upper bound termed as active nodes' path excluded upper bound (AB), is theoretically tighter than the upper bound proposed , which we call it the naive upper bound (NB). In order to

validate our theory, we run empirical experiments to compare our bound AB with the naive upper bound. We first extract a series of snapshot graphs from Mobile datasets by setting both time window and time difference to one hour. We run equivalent number of iterations in computing both AB and NB on the same node set with size k = 30 where propagation probabilities are set according to DWA model. The seed set is selected by Greedy algorithm that maximizes the influence under each snapshot. As is shown in Figure 9, our bound is consistently tighter than the naive bound proposed in as suggested by our theory. It should be noticed that the poor performance of NB under DWA model is due to the fact that sometimes NB fails to converge in Mobile network.

Upper Bound of Node Replacement Gain Module:

In this section, we illustrate the only mysterious part in our UBI algorithm, namely the computation of the upper bound of the replacement gain u;vs (S). Zhou et al. first use the upper bound on influence function to accelerate the greedy algorithm in influential seeds selection . we propose a tighter upper bound on the replacement gain by excluding the influence along paths, which include incoming edges to the seed set. We have shown previously how to compute a tighter bound on the replacement gain for one static network with a fixed seed set S. However, as network changes constantly, we need to update the upper bound according to the changes in propagation probability. Moreover, as we include new node into the seed set S, we also need to update the upper bound as the propagation probability matrix PG(S+T) also changes.

V. Conclusion

We explore a unique drawback, particularly authoritative Node following drawback, as Associate in Nursing extension of Influence Maximization drawback to dynamic networks, that aims at following a group of authoritative nodes dynamically specified the influence unfold is maximized at any moment. we have a tendency to propose Associate in Nursing economical algorithmic program UBI to unravel the INT drawback based mostly plan of the Interchange Greedy methodology.We utilize the boundary on node replacement gain to accelerate the method. Moreover, Associate in Nursing economical methodology for change the boundary is projected to handle the evolution of the network structure. in depth experiments on 3 real social networks show that our methodology outperforms progressive baselines in terms of each influence coverage and period. Then we have a tendency to propose UBI+ algorithmic program that improves the computation of the boundary and achieves higher influence unfold. we'd wish to generalize our UBI algorithmic program to trace authoritative nodes beneath the opposite wide adopted diffusion model, Linear Threshold model beneath dynamic networks. Moreover, it'll be fascinating if we are able to mix our work with . that's to trace a series of authoritative nodes wherever the diffusion method is additionally meted out beneath a dynamic network rather than the static exposure graph.

References

[1] F. Bonchi, “Influence propagation in social networks: A data mining perspective,” in WI-IAT, 2011.

[2] A. Goyal, F. Bonchi, and L. V. S. Lakshmanan, "A data-based approach to social influence maximization," Proc. VLDB Endow., 2011.
[3] Q. Jiang, G. Song, G. Cong, Y. Wang, W. Si, and K. Xie, "Simulated annealing based influence maximization in social networks," in AAAI, 2011.
[4] P. Domingos and M. Richardson, "Mining the network value of customers," in KDD, 2001.
[5] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in KDD, 2003.
[6] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in KDD, 2002.
[7] M. Kimura and K. Saito, "Tractable models for information diffusion in social networks," in PKDD, 2006.
[8] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. M. VanBriesen, and N. S. Glance, "Cost-effective outbreak detection in networks," in KDD, 2007.
[9] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in KDD, 2009.
[10] W. Chen, C.Wang, and Y.Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," in KDD, 2010.
[11] J. Coleman, H. Menzel, and E. Katz, Medical Innovations: A Diffusion Study. Bobbs Merrill, 1966.
[12] T. Valente, Network Models of the Diffusion of Innovations. Hampton Press, 1955.