International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 24th Feb 2018. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-02

Title: **AN EFFICIENT GREEDY APPROACH FOR NODE TRACKING IN DYNAMIC OSN'S**

Volume 08, Issue 02, Pages: 112–117.

Paper Authors

MS.V.HEMA PRIYA, K.SIREESHA

Vignan's Lara Institute of Technology & Science

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

AN EFFICIENT GREEDY APPROACH FOR NODE TRACKING IN DYNAMIC OSN'S

MS.V.HEMA PRIYA[1], K.SIREESHA[2]
Assistant Professor[1], Department of M.C.A ,Vignan's Lara Institute of Technology & Science
M.C.A Student[2],Department of M.C.A ,Vignan's Lara Institute of Technology & Science

Abstract:
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, we 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 static network, INT problem focuses on tracking a set of influential nodes that keeps maximizing the influence as the network evolves. Utilizing the smoothness of the evolution of the network structure, we propose an efficient algorithm, Upper Bound Interchange Greedy (UBI) and a variant, UBI+. Instead of constructing the seed set from the ground, we start from the influential seed set we find previously and implement node replacement to improve the influence coverage. Furthermore, by using a fast update method by calculating the marginal gain of nodes, our algorithm can scale to dynamic social networks with millions of nodes. Empirical experiments on three real large-scale dynamic social networks show that our UBI and its variants, UBI+ achieves better performance in terms of both influence coverage and running time.

Algorithm Implementation

A greedy algorithm, as the name suggests, **always makes the choice that seems to be the best at that moment**. This means that it makes a locally-optimal choice in the hope that this choice will lead to a globally-optimal solution.Assume that you have an objective function that needs to be optimized (either maximized or minimized) at a given point. A Greedy algorithm makes greedy choices at each step to ensure that the objective function is optimized. The Greedy algorithm has only one shot to compute the optimal solution so that it never goes back and reverses the decision.Greedy algorithms have some advantages and disadvantages:

1. It is quite easy to come up with a greedy algorithm (or even multiple greedy algorithms) for a problem.
2. Analyzing the run time for greedy algorithms will generally be much easier than for other techniques (like Divide and conquer). For the Divide and conquer technique, it is not clear whether the technique is fast or slow. This is because at each level of recursion the size of gets smaller and

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

the number of sub-problems increases.

3. The difficult part is that for greedy algorithms you have to work much harder to understand correctness issues. Even with the correct algorithm, it is hard to prove why it is correct. Proving that a greedy algorithm is correct is more of an art than a science. It involves a lot of creativity.

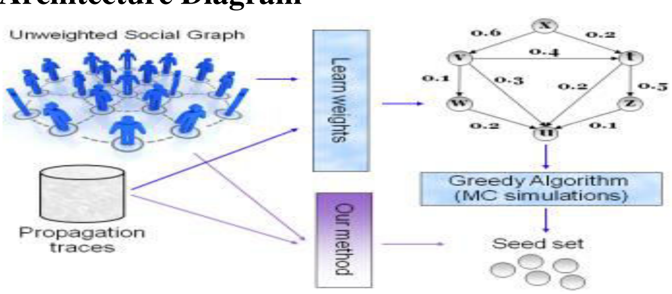
Existing system:

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 Facebook, 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.

Proposed system:

For real dynamic social network, it is unlikely to have abrupt and drastic changes in graph structure in a short period of time. As a result, the similarity in structure of graphs from two consecutive snapshots could lead to similar seed sets that maximize the influence under each graph. Based on the above idea, we propose UBI algorithm for the INT problem, in which we find the seed set that maximizes the influence under Gt+1 based on the seed set St we have already found for graph Gt. Instead of constructing the seed set for graph Gt+1 from the ground, we start with St and continually update by replacing the nodes in St to improve the influence coverage. Our algorithm first uses an initial set and several rounds of interchange heuristic to maximize the influence, as mentioned in the paper. So the interchange heuristic obviously works on a snapshot graph. When extended to the dynamic graph, our algorithm only needs to interchange for a few more rounds after each time window and can achieve a faster update. More detailed descriptions about how our method works on the snapshot graphs and dynamic networks will be presented in the next two subsections.

Architecture Diagram



Modules:

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

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

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.

Conclusion:

We explore a novel problem, namely Influential Node Tracking problem, as an extension of Influence Maximization problem to dynamic networks, which aims at tracking a set of influential nodes dynamically such that the influence spread is maximized at any moment. We propose an efficient algorithm UBI to solve the INT problem based idea of the Interchange Greedy method.We utilize the upper bound on node replacement gain to accelerate the process. Moreover, an efficient method for updating the upper bound is proposed to handle the evolution of the network structure. Extensive experiments on three real social networks show that our method outperforms state-of-the-art baselines in terms of both influence coverage and running time. Then we propose UBI+ algorithm that improves the computation of the upper bound and achieves better influence spread.

References

[1] K. Young, M.Pistner, J.O'Mara, andJ.Buchanan.Cyber-disorders: The mental health concern for the new millennium. Cyberpsychol. Behav., 1999.

[2] J. Block. Issues of DSM-V: internet addiction. American Journal of Psychiatry, 2008.

[3] K. Young. Internet addiction: the emergence of a new clinical disorder, Cyberpsychol. Behav., 1998.

[4] I.-H. Lin, C.-H. Ko, Y.-P. Chang, T.-L. Liu, P.-W. Wang, H.-C. Lin, M.-F. Huang, Y.-C. Yeh, W.-J. Chou, and C.-F. Yen. The association between suicidality and Internet addiction and activities in Taiwanese adolescents. Compr. Psychiat., 2014.

[5] Y. Baek, Y. Bae, and H. Jang. Social and parasocial relationships on social network sites and their differential relationships with users' psychological well-being. Cyberpsychol. Behav. Soc. Netw., 2013.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

[6] D. La Barbera, F. La Paglia, and R. Valsavoia. Social network and addiction. Cyberpsychol. Behav., 2009.
[7] K. Chak and L. Leung. Shyness and locus of control as predictors of internet addiction and internet use. Cyberpsychol. Behav., 2004.
[8] K. Caballero and R. Akella. Dynamically modeling patients health statefromelectronicmedicalrecords:atimeseriesapproach.KDD, 2016.
[9] L. Zhao and J. Ye and F. Chen and C.-T. Lu and N. Ramakrishnan. Hierarchical Incomplete multi-source feature learning for Spatiotemporal Event Forecasting. KDD, 2016.
[10] E. Baumer, P. Adams, V. Khovanskaya, T. Liao, M. Smith, V. Sosik, and K. Williams. Limiting, leaving, and (re)lapsing: an exploration of Facebook non-use practices and experiences. CHI, 2013.
[11] R. Jain and N. Abouzakhar. A comparative study of hidden markov model and support vector machine in anomaly intrusion detection. JITST, 2013.
[12] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li. User-level sentiment analysis incorporating social networks. KDD, 2011.
[13] R. Collobert, F. Sinz, J. Weston, and L. Bottou. Large scale transductive svms. JMLR, 2006.
[14] L. Leung. Net-generation attributes and seductive properties of the internet as predictors of online activities and internet addiction. Cyberpsychol. Behav. Soc. Netw., 2004.
[15] J. Cacioppo, J. Fowler, and N. Christakis. Alone in the crowd: the structure and spread ofloneliness in a large social network. J. Pers. Soc. Psychol., 2009.
[16] J. Kleinberg. Bursty and hierarchical structure in streams. KDD, 2002. [17] K.-L. Liu, W.-J. Li, and M. Guo. Emoticon smoothed language models for twitter sentiment analysis. AAAI, 2012.
[18] C. Andreassen, T. Torsheim, G. Brunborg, and S. Pallesen. Development of a Facebook addiction scale. Psychol. Rep., 2012.
[19] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi. On the evolution of user interaction in Facebook. WOSN, 2009.
[20] E. Ferrara, R. Interdonato, and A. Tagarelli. Online popularity and topical interests through the lens of Instagram. HT, 2014.
[21] M. Saar-Tsechansky and F. Provost. Handling missing values when applying classification models. JMLR, 2007.
[22] I. Witten and E. Frank. Data mining: practical machine learning tools and techniques with Java implementations. MorganKaufmann, San Francisco, 2000.
[23] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines, 2001.
[24] F. Chang, C.-Y. Guo, X.-R. Lin, and C.-J. Lu. Tree decomposition for large-scale SVM problems. JLMR, 2010.
[25] P. Comon, X. Luciani, and A. L. D. Almeida. Tensor decompositions, alternating least squares and other tales. Journal of Chemometrics, 2009.
[26] E. Acar, D. M. Dunlavy, and T. G. Kolda. A scalable optimization approach for fitting canonical tensor decompositions.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Journal of Chemometrics, 2011.
[27] L.He,C.-T.Lu,J.Ma,J.Cao,L.Shen,andP.S.Yu.Joint community and structural hole spanner detection via harmonic modularity. KDD, 2016.
[28] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummad. Measuring userinfluenceon twitter:Themillion follower fallacy. ICWSM, 2010.
[29] K. Hayashi, T. Maehara, M. Toyoda, and K. Kawarabayashi. Realtime top-r topic detection on twitter with topic hijack filtering. KDD, 2015.
[30] J. Gill and G. King. What to do when your Hessian is not invertible: Alternatives to model respecification in nonlinear estimation. Sociological Methods and Research, 2004.
[31] B. K˚agstr¨om and P. Poromaa. Distributed and shared memory block algorithms for the triangular Sylvester equation with sep−1 estimators. SIAM Journal on Matrix Analysis and Applications, 1992.
[32] L. R. Tucker. Some mathematical notes on three-mode factor analysis. Psychometrika, 1966.