Scientific Journal of Impact Factor (SJIF): 4.14
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 3, Issue 4, April -2016

Formulation Technique to Influence Maximization Problem as Query Processing to Distinguish Users

Piyush Kala[1], *Bhavesh Parbate*[2], *Rahul Kamble*[3], *Navneet Yadav*[4], Prof. Shubhra Mukherjee[5]

[1,2,3,4,5] *Department of Computer Engineering, Parvatibai G.Moze College of Engg. Wagholi, Pune,*

Abstract *— Impact extension is acquainted with augment the benefit of viral promoting in informal organizations. The shortcoming of impact expansion is that it doesn't recognize particular clients from others, regardless of the possibility that a few things can be helpful for the particular clients. For such things, it is a superior system to concentrate on boosting the impact on the particular clients. In this paper, we detail an impact boost issue as question handling to recognize particular clients from others. We demonstrate that the question handling issue is NP-hard and its target capacity is sub secluded. We propose a desire model for the estimation of the target capacity and a quick covetous based close estimation strategy utilizing the desire model. For the desire model, we explore a relationship of ways between clients. For the covetous technique, we work out a productive incremental overhauling of the negligible addition to our goal capacity. We lead trials to assess the proposed technique with genuine datasets, and contrast the outcomes and those of existing systems that are adjusted to the issue. From our trial results, the proposed strategy is no less than a request of extent speedier than the existing routines by and large while accomplishing high exactness.*

Keywords- *Graph algorithms, influence maximization, independent cascade model, social networks*

I. INTRODUCTION

The measure of spread of data is relentlessly expanded in online interpersonal organizations, for example, Facebook and Twitter. To utilize online interpersonal organizations as a promoting stage, there is loads of examination on the best way to utilize the proliferation of impact for viral advertising. One of the exploration issues is influence maximization (IMAX), which plans to discover k seed clients to amplify the spread of impact among clients in interpersonal organizations. It is ended up being a NP-hard issue by Kempe et al. Since they proposed an avaricious calculation for the issue, numerous analysts have proposed different heuristic routines.

Viral showcasing is one of the key utilizations of impact boost. In viral advertising, a thing that an advertiser needs to advance is diffused into informal communities "by overhearing people's conversations" correspondence. From the point of view of advertising, impact augmentation gives how to get the most extreme benefit from every one of the clients in an informal organization through viral showcasing. In any case, impact amplification is not generally the best technique for viral showcasing, on the grounds that there can be a few things that are helpful to just particular clients. These particular clients can be a couple individuals with a typical enthusiasm for a given thing, some or all individuals in a group, or some or all clients in a class. There is no restriction for being particular clients. For instance, consider an advertiser that is approached to advance a restorative item for ladies through viral showcasing. For the corrective item, the particular clients are female clients why should likely utilize it and male clients who wish to buy it as a present for female clients. For this situation, the advertiser does not should be worried about alternate clients in light of the fact that the restorative item is not helpful to them. Rather, it is a superior method to concentrate on augmenting the quantity of impacted particular clients, yet impact amplification has the shortcoming that it can't recognize them from alternate clients. The main method for taking care of such focuses with impact boost is making a homogeneous diagram with the objectives and executing impact expansion on the chart. On the other hand, the aftereffect of this methodology ought to be off base; on the grounds that there can be a few clients who are not targets but rather can unequivocally impact the objectives.

II. LITERATURE REVIEW

1.Maximizing the spread of influence through a social network

AUTHORS: D. Kempe, J. Kleinberg, and E. Tardos

This paper considers this problem in several of the most widely studied models in social network analysis. The optimization problem of selecting the most influential nodes is NP-hard here, and we provide the first provable approximation guarantees for efficient algorithms. Using an analysis framework based on sub modular functions, they show that a natural greedy strategy obtains a solution that is probably within 63% of optimal for several classes of models; our framework suggests a general approach for reasoning about the performance guarantees of algorithms for these types of influence problems in social networks. We also provide computational experiments on large collaboration networks, showing that in addition to their provable guarantees, our approximation algorithms significantly out-perform node-selection heuristics based on the well-studied notions of degree centrality and distance centrality from the field of social networks.

2.Labeled influence maximization in social networks for target marketing

AUTHORS: F.-H. Li, C.-T.Li, and M.-K. Shan

In this paper, focus on the target marketing. System propose the labeled influence maximization problem, which aims to find a set of seed nodes which can trigger the maximum spread of influence on the target customers in a labeled social network. They propose three algorithms to solve such labeled influence maximization problem. System first develops the algorithms based on the greedy methods of original influence maximization by considering the target customers. Moreover, system develop a novel algorithm, Maximum Coverage, whose central idea is to offline compute the pair wise proximities of nodes in the labeled social network and online find the set of seed nodes. This allows the marketers to plan and evaluate strategies online for advertised products. The experimental results on IMDb labeled social network show our methods can achieve promising performances on both effectiveness and efficiency.

3. Profit maximization over social networks

AUTHORS: W. Lu and L. Lakshmanan

In this work, distinguish between influence and adoption by explicitly modeling the states of being influenced and of adopting a product. They extend the classical Linear Threshold (LT) model to incorporate prices and valuations, and factor them into users' decision-making process of adopting a product. We show that the expected profit function under our proposed model maintains sub modularity under certain conditions, but no longer exhibits monotonicity, unlike the expected influence spread function. To maximize the expected profit under our extended LT model, we employ an unbudgeted greedy framework to propose three profit maximization algorithms. The results of our detailed experimental study on three real-world datasets demonstrate that of the three algorithms, PAGE, which assigns prices dynamically based on the profit potential of each candidate seed, has the best performance both in the expected profit achieved and in running time.

4. Mining the network value of customers.

AUTHORS: P. Domingos and M. Richardson, “

One of the major applications of data mining is in helping companies determine which potential customers to market to. If the expected profit from a customer is greater than the cost of marketing to her, the marketing action for that customer is executed. So far, work in this area has considered only the intrinsic value of the customer (i.e, the expected profit from sales to her). We propose to model also the customer's *network value:* the expected profit from sales to other customers she may influence to buy, the customers those may influence, and so on recursively. Instead of viewing a market as a set of independent entities, we view it as a social network and model it as a Markov random field. We show the advantages of this approach using a social network mined from a collaborative filtering database. Marketing that exploits the network value of customers---also known as viral marketing---can be extremely effective, but is still a black art. Our work can be viewed as a step towards providing a more solid foundation for it, taking advantage of the availability of large relevant databases.

5. Cost-effective outbreak detection in networks

AUTHORS: J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance

These seemingly different problems share common structure: Outbreak detection can be modeled as selecting nodes (sensor locations, blogs) in a network, in order to detect the spreading of a virus or information as quickly as possible. Paper presents a general methodology for near optimal sensor placement in these and related problems. System demonstrates that many realistic outbreak detection objectives (e.g., detection likelihood, population affected) exhibit the property of "sub modularity". They exploit sub modularity to develop an efficient algorithm that scales to large problems, achieving near optimal placements, while being 700 times faster than a simple greedy algorithm. They also derive online bounds on the quality of the placements obtained by any algorithm. Our algorithms and bounds also handle cases where nodes (sensor locations, blogs) have different costs.

III. PROPOSED SYSTEM

In this paper, we propose a new efficient expectation model for the influence spread of a seed set based on independent maximum influence paths (IMIP) among users. We also show that the new objective function of the new expectation model is sub modular. Based on the new expectation model, we present a method to efficiently process an IMAX query. The method consists of identifying local regions containing nodes that influence the target nodes of a query and approximating optimal seeds from the local regions as the result of the query. Identifying such local regions helps to reduce the processing time, when the number of targets in an IMAX query is small compared to the number of all nodes. To approximate optimal seeds, we use a greedy method based on the marginal gain to the new objective function. In addition, we present a method to incrementally update the marginal gain of each user to accelerate the greedy method. We identify the limitations of existing researches related to maximizing influence on specific targets. We formulate an influence maximization problem as query processing without predefined labels to address the limitations. We prove that the problem is NP-hard and that the objective function of the IMAX query problem is sub modular. Based on the sub modularity of the objective function, we present a greedy algorithm for IMAX query processing. We propose a new efficient expectation model for influence spread of a seed set. We show that the new objective function of the expectation model is sub modular. Based on the new expectation model, we propose a greedy-based approximation method to process an IMAX query with efficient incremental updating of the marginal gain of each user. We also propose an effective method to reduce the number of candidates for optimal seeds by identifying users who strongly influence targets from preprocessed data. We experimentally demonstrate that our identifying local influencing regions technique is very powerful and the proposed method is at least an order of magnitude faster than the comparison methods in most cases with high accuracy. Identifying local influencing regions makes the basic greedy algorithm.

IV. Mathematical Model

Let S is the Whole System Consist of

S= {I, P, O}

I = Input.

I = {U, Q, D}

U = User

U = {u1,u2….un}

Q = Query Entered by user

Q = {q1, q2, q3…qn}

D = Dataset.

P = Process:

P = {IMIP, CELF, PMIA, IRIE, CD Model}

IMIP: independent maximum influence paths

CELF: Cost-Effective Lazy Forward

PMIA: prefix excluding maximum influence arborescence

Step1: User enters the Query.

Step2: In comparison process following methods will be performed.

Step3: Independent maximum influence paths (**IMIP**):

We propose a new efficient expectation model for the influence spread of a seed set based on independent maximum influence paths (IMIP) among users.

Step4: **CELF++:** is an improved greedy algorithm exploiting Sub-modularity.

Step5: **PMIA** is a greedy-based algorithm based on maximum influence paths between nodes. In PMIA, parameter u is used to prune out maximum influence
paths having low influence .

Step6: **IRIE:** is one of recent algorithms for influence maximization.

Step7: **CD Model:** CD is the greedy method using the CD model. The CD model is a probabilistic model based on users' historical action logs. We use this method only for the experiment related to the actual influence spread.

Output: Finally the particular result will be shown to user as per his query.

V. SYSTEM ARCHITECTURE

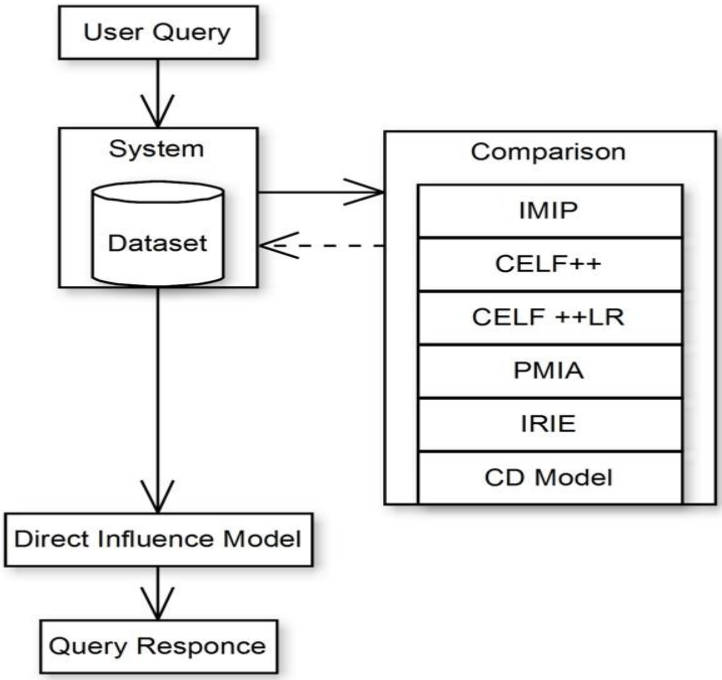

Fig.1 System Architecture

VI. RESULT

Add post

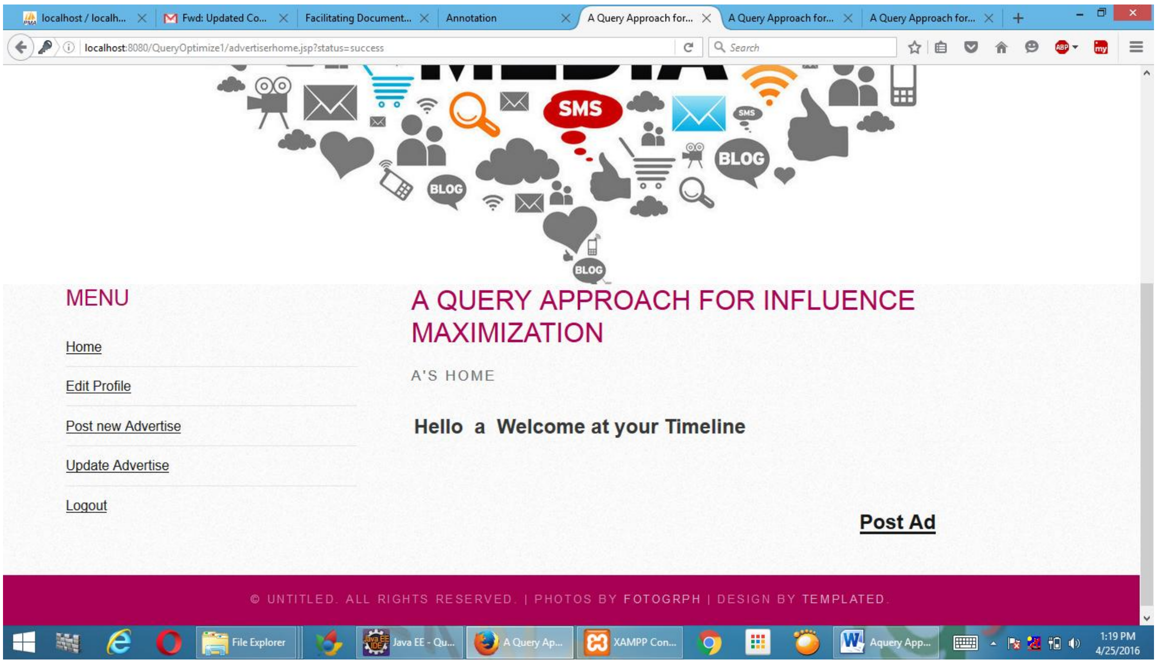

Fig2. Snapshot of add post

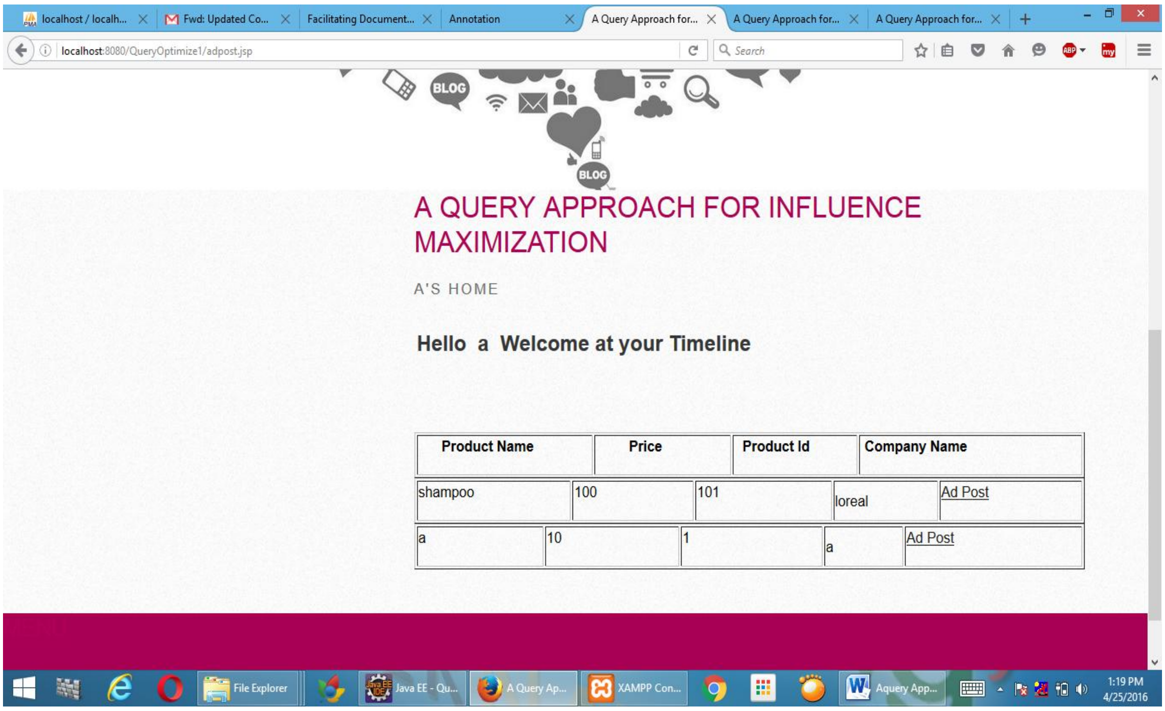

Fig3 Snapshot of product advertise

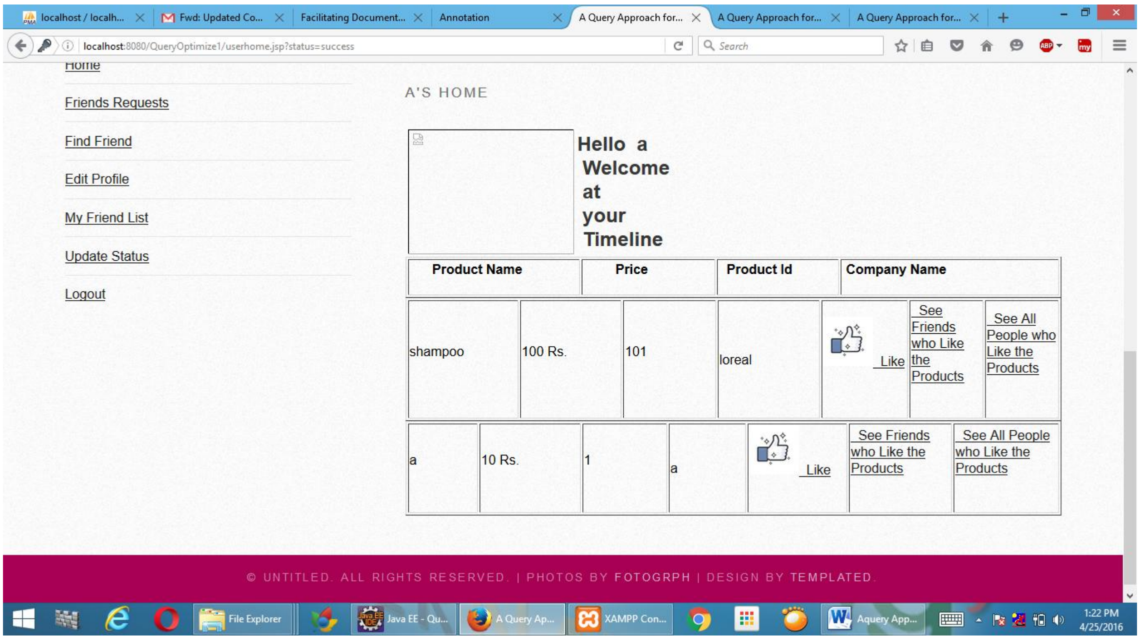

Fig 4 Snapshot of display product on user timeline

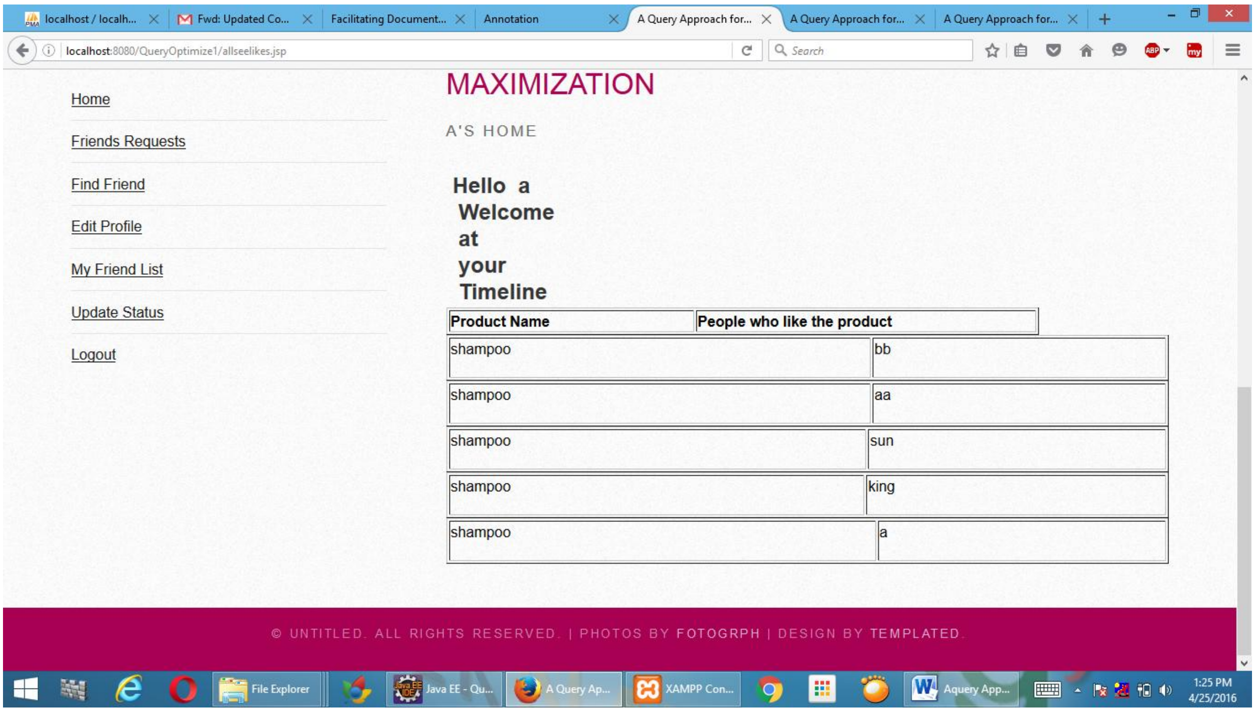

Fig5. Friends like product

VII. CONCLUSION

In this paper, we plan IMAX question preparing to expand the impact on particular clients in informal organizations. Since IMAX inquiry preparing is NP-hard and ascertaining its target capacity is #P-hard, we concentrate on the best way to surmised ideal seeds proficiently. To inexact the estimation of the goal capacity, we propose the IMIP model taking into account autonomy between ways. To prepare an IMAX question effectively, separating possibility for ideal seeds is proposed what's more, the quick voracious based guess utilizing the IMIP model. We tentatively show that our distinguishing neighborhood impacting locales procedure is viable and the proposed strategy is generally no less than a request of extent speedier than PMIA and IRIE with comparable exactness furthermore, the proposed strategy is generally six requests of extent speedier than CELF++ and the recognizing nearby impacting districts system makes CELF++ around 3.2 times speedier while accomplishing high precision.

ACKNOWLEDGMENT

We might want to thank the analysts and also distributers for making their assets accessible. We additionally appreciative to commentator for their significant recommendations furthermore thank the school powers for giving the obliged base and backing.

VIII. REFERENCES

[1] Jong-Ryul Lee and Chin-Wan Chung "A Query Approach for Influence Maximization on Specific Users in Social Networks"

[2] F.-H. Li, C.-T. Li, and M.-K. Shan, "Labeled influence maximization in social networks for target marketing," in Proc. IEEE 3rd Int. Conf. Privacy, Secur., Risk Trust, Int. Conf. Social Comput., 2011, pp. 560–563.

[3] W. Lu and L. Lakshmanan, "Profit maximization over social networks," in Proc. IEEE 12th Int. Conf. Data Mining, 2012, pp. 479–488.

[4] P. Domingos and M. Richardson, "Mining the network value of customers," in Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2001, pp. 57–66.

[5] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2007, pp. 420–429.

[6] A. Goyal, W. Lu, and L. V. Lakshmanan, "CELF++: Optimizing the greedy algorithm for influence maximization in social networks," in Proc. 20th Int. Conf. Companion World Wide Web, 2011, pp. 47–48.

[7] Y. Wang, G. Cong, G. Song, and K. Xie, "Community-based greedy algorithm for mining top-k influential nodes in mobile social networks," in Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2010, pp. 1039–1048.

[8] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2009, pp. 199–208.

[9] 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.

[10] Q. Jiang, G. Song, C. Gao, Y. Wang, W. Si, and K. Xie, "Simulated annealing based influence maximization in social networks," in Proc. 25th AAAI Conf. Artif. Intell., 2011, pp. 127–132.

[11] K. Jung, W. Heo, and W. Chen, "Irie: Scalable and robust influence maximization in social networks," in Proc. IEEE 12th Int. Conf. Data Mining, 2012, pp. 918–923.

