International Journal of Gender, Science and Technology
ISSN: 2040-0748 UGC Care Group I Journal.
Vol-12 Issue-02 Dec 2023

https://ijgst.com.2023.v12.i2.pp160-168

AN ENHANCED STRATEGY FOR SOCIAL NETWORK PROFIT MAXIMIZATION

G MALLIKHARJUNA RAO [1] THADIGIRI SOWNDARYA [2] MANDE PRAVEEN [3] S.V.KISHORE BABU [4]

[1] ASSOC.PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[2] ASST. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[3]ASST. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[4] ASSOC. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING,
[1,2,3,4] SRI MITTAPALLI COLLEGE OF ENGINEERING

Abstract-

A lot of time and energy has been spent studying viral marketing in the last ten years. The majority of previous research on seed selection in social networks ignores the case when group actions might provide a reward. The level of enthusiasm among participants is a good indicator of the activity's profitability. A same piece of information might elicit vastly varying levels of enthusiasm from various demographics. How can we maximise the projected total amount of profit by selecting the seed users in a social network given a profit function? There is no straightforward application of current methods to this issue since it is fundamentally different from the standard impact maximisation problem. Here we take a look at the issue of social network activity profit maximisation. First, we show that the basic greedy algorithm cannot approach the nondeterministic polynomial time-hard maximising activity profit issue within a constant factor. The degree to which a function deviates from submodularity is called its supermodular degree. Our technique is designed to produce an approximation ratio of $(1/(\Delta + 2))$ given that the social network has a supermodular degree. After that, we create an exchange-based method to boost the solution's quality even more. To get around the algorithms' heavy computing requirements, we also come up with a randomised variation method. Extensive experimental findings on three genuine benchmark data sets show that our algorithms outperform many baseline heuristics in terms of efficiency and effectiveness.

Keywords-- Activity profit maximization, approximation algorithms, social computing.

I. INTRODUCTION

Facebook, LinkedIn, ResearchGate, and many other online social networking platforms now include personal profiles held by an increasing number of users. The influence of online social networks in several fields, including viral marketing and the presidential election, has been hard to ignore due to the exponential growth of user-generated content on these platforms. Social media ads perform better than those on more conventional broadcast media, according to recent studies. Online social media have gradually become an integral part of people's everyday lives, both in terms of the information they disseminate and the impact they propagate. As a result, there is a wealth of literature on the topics of influence-based research challenges and influence-driven information technology. Prior research on viral marketing has mostly concentrated on themes pertaining to items utilised by a single individual. But there are items that can need more than one person to utilise them. For instance, in order to complete the most challenging levels of a cooperative video game like Monaco, two or more players are required. Online games that involve numerous players joining before the game begins, such as Texas hold 'em, are another good example. Products requiring two-way user interaction were the starting point for the research by Wang et al. [1]. Our focus in this research is on products that encourage

collaboration among several users. Think of a social network that follows a certain paradigm for the spread of information. With a positive integer k provided, the goal of the famous influence maximisation issue is to choose a collection of k seeds that will maximise the influence spread, which is defined as the number or anticipated number of active nodes. Instead of measuring the influence spread, the objective function in the activity maximisation issue stated in [1] assesses the overall "activity strength" or, alternatively, activity profit. We think of an action as a two-property event in order to account for goods with any amount of consumers. To start, there are at least two people involved in this. Second, money would come in if this happened. A group of people might do things like buy something or play a game together. Every action constitutes a hypergraph's hyperedge set because it may be symbolised by a collection of users. The earnings that various activities bring in might vary. We aim to optimise the overall profit, or the profit that active users are projected to contribute. According to [1], the IC model and the LT model do not provide submodular or supermodular objective functions, respectively, for the activity maximisation issue. In order to derive an approximation algorithm with theoretically assured performance, their approach makes use of a technique known as the sandwich method. In order to solve the activity profit maximisation issue, we shall use a novel approach in this study. Computational experiments will back up our findings, and we will also offer new theoretical improvements to the approach.

II. RELATED WORKS

When studying social networks, one of the primary goals should be to maximise one's influence [2]-[8]. The influence maximisation issue is NP-hard and the influence spread problem is #P-hard for many information diffusion models, particularly the IC model and the LT model. On the other hand, there exist randomised algorithms [9]-[11] that, given an input social network with n nodes and m edges, provide an approximation of $(1 - e{-}1 - \varepsilon)$ in time $O((m + n))$ with a probability of $1 - \varepsilon$. Even for arborescence directed into a root, the impact maximisation issue is NP-hard, according to an intriguing hypothesis by Bharathi et al. [12]. The IC model is supported by this hypothesis, as shown by Lu et al. [13]. Influence maximisation in the LT model may be solved in polynomial time, as shown by Wang et al. [14]. I had no idea that the IC and LT models might provide differing computing difficulties for the same issue before today. When it comes to viral marketing, the influence maximisation issue is really useful. pages 15–18. While the majority of studies have focused on advertising to a single user, Wang et al. [1] investigate a challenge related to promoting a product with two-user activities. A fundamental shift in the mathematical formulation's property, from submodular maximisation to monotone nonsubmodular maximisation, results from their disagreement on the notion of activity. Both of these are subfields of nonlinear optimisation by combinations. Both monotone submodular maximisation [19] and nonmonotone submodular maximisation

https://ijgst.com.2023.v12.i2.pp160-168

[20]-[22] have received much research in this field. Lately, however, many have started paying attention to monotone nonsubmodular study.Following is a synopsis of the many contributions made by us in this publication. 1) In order to maximise the projected total profit in a social network, we present a new MAP issue. Our goal captures interactions across numerous active users, which is a unique and original characteristic of our challenge. Our investigation of the information diffusion process's generalised multiuser interactions is the first of its kind. The traditional influence maximisation issue is a subset of our overall challenge. 2) The MAP problem's intricacy is examined. Additionally, we demonstrate that the problem's objective function does not exhibit any submodular or supermodular properties. Assuming the supermodular degree is constrained by Δ, we provide a roughly maximisation [1], [23]-[26]. We want to contribute to this area of

speaking procedure that produces an approximation ratio of $(1/(\Delta + 2))$. In order to make the solution even better, we create an algorithm that relies on exchanges. The third point is that we establish that it is #P-hard to compute the precise value of the objective function of the MAP problem given a seed set. We develop a randomised variation (RV) method to alleviate the computational load of the issue and take on this task. Incorporating this method into our algorithms has been beneficial, according to the experimental findings. 4) In order to test how well our algorithms operate, we use real-world benchmark data sets from social networks. Our algorithms outperform a few basic heuristics in terms of efficiency and efficacy, according to empirical assessment data.

.

III. PROPOSED METHOD

The IC model and the LT model are two of the most popular influence propagation models used in the field of influence maximisation research. In this study, we use the IC model to depict the social network's diffusion dynamics and a directed weighted graph G = (V, E) to depict the network's structure. Every node in the IC model may be in either the active or inactive state, and the influence probability (puv) is a weight that measures the likelihood that u will activate v once u becomes active. Any two nodes in a network may activate each other without affecting any other nodes. Events in the IC model take place in separate time periods. During time slot 0, which is the initial part of the propagation process, nodes outside of the seed set are not doing anything.A node has one opportunity to independently activate each of its inactive neighbours in time slot t+1 after being activated in time slot t, with the probability of the appropriate edge weight. One way to model this procedure is to imagine tossing a coin with a bias towards the side that shows heads. A successful activation attempt and the advantage are both signified by a head in the coin toss. In any other case, the activation attempt is deemed unsuccessful and the edge is proclaimed blocked. Until the diffusion process is complete, an activated node remains active. When it is no longer possible to activate any further nodes in the network, the procedure terminates.We use the notation "hyperedge e" with a collection of "head nodes" (He) and "tail nodes" (Te) to represent actions that involve more than one user. Take note that e becomes a simple edge when |He| = |Te| = 1. To avoid confusion, we refer to both simple

edges and hyperedges in the graph by the same unambiguous term: edge. This is because a simple edge is a subtype of hyperedge. In order to represent the information dispersion in a hypergraph, we use a basic expansion of the original IC model. Once every node in an edge's head set (He) becomes active in time slot t, all the nodes in the tail set (Te) have a chance to become active independently in time slot t+1. When further activations are no longer possible, the diffusion process comes to a halt. When Δ = 0, the submodularity of the function f(・) becomes apparent. Based on the average marginal benefit, we develop an IESG method for the nonsubmodular instance of the influence maximisation problem with limited supermodular degree.We begin with a null set S0, as specified in Algorithm 1. The difference between the approximate and original optimum solutions is still somewhat large, even if we can get an approximation ratio for MAP. The EIA that was developed with the help of the desirable MAP attributes is detailed below. We begin by obtaining an optimal criteria and discussing the optimisation condition of MAP. Next, we'll go over the M-convexity of the MAP viable area. Finally, the EIA is here. We isolate the optimisation condition first. Despite the fact that the greedy method IESG provides MAP with an approximation of the solution, there is still a significant difference between the achieved approximation and the theoretical ideal solution. There is still a need for further research into the basic topic of how to assess the potential for improvement of an approximation answer. This issue is addressed by the following theorem.

Algorithm 1 IESG

Input: a hyper-graph, G, and a constant, k.
Output: a set of seed nodes, S.

1: Initialize $i = 0$ and $S_0 = \emptyset$.
2: While $|S_i| < k$ do
3: Find out $\arg\max_{v \in V, D'_v \subseteq D^+_f(v)} [f(S_i \cup \{v\} \cup D'_v) - f(S_i \cup D'_v)]$, constrained by $|S_i \cup \{v\} \cup D'_v| < k$.
4: Update $S_{i+1} = S_i \cup \{v\} \cup D'_v$.
5: Update i to be $i + 1$.
6: Return $S = S_i$ as the set of seed nodes.

Algorithm 2 EIA

Input: a solution S_0 of influence function maximization problem with $|S_0| = k$.

Output: a set of non-improvable solution $\widehat{S}$.

1: Initialize $i = 0$ and $S = S_0$

2: Find out $v_R = \arg\min_{v \in S} \Delta_v f(S \backslash \{v\})$, let $\Delta(v_R) = \Delta_{v_R} f(S \backslash \{v_R\})$

3: Find out $v_A = \arg\max_{v \in V \backslash S} \Delta_v f(S \backslash \{v_R\})$ and let $\Delta(v_A) = \Delta_{v_A} f(S \backslash \{v_R\})$

4: If $\Delta(v_R) \geq \Delta(v_A)$, then S is non-improvable, stop; otherwise $S := S - v_R + v_A$ and go to step 2.

IV. RESULTS AND DISCUSSION

Here, we examine the IESG algorithm's approximation ratio. Then, to take things a notch further, we provide the EIA, or exchange improvement algorithm. The NP-hardness of the issue and the #P-hardness of the objective function value calculation make it impossible to solve the MAP problem in reality, even if the algorithms described in the previous sections have excellent approximation performance. It is still NP-hard to calculate the supermodular set. Here, we use a viable approach for the MAP issue that is based on the random technique. To assess the efficacy and efficiency of the suggested methods, we perform comprehensive experiments on benchmark datasets in this section. Table I provides an overview of the primary metrics from the datasets that were used in our research. Node degree and node supermodular degree distributions for all datasets are shown in Fig. 1. We evaluate three baseline algorithms and compare them to our own. In each iteration, Myopic chooses a node that increases the overall profit the most, and so on, until k seeds are chosen. Until k seeds are chosen, DegMax iteratively chooses a node with the greatest out-degree. Until k seeds are chosen, InfMax chooses a node based on the highest increase in the influence spread in each iteration. The algorithm just takes social impact into account, disregarding the distribution of profits. With k ranging from 2 to 16, the anticipated total profit is shown in Figs. 2 and 3 for different algorithms and baselines. The projected total profit from activities is shown on the y-axis, while the size of the seed set is represented on the x-axis.

TABLE I

CHARACTERISTICS OF SOCIAL NETWORK DATA SETS

Name	#Nodes	#Edges	Avg. degree
Facebook	899	72,821	165
arXiv	16,726	66,759	11.9
Epinions	22,166	353,546	33.5

https://ijgst.com.2023.v12.i2.pp160-168

It should be noted that Fig. 2 displays the outcomes of many algorithms with uniform activity profit distribution settings. Each action is specifically given a profit of one. The number of active edges between the nodes in this instance is precisely equal to the projected total profit. Since a bigger k improves the likelihood that seeds will affect more nodes and eventually result in more active behaviours, the projected total profit rises as k increases. In the uniform condition, EIA and IESG perform better than all other baselines for any k, as shown in Fig. 2(a)–(c). Additionally, as k rises, we see that the DegMax algorithm's difference from other methods widens. DegMax performs badly because it ignores social impact and simply makes use of the social network's structural qualities. Because it just takes into account the diffusion process and ignores the activity profit when choosing seeds, InfMax is unable to provide satisfactory results. The nodes in a node's supermodular set may enhance the node's predicted profit by a little amount, but Myopic only chooses one seed node at a time, ignoring the combinations of nodes that may trigger many more activities via hyperedges.As k approaches 12, the projected total profit produced by both IESG and EIA converges for the Facebook data set, as seen in Fig. 2(a). The main cause is that Facebook has a considerably denser network structure and fewer nodes than the other two data sets. It is thus feasible to affect almost every node in the social network using a minimal number of seed nodes. In comparison to IESG and EIA, Myopic performs similarly on this data set. The anticipated total profit generated by all algorithms on arXiv and Epinions continues to rise as k rises. As k grows, there is a greater disparity between Myopic and our suggested methods. By using the power of exchange and looking at a node's supermodular set to integrate the impact boosting from choosing many seeds at once, EIA and IESG are able to achieve a better solution without considering the node combinations. Figure 3 displays the outcomes of many algorithms for the activity profit distribution under the influence probability setting. In particular, the profit given to each action is equal to the impact likelihood linked to the relevant edge in the network. As anticipated, the estimated total profit rises with an increase in k. The findings are quite comparable to those in Fig. 2(a)–(c), as shown in Fig. 3(a)–(c). In every scenario, the EIA and IESG algorithms perform better than the baseline algorithms. The reason for this is because the baseline algorithms completely disregard the distribution of activity profits and only take into account the social network's structural characteristics or the influence diffusion process. This explains why the baseline algorithms often fail while our suggested algorithms consistently perform well. As shown by the predicted total profit, it was noteworthy that InfMax's outcomes performed better under the impact probability setting. InfMax chooses a seed node based on the current seed set that produces the biggest marginal increment of the influence spread in each iteration until k seed nodes are chosen. Consequently, a collection of very significant nodes is often chosen by InfMax to serve as the seed set.Edges with greater influence probabilities are more likely to become active and contribute to the diffusion of influence, and the activity profit for each edge is equal to the influence probability associated with the edge under the influence probability setting. The edges with a greater influence probability (also known as activity profit) become active and contribute to the overall profit at the conclusion of the diffusion process while the algorithm chooses the seed set that results in the highest influence spread.

.

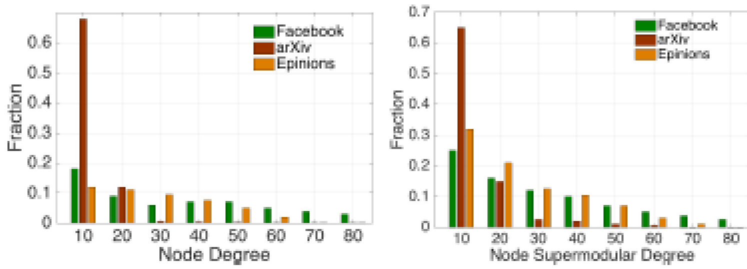

Fig. 1. Node degree and supermodular degree distribution.

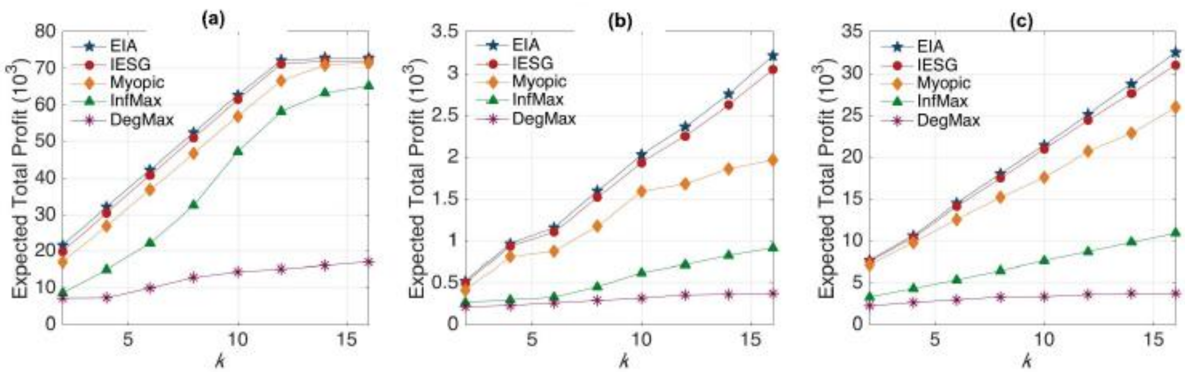

Fig. 2. Expected total profit versus seed set size produced by various algorithms under uniform profit setting. (a) Facebook. (b) arXiv. (c) Epinions.

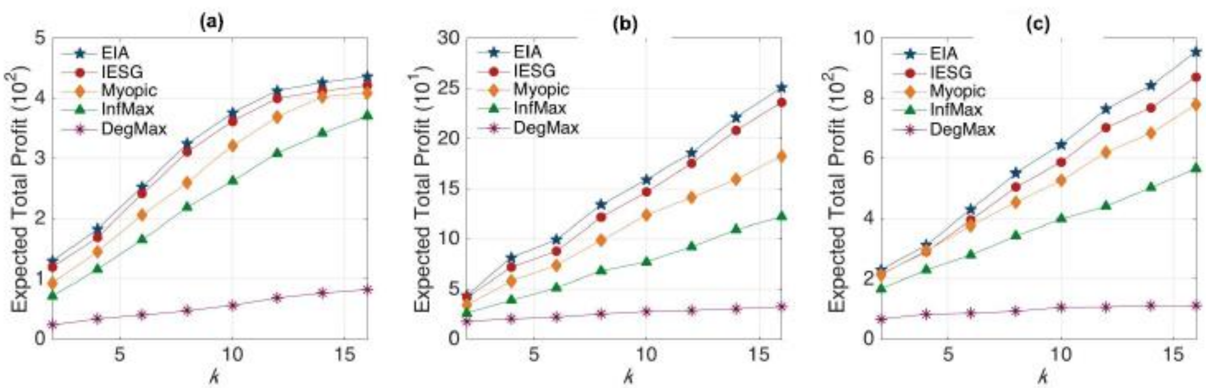

Fig. 3. Expected total profit versus seed set size produced by various algorithms under influence profit setting. (a) Facebook. (b) arXiv. (c) Epinions.

V. FUTURE SCOPE AND CONCLUSION

We examine a new and significant MAP issue in this research. We investigate the generalised multiuser interactions in information diffusion for the first time. As special examples, our subject involves a number of traditional influence maximization-related issues. If the supermodular degree is bounded by Δ, we offer an approximation approach with an approximate ratio of 1 Δ+2. To further enhance the solution's quality, we create an exchange-based algorithm. We provide an RV method to lessen the MAP problem's computational load

.

.

REFERENCES

[1] Z.Wang, Y. Yang, J. Pei, and E. Chen. (2016). "Activity maximization by effective information diffusion in social networks." [Online]. Available: https://ieeexplore.ieee.org/document/8010858

[2] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2003, pp. 137–146.

[3] W. Xu, Z. Lu, W. Wu, and Z. Chen, "A novel approach to online social influence maximization," *Social Netw. Anal. Mining*, vol. 4, no. 1, p. 153, 2014.

[4] 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.

[5] W. Chen, Y. Yuan, and L. Zhang, "Scalable influence maximization in social networks under the linear threshold model," in *Proc. IEEE 10th Int. Conf. Data Mining (ICDM)*, Dec. 2010, pp. 88–97.

[6] 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.

[7] Y. Zhang, Q. Gu, J. Zheng, and D. Chen, "Estimate on expectation for influence maximization in social networks," in *Proc. Adv. Knowl. Discovery Data Mining*, pp. 99–106, 2010.

[8] A. Goyal, F. Bonchi, and L. V. S. Lakshmanan, "A data-based approach to social influence maximization," *Proc. VLDB Endowment*, vol. 5, no. 1, pp. 73–84, Sep. 2011.

[9] Y. Tang, X. Xiao, and Y. Shi, "Influence maximization: Near-optimal time complexity meets practical efficiency," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2014, pp. 75–86.

[10] Y. Tang, Y. Shi, and X. Xiao, "Influence maximization in near-linear time: A martingale approach," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2015, pp. 1539–1554.

[11] H. T. Nguyen, M. T. Thai, and T. N. Dinh, "Stop-and-stare: Optimal sampling algorithms for viral marketing in billion-scale networks," in*Proc. Int. Conf. Manage. Data*, 2016, pp. 695–710.

[12] S. Bharathi, D. Kempe, and M. Salek, "Competitive influence maximization in social networks," in *Internet and Network Economics*. Berlin, Germany: Springer, 2007, pp. 306–311.

[13] Z. Lu, Z. Zhang, and W. Wu, "Solution of Bharathi–Kempe–Salek conjecture for influence maximization on arborescence," *J. Combinat. Optim.*, vol. 33, no. 2, pp. 803–808, 2017.

[14] A.Wang, W. Wu, and L. Cui, "On Bharathi–Kempe–Salek conjecture for influence maximization on arborescence," *J. Combinat. Optim.*, vol. 31, no. 4, pp. 1678–1684, 2016.

[15] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," in *Proc. 8th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2002, pp. 61–70.

[16] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral

marketing," *ACM Trans. Web*, vol. 1, no. 1, p. 5, 2007.

[17] 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.

[18] F. Bonchi, C. Castillo, A. Gionis, and A. Jaimes, "Social network analysis and mining for business applications," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 22, Apr. 2011.

[19] G. L. Nemhauser and L. A. Wolsey, "Best algo rithms for approximating the maximum of a submodular set function," *Math. Oper. Res.*, vol. 3, no. 3, pp. 177–188, 1978.

[20] U. Feige, V. S. Mirrokni, and J. Vondrák, "Maximizing non-monotone submodular functions," *SIAM J. Comput.*, vol. 40, no. 4, pp. 1133–1153, 2011.

[21] J. Lee, V. S. Mirrokni, V. Nagarajan, and M. Sviridenko, "Non-monotone submodular maximization under matroid and knapsack constraints," in *Proc. 41st Annu. ACM Symp. Theory Comput.*, 2009, pp. 323–332.

[22] M. Feldman, J. Naor, and R. Schwartz, "A unified continuous greedy algorithm for submodular maximization," in *Proc. IEEE 52nd Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2011, pp. 570–579.

[23] U. Feige and R. Izsak, "Welfare maximization and the supermodular degree," in *Proc. 4th Conf. Innov. Theor. Comput. Sci.*, 2013, pp. 247–256.

[24] M. Feldman and R. Izsak. (2014). "Constrained monotone function maximization and the supermodular degree." [Online]. Available: https://arxiv.org/abs/1407.6328

[25] W. Lu, W. Chen, and L. V. Lakshmanan, "From competition to complementarity: Comparative influence diffusion and maximization," *VLDB En dowment*, vol. 9, no. 2, pp. 60–71, 2015.

[26] W. Chen, T. Lin, Z. Tan, M. Zhao, and X. Zhou, "Robust influence maximization," in *Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 795–804.

[27] K. Murota, *Discrete Convex Analysis*. Philadelphia, PA, USA: SIAM, 2003.

[28] Vivek, Kolla, et al. "An Efficient Triple-Layered and Double Secured Cryptography Technique in Wireless Sensor Networks." 2021 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER). IEEE, 2021.

[29] Anusha, Pureti, T. Sunitha, and Mastan Rao Kale. "Detecting and Analyzing Emotions using Text stream messages." *ECS Transactions* 107.1 (2022): 16913