A Heuristic Mixed Model for Viral Marketing Cost Minimization in Social Networks

Ashis Talukder
Department of Computer Science and Engineering
Kyung Hee University
South Korea
email: ashis@khu.ac.kr

Do Hyeon Kim
Department of Computer Science and Engineering
Kyung Hee University
South Korea
email: doma@khu.ac.kr

Choong Seon Hong
Department of Computer Science and Engineering
Kyung Hee University
South Korea
email: cshong@khu.ac.kr

***Abstract*—The Influence maximization (IM) aims at estimating a small number of influential users that maximize the viral marketing profit whereas, the Reverse Influence Maximization (RIM) deals with the minimization of Viral Marketing (VM) cost in social networks. Here, the VM cost is measured by the minimum number of nodes that are required to activate seed nodes and the profit is defined by the maximum number of nodes influenced by seed users when they are initially activated. However, most of the existing works focus on the profit maximization without considering the VM cost. Thus, in this research, we introduce a Viral Marketing Cost (VMC) Minimization problem and propose a Heuristic Mixed (HM) approach under mixed Reverse Independent Cascade (RIC) and Reverse Linear Threshold (RLT) diffusion models. The proposed HM model employs the greedy technique along with a heuristic approach to optimize the VM cost. Moreover, our model resolves the challenging issues of the RIM problem more efficiently. Finally, we simulate our model using data-sets of two real social networks, and the result shows that our model outperforms the baseline and existing models.**

***Index Terms*—influence maximization; reverse influence maximization; viral marketing; viral marketing cost; social network.**

I. Introduction

Nowadays, social networks have become the ultimate platform for marketing especially, for Viral Marketing (VM) when Facebook, Twitter-like thousands of social networks have acquired huge popularity around the globe. Influence Maximization (IM), being such a viral marketing tool, has gained vast research interest in the recent decade [1]. In viral marketing, people are influenced by other impactful persons, $e.g.$, family members, friends, colleagues or some trend-setters in the *word-of-mouth* effect [2], [3].

The main purpose of the influence maximization is to find k-top influential nodes, which can maximize the influence diffusion in the network. The influence is measured by the maximum number of nodes that can be activated by the seed users when they are assumed to be activated initially. The remarkable pioneering work was conducted by Kempe *et al.* [2] in 2003. They proposed Linear Threshold (LT) and Independent Cascade (IC) models, which achieved enormous popularity. Thereafter, many methods are proposed in literature to solve the IM problem due to its versatile applications, $e.g.$, profit maximization [3], [4], [5], marketing campaign [6], contagion detection [7], rumor monitoring [8], [9], [10], [11], domain expert identification [12], online recommendation [13]. However, most of the state-of-the-art models suffer from a flaw that the seed nodes are assumed to be activated initially.

In many works, the viral marketing profit is maximized in social networks by maximizing the product adoption taking the product price into account [3] whereas, in many studies, the product adoption is maximized without considering the product price [14]. In both the studies, in order to activate seed users initially, some incentives such as free products or services are offered. However, seed nodes can also be activated by other influential users as influenced in the IM process. This gives rise to a new research direction called Viral Marketing Cost (VMC) optimization problem. However, most of the existing studies do not address the estimation of the VM cost, which is defined by the minimum number of influencer nodes required to activate seed nodes. Therefore, to compute the VM cost, Reverse Influence Maximization (RIM) models are proposed in which influence is diffused throughout the network in reverse order [15], [16], [17]. However, the existing RIM models are incapable of handling the challenging issues, which include setting stopping condition, handling three basic network component (BNC), and insufficient influence.

Therefore, in this paper, we introduce a RIM-based Viral Marketing Cost (VMC) minimization problem and propose a Heuristic Mixed (HM) model to estimate the optimal VM cost in social networks. The key contributions of this paper are stated below.

1) We introduce a VMC minimization problem along with a Heuristic Mixed (HM) solution model, which jointly employs the IC and LT models in reverse order as well as a greedy optimization technique.
2) We modify the traditional IC and LT models as Reverse IC (RIC) and Reverse LT (RLT) models, respectively, to

apply in reverse order for node activation in our model.

3) We incorporate a cost minimization heuristic, which is applied in the influence weight calculation and apply with greedy optimization to minimize the VM cost.
4) We simulate the proposed model using two real datasets of widely recognized social networks, and the result shows that the proposed model outperforms the existing models.

The rest of the paper is as organized: section II presents the study of the state-of-the-art models. We describe the system model, and the problem definition in section III and the proposed HM model along with theoretical approximation bound is stated in section IV. The performance analysis is provided in section IV, and finally, the concluding remarks are given in section VI. The abbreviations used in this paper are listed with elaborations in Appendix A.

II. Existing Study

The Influence Maximization which identifies impactful seed users for viral marketing in the social networks, was introduced by Domingos *et al.* [18] in 2001. They searched for influential users by the network value of a customer by considering the social network as a target market. However, the accuracy of their model was not remarkably high as compared to the actual one. Therefore, Kempe *et al.* [2] formulate the IM model with performance bound $(1 - \frac{1}{e})$. Their proposed Linear Threshold (LT) and Independent Cascade (IC) models gain huge acceptance in the IM research community, and many LT-based and IC-based models are proposed after that.

However, the major drawback of the basic greedy model described in [2] is that the influence propagation model is NP-Hard and the greedy algorithm is quadratic. Thus, Chen *et al.* propose Local Directed Acyclic Graph (LDAG) model [19] and Degree Discount heuristics technique under IC model [20] to address the NP-Hard issue and Leskovec *et al.* [7] propose the Cost Effective Lazy Forward (CELF) algorithm to resolve the complexity (quadratic) issue. Thereafter, Goyal *et al.* [21] propose CELF++ method which outperforms the CELF model by $35 - 55\%$. However, the above studies conduct viral marketing in social networks without addressing the VM cost estimation.

In most of the above works, either the seed nodes are assumed to be initially activated, or a free sample product or service is offered to activate seed nodes. Moreover, they do not consider the fact that seed nodes can be motivated by some other influential users likewise in the IM mechanism. Thus, the authors in [15] and [16] introduce Reverse Influence Maximization problem in which influence is disseminated in reverse direction to compute the VM cost. They propose the Random RIM (R-RIM), and Randomized Linear Threshold RIM (RLT-RIM) in their works. Moreover, the authors also identify some challenges which include handling three basic network components, stopping criteria, and insufficient influence and NP-Hardness issue. However, these challenges are not handled properly in their works.

TABLE I
LIST OF PARAMETERS

Symbol	Meaning
$G(V, E)$	Social network.
V	Set of social network users.
E	Social relationship among users.
D_v^{out}	Out-neighbor set of v.
D_v^{in}	In-neighbor set of v.
d_v^{out}	Out-degree of v, $d_v^{out} = \lvert D_v^{out}\rvert$.
d_v^{in}	In-degree of v, $d_v^{in} = \lvert D_v^{in}\rvert$.
S	The seed set.
k	Size of the seed set S.
$\Lambda(v)$	The partial VM cost set of node v.
$\Lambda(S)$	The VM cost set of all the nodes of S.
$\lambda(v)$	The VM cost of node v, $\lambda(v) = \lvert\Lambda(v)\rvert$.
$\lambda(S)$	The VM cost of all the nodes of S, $\lambda(S) = \lvert\Lambda(S)\rvert$.
A_{new}	Newly activated in-neighbors.
$A_{current}$	Newly activated in-neighbors for current node u.
A_{target}	All the in-neighbors for current node u.
A_v	Minimum number of nodes selected to activate v.
w_{uv}	Social influence of node u to v.
$\tilde{w_{uv}}$	Composite influence of node u to v.
h_{uv}	Heuristic influence of node u to v.
$\tilde{h}_{uv}$	Normalized heuristic influence of node u to v.
α	A priority parameter, $0 \leq \alpha \leq 1$.
p	Activation probability in the RIC model.
λ	The estimated VM cost.
λ^*	The optimal VM cost.
d	Average in-degree in G.
C	The complexity of the proposed algorithm.

Thus, in this paper, we propose a Heuristic Mixed (HM) model for VMC Minimization. The proposed method not only handles RIM challenges efficiently but also estimates the optimized VM cost which is better than that of existing RIM models.

III. System Model

We consider a scenario of Profit Maximization in a social network and in general, the profit can be maximized in two ways, either by maximizing the revenue or by minimizing the cost. The Influence Maximization deals with revenue maximization whereas, the Reverse Influence Maximization deals with cost minimization. In this paper, we consider the Viral Marketing Cost (VMC) minimization mechanism for social networks and thus, our goal is to minimize the VM cost.

To formulate the VMC minimization problem, let us consider a social network, $G(V, E)$ which has a set V, of social network users (also known as individuals or nodes) as well as a set E, of social bonds among users. For instance, if G is the Facebook social network then, any two nodes, u and $v \in V$ indicate two Facebook users and a link $(u, v) \in E$ indicates that u and v are friends in the Facebook. We consider the in-neighbors and out-neighbors sets of any node v to be D_v^{in} and D_v^{out}, respectively along with in-degree and out degree, $d_v^{in} = |D_v^{in}|$ and $d_v^{out} = |D_v^{out}|$, respectively.

In the viral marketing, the cost that is incurred to activate the seed users is considered as the Viral Marketing cost, which is the main focus of our study. Moreover, the VMC minimization problem ensures the activation of the seed nodes and estimates the users which influence them to be activated. The number of users who activate the seed nodes is termed as the VM cost and is denoted by $\lambda(S)$. The objective of the VMC minimization problem is to minimize the $\lambda(S)$ for a given seed set S.

Definition 1 (The VMC Minimization Problem). Given a social network $G(V, E)$ and a seed set S of size k, the RIM-based VMC minimization problem is to minimize the viral marketing cost $\lambda(S)$, defined by the minimum number of nodes that must be activated in order to activate all the seed nodes in S. □

IV. The Proposed Heuristic Mixed Model

In this section, we propose a Heuristic Mixed (HM) Model to address the VMC minimization problem. In the HM model, we jointly employ the modified Independent cascade (IC) model and the modified Linear threshold models (LT) [2] along with a greedy optimization technique.

A. The Working Principles of the HM Model

The working philosophy of the proposed HM model is explained properly in Fig. 2. The HM model calculates the optimal VM cost $\lambda(v)$, for each seed node $v \in S$ and then, aggregates them to generate the optimal VM cost, $\lambda(S)$. However, the HM model estimates the $\lambda(v)$ in two steps. In the first step, the HM model computes the partial VM cost $\lambda(u)$ for $u \in D_v^{in}$ by activating nodes using the Reverse IC model. Then, the most influential subset of D_v^{in} is selected by jointly using the greedy method and a heuristic technique along with the Reverse LT model.

1) Partial VM Cost Estimation: The process of the partial VM cost $\lambda(u)$-estimation is illustrated in the Fig. 2. We modify the traditional IC model to use in reverse order. Generally, the IC model finds how many nodes can be activated by a node u if it is activated initially. However, we determine how many modes are required to activate the node u by applying IC model in reverse order and name it as Reverse IC (RIC) model.

In the RIC model, initially, we assume that A_{new} holds the newly activated node at any hop, $A_{target}(u)$ are the nodes, which will be checked for the activation of the node u, $A_{current}$ includes the activated nodes at the current hop, and $\Lambda(u)$ stores the total nodes which contribute to the activation of u in all the previous hops to current hop. The process starts with the initialization:

$$A_{new} = \{u\} \tag{1}$$

$$\Lambda(u) = \{u\} \tag{2}$$

Then, for the activation of each $u \in A_{new}$, the target in-neighbors set, $A_{target}(u)$ is populated as:

$$A_{target}(u) = D_u^{in} \tag{3}$$

Next, each target node $w \in A_{target}(u)$ is given a single chance to activate the node u by tossing a biased coin with

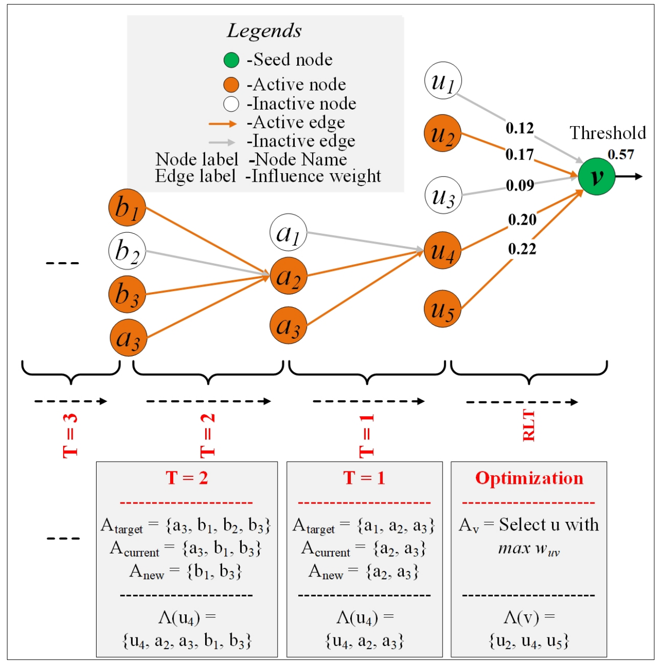


Fig. 1. The working philosophy of the Heuristic Mixed (HM) model. To estimate the $\Lambda(v)$, first, $\Lambda(u)$ is calculated for all $u \in D_v^{in}$. Say, for u_4, we start with, $\Lambda(u_4) = A_{new} = \{u_4\}$, $A_{target} = \{a_1, a_2, a_3\}$. At the hop, $T = 1$, the node $A_{current} = \{a_2, a_3\}$ are activated (say) by RIC model with a probability p and hence, we have, $A_{new} = \{a_2, a_3\}$ and $\Lambda(u_4) = \{u_4, a_2, a_3\}$ for the next hop. At $T = 2$, there is no node activation for a_3, since $D_{a_3}^{in} = \emptyset$. However, for the node a_2, we have $A_{target} = \{a_3, b_1, b_2, b_3\}$ (here, a_3 is given repeatedly since this is a graph, not a tree, we show the nodes hop-by-hop for better realization only). We assume that the nodes a_3, b_1, and b_3 are activated and thus, $A_{current} = \{a_3, b_1, b_3\}$. However, a_3 was previously activated and thus, we have, $A_{new} = \{b_1, b_3\}$ for the next hop. We also have, $\Lambda(u_4) = \{u_4, a_2, a_3, b_1, b_3\}$. Since there are no in-neighbors of the nodes in A_{new}, the estimation stoops at $T = 3$ with a partial VM cost $\Lambda(u_4) = \{u_4, a_2, a_3, b_1, b_3\}$. Similarly, we compute $\Lambda(u)$ for all $u \in D_v^{in}$. Finally, a set, $A_v \subset D_v^{in}$, of minimum number of nodes having maximum w_{uv} is chosen to activates v by the RLT model and we get, $\Lambda(v) = \cup_{u \in A_v} \Lambda(u)$.

some probability p. If the coin toss results in the head, we assume that the node w influences u to be activated and w is added to the set $A_{current}$. At the end of the hop, we update all these parameters for the next hop as:

$$A_{new} = A_{current} - \Lambda(u) \tag{4}$$

$$\Lambda(u) = \Lambda(u) \cup A_{current} \tag{5}$$

In the Eq. (4), the nodes which are explored in any previous hop are excluded and only the newly activated nodes are passed to the next hop. The same process is repeated for the next in-neighbor hop of the node u for each of the newly activated node w and so on. The RIC process terminates when there is no newly activated node at any hop, *i.e.*, $A_{new} = \emptyset$.

a) Learning Probability (p) for the RIC Model: Kempe *et al.* [2] suggest the values of p to be uniformly chosen either 0.01 or 0.1 whereas, in the Tri-valency model, p takes a value randomly from $\{0.1, 0.01, 0.001\}$ [19], [22], [23], [24]. However, we choose the model introduced by Kempe *et al.* [2].

2) The Optimization Model: To optimize the cost by the RLT model, we compute social influence w_{uv} of the node u to node v by degree centrality method [2]. Here, we introduce a heuristic that is employed along with greedy optimization.

We contribute a heuristic influence h_{uv} which is reverse proportional to the estimated $|\Lambda(u)|$ and indicates that a node with lower partial VM cost $|\Lambda(u)|$ has the higher probability to be the most influential. Thus, selecting the node with a higher value of h_{uv}, can reduce the VM cost. We combine the normalized value of heuristics influence $\tilde{h}_{uv}$ with the social influence w_{uv} to get composite optimized influence weight which is given by,

$$\tilde{w}_{uv} = \alpha \tilde{h}_{uv} + (1-\alpha) w_{uv} \tag{6}$$

where, $\alpha \in (0,1)$ is a priority parameter between the heuristic and the social influence.

We have composite influence weight and partial VM cost of each u and now, our target is to select a set $A_v \subset D_v^{in}$, of the least number of most inflectional in-neighbor u whose aggregated influence can activate v. This is accomplished by the RLT model in which an in-neighbors u with maximum $\tilde{w}_{uv}$ is taken greedily. Then, we estimate the desired VM cost set and VM cost as:

$$\Lambda(v) = \bigcup_{u \in A_v} \Lambda(u) \tag{7}$$

$$\Lambda(S) = \bigcup_{v \in S} \Lambda(v) \tag{8}$$

$$\lambda(S) = \big|\Lambda(S)\big|\,|. \tag{9}$$

B. The HM Algorithm

The proposed HM model is stated in Algorithm 1. The partial VM cost of a node $u \in D_v^{in}$ is computers in line $5-18$ in which the nodes activation is accomplished by the RIC model in lines $8-12$ whereas, mixed greedy and heuristic optimization is done in line 16. The final VM cost set and the VM cost are calculated in line 18 and line 20, respectively.

Theorem 1. *The VMC minimization problem under the HM model is NP-Hard.*

Proof. The optimization strategy used in the HM model to solve the VMC minimization problem is merely a Knapsack technique, and at each time, the model selects an in-neighbor u with maximum influence. Moreover, the RIC model used in the node activation process is just a variation of the traditional IC model under which the IM problem is NP-Hard [2]. Moreover, the Knapsack problem is a well-known NP-Hard [25] problem as well and hence, our VMC minimization problem under the HM model is also NP-Hard. □

C. Theoretical Performance Bound

Here, we provide a theoretical bound for the greedy optimization.

Theorem 2. *The proposed Heuristic Mixed (HM) Algorithm is a 2-approximation algorithm, i.e.*

$$\frac{\lambda^*}{\lambda} \geq \frac{1}{2} \tag{10}$$

Algorithm 1: The HM Algorithm

Input: $G(V,E), S$
Result: $\lambda(S)$

1 $\Lambda(S)\emptyset$;
2 **for** $v \in S$ **do**
3     $\Lambda(v) \leftarrow \emptyset$;
4     $A_{new} \leftarrow \{u\}$;
5   **while** $u \in A_{new}$ **do**
6         $A_{target} \leftarrow D_u^{in}$;
7         $A_{current} \leftarrow \emptyset$, $\Lambda(u) \leftarrow \{u\}$;
8   **while** $w \in A_{target}$ **do**
9   **if** *w is activated with probability p* **then**
10                $A_{current} \leftarrow A_{current} \cup \{w\}$;   `/* Node activated by RIC model */`
11   **end**
12   **end**
13        $A_{new} \leftarrow A_{new} - \Lambda(u) - \Lambda(S)$;   `/* Already activated node is excluded */`
14        $\Lambda(u) \leftarrow \Lambda(u) \cup A_{new}$;
15   **end**
16    $A_v \leftarrow$ A minimum set of $u \in D_v^{in}$ with max $\tilde{w}_{uv}$ to activate v;   `/* Greedy & heuristic optimization */`
17    $\Lambda(v) \leftarrow \cup_{y \in A_v} \Lambda(y)$;
18    $\Lambda(S) \leftarrow \Lambda(S) \cup \Lambda(v)$;   `/* The VM cost set */`
19 **end**
20 $\lambda(S) \leftarrow \big|\Lambda(S)\big|$;   `/* Final VM cost */`
21 *return* $\lambda(S)$;

Proof. The proof of the Theorem 2 follows from the proof of the Theorem 1, in which we have already proved that the proposed model is a variation of the Knapsack problem. Again, the Knapsack is a 2-approximation algorithm [26], and thus, the proposed HM model exhibits the same approximation ratio, that is,

$$\tfrac{\lambda^*}{\lambda} \geq \tfrac{1}{2}.$$ □

Now, for the running time derivation, let us assume that d is the average in-degree in G. Then, the running time of the HM algorithm is given by,

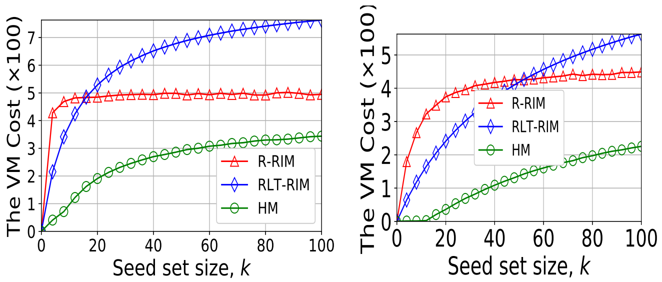
$$C \leq \underbrace{k}_{\text{line } 2}(\underbrace{d}_{\text{line } 5}(\underbrace{d}_{\text{line } 8-12} + \underbrace{d}_{\text{line } 16})) \approx O(kd^2). \tag{11}$$

V. Performance Evaluation

We evaluate the performance of the proposed HM model by using real datasets of two widely popular social networks. We present a comparative analysis of the proposed model with the existing models for both the datasets.

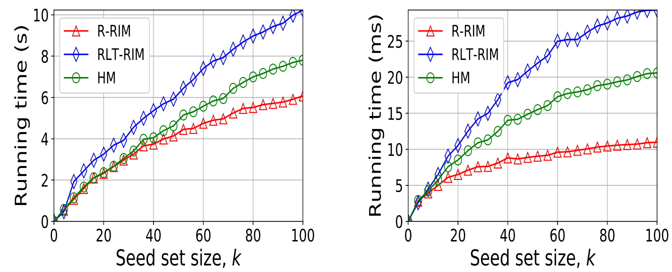
A. Data Collection

Two real datasets are collected from SNAP dataset collections [27] for the simulation of the proposed model as



(a) Epinions dataset. **(b)** Facebook dataset.

Fig. 2. The VM cost (in hundreds) for different values of k (1 to 100) for a) Epinions, and b) Facebook dataset.



(a) Epinions dataset. **(b)** Facebook dataset.

Fig. 3. The running time for different values of k (1 to 100) for a) Epinions, and b) Facebook dataset.

presented in Table II. The first dataset is from the Epinions[1] social network which is a *who-trust-whom* network having $75,879$ users and $508,837$ social relations among them. On the other hand, Facebook[2] is a *friendship* network which possesses $4,039$ nodes and $88,234$ edges.

TABLE II
DATASET STATISTICS

Social Networks	Nodes	Edges
Epinions	$75,879$	$508,837$
Facebook	$4,039$	$88,234$

B. Simulation Setup

In our simulation, a set of Python programs is executed on an Intel (R) Core (TM) i7-3537 CPU @ 2.00GHz, 2.50GHz machine with 4 GB RAM. We apply the *Monte Carlo (MC)* simulation [2] and the expected value of each the parameter is taken for comparison. For simplicity, the seed set S is generated randomly. The values of p is taken randomly from the set $\{0.1, 0.01\}$ [2]. We generate the threshold θ_v for each seed node by Heuristic Individual (HI) model [28] and assume $\alpha = 0.5$ for our simulation. We compare the results with that of existing R-RIM and RLT-RIM models [15], [16], [29].

C. Performance Analysis

In this section, we analyze the performance of the proposed HM algorithm concerning the estimated VM cost, the running time and the ability to meet RIM challenges.

1) The VM Cost: Fig. 2 depicts the relative VM cost estimated by the proposed HM algorithm along with the R-RIM and RLT-RIM models for Epinions and Facebook datasets. The R-RIM model computes the cost randomly from a tentative solution space whereas, the RLT-RIM selects a node randomly, aggregates its influence and then, checks for the seed node's threshold to activate it. On the other hand, the HM model uses an the RIC model to estimate the partial VM cost which is then minimized by the greedy approach. The empirical result shows that the proposed model has approximately $2-3$ times lower VM cost as compared to that of existing models for all the datasets and this is due to the greedy and heuristic optimization.

2) Running Time: The running time comparison of different models is shown in Fig. 3 for both the datasets. In the RLT-RIM algorithm, after selecting each in-neighbor node randomly, it integrates its influence and then, compares with the threshold of the seed node. If the seed node is not activated, the process continues. Thus, the RLT-RIM requires higher running time as compared to the rest two models. On the other hand, the R-RIM model selects seed influencer nodes randomly and thus, has the fastest running time. The simulation unveils that the running time of the HM model is sandwiched between two existing models for both the datasets. The Epinions network has a remarkably higher number of edges and thus, has a higher average degree (d) as compared to the Facebook network. Therefore, the running time for the Epinions dataset is higher than that of the Facebook dataset.

3) Handling RIM Challenges: The proposed HM model handles the RIM challenges in an efficient manner. Our algorithm gets rid of considering BNC issue by employing the RIC model. Again, The NP-Hardness is mitigated by the greedy model. Moreover, the node activation process by the RIC model continues until there is no newly activated node, that is, the stopping condition is set in a better way than the existing models. In the existing models, especially in the RLT-RIM model, the insufficient influence might occur at ever node whether it is a seed node or an intermediate in-neighbor node. However, in the case of the HM model, the insufficient influence might occur only with the seed node and not for any intermediate in-neighbor node and thus, has lower insufficient influence effect.

VI. CONCLUSION

In this paper, we propose a Heuristic Mixed (HM) model to solve the Viral Marketing Cost (VMC) minimization problem. We modify the Independent cascade (IC) model as Reverse IC (RIC) model to use it in reverse order for node activation process. In the final seed activation process, we use the Reverse LT (RLT) model which is a variation of the LT model. Further, we employ a heuristic and the greedy approach jointly to optimize the VM cost. We have shown that the VMC minimization problem is NP-hard and thus, deduce a

[1]https://snap.stanford.edu/data/soc-Epinions1.html
[2]https://snap.stanford.edu/data/egonets-Facebook.html

theoretical performance bound of the greedy solution. Finally, the experimental results show that the proposed model exhibits better VM cost along with good running time as well as meets the RIM challenges in an efficient manner.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2015-0-00567, Development of Access Technology Agnostic Next-Generation Networking Technology for Wired-Wireless Converged Networks).

*Professor Dr. CS Hong is the corresponding author.

References

[1] J.-R. Lee and C.-W. Chung, "A query approach for influence maximization on specific users in social networks," *IEEE Transactions on knowledge and data engineering*, vol. 27, no. 2, pp. 340–353, 2015.

[2] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003, pp. 137–146.

[3] W. Lu and L. V. Lakshmanan, "Profit maximization over social networks," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. IEEE, 2012, pp. 479–488.

[4] H. Zhang, H. Zhang, A. Kuhnle, and M. T. Thai, "Profit maximization for multiple products in online social networks," in *Computer Communications, IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on*. IEEE, 2016, pp. 1–9.

[5] D. T. Nguyen, S. Das, and M. T. Thai, "Influence maximization in multiple online social networks," in *Global Communications Conference (GLOBECOM), 2013 IEEE*. IEEE, 2013, pp. 3060–3065.

[6] Y. Mei, W. Zhao, and J. Yang, "Influence maximization on twitter: A mechanism for effective marketing campaign," in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–6.

[7] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance, "Cost-effective outbreak detection in networks," in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2007, pp. 420–429.

[8] A. Talukder, R. Kamal, A. K. Bairagi, M. G. R. Alam, S. F. Abedin, M. A. Layek, H. T. Nguyen, and C. S. Hong, "Rumors in the social network: Finding the offenders using influence maximization," in *Korean Computer Congress (KCC)*. KIISE, 2015, pp. 1214–1216.

[9] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 199–208.

[10] C. Budak, D. Agrawal, and A. El Abbadi, "Limiting the spread of misinformation in social networks," in *Proceedings of the 20th international conference on World wide web*. ACM, 2011, pp. 665–674.

[11] G. Tong, W. Wu, and D.-Z. Du, "Distributed rumor blocking with multiple positive cascades," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 468–480, 2018.

[12] J. Zhou, Y. Zhang, and J. Cheng, "Preference-based mining of top-k influential nodes in social networks," *Future Generation Computer Systems*, vol. 31, pp. 40–47, 2014.

[13] A. Goyal and L. V. Lakshmanan, "Recmax: Exploiting recommender systems for fun and profit," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 1294–1302.

[14] S. Bhagat, A. Goyal, and L. V. Lakshmanan, "Maximizing product adoption in social networks," in *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, 2012, pp. 603–612.

[15] A. Talukder, M. G. R. Alam, A. K. Bairagi, S. F. Abedin, M. A. Layek, H. T. Nguyen, and C. S. Hong, "A cost optimized influence maximization in social networks," in *2017 IEEE The 19th Asia-Pacific Network Operations and Management Symposium (APNOMS 2017)*. IEEE, 2017.

[16] A. Talukder, M. G. R. Alam, H. T. Nguyen, and C. S. Hong, "A cost optimized reverse influence maximization in social networks," in *2018 IEEE/IFIP Network Operations and Management Symposium (NOMS 2018)*. IEEE, 2018.

[17] A. Talukder, M. G. R. Alam, A. K. Bairagi, S. F. Abedin, H. T. Nguyen, and C. S. Hong, "The grim: Target marketing in social network," in *The International Symposium on Perception, Action, and Cognitive Systems (PACS 2016)*, 2016, pp. 24–25.

[18] P. Domingos and M. Richardson, "Mining the network value of customers," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2001, pp. 57–66.

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

[20] W. Chen, A. Collins, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei, and Y. Yuan, "Influence maximization in social networks when negative opinions may emerge and propagate." in *SDM*, vol. 11. SIAM, 2011, pp. 379–390.

[21] A. Goyal, W. Lu, and L. V. Lakshmanan, "Celf++: optimizing the greedy algorithm for influence maximization in social networks," in *Proceedings of the 20th international conference companion on World wide web*. ACM, 2011, pp. 47–48.

[22] K. Jung, W. Heo, and W. Chen, "Irie: Scalable and robust influence maximization in social networks," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. IEEE, 2012, pp. 918–923.

[23] A. Talukder and C. S. Hong, "Active reverse path based reverse influence maximization in social networks," in *Korea Software Congress (KSC 2017)*. KIISE, 2017, pp. 1203 – 1205.

[24] ——, "Epidemiological reverse influence maximization in social networks with negative influencing," in *Korean Computer Congress (KCC 2018)*. KIISE, 2018, pp. 1277–1279.

[25] E. Horowitz and S. Sahni, *Fundamentals of computer algorithms*. Computer Science Press, 1978.

[26] A. Caprara, H. Kellerer, U. Pferschy, and D. Pisinger, "Approximation algorithms for knapsack problems with cardinality constraints," *European Journal of Operational Research*, vol. 123, no. 2, pp. 333–345, 2000.

[27] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, jun 2014.

[28] A. Talukder, M. G. R. Alam, A. K. Bairagi, S. F. Abedin, H. T. Nguyen, and C. S. Hong, "Threshold estimation models for influence maximization in social network," in *Korean Institute of Information Scientists and Engineers (KIISE)*. KIISE, 2016, pp. 888–890.

[29] A. Talukder, A. K. Bairagi, D. H. Kim, and C. S. Hong, "Reverse path activation-based reverse influence maximization in social networks," *Journal of The Korean Institute of Information Scientists and Engineers (JOK)*, vol. 45, no. 11, pp. 1203–1209, 2018.

Appendix A

TABLE III
LIST OF ABBREVIATIONS.

Abbreviation	Elaboration
CELF	Cost Effective Lazy Forward
HI	Heuristic Individual Threshold model
HM	Heuristic Mixed model
IC	Independent Cascade model
IM	Influence Maximization
LDAG	Local Directed Acyclic Graph
LT	Linear Threshold model
RIC	Reverse Independent Cascade model
RIM	Reverse Influence Maximization
RLT	Reverse Linear Threshold model
RLT-RIM	Randomized RIM under LT model
R-RIM	Random RIM model
VM	Viral Marketing
VMC	Viral Marketing Cost