Popularity Ratio Maximization: Surpassing Competitors through Influence Propagation

HAO LIAO, Shenzhen University, China
SHENG BI, Shenzhen University, China
JIAO WU, Shenzhen University, China
WEI ZHANG, Shenzhen University, China
MINGYANG ZHOU, Shenzhen University, China
RUI MAO, Shenzhen University, China
WEI CHEN*, Microsoft Research, China

In this paper, we present an algorithmic study on how to surpass competitors in popularity by strategic promotions in social networks. We first propose a novel model, in which we integrate the Preferential Attachment (PA) model for popularity growth with the Independent Cascade (IC) model for influence propagation in social networks called PA-IC model. In PA-IC, a popular item and a novice item grab shares of popularity from the natural popularity growth via the PA model, while the novice item tries to gain extra popularity via influence cascade in a social network. The *popularity ratio* is defined as the ratio of the popularity measure between the novice item and the popular item. We formulate *Popularity Ratio Maximization (PRM)* as the problem of selecting seeds in multiple rounds to maximize the popularity ratio in the end. We analyze the popularity ratio and show that it is monotone but not submodular. To provide an effective solution, we devise a surrogate objective function and show that empirically it is very close to the original objective function while theoretically, it is monotone and submodular. We design two efficient algorithms, one for the overlapping influence and non-overlapping seeds (across rounds) setting and the other for the non-overlapping influence and overlapping seed setting, and further discuss how to deal with other models and problem variants. Our empirical evaluation further demonstrates that our proposed method consistently achieves the best popularity promotion compared to other methods. Our theoretical and empirical analyses shed light on the interplay between influence maximization and preferential attachment in social networks.

164

CCS Concepts: • **Information systems** → **Data mining**; *Computational advertising*; **Social networking sites**.

Additional Key Words and Phrases: graph algorithms, preferential attachment, popularity ratio maximization

---

*Contact author of this paper

Authors' addresses: Hao Liao, haoliao@szu.edu.cn, Shenzhen University, Shenzhen, China; Sheng Bi, Shenzhen University, Shenzhen, China, 2060271075@email.szu.edu.cn; Jiao Wu, Shenzhen University, Shenzhen, China, 1800271041@email.szu.edu.cn; Wei Zhang, Shenzhen University, Shenzhen, China, 2210275010@email.szu.edu.cn; Mingyang Zhou, Shenzhen University, Shenzhen, China, zmy@szu.edu.cn; Rui Mao, Shenzhen University, Shenzhen, China, mao@szu.edu.cn; Wei Chen, Microsoft Research, Beijing, China, weic@microsoft.com.

---

1 INTRODUCTION

Influence maximization (IM) is the problem of finding a subset of nodes (seed nodes) in a social network that could maximize the spread of influence [18]. It is a well-studied problem with applications on viral marketing, information propagation monitoring and control, etc (cf. [8, 23]). While most studies treat influence maximization as a stand-alone problem for viral marketing, in this paper, we want to explore the means of using influence maximization to boost the popularity of an item (an idea, a product, etc) and surpass a competitor in a natural growth environment.

Consider the following hypothetical motivating example. Alice opened her new restaurant, Café Alice, at a local shopping mall. However, a popular restaurant Bob's Kitchen is in the same mall. Without promotion, Alice would definitely lose many customers to Bob's Kitchen. Thus Alice wants to distribute free-dish coupons to selected people, hoping them to propagate the information about Café Alice and attract more people to her restaurant, boosting the popularity of her restaurant in a short period of time to catch up or even surpass Bob's Kitchen. Similar situations arise in online social media, such as a new blogger or podcaster trying to increase popularity through promotion in social networks. Different from pure viral marketing, popularity growth has a natural rich-get-richer effect independent of the social network, meaning that the popular item would naturally attract more customers or users. For example, customers typically check nearby restaurants' popularity ratings on mobile apps such as Yelp for restaurant selection, making more people visit the popular restaurant. This effect certainly gives a hard time for the new restaurant to catch up. Thus, incentive promotion by the new restaurant needs to incorporate this factor when deciding on viral marketing strategies.

This paper integrates the natural popularity growth with the incentive promotion into one coherent model to solve the above problem. In particular, the preferential attachment (PA) [1, 28, 36] is a well-known model for the natural rich-get-richer effect of popularity growth, while the independent cascade (IC) model [18] is a classical model for influence propagation. We integrate these two models into the novel PA-IC model to characterize the combined effect of natural popularity growth and incentive promotion. More specifically, in a multi-round setting, we model the popularity growth of a novice item and a popular item. We use d_t^p and d_t^n to denote the popularity of the popular and novice items respectively, at the end of round t. Each round, z number of customers pick from these two items with probability proportional to their popularity, realizing the preferential attachment model. Meanwhile, the novice item executes a promotion plan, selecting a set of seeds S_t in round t for T consecutive rounds, and the influence spread generated by the seed set S_t would add to the popularity measure of the novice item at the end of round t. We define $r_t = d_t^n/d_t^p$ as the *popularity ratio* in round t. The promotion task of the novice item is defined as the following *popularity ratio maximization (PRM)* problem: Given (a) the social network and its IC model parameters, (b) the initial popularity measures d_0^p and d_0^n, and (c) a promotion budget k, find k seeds and allocate them into T consecutive rounds such that the popularity ratio r_T at the end of round T is maximized. Depending on whether we allow overlapping seeds (OS) across rounds and whether repeated activation of the same nodes in different rounds are repeatedly counted (overlapping influence, OI), we further consider different settings such as the overlapping influence and non-overlapping seeds (OINS) and the non-overlapping influence and overlapping seeds (NIOS).

For both OINS and NIOS settings, we derive the formula for the popularity ratio and show that it is monotone but not submodular, indicating that direct optimization on this objective function may be complex. To provide an effective solution, we simplify the objective function and obtain a surrogate function and show that empirically it is very close to the original objective function while theoretically, it is monotone and submodular. Based on the reverse influence sampling

approach [3] and the Influence Maximization via Martingales(IMM) algorithm [32], we design PRM-IMM algorithms to solve this surrogate problem for the two settings, and prove their theoretical guarantees. We further discuss how to handle other models and problem variants, such as when the popular item also conducts promotion, and when the novice item wants to use the minimum budget to surpass the popular item in a given time limit. Our empirical evaluation demonstrates that the proposed PRM-IMM method always achieves the best popularity promotion compared to other methods.

One important point we demonstrate both analytically and empirically is that the seed allocation of the novice item is tightly dependent on the natural growth of the environment as well as the popularity of the popular item. This point indicates that the PRM task is quite different from the classical influence maximization — the promotion of the novice item has to consider the popularity growth and its relative position against the popular item, while the classical influence maximization only targets at increasing one's own influence spread or popularity.

In summary, our contributions include: (a) proposing the PA-IC model that integrates preferential attachment with independent cascade propagation model; (b) formulating the PRM problem and studying its properties; and (c) designing an efficient algorithm to solve the PRM problem, and discussing the extension to other problem variants. To the best of our knowledge, this is the first study that integrates preferential attachment and an influence propagation model into a coherent popularity growth model and provides an algorithmic study on such an optimization problems.

2 RELATED WORKS

Preferential Attachment. Yule firstly considered using the preferential attachment to explain the power-law distribution of flowering plants [36]. A clearer and more general development of how preferential attachment leads to a power law was given by Simon [28]. The Barabási-Albert (BA) model shows that the power-law degree distribution in the real networks can be produced by the combination of growth and preferential attachment [1]. Preferential attachment is now the standard model for the rich-get-richer effect.

Influence Maximization. Domingos and Richardson first study influence maximization (IM) [12, 27]. Then it is formulated as an optimization problem by Kempe et al. [18], who define the famous independent cascade and linear threshold models, showing that influence maximization under these models are NP-hard, and connecting the problem to monotone and submodular function maximization. Influence maximization has been extensively studied (see surveys [8, 23]). One direction is to improve the scalability of the algorithms [3, 9–11, 14, 17, 19, 21, 35]. The state-of-the-art scalable solution is based on the reverse influence sampling (RIS) approach first proposed by Borg et al. [3], and later improved and refined by a series of studies [6, 26, 31–33]. Our algorithms are also based on the RIS approach and is adapted from the IMM algorithm of [32].

Another related direction is competitive influence maximization for multiple items [4, 8, 15, 16, 22, 24]. However, all these studies focus on the competitive diffusion of multiple items in social networks. In contrast, the competition in our model stems from the popularity growth dictated by preferential attachment, and influence maximization is only a tool for the novice item to increase popularity. To the best of our knowledge, we are the first to consider popular growth via both the PA effect and network diffusion, and study their nontrivial interplay in an optimization problem.

3 MODEL AND PROBLEM DEFINITION

In this section, we introduce the PA-IC model, which integrates the preferential attachment (PA) mechanism for popularity growth with the independent cascade (IC) model for influence propagation. The PA mechanism models show two items, a popular item and a novice item, divide the shares of popularity from the natural growth of the customer base, while the novice item further utilizes

the IC model propagation to promote its popularity in order to catch up with the popular item. Based on the PA-IC model, we define the optimization problem of *popularity ratio maximization*, which characterizes how to allocate the promotional budget of the novice item to achieve the best promotion results.

3.1 PA-IC Model

Preferential attachment (PA) [1] is a well-known model to characterize the rich-get-richer effect in the growth of networks or popularity. The basic form of the PA model is as follows. Let d_t^i be the popularity measure of item i at time t. When a number of z users come at time t, each user selects item i with probability proportional to d_t^i. Thus popularity grows proportionally, i.e., $d_{t+1}^i = d_t^i + z \cdot d_t^i / \sum_j d_t^j$, where z is the natural growth parameter.

Beyond the natural growth of popularity governed by the PA mechanism, we would like to incorporate the influence propagation model for popularity promotion. Independent cascade (IC) model [18] is a classical diffusion model widely adopted in the information and influence propagation literature, and its parameters can be effectively learned in many applications (e.g. [2, 8]). In the IC model, a social network is modeled as a directed graph $G = (V, E, p)$, where V is the set of $N = |V|$ nodes representing individuals, E is the set of $M = |E|$ directed edges representing influence relationships between pairs of individuals, and $p : E \rightarrow (0, 1]$ gives the influence probability on every edge, i.e., $p(u, v)$ is the influence probability on edge $(u, v) \in E$. Nodes have two states, active and inactive. Nodes that have been activated will always remain active. The propagation process starts from a seed set $S \subseteq V$ in discrete time steps. At step $\tau = 0$, only the nodes in S are active. At step $\tau \geq 1$, a node u activated at step $\tau - 1$ attempts to activate each of its inactive out-neighbor v with success probability $p(u, v)$. Propagation terminates when no more nodes are activated. A key measure is the *influence spread* of the seed set S, denoted as $\sigma(S)$, which is the expected number of activated nodes at the end of propagation starting from the seed set S. The IC model has the following equivalent live-edge graph description. A random *live-edge graph* L is sampled from $G = (V, E, p)$ such that every edge $(u, v) \in E$ has an independent probability of $p(u, v)$ to be included in L. Then given a seed set S, the set of nodes activated in a stochastic diffusion from S is the set of nodes reachable from S in the live-edge graph L, denoted as $\Gamma(S, L)$. Therefore, we have $\sigma(S) = \mathbb{E}_L[|\Gamma(S, L)|]$.

We integrate the PA model with the IC model in the following PA-IC model to characterize both the natural growth and the promotional growth of the popularity. For simplicity, this paper focuses on two competing items: a popular item and a novice item, since all items grow proportionally in the PA model.

The popular item starts with a higher popularity measure, while the novice item starts with a lower one. If only natural growth is available, the novice item would never catch up with the popular item, and the gap would only be widened due to the rich-get-richer effect of the PA model. To catch up and surpass the popular item, the novice item needs to employ incentive promotion in the social network to increase its popularity. The technical description of the PA-IC model to cover the above aspects is given below.

In the PA-IC model, we model T rounds of promotion and influence propagation. Let d_t^p and d_t^n denote the popularity measure of the popular and novice items respectively, at the end of round t. In each round $t \in [T]$, some customers would naturally select between the novice and popular item. We use the natural growth parameter z to denote the number of customers. These z customers select between novice and popular items according to the PA model.

Besides the natural growth, in each round t, the novice item would select a seed set S_t for its promotion, and these seeds will propagate influence throughout the social network G following the
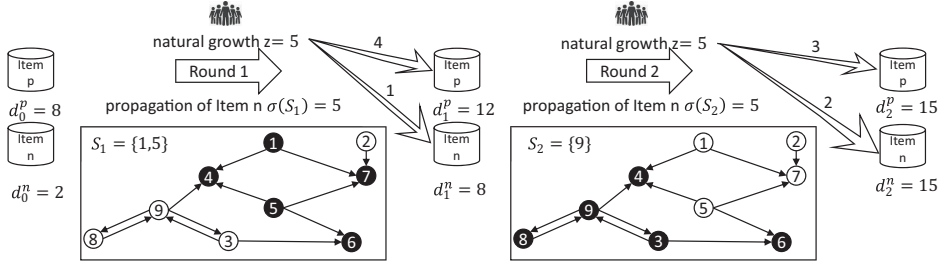
Fig. 1. A numerical example of the PA-IC model with the overlapping influence setting. In this example, we show the changes in popularity measure of two items, within two rounds. In this directed graph, every edge's activation probability is equal to 1. Black nodes is the "active" nodes in each round.

IC model. Note that in round t, the IC model may take multiple *steps* to propagate the promotion until the propagation ends. The expected number of influenced nodes $\sigma^{\square}(S_t)$ is also added to the popularity measure of the novice item, where the actual form of $\sigma^{\square}(S_t)$ depends on the setting we consider, and it will be explained shortly after Eq.(1) below.

The natural growth mechanism and propagation mechanism respectively model two different growth processes of an item's popularity in reality. Therefore, the natural growth mechanism and the propagation mechanism together give the following inductive formulation of the popularity growth of the two items.

$$d_t^p = d_{t-1}^p + \frac{z \cdot d_{t-1}^p}{d_{t-1}^p + d_{t-1}^n}, d_t^n = d_{t-1}^n + \sigma^{\square}(S_t) + \frac{z \cdot d_{t-1}^n}{d_{t-1}^p + d_{t-1}^n}. \tag{1}$$

In Eq.(1), the notation $\sigma^{\square}$ needs to be instantiated to a concrete quantity according to the actual setting used. We mainly consider two settings — the overlapping influence (OI) and the non-overlapping influence (NI) settings. In the overlapping influence (OI) setting, the influence spread between different rounds are allowed to be overlapped when they are added to the popularity measure. In this case, $\sigma^{\square}(S)$ is instantiated to $\sigma^{OI}(S)$, and is simply the standard influence spread $\sigma(S)$. In the non-overlapping Influence (NI) setting, the influence spread of S_t in round t cannot overlap with the influence spread in the previous rounds, and thus it is the marginal influence spread of S_t given the seed sets $S_1, \ldots, S_{t-1}$ of the previous rounds. Hence, in this case $\sigma^{\square}(S)$ is instantiated to $\sigma^{NI}(S_t|S_1, \ldots, S_{t-1})$, as defined below:

$$\sigma^{NI}(S_t|S_1 \cdots S_{t-1}) = \mathbb{E}_{L_1,\ldots,L_t} \left[\left\| \Gamma(S_t, L_t) \setminus \bigcup_{i=1}^{t-1} \Gamma(S_i, L_i) \right\| \right]. \tag{2}$$

where L_i, $i \in [t]$ is the random live-edge graph in round i. When the context is clear, we use $\sigma^{NI}(S_t)$ as a shorthand for $\sigma^{NI}(S_t|S_1 \cdots S_{t-1})$.

One remark to Eq.(1) is that both the natural growth and the social network promotion parts are represented in the expectation form. One may formulate them as random variables, and take the expectation in the end. The expectation form is easier to handle, and our empirical evaluation would show that such representation does not lose fidelity in terms of the solution quality.

Another remark concerns the distinction of the natural growth and promotion in our model. The natural growth follows the preferential attachment and is independent of the social network, while the promotion follows the influence diffusion model and is dependent on the social network. One may say that the natural growth may also generate influence diffusion in social networks. The

Table 1. Four settings of the PRM problem

	Overlapping Influence	Non-overlapping Influence
Overlapping Seeds	OIOS	NIOS
Non-overlapping Seeds	OINS	NINS

reason we do not include the social network diffusion for the natural growth part of the model is that (a) the natural growth mainly reflects the rich-get-richer effect and models the behavior that people gets the popularity information from online review platforms such as Yelp or direct offline observations (e.g., by observing the occupancy of the restaurants) and act accordingly, and this part is not related to social network diffusion; and (b) inclusion of the influence diffusion for the natural growth part will further complicate the model, making it less clear in demonstrating the interaction between the preferential attachment for natural growth and the influence diffusion for viral promotion.

Fig. 1 is a numerical example to illustrate our PA-IC model in the OI setting. There are two items in this social network: item p is the popular item with the initial popularity measure $d_0^p = 8$, item n is the novice item with the initial popularity measure $d_0^n = 2$, and the natural growth $z = 5$. All edges in the social network have probability 1. In round 1, according to the PA mechanism, the popularity measure of item p will increase by 4, and the popularity measure of item n will increase by 1. At the same time, the novice item chooses the seed set $S_1 = \{1, 5\}$ to promote their product and the influence spread $\sigma(S_1) = 5$. So the popularity measure of item n will additionally increase by 5. At the end of round 1, $d_1^p = d_0^p + \frac{z \cdot d_0^p}{d_0^p + d_0^n} = 12$, $d_1^n = d_0^n + \frac{z \cdot d_0^n}{d_0^p + d_0^n} + \sigma(S_1) = 8$. In round 2, according to the PA mechanism, the popularity measure of item p will increase by 3, and the popularity measure of item n will increase by 2. In this round, item n chooses the seed set $S_2 = \{9\}$ to promote their product, which generates influence spread $\sigma(S_2) = 5$. So at the end of round 2, $d_2^p = d_2^n = 15$, which means that the novice item has caught up with the popular item.

Finally, the model can be further extended to allow different natural growth count a_t for each round t, or allow the popular item to also have a viral promotion mechanism. We defer the discussion of these extensions and their impacts on our algorithm to Section 6.

3.2 Popularity Ratio Maximization

We define the *popularity ratio* between the novice and the popular item at the end of round t as $r_t = d_t^n / d_t^p$. Without the promotional mechanism, this ratio would not change. Thus, from the novice item's perspective, it wants to increase this ratio as much as possible and as soon as possible through promotion, but it certainly has budget constraints.

Technically, we model this as a popularity ratio maximization problem under a budget constraint. We use the pair (v, t) to denote that the novice item selects v in round t as a seed, i.e., $v \in S_t$, and we use $\mathcal{S} = S_1 \times \{1\} \cup S_2 \times \{2\} \cdots \cup S_T \times \{T\}$ to represent the overall allocation of seeds over T rounds. The novice item has a total budget k, restricting the total number of seeds that it can select, i.e., $|\mathcal{S}| \leq k$. Let the final popularity ratio after T rounds promotion be $r_T(\mathcal{S}) = d_T^n / d_T^p$.

Informally, popularity ratio maximization (PRM) is to find a seed allocation $\mathcal{S}$ with $|\mathcal{S}| \leq k$ for the novice item so that popularity ratio $r_T(\mathcal{S})$ at the end of round T is maximized. In the seed allocation $\mathcal{S}$, the seed set for different round may or may not be allowed to overlap, and both may be reasonable depending on the application scenario. Thus, together with the overlapping or non-overlapping influence settings, there are four settings as shown in Table 1, where OIOS refers

to overlapping influence and overlapping seeds and NINS refers to non-overlapping influence and non-overlapping seeds. The other two settings have been defined before. Formally, we define the PRM problem below based on these settings.

Definition 1 (Popularity Ratio Maximization). *Given (a) a social network* $G = (V, E, p)$, *(b) initial popularity measures* d_0^n *and* d_0^p *for the novice and popular items respectively, (c) total round number* T, *(d) natural growth count* z, *(e) budget* k, *the task of* popularity ratio maximization (PRM) *is to find an optimal seed set allocation* $\mathcal{S}^*$ *for the novice item in one of the four settings depicted in Table 1, such that the total number of seeds does not exceed* k, *and when the popularity measures evolve according to Eq.(1), the final popularity ratio* $r_T(\mathcal{S}^*)$ *is maximized. For the OINS setting, it is*

$$\mathcal{S}^* \in \underset{\mathcal{S} \subseteq V \times [T], |\mathcal{S}| \le k, S_t \cap S_{t'} = \emptyset \, \forall t \ne t'}{\mathrm{argmax}} r_T(\mathcal{S}).$$

Other settings can be similarly formulated.

Note that it may not be wise to use all the budget in the first round since too many seeds could generate redundant activations of the same node, which is only counted once in the popularity measure. On the other hand, spreading the budget evenly across all T rounds may not be a good choice either, since it would make the popularity growth of the novice item slower and thus harder to catch up with the popular item. Moreover, even though the objective is to increase the popularity of the novice item, the problem is still tightly related to the popularity of the popular item, because the PA mechanism links them together. Therefore, PRM is non-trivial both in considering seed allocation across multiple rounds and in considering the relationship with the popular item. Furthermore, PRM needs to consider the impact of different settings. In the following two sections, we will consider two representative settings — overlapping influence and non-overlapping seeds (OINS) and non-overlapping influence and overlapping seeds (NIOS). Other settings can be treated similarly.

PRM can be extended to other variants, and we will consider a number of them in Section 6.

4   RESULTS ON THE OINS SETTING

We study the OINS setting in this section. We first derive the exact formula for the objective function of the PRM problem. We show that the exact objective function is not submodular, implying a rather difficult optimization task. To tackle this problem, we provide a heuristic simplification of the objective function, resulting in a round-weighted influence objective function $\rho_T(\mathcal{S})$, and its corresponding surrogate optimization task called *round-weighted influence maximization (RWIM)*. We show that function $\rho_T(\mathcal{S})$ is monotone and submodular, which allows us to design an efficient approximation algorithm PRM-IMM based on the IMM algorithm [32]. The technical novelty of PRM-IMM includes properly defining the pair-wise reverse reachable set (PW-RR set) to estimate the seed set's influence in the different rounds with different weights, and calculating the number of PW-RR sets needed to satisfy the approximate ratio. Although our approximation guarantee is on the surrogate RWIM algorithm, our experimental evaluation will demonstrate that it also solves the original PRM problem with the best performance compared to other baselines.

4.1   Objective Function under the OINS Setting

The following lemma states the exact formula for the objective function $r_T^{OI}(\mathcal{S})$ of PRM under the overlapping influence (OI) setting, which is derived recursively.

LEMMA 1. *For the OINS setting, the popularity ratio function at the end of round T is:*

$$r_T^{OI}(\mathcal{S}) = (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t)}{d_0^n + d_0^p + z \cdot t + \sum_{i=1}^{t-1} \sigma(S_i)} \right) - 1, \tag{3}$$

where $r_0 = d_0^n / d_0^p$ is the initial popularity ratio, $\mathcal{S} = \bigcup_{t=1}^{T} S_t \times \{t\}$.

PROOF. We show Eq.(3) by deriving the recursive formula for $r_t, t \geq 1$. Suppose at the end of round $t-1$ the popularity ratio is r_{t-1}, and the popularity measures for the novice and popular items are d_{t-1}^n and d_{t-1}^p, respectively. By definition, we have

$$r_t + 1 = \frac{d_t^n + d_t^p}{d_t^p} = \frac{d_{t-1}^p + d_{t-1}^n + \sigma^{OI}(S_t) + z}{d_{t-1}^p + z \cdot \frac{d_{t-1}^p}{d_{t-1}^p + d_{t-1}^n}} \tag{4}$$

$$= \frac{d_{t-1}^p + d_{t-1}^n + \sigma^{OI}(S_t) + z}{d_{t-1}^p + d_{t-1}^n + z} \cdot \frac{d_{t-1}^p + d_{t-1}^n}{d_{t-1}^p}$$

$$= \left(1 + \frac{\sigma^{OI}(S_t)}{d_0^n + d_0^p + z \cdot t + \sum_{i=1}^{t-1} \sigma^{OI}(S_i)} \right) (r_{t-1} + 1), \tag{5}$$

where the second equality in (4) is from Eq.(1), and Eq.(5) uses the fact that $d_{t-1}^p + d_{t-1}^n$ is the total popularity measure of the two items at the end of round $t-1$, which is accumulated from $d_0^n + d_0^p$ by $t-1$ rounds of natural growth ($t \cdot z$) and the promotion effect ($\sum_{i=1}^{t-1} \sigma^{OI}(S_i)$). Eq. (3) can be obtained immediately with the above recursive formula and the fact that $\sigma^{OI}(S_i) = \sigma(S_i)$.   □

One feature of Eq. (3) is that the popularity ratio depends on the interplay among the parameters of the novice item (d_0^n), the popular item (d_0^p), the natural growth (z) and the total influence effect $\sigma(S_t)$, which makes the seed selection more complicated in the PRM task than the classical influence maximization task.

A set function $f : 2^V \to \mathbb{R}$ is monotone if $f(S) \leq f(Q)$ whenever $S \subseteq Q \subseteq V$, and f is submodular if, for any $S \subseteq Q \subseteq V$ and any $x \in V \backslash Q$, $f(S \cup \{x\}) - f(S) \geq f(Q \cup \{x\}) - f(Q)$. Submodularity means that the marginal value of an element with respect to a set S decreases as S grows. Maximizing a monotone and submodular set function with a cardinality constraint can be achieved by a simple greedy algorithm with an approximation ratio of $1 - 1/e$ [25]. Unfortunately, the exact popularity ratio function in Eq. (3) is not submodular, as shown below.

LEMMA 2. *The popularity ratio function $r_T^{OI}(\mathcal{S})$ is monotone but not submodular.*

The proof is omitted, and the non-submodularity is due to the multiplication of submodular functions in Eq.(3).

4.2 Round-Weighted Influence Maximization

Due to the non-submodularity of the popularity ratio function, it is challenging to design a good algorithm for the PRM problem directly. In this section, we simplify the popularity ratio function to obtain a surrogate function and show that this surrogate function is monotone and submodular.

As indicated in the proof sketch of Lemma 2, the complication of the original objective function of Eq.(3) is the series of multiplications, and thus this is the target of our simplification. In particular, our simplification consists of two steps: (a) expanding the multiplication series of Eq.(3) and only keeping the first-order terms; (b) removing the $\sigma(S_1), \ldots, \sigma(S_{T-1})$ in the denominator of each term left after step (a). Step (a) decreases the objective value, while step (b) increases the objective

value, compensating the effect of step (a). Our experimental evaluation (Section 7.2) shows that the surrogate function we develop is reasonably close to the original function in all the test cases. The following is the formula for the surrogate function (ignoring the constant factor $r_0 + 1$ and the additive terms in Eq.(3)):

$$\rho_T^{OI}(\mathcal{S}) = \frac{\sigma(S_1)}{d_0^n + d_0^p + z} + \frac{\sigma(S_2)}{d_0^n + d_0^p + 2z} + \cdots + \frac{\sigma(S_T)}{d_0^n + d_0^p + T \cdot z}.$$

In the above formula, we can set

$$w_t = \frac{1}{d_0^n + d_0^p + t \cdot z} \tag{6}$$

as the weight for $\sigma(S_t)$ for every round t, and turn the objective function into the following more general objective function:

$$\rho_T^{OI}(\mathcal{S}) = \sum_{t=1}^{T} w_t \cdot \sigma^{OI}(S_t) = \sum_{t=1}^{T} w_t \cdot \sigma(S_t). \tag{7}$$

Function $\rho_T^{OI}(\mathcal{S})$ means that we need to find an allocation of seeds of T rounds and each round $t \in [T]$ has its own importance weight w_t, which is non-negative and decreasing, i.e., $w_1 \geq w_2 \geq \cdots \geq w_T \geq 0$. We call $\rho_T^{OI}(\mathcal{S})$ the *round-weighted influence function*.

The relationship between the popularity ratio function and round-weighted influence function is $r_T^{OI}(\mathcal{S}) \approx (1 + \rho_T^{OI}(\mathcal{S}))(r_0 + 1) - 1$. The surrogate function is summarized below as the round-weighted influence maximization problem.

Definition 2 (Round-weighted Influence Maximization in the OINS Setting). *Given (a) a social network $G = (V, E, p)$, (b) initial popularity measures d_0^n and d_0^p for the novice and popular items respectively, (c) total round number T, (d) natural growth count z, (e) budget k, (f) the non-negative and non-increasing weight sequence $w_1, w_2, \ldots, w_T$, the task of* round-weighted influence maximization (RWIM) *in the OINS setting is to find an optimal seed set allocation $\mathcal{S}^*$ for the novice item under the constraints that the seeds selected for each round are disjoint, and the total number of seeds does not exceed k, such that the round-weighted influence $\rho_T^{OI}(\mathcal{S}^*)$ is maximized. That is,*

$$\mathcal{S}^* \in \underset{\mathcal{S} \subseteq V \times [T], |\mathcal{S}| \leq k, S_t \cap S_{t'} = \emptyset \forall t \neq t'}{\operatorname{argmax}} \rho_T^{OI}(\mathcal{S}).$$

Lemma 3. *For any influence graph $G = (V, E, p)$, under the PA-IC model, the round-weighted influence function $\rho_T^{OI}(\mathcal{S})$ defined in Eq.(7) is monotone and submodular.*

4.3 Pair-wise Reverse Reachable Sets

State-of-the-art IM algorithms, including IMM, use the reverse influence sampling (RIS) approach governed by reverse reachable (RR) sets. An RR set is a random set of nodes sampled from the graph by (a) first selecting a node v uniformly at random from the graph, and (b) simulating the reverse propagation of the model (e.g., IC model) and adding all visited nodes into the RR set. The main property of a random RR set R is that: influence spread $\sigma(S) = N \cdot \mathbb{E}\left[\mathbb{I}\{S \cap R \neq \emptyset\}\right]$ for any seed set S, where $\mathbb{I}$ is the indicator function, N is the number of nodes in graph G. After finding a large enough number of RR sets, the original influence maximization problem is turned into a k-max coverage problem — finding a set of k nodes that covers the most number of RR sets, where a set S covers an RR set R means that $S \cap R \neq \emptyset$. All RIS algorithms have the following two steps: 1. Generate a sufficiently large set of random RR sets, called Sampling step. 2. Find k nodes that cover the most number of RR sets, called NodeSelection step.

We now introduce our adaptation of the RR set to the PRM setting, particularly, the pair-wise reverse reachable (PW-RR) set.

PW-RR set and its generation process. Since the RR set is not designed to estimate the influence of nodes in different rounds, we define the pair-wise reverse reachable (PW-RR) set for the PRM problem and denote it as $R^{(t)}$. PW-RR can distinguish between different rounds of RR set, and it also corresponds to different weights for different rounds of RR sets so that the round-weighted influence of nodes in different rounds can be calculated. A (random) pair-wise RR set $R^{(t)}$ is an RR set R rooted at a node picked uniformly at random from V, and t is picked uniformly at random from $[T]$, and $R^{(t)} = R \times \{t\}$.

Property of PW-RR set The main property of random PW-RR set $R^{(t)}$'s is that we can use them to estimate the round-weighted influence spread $\rho_T^{OI}(S)$.

LEMMA 4. *For any set of node-round pairs S, we have*

$$\rho_T^{OI}(S) = N \cdot T \cdot \mathbb{E}\left[w_t \cdot \mathbb{I}\left[S \cap R^{(t)} \neq \emptyset\right]\right], \tag{8}$$

where $R^{(t)}$ is a random PW-RR set, and the expectation is taken from the randomness of (a) the root of a PW-RR set $R^{(t)}$ uniformly chosen at random, (b) the randomly generated RR set from the root, and (c) round number t uniformly chosen at random.

PROOF. The lemma is shown by the following:

$$\mathbb{E}\left[w_t \cdot \mathbb{I}\left[S \cap R^{(t)} \neq \emptyset\right]\right] = \frac{1}{T}\sum_{t=1}^{T} w_t \cdot \mathbb{E}\left[\mathbb{I}\{S_t \cap R \neq \emptyset\}\right] \tag{9}$$

$$= \frac{1}{T}\sum_{t=1}^{T} w_t \cdot \frac{1}{N}\sigma\left(S_t\right) = \frac{1}{N \cdot T}\rho_T^{OI}(S), \tag{10}$$

where the R in Eq.(9) is a random RR set as defined at the beginning of this section, the first equality in (10) is by the property of the RR set as given at the beginning of this section, and the second equality in (10) is by Eq.(7).

□

Lemma 4 reflects the overlapping influence (OI), and it is clearly seen in Eq.(10) where influence spread of different rounds are summed up, allowing overlapping influence.

4.4 Popularity Ratio Maximization IMM

We could base the key formula of Eq.(8) to estimate the overlapping influence spread $\rho_T^{OI}(S)$. Let $\mathcal{R}$ be a collection of θ PW-RR sets we generated independently, $\mathcal{R} = \{R_1^{(t_1)}, R_2^{(t_2)}, \ldots, R_\theta^{(t_\theta)}\}$. Define random variable $Y_i^{\mathcal{R}}(S) = w_{t_i} \cdot \mathbb{I}\left[S \cap R_i^{(t_i)} \neq \emptyset\right]$. We can estimate $\rho_T^{OI}(S)$ using the following estimator $\hat{\rho}_T^{OI}(S, \mathcal{R})$:

$$\hat{\rho}_T^{OI}(S, \mathcal{R}) = \frac{N \cdot T}{\theta}\sum_{i=1}^{\theta} Y_i^{\mathcal{R}}(S).$$

Estimator $\hat{\rho}_T^{OI}(S, \mathcal{R})$ can be viewed as a weighted coverage function on a bipartite graph with $V \times [T]$ on the one side and PW-RR sets $\mathcal{R}$ on the other side.

Our PRM-IMM algorithm follows the IMM structure [32] and consists of two components: PRM-NodeSelection for selecting top k pairs from the above bipartite graph, and PRM-Sampling for generating enough PW-RR set samples. We now explain the main adaption of these two components to our PRM settings.

Algorithm 1 PRM-NodeSelection-OINS

Input: PW-RR sets $\mathcal{R}$, budget k, weights $(w_t)_{t \in [T]}$ (Eq.(6))

Output: seed set $\hat{\mathcal{S}}^g$

1: $RR[(v,t)] = \{R \in \mathcal{R}|(v,t) \in R\}$, $c[(v,t)] = w_t \cdot |RR[(v,t)]|$, $\forall(v,t) \in V \times [T]$ /* $RR[(v,t)]$'s and $c[(v,t)]$'s can be constructed during the generation of $\mathcal{R}$ */

2: for all $R \in \mathcal{R}, t \in [T]$, $covered[R] = false$

3: $\mathcal{S} = \emptyset$ /* the element of $\mathcal{S}$ is (v,t) */

4: **for** $i = 1$ to k **do**

5:    $(v,t) = \text{argmax}_{(v,t) \in (V \setminus \mathcal{S}_{\text{node}}) \times [k]}\, c[(v,t)]$

6:    /*$\mathcal{S}_{node}$ is the node set of $\mathcal{S}$ */

7:    $\mathcal{S} = \mathcal{S} \cup \{(v,t)\}$

8:    **for** $R \in RR[(v,t)] \wedge covered[R] == false$ **do**

9:       $covered[R] = true$

10:      **for** $(u,t) \in R \wedge u \neq v$ **do**

11:         $c[(u,t)] = c[(u,t)] - w_t$

12:      **end for**

13:    **end for**

14: **end for**

15: **return** $\hat{\mathcal{S}}^g = \mathcal{S}$

PRM-NodeSelection-OINS Algorithm 5 provides the pseudocode for the PRM-NodeSelection under the OINS setting, which is patterned from the similar algorithms in [7, 32]. It mainly implements a greedy selection of k pairs to cover the PW-RR sets. Several points worth mentioning are: (a) the greedy algorithm works on PW-RR sets for (v,t) pairs, as described in Section 4.3; (b) the overlapping influence (OI) aspect is reflected in the PW-RR set design and Lemma 4; and (c) the non-overlapping seed (NS) aspect is reflected in line 5, where nodes already selected, $\mathcal{S}_{\text{node}}$, are removed from further consideration in seed selection.

Due to the NS requirement, the feasible solutions for the bipartite coverage problem actually form a partition matroid.[1] The greedy algorithm on the partition matroid gives a $1/2$ approximation of the optimal solution.[2]

PRM-Sampling The sampling procedure also follows the IMM structure (Algorithm 2), in that it uses iterative halving to estimate a lower bound LB of the optimal solution OPT, and then use LB to determine the final number of PW-RR sets needed. The parameters α, β are defined below:

$$\alpha = \sqrt{l \ln N + \ln 4}, \quad \beta = \sqrt{\frac{1}{2} \cdot \left(\ln \binom{N}{k} + \alpha^2 + k \ln T \right)}. \tag{11}$$

There are a few differences compared to the classical IMM. First, the θ_i (line 4) and $\tilde{\theta}$ (line 12) have additional coefficients w_1 and T in the numerator. Factor T is because we allocate seeds in T rounds. Factor w_1 is technically because random variable $Y_i^{\mathcal{R}}(\mathcal{S})/w_1$ is bounded within $[0, 1]$, so

---

[1]A matroid $(E, \mathcal{I})$ on a finite ground element set E is a collection $\mathcal{I}$ of subsets of E, each called an independent set, with the following two properties: (a) a subset of an independent set is still an independent set; and (b) for two independent sets $I_1, I_2 \in \mathcal{I}$, if $|I_1| < |I_2|$, then there is at least one element $e \in I_2 \setminus I_1$ such that $I_1 \cup \{e\}$ is also an independent set. A partition matroid is a matroid where the ground set E is partitioned into disjoint sets $E_1, E_2, \ldots, E_s$, and each E_i has a capacity constraint b_i such that for all $I \in \mathcal{I}, |I \cap E_i| \leq d_i$.

[2]We are aware that there is a polynomial-time algorithm that achieves $(1 - 1/e)$-approximation for submodular maximization under a matroid constraint [5], but the algorithm is highly inefficient, and thus we choose the simple greedy algorithm for our implementation.

Algorithm 2 PRM-IMM-OINS algorithm

Input: directed graph $G = (V, E)$, round number T, budget k,
weights $(w_t)_{t \in [T]}$ (Eq.(6)), approximate ratio parameter ε,
Error probability parameter l.

Output: $\hat{S}^g$

1: $\mathcal{R} = \emptyset, LB = w_1, \varepsilon' = \sqrt{2}, \theta_0 = 0$
2: **for** $i = 1$ to $\lfloor \log_2 N \rfloor - 1$ **do**
3:      $x_i = \frac{\sum_{t=1}^{T} w_t \cdot N}{2^i}$
4:      $\theta_i = \left\lceil \frac{w_1 \cdot N \cdot T \left(2 + \frac{2}{3} \varepsilon'\right) \left(\ln \log_2 N + 2\beta^2\right)}{\varepsilon'^2 x_i} \right\rceil$
5:      independently generate $\theta_i - \theta_{i-1}$ PW-RR sets into $\mathcal{R}$:
6:      $\mathcal{S} = \text{PRM-NodeSelection-OINS}(\mathcal{R}, k, (w_t)_{t \in [T]})$
7:      **if** $\hat{\rho}_T(\mathcal{S}, \mathcal{R}) \geq (1 + \varepsilon') x_i$ **then**
8:          $LB = \frac{\hat{\rho}_T(\mathcal{S}, \mathcal{R})}{1 + \varepsilon'}$
9:          **break**
10:     **end if**
11: **end for**
12: $\tilde{\theta} = \left\lceil \frac{2 w_1 \cdot N \cdot T \left(\frac{1}{2} \alpha + \beta\right)^2}{LB \cdot \varepsilon^2} \right\rceil$ /* α and β is defined in Eq.(11).*/
13: clear the $\mathcal{R}$, and regenerate $\tilde{\theta}$ PW-RR sets into $\mathcal{R}$
14: $\hat{S}^g = \text{PRM-NodeSelection-OINS}(\mathcal{R}, k, (w_t)_{t \in [T]})$
15: return $\hat{S}^g$

---

that we can apply the Chernoff bound on it to obtain the desired result. Second, the parameter β also adds an extra $k \ln T$, because in our problem, the number of possible solutions is upper bounded by $\binom{N}{k} \cdot T^k$, which means that k nodes are selected among N nodes and each selected node may be placed in any of the T rounds. Of course, our approximation guarantee is also changed from $1 - 1/e - \varepsilon$ to $1/2 - \varepsilon$, because we are dealing with a more general partition matroid constraint. The regeneration of PW-RR sets in line 13 ensures that all PW-RR sets in $\mathcal{R}$ are mutually independent [6].

The following theorem summarizes the correctness and time complexity of our algorithm PRM-IMM.

THEOREM 5. *Let $\hat{S}^g$ be the output of PRM-IMM-OINS and S^* be the optimal solution of round-weighted influence maximization problem. We have that with the probability at least $(1 - \frac{1}{N^l})$,*

$$\rho_T^{OI}(\hat{S}^g) \geq \left(\frac{1}{2} - \varepsilon\right) \rho_T^{OI}(S^*),$$

where $\varepsilon > 0, l > 0$. The expected running time of PRM-IMM-OINS is $O((k + l)(M + N)T \log(N \cdot T) / \varepsilon^2)$.

The proof of the theorem follows the analysis structure of IMM [32], and utilizes the key result in Lemma 4.

We can see that PRM-IMM-OINS provides the approximation guarantee while still runs in time near linear to the graph size. Compared to the original IMM, the running time is increased by a factor of T, caused by the spreading of search of seeds across T rounds rather than in one round.

5 RESULTS ON THE NIOS SETTING

We now turn to the non-overlapping influence with overlapping seeds (NIOS) setting. The NI setting causes important changes to the objective function and the algorithm design, as we will describe in this section.

5.1 Objective Function under the NIOS Setting

Similar to Eq.(3) for the OINS setting, we have

$$r_T^{NI}(\mathcal{S}) = (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma^{NI}(S_t)}{d_0^n + d_0^p + z \cdot t + \sum_{i=1}^{t-1} \sigma^{NI}(S_i)}\right) - 1.$$

The main difference is that it uses the marginal influence $\sigma^{NI}(S_t)$ as defined in Eq.(2), because influence cannot be overlapping under NIOS. With a similar heuristic simplification, we obtain the round-weighted influence function with w_t defined in (6):

$$\rho_T^{NI}(\mathcal{S}) = \sum_{t=1}^{T} w_t \cdot \sigma^{NI}(S_t).$$

One can show that $\rho_T^{NI}(\mathcal{S})$ has the following equivalent form:

$$\rho_T^{NI}(\mathcal{S}) = \sum_{v \in V} \mathbb{E}_{L_1,\ldots,L_T} \left[\max_{t \in [T]} w_t \cdot \mathbb{I}\{v \in \Gamma(S_t, L_t)\}\right]. \tag{12}$$

This is because for any node $v \in V$, it may be influenced by some pair $(u, t) \in \mathcal{S}$ in round t, and when multiple seed-round pairs occur, the one with the largest weight among these pairs is taken to calculate the final influence to v. Based on Eq.(12), one can show that $\rho_T^{NI}(\mathcal{S})$ is monotone and submodular.

5.2 Efficient Algorithm for NIOS setting

The PRM-IMM-NIOS algorithm follows the same structure as the PRM-IMM-OINS algorithm, but due to the non-overlapping influence requirement and the round-based weights nature, the RR set and the node selection procedure has significant difference, as we now explain.

For the RR set, for each randomly selected root $v \in V$, we need to generate T RR sets rooted at v, denoted as $R_{v,t}, t \in [T]$ in pair notation. The multi-round RR set (MR-RR set) is the union of all T RR sets, as $R_v = \bigcup_{t=1}^{T} R_{v,t}$. This is because in the NI setting a node v is only influenced once in all T rounds. The MR-RR set is similar to the one used in [29], but here different rounds have different weights, leading to a more complicated node selection procedure.

In the PRM-NodeSelection-NIOS procedure (Algorithm 3), we maintain a variable t_R for each MR-RR set R, and it denotes the lowest round number t in which the root v_R of R has been influenced by some already selected seed nodes. Initially, $t_R = T+1$, since no seed has been selected and thus no node is influenced. We also use variable $c[(u, t)]$ to store the remaining marginal influence of pair (u, t) on all MR-RR sets. Variables t_R and $c[(u, t)]$ are used to keep track of the marginal influence to v_R when some additional pair (v, t) is selected as a seed (lines 7–16): when $t < t_R$, this means that the root v_R is now influenced at the lowest round t, and thus for any other pair $(u, t') \in R$ with $t' < t_R$, if $t' \geq t$, it no long generates any marginal influence to v_R (the marginal influence is reduced by $w_{t'} - w_{t_R}$, line 10); and if $t' < t$, the remaining marginal influence (u, t') could generate to v_R through R is $w_{t'} - w_t$ (the marginal influence is reduced by $w_t - w_{t_R}$, line 12). This makes this node selection procedure more complicated than the OINS setting. Note that in line 5, we select the next seed from $V \times [k] \setminus \mathcal{S}$, reflecting that we allow overlapping seeds in different rounds.

Algorithm 3 PRM-NodeSelection-NIOS

Input: the set of MR-RR sets $\mathcal{R}$, seeds budget k,

weights $(w_t)_{t \in [T]}$ (Eq.(6)). particularly, the value of w_{T+1} is 0.

Output: k size seed set $\hat{\mathcal{S}}^g$.

1:   $RR[(v,t)] = \{R \in \mathcal{R} | (v,t) \in R\}$, $c[(v,t)] = w_t \cdot |RR[(v,t)]|$, $\forall (v,t) \in V \times [T]$ /* $RR[(v,t)]$'s
     and $c[(v,t)]$'s can be constructed during the generation of $\mathcal{R}$ */

2:   for all $R \in \mathcal{R}$, $t_R = T + 1$; $w_{T+1} = 0$

3:   $\mathcal{S} = \emptyset$ /* the element of $\mathcal{S}$ is (v,t) */

4:   **for** $i = 1$ to k **do**

5:     $(v,t) = \text{argmax}_{(v,t) \in V \times [k] \setminus \mathcal{S}} c[(v,t)]$

6:     $\mathcal{S} = \mathcal{S} \cup (v,t)$

7:     **for** $R \in RR[(v,t)] \wedge t < t_R$ **do**

8:       **for** $(u,t') \in R \wedge (u,t') \neq (v,t) \wedge t' < t_R$ **do**

9:         **if** $t' \geq t$ **then**

10:          $c[(u,t')] = c[(u,t')] - (w_{t'} - w_{t_R})$

11:         **else**

12:          $c[(u,t')] = c[(u,t')] - (w_t - w_{t_R})$

13:         **end if**

14:       **end for**

15:       $t_R = t$

16:     **end for**

17: **end for**

18: **return** $\hat{\mathcal{S}}^g = \mathcal{S}$

In the sampling part, similar to OINS setting, we can iteratively guess the lower bound x_i by halving and verifying through NodeSelection. We only need to modify parameter β to $\sqrt{\frac{1}{2} \cdot \left(\ln \binom{NT}{k} + \alpha^2 \right)}$, and the $\binom{NT}{k}$ is due to selecting k pairs from all NT pairs in the overlapping seeds setting. All other aspects of sampling part remains the same, and thus we ignore the pseudocode here. For the other two problems mentioned in Section 3.2, NINS and OIOS, we can plug in different components in OINS and NIOS to achieve the goal, and we omit the details here.

6 OTHER VARIANTS OF PRM PROBLEM

In this section, we show that our object function analysis and algorithm design are robust in that they can be adapted to a number of variants of the PRM problem.

6.1 Per-round Budget

In the PRM definition (Definition 1), the budget k is over all T rounds. One variant is that every round has its own budget k, similar to [29]. It also has four settings as in Table 1. Among them, the OIOS setting is the same as solving the traditional IM problem and then using the same seed set for all rounds. Other settings can be solved following the similar technique as the overall budget case, but when selecting next node-round pair as a seed, it needs to respect the per-round budget rather than the overall budget.

6.2 Variable Natural Growth Count

We can relax the assumption that the natural growth count z is constant for all rounds to allow a natural growth vector $z = [z_1, z_2, \cdots, z_T]$ with z_t denoting the natural customer count in round t. Then, similar to the derivation process in Section 4.1, we can obtain the formula for $r_T^{OI}(\mathcal{S})$ as follows:

$$r_T^{OI}(\mathcal{S}) = (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t)}{d_0^n + d_0^p + \sum_{i=1}^{t} z_i + \sum_{i=1}^{t-1} \sigma(S_i)} \right) - 1.$$

The derivation of the surrogate function $\rho_T^{OI}(\mathcal{S})$ and the rest algorithm design is the same, and we just need to redefine the weight $w_t = 1/(d_0^n + d_0^p + \sum_{i=1}^{t} z_i)$.

6.3 Popular Item Engages in Promotion

Our PRM problem considers passive popular item that does not engage in promotion. When popular item also engages in promotion, the situation is more complicated. Here we assume that the promotion of the popular item in all rounds are known, and the promotional result on round t is given as $p_t, t \in [T]$.

The case of unknown popular item promotions would lead the problem into the game-theoretic setting, which is beyond the scope of this paper. When $p_1, \ldots, p_T$ are known, in the OINS setting, the close form of the objective function $r_T^{OI}(\mathcal{S})$ is not easily derived as in Lemma 1. Instead we can obtain its upper and lower bound as follows.

$$r_T \leq (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t)}{d_0^n + d_0^p + zt + \sum_{i=1}^{t-1}(\sigma(S_i) + p_i)} \right) - 1, \tag{13}$$

$$r_T \geq (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t) - p_t}{d_0^n + d_0^p + zt + \sum_{i=1}^{t-1}(\sigma(S_i) + p_i) + p_t} \right) - 1. \tag{14}$$

Eq.(13) can be viewed as treating p_t as the additional natural growth in round t, and thus disfavor of the popular item and in favor of the novice item. Eq.(14) works when $d_p^t \geq d_n^t$, i.e. popular item is still the majority in popularity. These bounds lead to two weight settings:

$$\overline{w}_t = \frac{1}{d_0^n + d_0^p + t \cdot z + \sum_{i=1}^{t} p_i}, \quad \underline{w}_t = \frac{1}{d_0^n + d_0^p + t \cdot z + \sum_{i=1}^{t} p_i + p_t}.$$

Then following the sandwich approximation strategy [24], we can apply the two sets of weights to the corresponding PRM-IMM algorithm, and obtain two seed allocations, and compare their results to find the better one. The approximation ratio for the surrogate objective would have an extra factor of $\max_{t \in T} \underline{w}_t / \overline{w}_t$.

6.4 Two Other Optimization Objectives

Beyond the fundamental problem of maximizing the popularity ratio, it is natural to investigate two other variants of the PRM problem in surpassing the competitor: minimizing seed budget and minimizing the number of rounds needed.

- Seed Minimization: Given a round budget T, minimize the number of seeds in $\mathcal{S}$ such that the final popularity ratio is at least 1, i.e., $r_T(\mathcal{S}) \geq 1$.
- Round Minimization: Given a seed budget k, minimize the round number T such that the final popularity ratio is at least 1, i.e., $r_T(\mathcal{S}) \geq 1$.

Note that $r_T(\mathcal{S}) \geq 1$ means that the novice item has successfully caught up or surpassed the popular item at the end of the promotional period. For the seed minimization problem, we first find

Table 2. Datasets

	DM	LastFM	FX01/02	HEPT	Yelp	DBLP	LJ
nodes	0.6K	1.8K	29.3K	27.7K	110K	650K	4800K
edges	3.3K	12.7K	212.6K	352.8K	950K	1900K	42800K
avg.degree	4.96	13.44	7.24	12.70	15.92	3.04	8.83

Table 3. Parameter Settings

	DM	LastFM	FX01	FX02	HEPT	Yelp	DBLP	LJ
r_0	0.25	0.067	0.13	0.067	0.25	0.02	0.025	0.025
d_0	900	1600	1700	1600	5000	51000	20500	205000
z	10	100	50	100	150	2000	2500	25000

the range of minimum budget by doubling the budget from 1, and then determine the minimum budget that would make $r_T \geq 1$ by a binary search on the budget, and for each budget tried, we use the PRM-IMM algorithms developed. Same binary search idea can be applied to the round minimization problem.

6.5 Multiple Items

In this section, we discuss the PRM problem when there are multiple items in the system.

The Novice Item View Point, with Multiple Promotional Items We first focus on the view point from the targeted novice item, which is trying to increase its popularity compared with other items. Other competing items may or may not have their own promotions. We denote other items as items $1, \ldots, s$, and the popularity of item $j \in [s]$ at the end of round t is d_t^j.

We do not need to distinguish between popular and novice items among other items, and only need to denote whether an item conducts promotion or not. Each item $j \in [s]$ may conduct promotion p_t^j in round t, and $p_t^j = 0$ means no promotion. The popularity ratio of our targeted novice item, r_t, is defined as $r_t = d_t^n / (\sum_{j=1}^{s} d_t^j)$. The following theorem demonstrates that, even in the multiple-item case, we are able to derive the upper and lower bounds for the popularity ratio r_T, similar to the results presented in Section 6.3.

THEOREM 6. *In the setting of multiple items with promotions, we can bound r_T as follows:*

$$r_T \leq (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t)}{\zeta} \right) - 1, \tag{15}$$

$$r_T \geq (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t) - \sum_{j=1}^{s} p_t^j}{\zeta + \sum_{j=1}^{s} p_t^j} \right) - 1, \tag{16}$$

where $\zeta = d_0^n + \sum_{j=1}^{s} (d_0^j) + zt + \sum_{i=1}^{t-1} (\sigma(S_i) + \sum_{j=1}^{s} p_i^j)$, and Inequality (16) holds when $d_t^n \leq \sum_{j=1}^{s} d_t^j$, for $t < T$.

The implication of the above theorem is that, from the perspective of the targeted novice item, all other items can be treated as one single composite item — as long as we know the combined promotional effect of all other items, namely $\sum_{j=1}^{s} p_t^j$, we can treat all other items as one single item and compute the promotion to the novice item in the same way as described in Section 6.3. Moreover, it is easy to verify that if none of the other items conduct promotion, i.e. $p_t^j = 0$ for all

t's and j's, we can treat all other items as a single composite item and derive the exact formula for r_T as in Lemma 1. The main reason for the above property is that the preferential attachment mechanism remains the proportional allocation whether the novice item is facing one other item or many other items.

This further extends our understanding of popularity ratio maximization in cases where multiple items (popular or not, having promotions or not) co-exist in the system.

The Game Theoretic Point of View The above is from one novice item's view point, on how to allocate seeds to maximize its popularity ratio when other items' promotions are known. When items' promotions are unknown to one another, we need to study the dynamic of the overall system, and we enter the game-theoretic setting. Here, each item's strategy is the allocation of seeds over T rounds under its budget, and its utility is its popularity ratio. The algorithm we designed before can be viewed as the algorithm for each individual item (player) to achieve its best response, or approximate best response. Collectively, we can study the best-response dynamic and the equilibrium behavior. Alternatively, one may also map our basic PA-IC model into the Stackelberg game model [13], where the novice item and the popular make moves one after another based on the other side's previous moves. However, the full study of the game-theoretic aspect under our PA-IC model and the PRM problem setting would be too complicated to be included in this paper, and we will leave the detailed formulation and the technical study as a future work item.

7 EXPERIMENTS

In this section, we empirically validate our algorithms on eight real-world networks.

7.1 Experiment Setup

Data description We apply eight real-world network datasets in our experiments, with basic statistics shown in Table 2. The DM dataset is a collaboration network where nodes represent the data-mining researchers and edges represent co-authorships [30]. The FX01/02 datasets are from the Flixster, a network of American social movie discovery services [2]. Each node is a user and a directed edge from node u to v is formed if v rates one movie shortly after u does so on the same movie. FX01/02 represent two networks on two different movie topics [2], The NetHEPT dataset is an academic collaboration network from the "High Energy Physics Theory" section of arXiv from 1991 to 2003, where nodes represent the authors and each edge represents one paper co-authored by two nodes [9]. LastFM is a UK-based music site that focuses on online radio and music communities [3]. We use the friend relationship on LastFM to build a user relationship social network. The Yelp dataset is released by yelp officially[4]. Yelp.com is the largest review site in the United States. We use the friend relationship data in Yelp to build a user's social network. DBLP is another academic collaboration network extracted from the online archive DBLP (dblp.uni-trier.de) and used for influence studies [34]. LiveJournal, denoted as LJ, is a free online community with almost 10 million members; a significant fraction of these members are highly active. It is obtained from Stanford's SNAP project [20].

Metrics We consider the following three metrics, reflecting what one could consider in practice:

- **Popularity ratio.** The popularity ratio at the end of promotions.
- **Surpass budget.** Given a fixed round T, the minimum budget needed for the novice item to surpass the popular item.
- **Surpass time.** Given a fixed budget k, the minimum round number when the novice item surpasses the popular item.

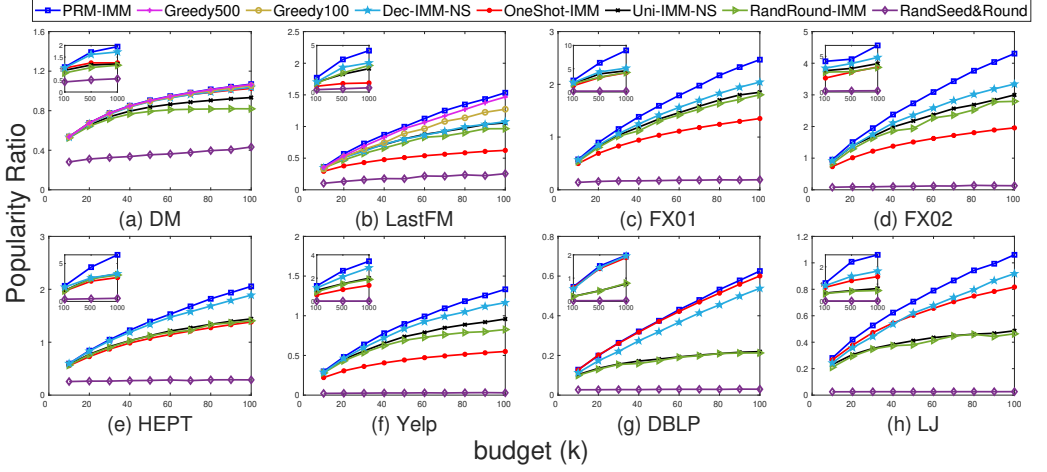
[3]http://www.last.fm/

[4]http://www.yelp.com

Fig. 2. The popularity ratio vs. budgets for different algorithms in OINS setting.

Parameter setup and reproducibility For DM and FX datasets, we use the edge weights learned in their respective studies [2, 30]. For the remaining datasets, we use the following "Weighted Cascade" policy adopted in [18] and often used in other influence maximization studies (e.g. [10, 32, 34]): for each edge (u, v), $p(u, v)$ is set to be $\frac{1}{N_{in}(v)}$, where $N_{in}(v)$ is the in-degree of v. Every PRM instance is also determined by the parameters r_0, d_0, z. We set proper parameters for each dataset, as shown in Table 3, to reflect different scenarios and match with the network size. Specifically, the settings of d_0 and z are mainly related to the size of the network. When the network size is large, promotion may generate large popularity from the social network. Thus, to match the natural growth parameter and to avoid promotion dominating the natural growth, we need to adjust d_0 and z accordingly to compensate for the large popularity generated by promotion. Parameter r_0 further covers different scenarios among the datasets.

We further vary these parameters and study their effects (Fig.4 and Fig.5). Following [32], we set the error probability parameter $l = 1$, approximation ratio parameter $\varepsilon = 0.1$ in PRM-IMM. For implementation details see the source code and proof appendix in this repository[5].

Baselines There is no existing baseline for the newly proposed PRM problem. Thus, we compare with the following reasonable heuristic baselines.

- **OneShot-IMM**: The largest k nodes are picked by the IMM algorithm, and all are placed in round 1.
- **Uni-IMM-NS/Uni-IMM-OS**: The NS version selects the largest k nodes by IMM and places them in all rounds separately with equal size in the order of the greedy selection. The OS version simply repeats the first round seeds in later rounds (for NIOS tests).
- **RandRound-IMM**: Pick the largest k nodes by the IMM algorithm and put them into T rounds randomly.
- **RandSeed&Round**: Randomly select k nodes and randomly place these nodes in each round.
- **Dec-IMM-NS/Dec-IMM-OS**: The NS version picks the largest k nodes by IMM and places seed nodes in descending order. The method of generating the decreasing sequence is to use one-fifth of the budget in the first round and one-fifth of the remaining budget in each subsequent round. The OS version repeats the seeds in the first round with the same budget as the NS version.

---

[5]https://github.com/Complex-data/PRM

Table 4. Running Time (OINS setting, $T = 100$, $k = 100$, results in seconds, average over 5 runs)

	IMM	PRM-IMM	Greedy100	Greedy500
DM	0.067	13.7	397.8	3489
LastFM	1.24	82.86	11976	53681
FX01	4.07	437	-	-
FX02	3.42	432	-	-
HEPT	2.39	399	-	-
Yelp	44.36	2959	-	-
DBLP	47.46	1176	-	-
LJ	371	15026	-	-

- **Greedy**: Directly apply the greedy algorithm, using simulations to estimate marginal gains and do not use the RIS approach.

The first five baselines are all heuristics with no theoretical guarantee. Algorithm Greedy has the theoretical guarantee as PRM-IMM, but it is very slow in running simulations, and we will only show its result on the two small datasets DM and LastFM.

7.1.1 Platform. We implement all algorithms in C++, compiled in Visual Studio 2019, and run our tests on a computer with 3.6GHz Intel(R) Core(TM) i9-9900K CPU (16 cores), 128G memory, and Windows 10 professional (64 bits).

7.2 Results

OINS main results. Fig. 2 shows the result of popularity ratio vs. budget on all datasets and all algorithms in the OINS setting. The results are averaged over 10 independent runs. Note that the y-axis corresponds to the original objective function (Eq. (3)), not the surrogate function (Eq. (7)). It is obvious that PRM-IMM outperforms the baseline algorithms with a large margin in many cases. More specifically, RandomSeed&Round is much worse than other algorithms, showing that randomly giving out promotions should be avoided if possible. Uniform-IMM and RandomRound-IMM have similar performance since both spread the seeds more or less evenly among the T rounds, and in many cases such even spreading results in poor performance. OneShot-IMM is at the opposite extreme, only allocating seeds in the first round. It has unstable performance, and in most cases it performs poorly, except for the DM and DBLP datasets.

Dec-IMM in general performs better than the OneShot/Uniform/RandomRound versions, but it still has a clear gap with our algorithm PRM-IMM. This is because all these heuristics cannot adjust to different network structure and diffusion characteristics of the networks (such as whether important seeds are likely to have overlapping influence), while our PRM-IMM automatically balances the early promotion and the influence redundancy among seeds.

Greedy should perform equally well as PRM-IMM, since PRM-IMM is just a more efficient way of implementing Greedy. But Greedy is very time consuming — as shown in Table 4, Greedy with 100 and 500 simulations per estimation is already several orders of magnitude slower than PRM-IMM, and it takes Greedy500 almost 15 hours to run on the 1.8K node LastFM network. Thus Greedy is not scalable and we cannot run it on other datasets except for DM and LastFM. Even for LastFM, Fig. 2(b) shows that Greedy500 and Greedy100 are not as competitive as PRM-IMM, due to the insufficient number of simulations.

Table 5. Running Time (OINS setting, $k = 500, 1000$ results in seconds, average over 5 runs)

	$k = 500$		$k = 1000$	
	IMM	PRM-IMM	IMM	PRM-IMM
DM	0.13	1.16	0.13	1.98
LastFM	0.54	48.59	0.61	48.59
FX01	11.47	138.62	16.51	229.9
FX02	10.52	137.52	16.25	253.48
HEPT	15.11	67.83	13.79	106.93
Yelp	20.36	1639.34	24.51	2715
DBLP	332.46	1317.18	383.73	1939.36
LJ	2505	5607	4251	10760

The top left corner of each subfigure in Fig.2 shows the result of different algorithms under larger seed budget of $k = 500, 1000$. It can be seen that PRM-IMM still performs better than other baselines for larger k's.

Table 4 shows that the running time of the algorithms with $k = 100$. Note that Random Seed&Round is a trivial baseline with almost instant running time, but its quality is far worse than all others as reported in Fig.2, and thus we do not include it in the running time table. For all other baselines except Greedy, they call IMM to select seeds first, and then do various simple allocation of seeds over rounds. Thus their running time is dominated by the IMM running time, which is reported in the table. The result shows that (a) Greedy100/Greedy500 is two to three orders of magnitude slower than PRM-IMM, and thus is not scalable at all; (b) PRM-IMM is around 100 times slower than the pure IMM algorithm (except for DBLP and LJ), which aligns with our theoretical analysis showing that our running time has an extra T factor. Note that for the two large datasets DBLP and LJ, we reduce its round input T to save time, since the results from smaller budgets show that only the first few rounds get allocated for seeds.

Table 5 shows the running time of PRM-IMM vs. IMM with $k = 500, 1000$. For this test, we optimize for T, meaning that we set T to a small value, and if the resulting allocation of PRM-IMM does not use all rounds, we stop the algorithm; otherwise, we double the T value and retry the algorithm. We can see that PRM-IMM can still scale to larger seed budgets and large networks.

Overall, our OINS results clearly show that our PRM-IMM algorithm outperforms all simple heuristics, while it runs much more efficiently than the greedy algorithm. PRM-IMM achieves a good balance between quality and efficiency.

Model justification We now justify two choices we made in our model formulation. First, Eq.(1) is based on the average popularity growth in each round rather than the random popularity growth, as we discussed in Section 3.1. We compare the effect of the average growth with the random growth, using the seed allocation selected by PRM-IMM with the budget ranging from 10 to 100. Fig. 3(a) shows that the results of the two effects are very close in four datasets (the other four have similar results), justifying our use of the average effect in Eq.(1).

Next, we compare the original objective function of Eq. (3) with the surrogate round-weighted influence function of Eq. (7). We use seed allocations selected by PRM-IMM for this comparison. Fig. 3(b) shows that the two functions agree with each other very well, with average derivation less than 5% for four datasets (and the other four have similar results). Hence, using the surrogate function is a reasonable choice, even though our PRM is a heuristic with respect to the original objective.

Parameter analysis in PRM We now analyze the effect of the initial total popularity measure
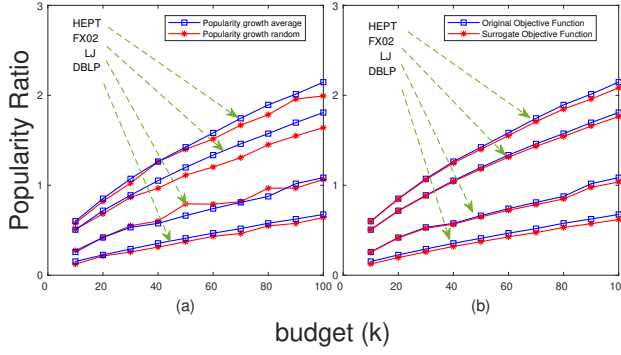
Fig. 3. Model justification. (a) Comparing between the random growth (red) and average growth (blue); (b) Comparing between the original objective function and the surrogate function.
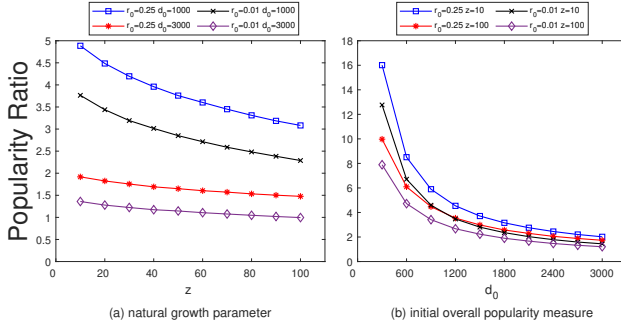

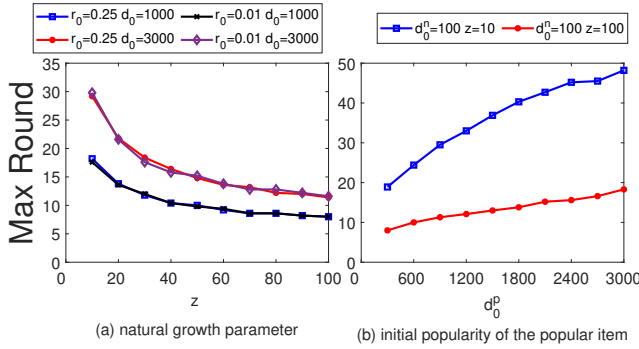Fig. 4. Parameter analysis for PRM under FX01.

Fig. 5. Different parameter's impact on seeds allocation, in the FX01 dataset.

d_0 and natural growth count z on the PRM-IMM algorithm. We run PRM-IMM to select 100 seeds for this analysis. Fig. 4 shows the result for FX01, and results on other datasets are similar.

Fig. 4(a) shows that when z increases, the popularity ratio decreases, indicating that the larger the natural growth of the popularity measure, the more difficult it is for the novice item to catch up with the popular item. Fig. 4(b) also shows the decreasing trend with respect to d_0. This is because
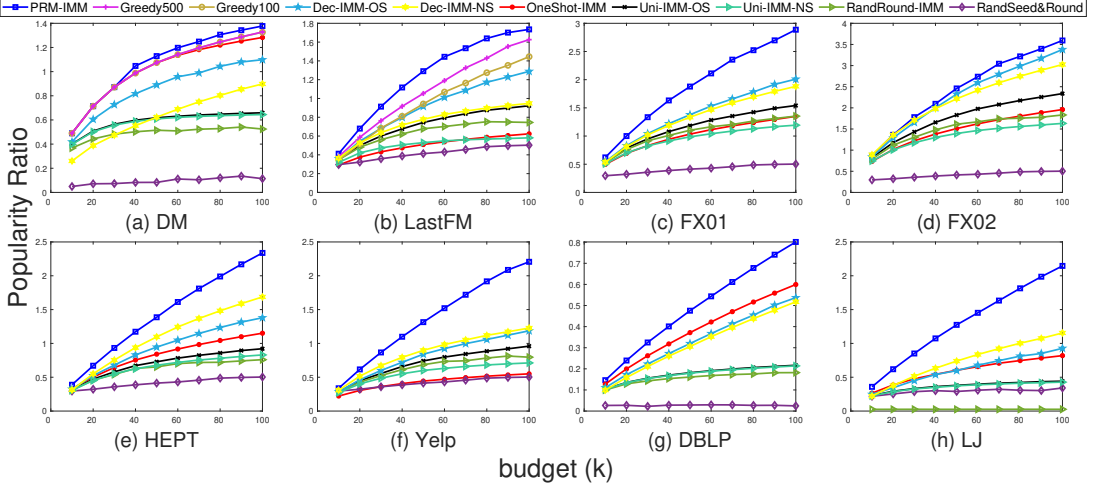
Fig. 6. The popularity ratio vs. budgets for different algorithms in NIOS setting

Table 6. Running Time (NIOS setting, $T = 100$, $k = 100$, results in seconds, average over 5 runs)

	IMM	PRM-IMM	greedy-100	greedy-500
DM	0.067	342.8	1274.8	3489
LastFM	1.24	1258	51034	249656
FX01	4.07	4396	-	-
FX02	3.42	6163	-	-
HEPT	2.39	8056	-	-
Yelp	44.36	39582	-	-
DBLP	47.46	1395	-	-
LJ	371	76396	-	-

when d_0 is larger while r_0 remains the same, the gap between the popular and novice item becomes larger and it is more difficult for the noice item to catch up.

Among the parameters, z and d_0^p (initial popularity of the popular item) are not directly related to the novice item. Thus, we further investigate how z and d_0^p would affect the seed allocation choice of the novice item in Fig. 5. Y-axis "max round" means the maximum number of rounds used in the allocation. Fig. 5 shows that when z increases, the max round decreases, and when d_0^p increases, the max round also increases. This shows the interesting adaptive property of PRM-IMM: when the natural growth count per round increases, the algorithm tries to put more seeds earlier, otherwise natural growth in each round will cause more people going to the popular item and making it harder to catch up; but if only the initial popularity of the popular item increases, the algorithm can use more rounds to achieve a better catch-up result in the end.

Results on NIOS The popularity ratio and running time results for NIOS are shown in Fig. 6 and Table 6, respectively. For this test, we further include baselines Uni-IMM-OS and Dec-IMM-OS to consider seed overlaps. Overall, the results are similar to those of OINS, and our PRM-IMM algorithm always performs the best among all baselines, and it runs significantly faster than the
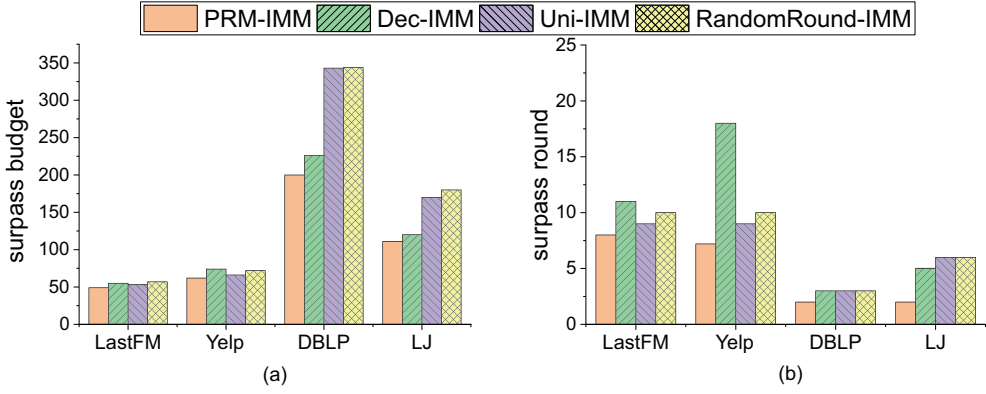
Fig. 7. (a) Fix round $T = 20$, find the smallest budget to make the popularity ratio exceed 1. (b) Fix budget (LastFM, $k = 55$; Yelp, $k = 75$; DBLP, $k = 230$; LJ, $k = 125$), make the popularity ratio exceed 1 in as few rounds as possible.
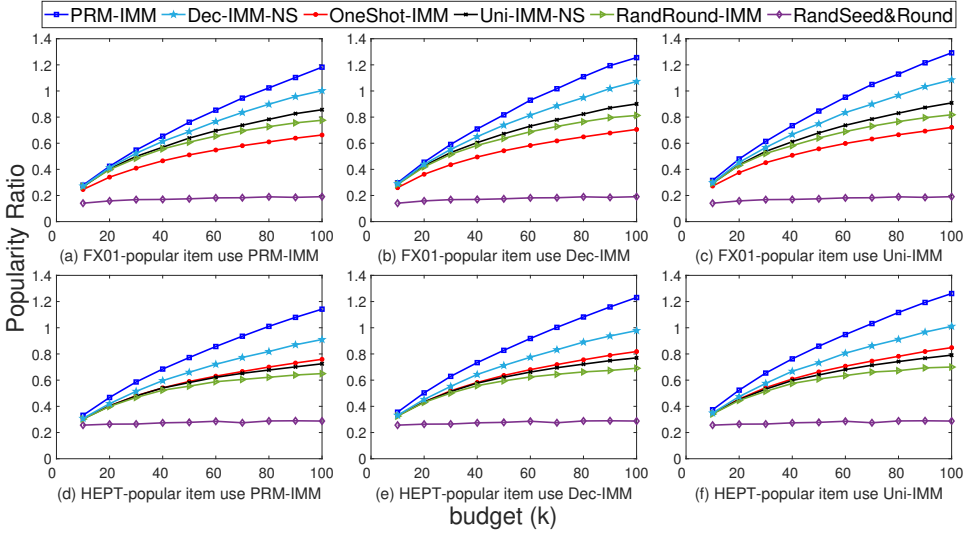

Fig. 8. The popularity ratio vs. budgets for different algorithms when the popular item has promotion.

simple greedy algorithm. The performance gap between our algorithm and other baselines is even larger in many cases compared to OINS. This is perhaps because the NIOS setting demands more sophisticated seed allocation. For example, we notice that seed overlaps vary with the networks. The OS specific baselines in many cases perform not as well as their NS counterparts, indicating that it is not easy to adjust proper overlaps among seeds. Compared to OINS, our PRM-IMM algorithm runs slower, and this is mainly because each MR-RR set for NIOS is much larger than the PW-RR set for OINS.

Alternative Objectives. Finally, we evaluate the two alternative objectives discussed in Section 6.4. Since RandomSeed&Round is clearly worse than other algorithms, while OneShot-IMM cannot assign seeds to multiple rounds, we do not include them as baselines in this part. We apply the idea in Section 6.4 to conduct these experiments for the OINS setting.

Fig.7 shows the results on four datasets, and the results on the other four datasets are similar. Fig.7(a) shows the result for the Seed Minimization problem, in which we fix $T = 20$, and then use different algorithms to find the minimum budget that can make the popular ratio exceed 1. Fig.7(b) shows the result for the Round Minimization problem. For this problem, we need to set different budgets for different datasets based on our result from Fig.7(a) to balance the influence of T and k.

The results in both figures show that our algorithm can achieve the best results compared to all other baselines.

Popular item having promotion. In this section, we evaluate different algorithms in the situation when the novice item knows the promotion plan of the popular item. We assume that the popular item uses a budget of 30 and uses three different strategies: PRM-IMM, Dec-IMM, Uni-IMM, to select seeds. Then the novice item uses PRM-IMM and other baselines to select seed nodes and evaluate the final popularity ratio. The results on FX01 and HEPT dataset are shown in Fig.8. We can see that even when the popular item has promotions, our PRM-IMM still performs well and beats all the baselines. Moreover, the end popularity ratio increases when the popular item promotion moves from PRM-IMM algorithm to Dec-IMM and then Uni-IMM, suggesting that the novice item can achieve better promotion result when the popular item uses a less competitive promotion strategy.

8 FUTURE WORK

One interesting future direction is to study adaptive strategies, where the seed selection of later rounds depends on the actual promotional results of the earlier rounds. Another direction is to consider the delayed effect of promotion, where social network promotion may take multiple rounds to generate popularity growth. One may also look into the model where the promotion of the popular item and the novice item also competes in the social network.

9 ACKNOWLEDGMENTS

REFERENCES

[1] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439 (1999), 509–512. https://doi.org/10.1126/science.286.5439.509

[2] Nicola Barbieri, Francesco Bonchi, and Giuseppe Manco. 2012. Topic-Aware Social Influence Propagation Models. In *12th IEEE International Conference on Data Mining, ICDM 2012, Brussels, Belgium, December 10-13, 2012*. IEEE Computer Society, 81–90. https://doi.org/10.1109/ICDM.2012.122

[3] Christian Borgs, Michael Brautbar, Jennifer T. Chayes, and Brendan Lucier. 2014. Maximizing Social Influence in Nearly Optimal Time. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*. SIAM, 946–957. https://doi.org/10.1137/1.9781611973402.70

[4] Ceren Budak, Divyakant Agrawal, and Amr El Abbadi. 2011. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India, March 28 - April 1, 2011*. ACM, 665–674. https://doi.org/10.1145/1963405.1963499

[5] Gruia Călinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. 2011. Maximizing a Monotone Submodular Function Subject to a Matroid Constraint. *SIAM J. Comput.* 40, 6 (2011), 1740–1766. https://doi.org/10.1137/080733991

[6] Wei Chen. 2018. An Issue in the Martingale Analysis of the Influence Maximization Algorithm IMM. In *Computational Data and Social Networks - 7th International Conference, CSoNet 2018, Shanghai, China, December 18-20, 2018, Proceedings (Lecture Notes in Computer Science, Vol. 11280)*. Springer, 286–297. https://doi.org/10.1007/978-3-030-04648-4_24

[7] Wei Chen. 2020. *Network Diffusion Models and Algorithms for Big Data (in Chinese)*. Post and Telecom Press.

[8] Wei Chen, Laks V. S. Lakshmanan, and Carlos Castillo. 2013. *Information and Influence Propagation in Social Networks*. Morgan & Claypool Publishers. https://doi.org/10.2200/S00527ED1V01Y201308DTM037

[9] Wei Chen, Chi Wang, and Yajun Wang. 2010. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 25-28, 2010*. ACM, 1029–1038. https://doi.org/10.1145/1835804.1835934

[10] Wei Chen, Yajun Wang, and Siyu Yang. 2009. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009*, John F. Elder IV, Françoise Fogelman-Soulié, Peter A. Flach, and Mohammed Javeed Zaki (Eds.). ACM, 199–208. https://doi.org/10.1145/1557019.1557047

[11] Wei Chen, Yifei Yuan, and Li Zhang. 2010. Scalable Influence Maximization in Social Networks under the Linear Threshold Model. In *ICDM 2010, The 10th IEEE International Conference on Data Mining, Sydney, Australia, 14-17 December 2010*. IEEE Computer Society, 88–97. https://doi.org/10.1109/ICDM.2010.118

[12] Pedro M. Domingos and Matthew Richardson. 2001. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, San Francisco, CA, USA, August 26-29, 2001*. ACM, 57–66. https://doi.org/10.1145/502512.502525

[13] D. Fudenberg and J. Tirole. 1993. *Game Theory*. MIT Press.

[14] Amit Goyal, Wei Lu, and Laks V. S. Lakshmanan. 2011. SIMPATH: An Efficient Algorithm for Influence Maximization under the Linear Threshold Model. In *11th IEEE International Conference on Data Mining, ICDM 2011, Vancouver, BC, Canada, December 11-14, 2011*. IEEE Computer Society, 211–220. https://doi.org/10.1109/ICDM.2011.132

[15] Sanjeev Goyal and Michael J. Kearns. 2012. Competitive contagion in networks. In *Proceedings of the 44th Symposium on Theory of Computing Conference, STOC 2012, New York, NY, USA, May 19 - 22, 2012*. ACM, 759–774. https://doi.org/10.1145/2213977.2214046

[16] Xinran He, Guojie Song, Wei Chen, and Qingye Jiang. 2012. Influence Blocking Maximization in Social Networks under the Competitive Linear Threshold Model. In *Proceedings of the Twelfth SIAM International Conference on Data Mining, Anaheim, California, USA, April 26-28, 2012*. SIAM / Omnipress, 463–474. https://doi.org/10.1137/1.9781611972825.40

[17] Kyomin Jung, Wooram Heo, and Wei Chen. 2012. IRIE: Scalable and Robust Influence Maximization in Social Networks. In *12th IEEE International Conference on Data Mining, ICDM 2012, Brussels, Belgium, December 10-13, 2012*. IEEE Computer Society, 918–923. https://doi.org/10.1109/ICDM.2012.79

[18] David Kempe, Jon M. Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*. ACM, 137–146. https://doi.org/10.1145/956750.956769

[19] Jinha Kim, Seung-Keol Kim, and Hwanjo Yu. 2013. Scalable and parallelizable processing of influence maximization for large-scale social networks?. In *29th IEEE International Conference on Data Engineering, ICDE 2013, Brisbane, Australia, April 8-12, 2013*. IEEE Computer Society, 266–277. https://doi.org/10.1109/ICDE.2013.6544831

[20] Jure Leskovec and Andrej Krevl. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. http://snap.stanford.edu/data.

[21] Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie S. Glance, and Matthew Hurst. 2007. Patterns of Cascading Behavior in Large Blog Graphs. In *Proceedings of the Seventh International Conference on Data Mining* (Minneapolis, Minnesota) *(SIAM '07)*. SIAM, 551–556. https://doi.org/10.1137/1.9781611972771.60

[22] Yanhua Li, Wei Chen, Yajun Wang, and Zhi-Li Zhang. 2013. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. In *Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy, February 4-8, 2013*. ACM, 657–666. https://doi.org/10.1145/2433396.2433478

[23] Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. 2018. Influence Maximization on Social Graphs: A Survey. *IEEE Trans. Knowl. Data Eng.* 30, 10 (2018), 1852–1872. https://doi.org/10.1109/TKDE.2018.2807843

[24] Wei Lu, Wei Chen, and Laks V. S. Lakshmanan. 2015. From Competition to Complementarity: Comparative Influence Diffusion and Maximization. *Proc. VLDB Endow.* 9, 2 (2015), 60–71. https://doi.org/10.14778/2850578.2850581

[25] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. 1978. *An analysis of the approximations for maximizing submodular set functions*. Mathematical Programming.

[26] Hung T. Nguyen, My T. Thai, and Thang N. Dinh. 2016. Stop-and-Stare: Optimal Sampling Algorithms for Viral Marketing in Billion-scale Networks. In *Proceedings of the 2016 International Conference on Management of Data, SIGMOD Conference 2016, San Francisco, CA, USA, June 26 - July 01, 2016*, Fatma Özcan, Georgia Koutrika, and Sam Madden (Eds.). ACM, 695–710. https://doi.org/10.1145/2882903.2915207

[27] Matthew Richardson and Pedro M. Domingos. 2002. Mining knowledge-sharing sites for viral marketing. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada*. ACM, 61–70. https://doi.org/10.1145/775047.775057

[28] Herbert A Simon. 1955. On a class of skew distribution functions. *Biometrika* 42, 3/4 (1955), 425–440.

[29] Lichao Sun, Weiran Huang, Philip S. Yu, and Wei Chen. 2018. Multi-Round Influence Maximization. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2018, London, UK, August 19-23, 2018*. ACM, 2249–2258. https://doi.org/10.1145/3219819.3220101

[30] Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. 2009. Social influence analysis in large-scale networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009*. ACM, 807–816. https://doi.org/10.1145/1557019.1557108

[31] Jing Tang, Xueyan Tang, Xiaokui Xiao, and Junsong Yuan. 2018. Online Processing Algorithms for Influence Maximization. In *Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference 2018, Houston, TX, USA, June 10-15, 2018*. ACM, 991–1005. https://doi.org/10.1145/3183713.3183749

[32] Youze Tang, Yanchen Shi, and Xiaokui Xiao. 2015. Influence Maximization in Near-Linear Time: A Martingale Approach. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, Melbourne, Victoria, Australia, May 31 - June 4, 2015*. ACM, 1539–1554. https://doi.org/10.1145/2723372.2723734

[33] Youze Tang, Xiaokui Xiao, and Yanchen Shi. 2014. Influence maximization: near-optimal time complexity meets practical efficiency. In *International Conference on Management of Data, SIGMOD 2014, Snowbird, UT, USA, June 22-27, 2014*. ACM, 75–86. https://doi.org/10.1145/2588555.2593670

[34] Chi Wang, Wei Chen, and Yajun Wang. 2012. Scalable influence maximization for independent cascade model in large-scale social networks. *Data Min. Knowl. Discov.* 25, 3 (2012), 545–576. https://doi.org/10.1007/s10618-012-0262-1

[35] Yu Wang, Gao Cong, Guojie Song, and Kunqing Xie. 2010. Community-based greedy algorithm for mining top-K influential nodes in mobile social networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 25-28, 2010*. ACM, 1039–1048. https://doi.org/10.1145/1835804.1835935

[36] George Udny Yule. 1925. II.—A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FR S. *Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character* 213, 402-410 (1925), 21–87.

Table 7. Notations

Notation	Description
$G = (V, E)$	a social network G with a node set V and an edge set E
N, M	the numbers of nodes and edges in G, respectively
d_t^n	the popularity measure of Novice item at the end of round
d_t^p	the popularity measure of Popular item at the end of round
k	number of seeds to be selected
τ	time step index, each time step is one step in one round of IC model p
t	round index, each round is a promotional round
T	Total rounds
S_t	seed set of round t
$\mathcal{S} = \bigcup_{t=1}^{T} S_t \times \{t\}$	pair set, a pair (u, t) is a node u at round t
$\mathcal{S}^*$	optimal pair set
$r_T(\mathcal{S})$	the popularity ratio at the end of round T
r_0	the initial popularity ratio
$p(u, v)$	the probability of node u active v
z	total natural growth of popularity measure
$\sigma(S)$	the influence spread of seed set S
$w = (w_1, w_2 ... w_T)$	the weight vector, w_t is the weight of round t
$\rho_T(\mathcal{S})$	weighted overlapping influence spread
$\rho^{(t)}(S_t)$	weighted overlapping influence spread of round t
$R, R(v)$	one RR set, $R(v)$ is the RR set rooted at node v
$R^{(t)}, R^{(t)}(v)$	one PW-RR set, $R^{(t)} = R \times \{t\}$, $R^{(t)}(v)$ is the RR set rooted at pai
$\mathcal{R}$	the set of pair-wised RR set
θ	the number of PW-RR sets that need to be generated
l	Error probability parameter
ε	Approximate ratio parameter
$b, Q, i, j, x_i, \alpha, \beta, OPT, \delta, LB, Pr, \omega, \Omega$	parameters in derivation

APPENDIX

10 NOTATIONS

11 MONOTONE AND SUBMODULAR OF POPULARITY RATIO FUNCTION

PROOF. It is proved that the expected influence spread set function σ is monotone and submodular, we don't show the proof process here. So we only need to prove that the popularity ratio function is monotone with respect to the function $\sigma(S_i), 0 < i < T$. We can prove the monotone of the

popularity ratio function. For simplicity, we denote $d_{t-1}^n + d_{t-1}^p + a$ as d_t, denote $\sigma(S_i)$ as x_i, $0 < i < T$.

$$
\frac{\partial r_T}{\partial x_i} = (r_0 + 1) \prod_{t=1}^{i-1} \left(1 + \frac{x_t}{d_t}\right) \left[\frac{1}{d_i} \prod_{t=i+1}^{T} \left(1 + \frac{x_t}{d_t}\right)\right.
$$

$$
\left. + \sum_{t=i+1}^{T} \frac{-x_t}{d_t^2} \prod_{s=i, s\neq t}^{T} \left(1 + \frac{x_s}{d_s}\right)\right]
$$

$$
= r_T \cdot \left[\frac{a}{d_{i+1}(d_i + x_i)} + \cdots + \frac{a}{d_T(d_{T-1} + \sigma(S_{T-1}))}\right.
$$

$$
\left. + \frac{1}{d_T + x_T}\right] > 0
$$

For any $0 < i < T$, $\frac{\partial r_T}{\partial x_i} > 0$, the popularity ratio function is monotone.

However the popularity ratio function does not satisfy submodular, and we will illustrate this property with a counter example below.

Consider the simplest case that social network has only three nodes u, v, w. No edges between nodes, the initial popularity measure of Novice item is $d_0^n = 1$ and the initial popularity measure of Popular item is $d_0^p = 2$, the increment of popularity measure of each round $a = 1$. In this case, the original

$$
r_T(S) = (r_0 + 1) \prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t)}{d_0^n + d_0^p + a \cdot t + \sum_{i=1}^{t-1} \sigma(S_i)}\right) - 1
$$

where T is at most 3, $r_0 = \frac{d_0^n}{d_0^p} = \frac{1}{2}$.

$$
r_T(S)
$$
$$
= \frac{3}{2} * (1 + \frac{\sigma(S_1)}{d_0^n + d_0^p + a}) * (1 + \frac{\sigma(S_2)}{d_1^n + d_1^p + a}) * (1 + \frac{\sigma(S_3)}{d_2^n + d_2^p + a}) - 1
$$

where $d_0^n + d_0^p + a = 4$, $d_1^n + d_1^p + a = 5 + \sigma(S_1)$, $d_1^n + d_1^p + a = 6 + \sigma(S_1) + \sigma(S_2)$.

$$
r_T(S)
$$
$$
= \frac{3}{2} * (1 + \frac{\sigma(S_1)}{4}) * (1 + \frac{\sigma(S_2)}{5 + \sigma(S_1)}) * (1 + \frac{\sigma(S_3)}{6 + \sigma(S_1) + \sigma(S_2)}) - 1
$$

Now the two pair set $S \subset Q$, $S = \{(u,1)\}$, $Q = \{(u,1),(v,1)\}$, and a pair $b = (w,2)$, Clearly $b \notin S$.

$$
r_T(S) = 0.875, r_T(S \cup \{b\}) = 1.1875
$$
$$
r_T(Q) = 1.25, r_T(Q \cup \{b\}) = 1.5714
$$

$$
r_T(Q \cup \{b\}) - r_T(Q) = 0.3214, r_T(S \cup \{b\}) - r_T(S) = 0.3125
$$

$$
r_T(Q \cup \{b\}) - r_T(Q) > r_T(S \cup \{b\}) - r_T(S)
$$

In this case, the marginal value of Q is larger than S. So the set popularity ratio function is not submodular.   □

12 SIMPLIFICATION PROCESS

That is:

$$\prod_{t=1}^{T} \left(1 + \frac{\sigma(S_t)}{d_0^n + d_0^p + a \cdot t + \sum_{i=1}^{t-1} \sigma(S_i)} \right)$$

$$\frac{\sigma(S_1)}{d_0^n + d_0^p + a} + \frac{\sigma(S_2)}{d_0^n + d_0^p + 2a + \sigma(S_1)} + \cdots + \frac{\sigma(S_T)}{d_0^n + d_0^p + T \cdot a + \sum_{i=0}^{T-1} \sigma(S_i)}$$

$$\frac{\sigma(S_1)}{d_0^n + d_0^p + a} + \frac{\sigma(S_2)}{d_0^n + d_0^p + 2a} + \cdots + \frac{\sigma(S_T)}{d_0^n + d_0^p + T \cdot a}$$

Combining the above simplified process, it can be noted that

$$\rho_T(\mathcal{S}) = \frac{\sigma(S_1)}{d_0^n + d_0^p + a} + \frac{\sigma(S_2)}{d_0^n + d_0^p + 2a} + \cdots + \frac{\sigma(S_T)}{d_0^n + d_0^p + T \cdot a}$$

is our weighted overlapping influence function, where d_0^n, d_0^p, a are our predefined parameters. Thus we can denote the weighted overlapping influence function as

$$\rho_T(\mathcal{S}) = \sum_{t=1}^{T} w_t \cdot \sigma(S_t) \tag{17}$$

13 PW-RR GENERATION PROCESS

Algorithm 4 PW-RR generation process.

Input: directed graph $G = (V, E)$, IC model, Max round T
Number of PW-RR set θ
Output: the set of PW-RR set $\mathcal{R}$

1: $\mathcal{R} = \emptyset$
2: **for** $0 < \theta$ **do**
3:     $\theta = \theta - 1$
4:     Generate an RR set R for a random node $v \in V$
5:     choose a round t uniformly at random from $[T]$
6:     put the RR set R and round t together as $R^{(t)}$
7:     $\mathcal{R} = \mathcal{R} \cup \{R^{(t)}\}$
8: **end for**
9: return $\mathcal{R}$

---

14 PROOF OF LEMMA 4

PROOF. The randomness of $Y(\mathcal{S})$ is from two aspect: (1) the root of a PW-RR set is uniformly random choose, (2) the round t of the PW-RR set is uniformly random choose.

$$\mathbb{E}\left[Y(\mathcal{S})\right]$$

$$= \frac{1}{T} \sum_{t=1}^{T} w_t \cdot \mathbb{E}\left(\mathbb{I}\{\mathcal{S} \cap \mathcal{R} \neq \emptyset\}\right)$$

$$= \frac{1}{T} \sum_{t=1}^{T} w_t \cdot \Pr\left\{S_t \cap R^{(t)} \neq \emptyset\right\}$$

$$= \frac{1}{T} \sum_{t=1}^{T} w_t \cdot \frac{1}{N} \sum_{v \in V} \Pr\left\{S_t \cap R^{(t)}(v) \neq \emptyset\right\}$$

$$= \frac{1}{T} \sum_{t=1}^{T} w_t \cdot \frac{1}{N} \sum_{v \in V} ap\left(S_t, v\right)$$

$$= \frac{1}{T} \sum_{t=1}^{T} w_t \cdot \frac{1}{N} \sigma\left(S_t\right)$$

For any seed set $S_t \in V$, any node $v \in V$, the probability that the seed set S_t activates node v with probability $ap(S_t, v)$. $ap(S_t, v)$ is the probability that S_t have an intersaction with a random RR set $R(v)$ rooted from node v. i.e. $ap(S, v) = \Pr\{S \cap R(v) \neq \emptyset\}$

$\square$

15 NODESELECTION ALGORITHM

Algorithm 5 provides the pseudo-code for the PRM-NodeSelection. We select the pair (v, t) with the largest marginal benefit in the line 4 to add to the set $\mathcal{S}$, and then subtract the part of this node that overlaps with other nodes in lines 8-12.

The input of PRM-NodeSelection includes $\mathcal{R}$, seeds budget k. Other data structure such as c and $\mathcal{RR}$ can help us to greedily select k pairs that maximize the $\hat{\rho}_T(\mathcal{S}, \mathcal{R})$ (line 4).

Algorithm 5 PRM-Nodeselection

Input: the set of PW-RR sets $\mathcal{R}$, seeds budget k
Output: k size seed set $\hat{\mathcal{S}}^g$.

1: **for** all $R \in \mathcal{R}, t \in [T], covered[t][R] = false$ **do**
2:     $\mathcal{S} = \emptyset$ /* the element of $\mathcal{S}$ is (v, t) */
3:     **for** $i = 1$ to k **do**
4:        $(v, t) = \mathrm{argmax}_{(v,t) \in (V \setminus \mathcal{S}_{\mathrm{node}}) \times [k]} \{w_t \cdot c[t][v]\}$
5:        /*where $\mathcal{S}_{node}$ is the node set of $\mathcal{S}$ */
6:        $\mathcal{S} = \mathcal{S} \cup (v, t)$
7:        **for** $R \in RR[t][v] \wedge covered[t][R] == flase$ **do**
8:           $covered[R] = true$
9:           **for** $R \in RR[t][v] \wedge u \neq v$ **do**
10:              $c[t][u] = c[t][u] - 1$
11:          **end for**
12:       **end for**
13:    **end for**
14: **end for**
15: **return** $\hat{\mathcal{S}}^g = \mathcal{S}$

---

16 ESTIMATE THE NUMBER OF PW-RR

We first give a general conclusion (Theorem 7) to show how the greedy solution obtained by the PRM-NodeSelection approaches the optimal solution of the weighted overlapping influence maximization problem when the $\hat{\rho}_T(\mathcal{S}, \mathcal{R})$ itself satisfies monotone submodular (It is easy to proof that $\hat{\rho}_T(\mathcal{S}, \mathcal{R})$ is submodular and nondecreasing with respect to $\mathcal{S}$).

We denote the random estimation of $\rho_T(\mathcal{S})$ as $\hat{\rho}_T(\mathcal{S}, \omega)$, where $\omega \in \Omega$ is a sample in random space Ω. $\mathcal{S}^*$ is the optimal solution of $\rho_T(\mathcal{S})$, $OPT = \rho_T(\mathcal{S}^*)$. $\hat{\mathcal{S}}^g(\omega)$ is the greedy result of $\hat{\rho}_T(\cdot, \omega)$. For $\varepsilon > 0$, we say a solution $\mathcal{S}$ is bad, if $\rho_T(\mathcal{S}) < (1/2 - \varepsilon) \cdot OPT$.

THEOREM 7. *for any $\varepsilon > 0$, $\varepsilon_1 \in (0, 2\varepsilon)$, $\delta_1, \delta_2 > 0$, if:*

(a) $\Pr_{\omega \sim \Omega} \{\hat{\rho}_T(\mathcal{S}^*, \omega) \geq (1 - \varepsilon_1) \cdot OPT\} \geq 1 - \delta_1$
(b) for any bad $\mathcal{S}$,

$\Pr_{\omega \sim \Omega} \left\{\hat{\rho}_T(\mathcal{S}, \omega) \geq \frac{1}{2}(1 - \varepsilon_1) \cdot OPT\right\} \leq \frac{\delta_2}{T^k \cdot \binom{N}{k}}$

(c) for all $\omega \sim \Omega$, $\hat{\rho}_T(\mathcal{S}, \omega)$ is monotone and submodular with respect to $\mathcal{S}$.

So that, $\Pr_{\omega \sim \Omega} \left\{\rho_T\left(\hat{\mathcal{S}}^g(\omega)\right) \geq \left(\frac{1}{2} - \varepsilon\right) \cdot OPT\right\} \geq 1 - \delta_1 - \delta_2$

PROOF. Because $\hat{\rho}_T(\mathcal{S}, \omega)$ is monotone and submodular with respect to $\mathcal{S}$, due to the property of partition matroid, we know the greedy solution returns a 1/2-approximate solution.

$$\hat{\rho}_T\left(\hat{\mathcal{S}}^g(\omega), \omega\right) \geq \frac{1}{2}\hat{\rho}_T(\mathcal{S}^*, \omega)$$

With (a), we have at least $1 - \delta_1$ probability that:

$$\hat{\rho}_T\left(\hat{\mathcal{S}}^g(\omega), \omega\right) \geq \frac{1}{2}(1 - \varepsilon_1) \cdot OPT$$

In T round, the number of the k size seed set is at most $T^k \cdot \binom{N}{k}$. The probability that every bad $\mathcal{S}$ satisfy (b) is less than $\frac{\delta_2}{T^k \cdot \binom{N}{k}}$, so the probability that existing a $\mathcal{S}$ to make $\rho_T(\mathcal{S}, \omega) \geq \frac{1}{2}(1 - \varepsilon_1) \cdot OPT$ is at most δ_2.

So $\Pr_{\omega \sim \Omega} \left\{\rho_T\left(\hat{\mathcal{S}}^g(\omega)\right) \geq \left(\frac{1}{2} - \varepsilon\right) \cdot OPT\right\} \geq 1 - \delta_1 - \delta_2$ □

We will use the concentration inequality to find out how much θ is sufficient to satisfy the conditions (a) and (b) in Theorem 7. For all subsequences of length θ, $\mathcal{R}[\theta]$ in the probability space Ω, each PW-RR set is also independent of each other, so we can use Chernoff bounds of independent sequences to analyze, which is more simple and intuitive[6].

THEOREM 8. *For any $\varepsilon > 0$, $\varepsilon_1 \in (0, 2\varepsilon)$, $\delta_1, \delta_2 > 0$:*

$$\theta^{(1)} = \frac{2w_1 N \cdot T \cdot \ln \frac{1}{\delta_1}}{OPT \cdot \varepsilon_1^2}, \theta^{(2)} = \frac{w_1 N \cdot T \cdot \ln\left(\frac{T^k \cdot \binom{N}{k}}{\delta_2}\right)}{OPT \cdot \left(\varepsilon - \frac{1}{2}\varepsilon_1\right)^2}$$

For any fixed $\theta > \theta^{(1)}$, $\Pr_{\omega \sim \Omega} \{\hat{\rho}_T(\mathcal{S}^*, \omega) \geq (1 - \varepsilon_1) \cdot OPT\} \geq 1 - \delta_1$
For any fixed $\theta > \theta^{(2)}$, any bad $\mathcal{S}$,

$$\Pr_{\omega \sim \Omega} \left\{\hat{\rho}_T(\mathcal{S}, \omega) \geq \frac{1}{2}(1 - \varepsilon_1) \cdot OPT\right\} \leq \frac{\delta_2}{T^k \cdot \binom{N}{k}}$$

Now we discuss the setting of parameters $\varepsilon_1, \delta_1, \delta_2$. The Settings of these parameters are not unique, and the method we describe below follows the settings in the original IMM algorithm. According to Theorem 8, assuming that OPT is known, the target of these parameters is to make the output of greedy solution $\hat{S}^g$ is the $1/2 - \varepsilon$ approximation of optimal solution with probability at least $1 - 1/(2N^l)$. The high probability of $1 - 1/(2N^l)$ was achieved because, in the next step, we would use the same high probability of $1 - 1/(2N^l)$ to obtain a better lower-bound estimate of OPT. Thus, the correctness of the overall algorithm would be guaranteed for an assignment with a high probability of $1 - 1/(N^l)$. The following corollary give the setting of the parameters.

Corollary 1. *Set* $\delta_1 = \delta_2 = \frac{1}{4n^l}, \varepsilon_1 = \varepsilon \cdot \frac{\alpha}{\frac{1}{2}\alpha+\beta}$

$$\alpha = \sqrt{l \ln N + \ln 4}, \beta = \sqrt{\frac{1}{2} \cdot \left(\ln \binom{N}{k} + l \ln N + \ln 4 + k \ln T \right)}$$

For any fixed $\theta > \frac{2N \cdot T \cdot \left[\frac{1}{2}\alpha+\beta \right]^2}{\varepsilon^2 \cdot OPT}$, *if the input of PRM-NodeSelection is* $\mathcal{R}_0[\theta], \mathcal{R}_0 \sim \Omega$,*the probability that PRM-NodeSelection's output* $\hat{S}^g(\mathcal{R}_0[\theta])$ *is the* $(1/2 - \varepsilon)$ *approximation of the optimal solution is at least* $1 - \frac{1}{2N^l}$.

17 PROOF OF THEOREM 7

For any $\varepsilon > 0, \varepsilon_1 \in (0, 2\varepsilon), \delta_1, \delta_2 > 0$:

$$\theta^{(1)} = \frac{2w_1 N \cdot T \cdot \ln \frac{1}{\delta_1}}{OPT \cdot \varepsilon_1^2}, \theta^{(2)} = \frac{w_1 NT \cdot \ln \left(\frac{T^k \cdot \binom{N}{k}}{\delta_2} \right)}{OPT \cdot \left(\varepsilon - \frac{1}{2}\varepsilon_1 \right)^2}$$

For any fixed $\theta > \theta^{(1)}$, $\Pr_{\omega \sim \Omega} \{ \hat{\rho}_T(S^*, \omega) \geq (1 - \varepsilon_1) \cdot OPT \} \geq 1 - \delta_1$
For any fixed $\theta > \theta^{(2)}$, every bad S,

$$\Pr_{\omega \sim \Omega} \left\{ \hat{\rho}_T(S, \omega) \geq \frac{1}{2} (1 - \varepsilon_1) \cdot OPT \right\} \leq \frac{\delta_2}{T^k \cdot \binom{N}{k}}$$

PROOF. When $\theta > \theta^{(1)}$ Notice that:

$$\hat{\rho}_T(S, \mathcal{R}) = \frac{N \cdot T}{\theta} \sum_{j=1}^{\theta} Y_j^{\mathcal{R}}(S)$$

$$\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_T \left(\mathcal{S}^*, \mathcal{R}_0 \right) < (1 - \varepsilon_1) \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \frac{N \cdot T}{\theta} \cdot \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0} \left(\mathcal{S}^* \right) < (1 - \varepsilon_1) \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0} \left(\mathcal{S}^* \right) < \frac{\theta}{N \cdot T} \cdot (1 - \varepsilon_1) \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0} \left(\mathcal{S}^* \right) - \theta \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{N \cdot T} < \frac{\theta (1 - \varepsilon_1) OPT}{N \cdot T} - \theta \frac{\rho_T \left(\mathcal{S}^* \right)}{N \cdot T} \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0} \left(\mathcal{S}^* \right) - \theta \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{N \cdot T} < -\varepsilon_1 \cdot \left(\theta \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{N \cdot T} \right) \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} \frac{Y_j^{\mathcal{R}_0} \left(\mathcal{S}^* \right)}{w_1} - \theta \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{w_1 N \cdot T} < -\varepsilon_1 \cdot \left(\theta \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{w_1 N \cdot T} \right) \right\}$$

$$\leq \exp \left(-\frac{\varepsilon_1^2}{2} \theta \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{w_1 N \cdot T} \right)$$

$$\leq \exp \left(-\frac{\varepsilon_1^2}{2} \cdot \frac{2 w_1 N \cdot T \cdot \ln \frac{1}{\delta_1}}{OPT \cdot \varepsilon_1^2} \cdot \frac{\rho_T \left(\mathcal{S}^* \right)}{N \cdot T} \right) = \delta_1$$

When $\theta > \theta^{(2)}$, set $\varepsilon_2 = \varepsilon - \frac{1}{2}\varepsilon_1$, $\rho_T(\mathcal{S}) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT$

$$\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta} \left(\mathcal{S}, \mathcal{R}_0 \right) \geq \frac{1}{2} (1 - \varepsilon_1) \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \frac{N \cdot T}{\theta} \cdot \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0}(\mathcal{S}) \geq \frac{1}{2} (1 - \varepsilon_1) \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0}(\mathcal{S}) - \theta \cdot \frac{\rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\theta}{N \cdot T} \left[\frac{1}{2} (1 - \varepsilon_1) \cdot OPT - \rho_T(\mathcal{S}) \right] \right\}$$

$$/*\rho_T(\mathcal{S}) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT */$$

$$\leq \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} Y_j^{\mathcal{R}_0}(\mathcal{S}) - \theta \cdot \frac{\rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\theta}{N \cdot T} \cdot \varepsilon_2 \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta} \frac{Y_j^{\mathcal{R}_0}(\mathcal{S})}{w_1} - \theta \cdot \frac{\rho_T(\mathcal{S})}{w_1 N \cdot T} \geq \left(\varepsilon_2 \cdot \frac{OPT}{\rho_T(\mathcal{S})} \right) \cdot \theta \cdot \frac{\rho_T(\mathcal{S})}{w_1 N \cdot T} \right\}$$

$$\leq \exp \left(-\frac{\left(\varepsilon_2 \cdot \frac{OPT}{\rho_T(\mathcal{S})} \right)^2}{2 + \frac{2}{3} \left(\varepsilon_2 \cdot \frac{OPT}{\rho_T(\mathcal{S})} \right)} \cdot \theta \cdot \frac{\rho_T(\mathcal{S})}{w_1 N \cdot T} \right)$$

$$\leq \exp \left(-\frac{\varepsilon_2^2 \cdot OPT^2}{2\rho_T(\mathcal{S}) + \frac{2}{3}\varepsilon_2 \cdot OPT} \cdot \theta \cdot \frac{1}{w_1 N \cdot T} \right)$$

$$\leq \exp\left(-\frac{\varepsilon_2^2 \cdot OPT^2}{2\left(\frac{1}{2} - \varepsilon\right) \cdot OPT + \frac{2}{3}\left(\varepsilon - \frac{1}{2}\varepsilon_1\right) \cdot OPT} \cdot \theta \cdot \frac{1}{w_1 N \cdot T}\right)$$

$$\leq \exp\left(-\frac{\left(\varepsilon - \frac{1}{2}\varepsilon_1\right)^2 \cdot OPT^2}{OPT} \cdot \frac{w_1 N \cdot T \cdot \ln\left(\frac{T^k \cdot \binom{N}{k}}{\delta_2}\right)}{OPT \cdot \left(\varepsilon - \frac{1}{2}\varepsilon_1\right)^2} \cdot \frac{1}{w_1 N \cdot T}\right)$$

$$= \frac{T^k \cdot \binom{N}{k}}{\delta_2}$$

$\square$

Theorem 9. *The probability of $LB \leq OPT$ is at least $1 - \frac{1}{2N^l}$, which means that the probability of*
$\tilde{\theta} \geq \frac{2nt \cdot \left[\frac{1}{2} \cdot \alpha + \beta\right]^2}{\varepsilon^2 \cdot OPT}$ *is at least $1 - \frac{1}{2N^l}$.*

18 PROOF OF THEOREM 9

Theorem 10. *For any $i = 1, 2, \cdots, \lfloor \log_2 N \rfloor - 1$,*

(1) if $x_i = \frac{\sum_1^k w_t \cdot N}{2^i} > OPT$, the probability of $\hat{\rho}_{T\theta_i}\left(\mathcal{S}_i, \mathcal{R}_0[\theta_i]\right) \geq \left(1 + \varepsilon'\right) \cdot x_i$ is at most $\frac{1}{2N^l \log_2 N}$.

(2) if $x_i = \frac{\sum_1^k w_t \cdot N}{2^i} \leq OPT$, the probability of $\hat{\rho}_{T\theta_i}\left(\mathcal{S}_i, \mathcal{R}_0[\theta_i]\right) \geq \left(1 + \varepsilon'\right) \cdot OPT$ is at most $\frac{1}{2N^l \log_2 N}$.

Proof. For any k size seed set $\mathcal{S}$.

$$\hat{\rho}_{T\theta_i}\left(\mathcal{S}, \mathcal{R}_0[\theta_i]\right) = \frac{N \cdot T \cdot \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S})}{\theta_i}$$

$$\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_i}\left(\mathcal{S}, \mathcal{R}_0\right) \geq (1 + \varepsilon') \cdot x_i \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \frac{N \cdot T}{\theta_i} \cdot \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) \geq (1 + \varepsilon') \cdot x_i \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\theta_i\,(1 + \varepsilon')\,x_i}{N \cdot T} - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \right\}$$

$/*$ because $x_i > OPT \geq \rho_T(\mathcal{S}) */$

$$\leq \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\theta_i}{N \cdot T} \varepsilon' \cdot x_i \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\varepsilon' \cdot x_i}{\rho_T(\mathcal{S})} \cdot \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} \frac{Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S})}{w_1} - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{w_1 N \cdot T} \geq \frac{\varepsilon' \cdot x_i}{\rho_T(\mathcal{S})} \cdot \frac{\theta_i \cdot \rho_T(\mathcal{S})}{w_1 N \cdot T} \right\}$$

$$\leq \exp\left(- \frac{\left(\frac{\varepsilon' \cdot x_i}{\rho_T(\mathcal{S})} \right)^2}{2 + \frac{2}{3}\left(\frac{\varepsilon' \cdot x_i}{\rho_T(\mathcal{S})} \right)} \cdot \frac{\theta_i \cdot \rho_T(\mathcal{S})}{w_1 N \cdot T} \right) \quad /* \text{ chernoff bound } */$$

$$= \exp\left(- \frac{(\varepsilon' \cdot x_i)^2}{2\rho_T(\mathcal{S}) + \frac{2}{3}(\varepsilon' \cdot x_i)} \cdot \frac{\theta_i}{w_1 N \cdot T} \right)$$

$$\leq \exp\left(- \frac{\varepsilon'^2 \cdot x_i}{2 + \frac{2}{3}\varepsilon'} \cdot \frac{\theta_i}{w_1 N \cdot T} \right)$$

$$\leq \exp\left(- \frac{\varepsilon'^2 \cdot x_i}{2 + \frac{2}{3}\varepsilon'} \right.$$
$$\cdot \frac{w_1 N \cdot T \left(2 + \frac{2}{3}\varepsilon' \right) \left(\ln T^k + \ln \binom{N}{k} + l \ln N + \ln 2 + \ln \log_2 N \right)}{\varepsilon'^2 x_i}$$
$$\left. \cdot \frac{1}{w_1 N \cdot T} \right)$$

$$= \frac{1}{2 T^k \cdot \binom{N}{k} N^l \log_2 N}$$

For any k size set $\mathcal{S}$.

$$\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_i}(\mathcal{S}, \mathcal{R}_0) \geq (1 + \varepsilon') \cdot OPT \right\}$$

$$= \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \frac{N \cdot T}{\theta_i} \cdot \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) \geq (1 + \varepsilon') \cdot OPT \right\}$$

$$\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\theta_i}{N \cdot T}(1 + \varepsilon') \cdot OPT - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \right\}$$

$$/* OPT \geq \rho_T(\mathcal{S}) */$$

$$\leq \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\theta_i}{N \cdot T} \varepsilon' \cdot OPT \right\}$$

$$\leq \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S}) - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \geq \frac{\varepsilon' \cdot OPT}{\rho_T(\mathcal{S})} \cdot \frac{\theta_i \cdot \rho_T(\mathcal{S})}{N \cdot T} \right\}$$

$$\leq \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \sum_{j=1}^{\theta_i} \frac{Y_j^{\mathcal{R}_0[\theta_i]}(\mathcal{S})}{w_1} - \frac{\theta_i \cdot \rho_T(\mathcal{S})}{w_1 N \cdot T} \geq \frac{\varepsilon' \cdot OPT}{\rho_T(\mathcal{S})} \cdot \frac{\theta_i \cdot \rho_T(\mathcal{S})}{w_1 N \cdot T} \right\}$$

$$\leq \exp \left(-\frac{\left(\frac{\varepsilon' \cdot OPT}{\rho_T(\mathcal{S})} \right)^2}{2 + \frac{2}{3} \left(\frac{\varepsilon' \cdot OPT}{\rho_T(\mathcal{S})} \right)} \cdot \frac{\theta_i \cdot \rho_T(\mathcal{S})}{w_1 N \cdot T} \right)$$

$$= \exp \left(-\frac{(\varepsilon' \cdot OPT)^2}{2\rho_T(\mathcal{S}) + \frac{2}{3}(\varepsilon' \cdot OPT)} \cdot \frac{\theta_i}{w_1 N \cdot T} \right)$$

$$\leq \exp \left(-\frac{\varepsilon'^2 \cdot OPT}{2 + \frac{2}{3}\varepsilon'} \cdot \frac{\theta_i}{w_1 N \cdot T} \right)$$

$$\leq \exp \left(-\frac{\varepsilon'^2 \cdot x_i}{2 + \frac{2}{3}\varepsilon'} \cdot \frac{\theta_i}{w_1 N \cdot T} \right)$$

$$\leq \exp \left(-\frac{\varepsilon'^2 \cdot x_i}{2 + \frac{2}{3}\varepsilon'} \right.$$

$$\cdot \frac{w_1 N \cdot T \left(2 + \frac{2}{3}\varepsilon' \right) \left(\ln T^k + \ln \binom{N}{k} + l \ln N + \ln 2 + \ln \log_2 N \right)}{\varepsilon'^2 x_i}$$

$$\left. \cdot \frac{1}{w_1 n T} \right)$$

$$= \frac{1}{2T^k \cdot \binom{N}{k} N^l \log_2 N}$$

With the union bound $\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_i}(\mathcal{S}_i, \mathcal{R}_0) \geq (1 + \varepsilon') \cdot OPT \right\} \leq \frac{1}{2N^l \log_2 N}$

$\square$

With Theorem 10 we know that LB is a lower bound of OPT with high probability, so θ satisfy the Corollary 1. Further we can know that the probability of $LB < OPT$ is at least $1 - 1/2N^l$.

PROOF. Set $LB_i = \frac{\hat{\rho}_{T\theta_i}(\mathcal{S}_i, \mathcal{R}_0)}{(1+\varepsilon')}$. When $OPT \geq x_{\lfloor \log_2 N \rfloor - 1}$.Set $i \geq 1$ is the smallest index to make $OPT \geq x_i$.

For any $i' \leq i - 1$, $OPT < x_{i'}$,

For any $i'' > i - 1$, $OPT \geq x_{i''}$,

We define the event ε as: for any $i' \leq i - 1$, $\hat{\rho}_{T\theta_{i'}}(\mathcal{S}_{i'}, \mathcal{R}_0) < (1 + \varepsilon') x_{i'}$, and for any $i'' \geq i$, $\hat{\rho}_{T\theta_{i'}}(\mathcal{S}_{i'}, \mathcal{R}_0) \geq (1 + \varepsilon') x_{i'}$

Notice that $i', i'' \geq 1$, so when $i = 1$, i' is not exist. The part of Event ε about i' is true. Event ε is the event that we expected. Because as Event ε happens, $LB = LB_i$ or $LB = 1$. So Event ε indicate that $LB \leq OPT$ so the upper bound of Event ε not happen is that:

$$\Pr_{\mathcal{R}_0 \sim \Omega} \{\neg \varepsilon\} \leq \sum_{i'=1}^{i-1} \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_{i'}}(\mathcal{S}_{i'}, \mathcal{R}_0) \geq (1 + \varepsilon')x_{i'} \right\} +$$
$$\sum_{i''=i-1}^{\lfloor \log_2 N \rfloor - 1} \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_{i''}}(\mathcal{S}_{i''}, \mathcal{R}_0) \geq (1 + \varepsilon')OPT \right\}$$

With the above, we know that $\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_{i'}}(\mathcal{S}_{i'}, \mathcal{R}_0) \geq (1 + \varepsilon')x_{i'} \right\} \leq \frac{1}{2N^l \log_2 N}$. And $\Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \hat{\rho}_{T\theta_{i''}}(\mathcal{S}_{i''}, \mathcal{R}_0) \geq (1 + \frac{1}{2N^l \log_2 N} \cdot \text{So } \Pr_{\mathcal{R}_0 \sim \Omega} \{\neg \varepsilon\} \leq \frac{1}{2N^l}.$

When $OPT < x_{\lfloor \log_2 n \rfloor - 1}$. $LB = 1$ with probability at least $1 - \frac{1}{2N^l}$. $\square$

For any $\varepsilon > 0, l > 0$, PRM-IMM Guarantees that $\hat{\mathcal{S}}^g$ is the $\frac{1}{2} - \varepsilon$ approximation of OPT with probability at least $1 - \frac{1}{N^l}$. Define the Event ε as the $LB \leq OPT$, and put the $\mathcal{R}'_0[\tilde{\theta}]$ to PRM-NodeSelection to get the seed set $\hat{\mathcal{S}}^g$ is the $\frac{1}{2} - \varepsilon$ approximation of the PRM problem.

$$\rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}'_0[\tilde{\theta}] \right) \right) \geq \left(\frac{1}{2} - \varepsilon \right) \cdot OPT$$

with union bound:

$$\Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}'_0 \sim \Omega} \{\neg \varepsilon\} \leq \Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}'_0 \sim \Omega} \{LB > OPT\}$$
$$+ \Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}'_0 \sim \Omega} \left\{ LB \leq OPT \wedge \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}'_0[\tilde{\theta}] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$$

And $\Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}'_0 \sim \Omega} \{LB > OPT\} = \Pr_{\mathcal{R}_0 \sim \Omega} \{LB > OPT\} \leq \frac{1}{2N^l}.$

Now we know that:
$\Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}'_0 \sim \Omega} \left\{ LB \leq OPT \wedge \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}'_0[\tilde{\theta}] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$

When $LB \leq OPT, \tilde{\theta} \geq \frac{2N \cdot \left[\frac{1}{2} \cdot \alpha + \beta \right]^2}{\varepsilon^2 \cdot OPT}$, and $\tilde{\theta} \leq \frac{2N \cdot \left[\frac{1}{2} \cdot \alpha + \beta \right]^2}{\varepsilon^2}$.

$$\theta_{\min} = \left\lceil \frac{2N \cdot \left[\frac{1}{2} \cdot \alpha + \beta\right]^2}{\varepsilon^2 \cdot OPT} \right\rceil, \theta_{\max} = \left\lfloor \frac{2N \cdot \left[\frac{1}{2} \cdot \alpha + \beta\right]^2}{\varepsilon^2} \right\rfloor. \tilde{\theta} \text{ is a integer range from } \theta_{min} \text{ to } \theta_{max}.$$

$$\Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}_0' \sim \Omega} \left\{ LB \leq OPT \wedge \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}_0'[\tilde{\theta}] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$$

$$\leq \Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}_0' \sim \Omega} \left\{ \tilde{\theta} \geq \theta_{\min} \wedge \tilde{\theta} \leq \theta_{\max} \wedge \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}_0'[\tilde{\theta}] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$$

$$= \sum_{\theta = \theta_{\min}}^{\theta_{\max}} \Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}_0' \sim \Omega} \left\{ \tilde{\theta} = \theta \wedge \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}_0'[\tilde{\theta}] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$$

$$= \sum_{\theta = \theta_{\min}}^{\theta_{\max}} \Pr_{\mathcal{R}_0 \sim \Omega \mathcal{R}_0' \sim \Omega} \left\{ \tilde{\theta} = \theta \wedge \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}_0'[\theta] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$$

/* because $\mathcal{R}_0$ is independent with $\mathcal{R}_0'$ * /

$$= \sum_{\theta = \theta_{\min}}^{\theta_{\max}} \Pr_{\mathcal{R}_0 \sim \Omega} \{ \tilde{\theta} = \theta \} \cdot \Pr_{\mathcal{R}_0' \sim \Omega} \left\{ \rho_T \left(\hat{\mathcal{S}}^g \left(\mathcal{R}_0'[\theta] \right) \right) < \left(\frac{1}{2} - \varepsilon \right) \cdot OPT \right\}$$

$$\leq \sum_{\theta = \theta_{\min}}^{\theta_{\max}} \Pr_{\mathcal{R}_0 \sim \Omega} \left\{ \tilde{\theta} = \theta \right\} \cdot \frac{1}{2N^l}$$

$$= \frac{1}{2N^l}$$

So $\Pr_{\mathcal{R}_0 \sim \Omega, \mathcal{R}_0' \sim \Omega} \{\neg \varepsilon\} \leq \frac{1}{N^l}$, which means that with probability at least $1 - \frac{1}{N^l}$, the output of PRM-IMM $\hat{\mathcal{S}}^g$ is the $\frac{1}{2} - \varepsilon$ approximation of OPT.

19 TIME COMPLEXITY

The time complexity of RPM-IMM is $O((k + l)(M + N)T \log(NT)/\varepsilon^2)$.

PROOF. We use the Martingale theorem in the IMM algorithm[32] to estimate the time complexity of the PRM-IMM algorithm. Then the time complexity of the IMM algorithm is $O(\mathbb{E}[\bar{\theta} + \tilde{\theta}] \cdot (EPT + 1))$, where $\mathbb{E}[\bar{\theta} + \tilde{\theta}]$ is the overall number of PW-RR sets needed to be generated.

$$\mathbb{E}[\bar{\theta} + \tilde{\theta}] \leq \frac{8 \left(\lambda^* + \lambda' \right) \cdot (1 + \varepsilon')^2}{\cdot OPT} + 2,$$

where

$$\lambda^* = \frac{4w_1 NT \cdot (\cdot \alpha + \beta)^2}{\varepsilon^2}$$

$$\lambda' = \frac{w_1 NT \cdot \left(2 + \frac{2}{3}\varepsilon' \right) \cdot \left(\ln \binom{N}{k} + \ell \ln N + \ln 2 + \ln \log_2 N + \ln T^k \right)}{\varepsilon'^2},$$

and α and β is defined in section 5.

Therefore, $\mathbb{E}[\bar{\theta} + \tilde{\theta}] = O \left(\frac{w_1 (k+l) NT \log NT}{OPT \varepsilon^2} \right)$. And $EPT = \frac{M}{N} \cdot \mathbb{E}[\sigma(v)]$ is the expected running time of generating a PW-RR set. Because $\mathbb{E}[\sigma(v)] \leq \frac{OPT}{w_1}$, so the expected running time is:

$$O \left(\frac{(k + l)(N + M)T \log NT}{\varepsilon^2} \right)$$

□