1-1-2023

Maximizing External Action with Information Provision over Multiple Rounds in Online Social Networks

Masaaki Miyashita
Soka University

Norihiko Shinomiya
Soka University

Daisuke Kasamatsu
Soka University

Genya Ishigaki
San Jose State University, genya.ishigaki@sjsu.edu

Recommended Citation
Masaaki Miyashita, Norihiko Shinomiya, Daisuke Kasamatsu, and Genya Ishigaki. "Maximizing External Action with Information Provision over Multiple Rounds in Online Social Networks" *IEICE Transactions on Information and Systems* (2023): 847-855. https://doi.org/10.1587/transinf.2022DAP0007

PAPER *Special Section on Data Engineering and Information Management*

Maximizing External Action with Information Provision Over Multiple Rounds in Online Social Networks

Masaaki MIYASHITA[†], ***Member***, **Norihiko SHINOMIYA**[†a)], ***Senior Member***, **Daisuke KASAMATSU**[†], ***Member***, ***and*** **Genya ISHIGAKI**[††], ***Nonmember***

SUMMARY Online social networks have increased their impact on the real world, which motivates information senders to control the propagation process of information to promote particular actions of online users. However, the existing works on information provisioning seem to oversimplify the users' decision-making process that involves information reception, internal actions of social networks, and external actions of social networks. In particular, characterizing the best practices of information provisioning that promotes the users' external actions is a complex task due to the complexity of the propagation process in OSNs, even when the variation of information is limited. Therefore, we propose a new information diffusion model that distinguishes user behaviors inside and outside of OSNs, and formulate an optimization problem to maximize the number of users who take the external actions by providing information over multiple rounds. Also, we define a robust provisioning policy for the problem, which selects a message sequence to maximize the expected number of desired users under the probabilistic uncertainty of OSN settings. Our experiment results infer that there could exist an information provisioning policy that achieves nearly-optimal solutions in different types of OSNs. Furthermore, we empirically demonstrate that the proposed robust policy can be such a universally optimal solution.

key words: *online social networks, decision-making, external behavior, information integration theory*

1. Introduction

In recent years, the rapid spread of social media has led to the formation of huge online social networks (OSNs) involving not only individuals but also business organizations and governments. The propagation of information in OSNs has the potential strongly to impact the real world because it increases the visibility of the information in an unconventional way. For example, an obscure product might become the center of attention after a night [1], or a protest against a controversial policy could make the government to refine it [2]. In contrast, the spread of misinformation and false rumors can cause social confusion by promoting inappropriate behaviors to an anonymous crowd, and the spread of false perceptions can lead to reputational damage and counterproductive political decisions.

The potential social impacts of OSNs motivate practices to control the propagation process by tuning information that users receive. More precisely, several methods have been proposed to increase the positive effects or suppress the negative effects of the diffusion of misinformation [3]–[5]. The viral marketing, where marketers advertise their products to promote buying behaviors, is a classical example. Other examples include awareness campaigns in OSNs for prosocial behavior such as health promotion and environmental conservation activities. Another more recent example is public relations for infection prevention practices and vaccination against COVID-19 by several organizations. This paper investigates information provision strategies that aim to promote particular actions of OSN users outside of OSNs.

Influence diffusion models in social networks, including OSNs, focus on the process of information propagation facilitated by the interactions among users [3], [4]. In this sense, the "influence" defined in the existing works induces changes in user behaviors *inside* of social networks. For example, an OSN user gets motivated to share a post promoting a product with their followers. However, our interests lie in the influence toward user behaviors *beyond* social networks, where the users receive information. With the previous example, we want to understand if the users who shared the promotional post actually purchase the product. When revisiting the existing works with this new definition of influence, the existing works seem to oversimplify the decision-making process that involves information reception, actions inside of social networks, and actions outside of social networks. For example, studies [6]–[10] consider product adoption, but all studies define the interaction between each person and product adoption as tightly coupled. In other words, someone who is influenced in a social network is automatically considered as a person who takes the associated action outside of the social network. Therefore, we propose a new information diffusion model that distinguishes user behaviors inside and outside of OSNs. For simplicity, we name those behaviors internal and external behaviors (actions).

In a situation where the two types of actions, internal and external, are independent, information providers need to have a strategy that maximizes the number of users taking the external action. Since a user does not take both of the actions unless it receives the information, it is essential to provide the information spreadable, i.e., that urges the user to take the internal action. If the tendencies to diffuse and to take an external action are similar, it is possible to increase

Manuscript received June 20, 2022.
Manuscript revised November 15, 2022.
Manuscript publicized February 3, 2023.
[†]The authors are with Graduate School of Engineering, Soka University, Hachioji-shi, 192–8577 Japan.
[††]The author is with Department of Computer Science, San José State University, San José, CA, USA, 95192.
a) E-mail: shinomi@ieee.org
DOI: 10.1587/transinf.2022DAP0007

the number of users who take the external action by simply providing information that affects users to diffuse. On the contrary, if the two tendencies are conflicted, information expanding its diffusion will not obtain more users who take the external action, and information promoting the intention of the external action will not grow diffusion, so these users will not increase.

In psychology, the behavior changes by information reception is called attitude change that is generalized as a transition of internal states induced by external stimuli [11]. Depending on the timing and the strength of the stimuli in an OSN, such as information propagation over multiple rounds, it is possible to further facilitate the external action. Here, given the discrete and feasible variations of information, a goal of information providers is to find a sequence of the information to get more users to take the external action among all permutations of the information variations. However, even if the information variation is simple, the optimal sequence of provided information is not trivial due to the complexity of OSNs.

Therefore, this paper discusses an optimization problem to maximize the number of users who take the external action after the information retrieval in an OSN. To distinguish the internal and external actions of users, we define a new information diffusion model that represents a situation where an information provider tries to promote a certain external action by providing information in multiple rounds. Our experimental results infer that there could exist an information provisioning policy that achieves nearly-optimal solutions in different types of OSNs. Furthermore, we empirically demonstrate that the proposed robust policy, which selects a message sequence to maximize the expected number of users who take the external action under the probabilistic uncertainty of OSN settings, is a competitive candidate for such a universally optimal solution.

2. Related Works

2.1 Diffusion Models

The Linear Threshold (LT) and the Independent Cascade (IC) models are basic diffusion models [12]. In the LT model, a node is active if the sum of the weights given by the neighboring activated nodes is over its threshold. In the IC model, every active node tries to activate inactivated neighbors once, according to a probability given on the edge of them. From the perspective of information diffusion, the LT model is a receiver-centric model in which the timing of information receipt depends on the receiver, while the IC model is a sender-centric model in which the timing depends on the sender [3]. The IC model seems more appropriate than the LT model for representing the user's response to the information received from each neighbor. Both of the models cannot distinguish between the reception and dissemination of information by a user from the definitions of activation. Thus, we define a new diffusion model to capture the effect of received information on dissemination behavior.

2.2 Influence Maximization

The Influence Maximization (IM) problem is to maximize the total number of active nodes by selecting a given number of nodes as a seed set for general diffusion models [4], [12]. The objective function of this problem takes the expected value of the total number of active nodes due to stochastic diffusion. The problem focuses on the starting point of the diffusion but does not consider what content to diffuse. The difference is that IM selects a set of users for a single diffusion, while our problem finds information content for each diffusion over multiple rounds.

2.3 Applied Research

Our model takes account of (1) an external action of social networks, (2) internal states of users' behavior, and (3) the effect of diffusion content (information). Many research projects proposed applying diffusion models and the IM problem, but there seems to be no research satisfying all the points to the best of our knowledge. We now describe some research related to each of them.

Two studies [8], [9] distinguish product adoption as an external action from interaction among users. Bhagat et al. aim to maximize the number of adopters of a product [8]. In their model, each activated user tries to adopt the product. It also has some influence on others despite the success of its adoption. Liang et al. proposed to maximize the revenue obtained from product adoption by multiple advertising a competitive product of an enterprise for OSNs [9]. Users who have not purchased any products purchase a promoted product if its price is within their budget, but they would promote the product whether or not they adopt it.

Hudson et al. suggested opinion maximization with an information diffusion model in OSNs that captures own personalities and opinions, which correspond to users' internal states [13].

Barbieri et al. proposed a topic-aware model to capture the effect of a generic item that is the subject of a viral marketing campaign on the influence of users [14].

Teng et al. aim to maximize product adoption based on user preferences for a product and the characteristics of products that change the preferences [10]. The preferences and the characteristics correspond to users' internal states and the effects of diffusion content, respectively.

3. Model Definition

In this section, we define a decision-making model and an information diffusion model to describe the internal and external actions.

3.1 Model Concept

Social media like Twitter and Facebook basically consist of follower-followee relationships among their users, where users *follow* others to obtain information from them. A *follower* of a user is who follows the user, and a *followee* of a user who a user follows.

Considering social media comprises an *information provider (IP)* and *OSN users* (hereafter abbreviated as users) with the follow-followee relationships, we assume that an OSN follows the conditions below.

External behavior

- An IP has a service outside of an OSN, and users act as consumers of the service.
- Every user can consume the service only once. (We call this consumption an *external action*.)
- An IP can measure the consumption of the service but cannot determine which general users consumed it.

Internal behavior

- An IP posts *messages* to publicize or promote her service on an OSN.
- An IP supplies each message after the spread of a previous message stops. (We call the period from the posting of a message to the stop of the spread of the message a *round*.)
- Users can share each post of an IP only once. (We name this sharing an *internal action*.)
- Any users do not post their own messages.
- When a user shares a post, it is conveyed to the user's followees who have not received it before the sharing.
- Users voluntarily view a message communicated by their followees.

Decision-making of internal/external action

- Users have intentions for the internal action (sharing posts of an IP) and the external action (the consumption of the service of an IP).
- Users decide whether to take each action or not on the change of its intention.
- A message affects users who receive it to change the intentions of both actions.
- An IP can freely post messages having some degree of the effect, but cannot post messages strongly affecting both of the internal and external actions.

Although real social media like Twitter and Facebook have various user operations: creating a post, liking a post, replying to a post (replying/commenting) and sharing a post (retweeting/sharing), we assume that users do not create their own messages and an IP does not share messages from users in order to focus on the effect of IP's messages on users.

Here is an example for the model concept: the promotion of a COVID-19 vaccination by Minister of Health, Labour and Welfare (MHLW) of the Japanese government. Table 1 shows the concrete examples corresponding to each component of the concept.

Table 1 MHLW Promotion of COVID-19 vaccinations

Component	Example
IP	MHLW
OSN users	OSN users residing in Japan
Type of messages	Promotion of COVID-19 vaccinations
Internal action	Sharing posts of MHLW
External action	Vaccination

3.2 User's Decision-Making Formulation

3.2.1 Information Integration Theory

Information Integration Theory (IIT), one of the approaches to attitude change in psychology, models the response of an individual to a series of information [15]. Its capability of mathematically representing user behaviors enables us to formulate our optimization problem about the information propagation process in OSNs. A basic model of IIT is proposed as follows:

$$R = w_0 s_0 + \sum_{i=1} w_i s_i, \tag{1}$$

where R is the outcome of the response, $w_0 s_0$ represents an initial opinion, and w_i and s_i denote a weight and a stimulus of the i-th piece of information, respectively. A weight indicates the degree of importance of its opinion compared with an initial opinion and the opinions of other pieces. A stimulus means the opinion of its piece. The model is called a weighted averaging model if $\sum_{i=0} w_i = 1$, or an additive model otherwise.

This model cannot show an intermediate opinion in the integration because the second term of the right-hand side of Eq. (1) stands for the integration of a series of information. Therefore, a variant of IIT, the *serial integration model* [16], is proposed as follows:

$$r_i = w_i s_i + (1 - w_i) r_{i-1}, \tag{2}$$

where a response r_i indicates an opinion changed by a i-th stimulus s_i and a prior opinion r_{i-1}. Jacoby et al. tested the serial model by measuring variations of the cognition of subjects who received successive information. It is shown that the mathematical properties of the model are consistent with the experimental results [17]. The serial model, which is based on IIT, enables us to formally represent the users' opinions (or behaviors) that are constantly influenced by a sequence of information received by the users.

3.2.2 Model Formulation

Suppose that the decision-making of the internal and external actions of OSN users follow the serial model, i.e., the following correspondences can be considered:

- messages of an IP, and pieces of information,

- the intention of each action, and the opinion about taking the action or not,
- the effect of a message on each action, and the stimulus of information in the serial model.

In order to formulate the correspondences, we assume the conditions as follows:

- Suppose each user has a constant weight as the sensitivity to a message,
- Denote the opinions (responses and stimuli) of an action as polar values in the range of $[-1, 1]$; positive numbers mean taking the action and values less than or equal to zero do not,
- Let a response value be the *utility* to make a decision of an action.

Definition 1 (Decision-making model)**.** Let M be a set of available messages for an IP. A user decision-making model is a tuple $F = (u_0, w, s)$ where $u_0 \in [-1, 1]$ is an initial value of utility, $s : M \to [-1, 1]$ is a function that maps a message to a stimulus, and $w \in [0, 1]$ is a weight.

Because the sequence of responses $\{r_i\}$ represents as a history of an individual's opinion updating in Eq. (2), we define utility updating by using the equation. From the conditions of the above model concept, it is derived that a user receives at most a message in a round. Then, the utility updating of a user is formulated as follows, representing the user's receipt messages as all messages that an IP posts, and the truth values of the user's reception of each message in a certain range of rounds.

Definition 2 (Utility updating)**.** Let M be a set of available messages for an IP and let $\{m_t\} := \{m_1, \cdots, m_t\}$ be a message sequence given by an IP from the first round to the t-th round, where $m_k \subseteq M$ is a message supplied in the k-th round. Let $\{b_t\}$ be a receipt sequence in which each component b_k is a truth value denoting whether a k-th message is received. The utility updating of a decision-making model $F = (u_0, w, s)$ in the t-th round is designated as a function u_F, which maps a message sequence $\{m_t\}$ and a receipt sequence $\{b_t\}$ to a utility value, defined as follows:

$$u_F(\{m_t\}, \{b_t\}) = \begin{cases} u_0 & t = 0, \\ f(u_F(\{m_{t-1}\}, \{b_{t-1}\}), m_t, w, s) & b_t \text{ is true}, \\ u_F(\{m_{t-1}\}, \{b_{t-1}\}) & \text{otherwise}, \end{cases} \tag{3}$$

where

$$f(x, m, w, s) = wx + (1 - w)s(m). \tag{4}$$

Property 1. Given a decision-making model $F = (u_0, w, s)$ and an i.i.d. sequence $\{\mathcal{M}_t\}$ where any variable $\mathcal{M}_k$ takes a message in the k-th round. Let $\bar{u}_F(t)$ be the expected utility $\bar{u}_F(t) = \mathbb{E}[u_F(\{\mathcal{M}_t\}, \{b_t\})]$ in the t-th round. Suppose that all the values of a receipt sequence $\{b_t\}$ are true, $\bar{u}_F(t) > 0$ holds if and only if $t > \log_{1-w} \frac{\bar{s}}{\bar{s}-u_0}$ where $\bar{s} = \mathbb{E}[s(\mathcal{M}_t)]$.

Define a metric of an activity for a decision-making model to represent differences among users in the level of activity for an action.

Definition 3 (Metric of user activity)**.** Given a decision-making model $F = (u_0, w, s)$, and an i.i.d. sequence $\{\mathcal{M}_t\}$. Let μ_F be a metric of the user action of F defined as follows:

$$\frac{1}{\mu_F} = \log_{1-w} \frac{\bar{s}}{\bar{s} - u_0}, \tag{5}$$

where $\bar{s} = \mathbb{E}[s(\mathcal{M}_t)]$.

From Property 1, $\frac{1}{\mu_F}$ represents the lower bound of the round to take an action if a user receive a message at random, so the less value, the fewer messages for an action. Thus, the reciprocal μ_F can be used as a measure of the user activity of F.

While the internal action can be taken at each round, the external action can only be taken once during all rounds. These conditions can be described in predicate logic by using a decision-making model F as follows.

$$\Phi_F(\{m_t\}, \{b_t\}) = [u_F(\{m_t\}, \{b_t\}) > 0], \tag{6}$$

$$\Psi_F(\{m_t\}, \{b_t\}) = (\forall t' < t)\, [u_F(\{m_{t'}\}, \{b_{t'}\}) \leq 0] \wedge u_F(\{m_t\}, \{b_t\}) > 0. \tag{7}$$

3.3 Diffusion Model Formulation

Users do not always receive messages sent to them because users receiving messages depends on their own uses of the media. We then express the uncertainty of the receipt as a probability named *receipt probability*. We define a diffusion model with a OSN topology, an IP, OSN users, decision-making models of internal and external actions of users, and receipt probabilities.

Definition 4 (Diffusion model)**.** Given a message set M, a diffusion model is a tuple $D = (G, \sigma, p, \{S_i\}_U, \{X_i\}_U)$ where all the components are described in Table 2.

Table 2 Components of a diffusion model

Component	Description
G	A directed graph of a pair (V, E) where V is a OSN user set and each edge $(i, j) \in E \subseteq V \times V$ is the direction of message communication from i to j.
σ	An IP node in G where $\sigma \in V$; let $U = V \setminus \{\sigma\}$ be a user set except IP.
p	A function that maps a node to a receipt probability.
$\{S_i\}_U$	A family of decision-making models of the internal action indexed by U; i's decision-making model denotes $S_i = (u^S_{0,i}, w^S_i, s^S_i)$.
$\{X_i\}_U$	A family of decision-making models of the external action indexed by U; i's decision-making model denotes $X_i = (u^X_{0,i}, w^X_i, s^X_i)$.

In order to formulate the number of users who take the external action over rounds, we describe the set of users who receive messages in each round and the set of users who take the external action by using a diffusion model. Given a diffusion model $D = (G, p, \sigma, \{S_i\}_U, \{X_i\}_U)$, let $\{m_t\}$ be a message sequence supplied by σ and let $\{B_t\}$ denote a sequence in which any B_k is a receipt outcome of all users in the k-th round with p. Let R_D correspond to a set of users who receive a message m_t at the t-th round. R_D is formulated as follows:

$$R_D(\{m_t\}, \{B_t\}) = \bigcup_{l=0} r_{D,l}(\{m_t\}, \{B_t\}), \tag{8}$$

where

$$\begin{aligned} & r_{D,l}(\{m_t\}, \{B_t\}) \\ & = \begin{cases} g_{D,\sigma}(\{m_t\}, \{B_t\}) & \text{if } l = 0, \\ \bigcup_{i \in r_{D,l-1}(\{m_t\},\{B_t\})} g_{D,i}(\{m_t\}, \{B_t\}) & \\ \quad \setminus \bigcup_{l'=0}^{l-1} r_{D,l'}(\{m_t\}, \{B_t\}) & \text{if } l > 0, \end{cases} \end{aligned} \tag{9}$$

$$\begin{aligned} & g_{D,i}(\{m_t\}, \{B_t\}) \\ & \quad = \{ j \in V \mid \Phi_{S_j}(\{m_t\}, \{B_{j,t}\}) \wedge (i, j) \in E \}. \end{aligned} \tag{10}$$

$r_{D,l}$ returns the set of users who received a message at step l in a round, which is empty if the diffusion is stopped. R_D joints these user sets that are non-empty.

Also, let Y_D correspond to a set of users who take the external action at the t-th round. Then, Y_D is expressed as the equation below:

$$\begin{aligned} & Y_D(\{m_t\}, \{B_t\}) \\ & \quad = \{ i \in R_D(\{m_t\}, \{B_t\}) \mid \Psi_{X_i}(\{m_t\}, \{B_{i,t}\}) \}, \end{aligned} \tag{11}$$

By Eq. (7), sets $Y_D(\{m_1\}, \{B_1\}), \cdots, Y_D(\{m_t\}, \{B_t\})$ are mutually disjoint. Thus, the total number of users who take the external action over t rounds is denoted as the sum of the size of $Y_D(\{m_k\}, \{B_k\})$ for $1 \leq k \leq t$.

$$\xi_D(\{m_t\}, \{B_t\}) = \sum_{k=1}^{t} |Y_D(\{m_k\}, \{B_k\})|. \tag{12}$$

Note that we assume that an IP can obtain the step l when $r_{D,l}$ returns the empty set and the value ξ_D takes, although how to measure these values is beyond the scope of this paper.

4. Problem Formulation

We formulate the problem of maximizing the number of users who take the external action with information provision over multiple rounds.

4.1 Problem Definition

Since $\{B_t\}$ is a outcome of taken by the function ξ_D is a stochastic value (not determined), We define an objective function that takes the expected value of the total number of users who take the external action as follows:

$$\omega_D(\{m_t\}) = \mathbb{E}\left[\xi_D(\{m_k\}, \{\mathcal{B}_k\})\right], \tag{13}$$

where $\{\mathcal{B}_t\}$ denotes an i.i.d. sequence in which each random variable $\mathcal{B}_k$ takes a receipt outcome of all users in the k-th round based on p.

Hereinafter, we simply say "the external action" to refer to the total number of users who take the external action.

Problem 1 (External action maximization over multiple rounds (AM))**.** Given a message set M, a diffusion model $D = (G, p, \sigma, \{S_i\}_U, \{X_i\}_U)$ and an integer value T, AM finds a message sequence $\boldsymbol{m}_D^*$ of length T to maximize the expected value of the external action for T rounds, i.e., $\boldsymbol{m}_D^* = \arg\max_{\{m_T\} \in M^T} \omega_D(\{m_T\})$.

By Eq. (13), any element of M^T corresponds to a real value, and M^T is finite. Hence, there clearly exists a solution of AM.

4.2 Problem Solving

For any OSN, we assume that there exists a unique diffusion model corresponding to the OSN and that an IP estimates the model by observing the target OSN. Let $\mathcal{D}$ be a random variable for possible diffusion models that an IP takes.

If an IP can always obtain a correct diffusion model, that is, if it knows the realization of $\mathcal{D}$, then it implements a policy $\pi : D \mapsto \boldsymbol{m}_D^*$. Because the policy takes the expected optimal value, π is a *optimal policy*. However, in reality, due to uncertain factors, such as the internal characteristics of OSN users, it is not possible for an IP to accurately infer a diffusion model. Therefore, we show the following applicable policy to any model.

Definition 5 (Robust AM policy (RAM))**.** Given a message set M, a random variable of diffusion models $\mathcal{D}$ and an integer value T, RAM is a policy to apply a message sequence $\hat{\boldsymbol{m}}$, a most effective solution of the feasible solutions, i.e., $\hat{\boldsymbol{m}} = \arg\max_{\{m_T\} \in M^T} \mathbb{E}[\omega_{\mathcal{D}}(\{m_T\})]$.

In general, it is hard to compute the RAM policy because the size of all combinations of all possible outcomes of $\{\mathcal{B}_t\}$ and $\mathcal{D}$ in calculating the expectations in the expression may be enormous. Therefore, it is necessary to approximate this policy for an IP to execute it.

5. Experimental Setup

5.1 Probability Distribution of Diffusion Models

The conditions of messages and a diffusion model are defined as follows:

1. $M = \{m_0, m_1, m_2\}$.
2. For any user $i \in U$, set $s_i^S(m_0) = 0.8$, $s_i^S(m_1) = 0.5$, $s_i^S(m_2) = 0.2$, $s_i^X(m_0) = 0.2$, $s_i^X(m_1) = 0.5$ and $s_i^X(m_2) = 0.8$.
3. For any user $i \in U$, set $p(i) = 0.5$, $u_{0,i}^S = -1$, $u_{0,i}^X = -1$.

Table 3 Undirected random graphs

Model	Number of nodes	Parameter assignments #1	#2	#3
Barabási–Albert model	500	$m = 2$	$m = 4$	$m = 8$
	1000	$m = 2$	$m = 5$	$m = 11$
	1500	$m = 2$	$m = 6$	$m = 14$
Watts–Strogatz model ($q = 0.5$)	500	$k = 6$	$k = 10$	$k = 18$
	1000	$k = 6$	$k = 14$	$k = 28$
	1500	$k = 8$	$k = 16$	$k = 34$
random regular graph	500	$d = 5$	$d = 10$	$d = 17$
	1000	$d = 7$	$d = 13$	$d = 25$
	1500	$d = 7$	$d = 15$	$d = 31$

4. Randomly sample σ from V with a uniform distribution.

5. For any decision-making model $F \in \{S_i\}_U \cup \{X_i\}_U$, we assume the following:

   a. μ_F has the Pareto distribution Par$(\frac{1}{\beta T}, \frac{\gamma}{\beta-\gamma})$ where $\beta = 0.8$ and $\gamma = 0.6$.

   b. Set w of F to a constant function that takes a certain value, which is derived from Eq. (5).

Items 1 and 2 assume that any message makes two stimuli (effects) of the internal and external decisions incompatible. Then, the message m_0 works more strongly for the internal action than the external action, m_2 works more strongly for the external action than the internal action, and m_1 works intermediately between both. For item 5, we make this assumption because it is known that the amount of activity of OSN users follows a power law [18]. β and γ in the parameters of the Pareto distribution satisfy $\sup \frac{1}{\mu_F} = \beta T$, $\mathbb{E}[\frac{1}{\mu_F}] = \gamma T$ by the definition of the distribution.

5.2 Network Models and Datasets

Table 3 lists the models of the random undirected graphs, the number of nodes, and the parameter variations of each model. There are three models, Barabási–Albert model (BA), Watts–Strogatz model (WS), and random regular graph (RR), with three patterns of 500, 1000, and 1500 nodes, respectively. The parameters are BA: the number of edges m to attach from a new node to existing nodes, WS: the number of nearest neighbors k to which each node is connected, and RR: The degree d of each node, assigned the values in the table. Note that WS has as a parameter the probability q of rewriting each edge, which is set to 0.5. Three patterns #1, #2 and #3 of parameter assignments are provided for each model, for each the number of nodes. The parameters of the models are set to values such that the sample mean of the average path length for each model is closest to 4.0, 3.0, and 2.5, respectively. The greater m, k, and d, the greater the number of edges and the density. All the models satisfy small-worldness, and only BA satisfies scale-freeness. We generate ten graphs for all 27 patterns in the random graphs. Table 4 shows the datasets based on real data from Facebook, Twitter [19], Hamsterster [20] and LastFM [19], [21].

Table 4 Real OSN datasets

OSN	Direction	Number of nodes	Number of edges
Facebook	Undirected	4,039	88,234
Twitter	Directed	81,306	1,768,149
Hamsterster	Undirected	2,426	16,631
LastFM	Undirected	7,624	27,806

5.3 Baseline Policy

We use Uniform Random Sampling (URS) as the baseline to evaluate the RAM policy. The expected value of the external action taken by URS is $\mathbb{E}[\omega_{\mathcal{D}}(\{\mathcal{M}_T\})]$ where $\{\mathcal{M}_T\}$ refers to the random variable of message sequences that follow a uniform distribution.

5.4 Expected Value Approximation

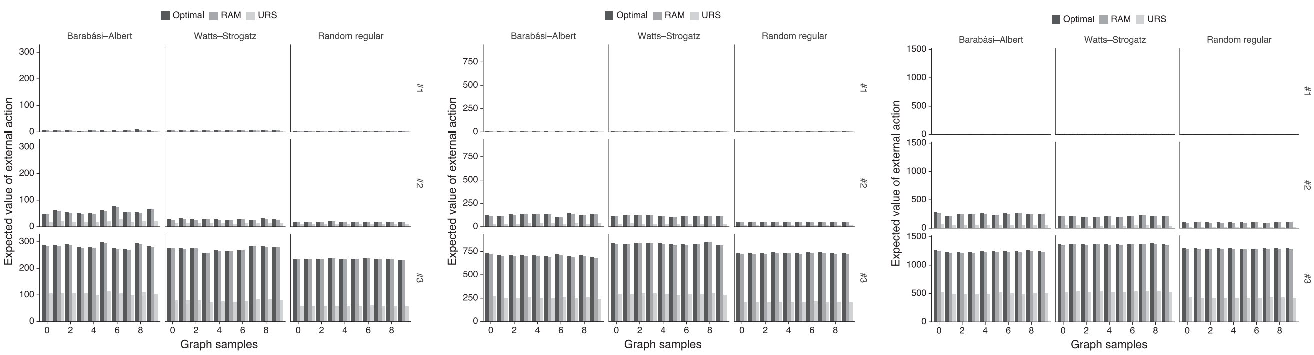
The expected values on $\{\mathcal{B}_t\}$ and $\mathcal{D}$ are approximated by a Monte Carlo method. The sample size of the former is 100. In the latter, the size of IPs is 16, and the size of decision-making models is 48 for internal and external actions of all users. The number of rounds is set to $T = 5$, and the optimal policy π, the RAM policy, and the URS $\{\mathcal{M}_T\}$ are calculated by enumerating all combinations of message sequences.

6. Experimental Results

Figure 1 illustrates the average value of the external action obtained by the optimal policy, the RAM policy, and the URS for each random graph, and Fig. 2 displays the ratio of the RAM policy and the URS to the value to the optimal policy. The x-axis of each cell shows the index of the generated sample graphs. Comparing the rows in Fig. 1, we can see that the values are higher in the order of #1, #2 and #3. This means that the shorter the average path length, the more likely the message is to spread. On the other hand, the comparison of rows in Fig. 2 shows that the RAM ratio has a higher value regardless of the assigned parameters, but the URS ratio varies greatly depending on the parameters.

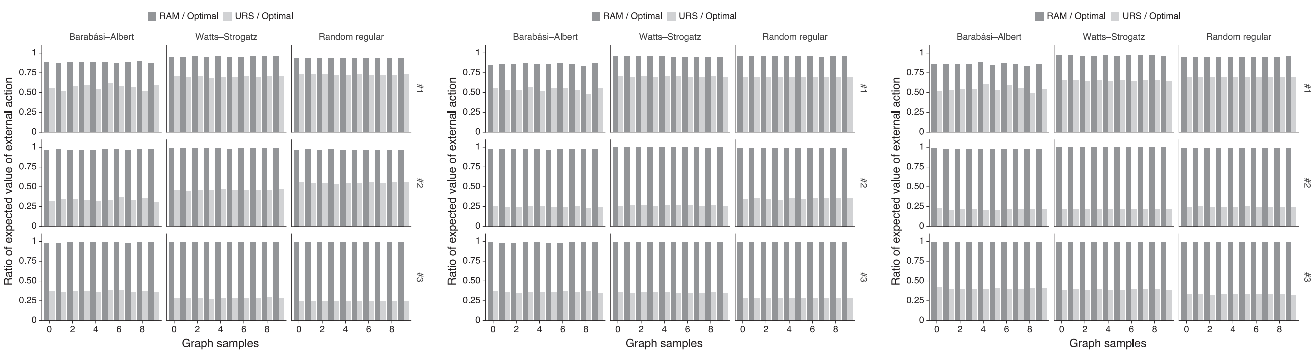
Figure 3 shows the mean value of the three policies obtained for the Facebook, Twitter, Hamsterster and LastFM datasets and the ratio of the mean values to the optimal policies. As can be seen in the figure, all the ratios of RAM take values greater than 0.95. Also, for the ratio of URS, Facebook is close to 0.45, Twitter is almost 0.25, Hamsterster is close to 0.35, and LastFM is almost 0.30.

These results provide an interesting implication that many graph parameters, which seem to have influences on the process of information dissemination, are not major contributing factors to determine optimal message ordering. This implication further suggests the existence of a universally nearly-optimal policy for a certain class of graphs. Specifically, our experiments empirically demonstrate that the differences in the size, density, scale-freeness, and graph



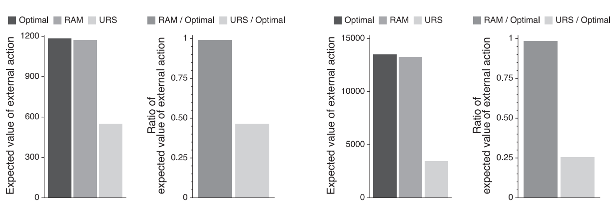
(**a**) Number of nodes: 500 (**b**) Number of nodes: 1000 (**c**) Number of nodes: 1500

Fig. 1 Expected value of external action

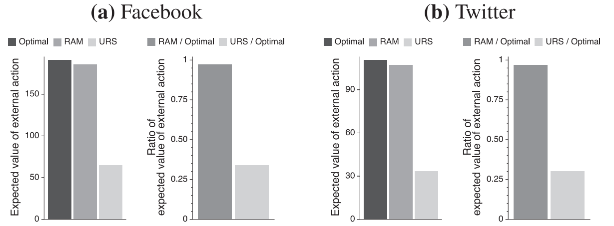


(**a**) Number of nodes: 500 (**b**) Number of nodes: 1000 (**c**) Number of nodes: 1500

Fig. 2 Ratio of expected value of external action to optimal policy



(**a**) Facebook (**b**) Twitter



(**c**) Hamsterster (**d**) LastFM

Fig. 3 Results of real OSN datasets

generation algorithms of OSNs do not affect the optimal sequence of messages as long as the OSNs have the small-world property. It is also suggested that the RAM policy is a promising prospect of such a universally optimal solution for the small-world OSNs.

These insights could be beneficial to IPs using a dynamic strategy aimed at maximizing the expected number of users who take the external action. It seems difficult for these IPs to implement such a strategy because they must take into account complex factors such as OSN states and state changes. Although, IPs could execute a nearly-optimal policy without the factors by using the RAM policy.

7. Conclusion

This paper defined a decision-making model of behavior for receiving information based on a psychological perspective, a diffusion model for internal and external actions of users, and the AM problem to maximize the number of users who take the external action by supplying messages over multiple rounds. We also formulated the RAM policy, an approximation of the optimal policy that takes into account the uncertainty of OSNs. Experimental results imply that many graph parameters, which seem to influence on the process of information diffusion, do not contribute to determine optimal message ordering.

Future works of this study include to expand the coverage of the RAM policy by varying fixed experimental parameters and to improve the computation efficiency of the policies. Because this experiment assigns one pattern of values to the initial utility distribution of OSN users and the available set of messages, it is necessary additionally to examine whether the RAM policy is effective even when the initial utility is varied or when the message variations are changed or increased. Moreover, since the experiment inefficiently calculates the optimal policy and the RAM policy by using brute-force search for message sequences, it is necessary to devise a way to limit the feasible space of message sequences to improve the computation efficiency.

References

[1] J. Syaifullah, M. Syaifudin, M.U. Sukendar, and J. Junaedi, "Social Media Marketing and Business Performance of MSMEs During the COVID-19 Pandemic," The Journal of Asian Finance, Economics and Business, vol.8, no.2, pp.523–531, 2021.

[2] E. Zhuravskaya, M. Petrova, and R. Enikolopov, "Political Effects of the Internet and Social Media," Annual Review of Economics, vol.12, no.1, pp.415–438, 2020.

[3] A. Guille, H. Hacid, C. Favre, and D.A. Zighed, "Information diffusion in online social networks: a survey," SIGMOD Rec., vol.42, no.2, pp.17–28, July 2013.

[4] Y. Li, J. Fan, Y. Wang, and K.-L. Tan, "Influence Maximization on Social Graphs: A Survey," IEEE Trans. Knowl. Data Eng., vol.30, no.10, pp.1852–1872, Oct. 2018.

[5] A. Zareie and R. Sakellariou, "Minimizing the spread of misinformation in online social networks: A survey," Journal of Network and Computer Applications, vol.186, 103094, July 2021.

[6] W. Lu and L.V.S. Lakshmanan, "Profit Maximization over Social Networks," 2012 IEEE 12th Inter. Conf. on Data Min., Brussels, Belgium, pp.479–488, Dec. 2012.

[7] P. Banerjee, W. Chen, and L.V.S. Lakshmanan, "Maximizing Welfare in Social Networks under A Utility Driven Influence Diffusion model," Proc. 2019 Int. Conf. on Manag. of Data, Amsterdam, The Netherlands, pp.1078–1095, June 2019.

[8] S. Bhagat, A. Goyal, and L.V.S. Lakshmanan, "Maximizing product adoption in social networks," Proc. fifth ACM inter. conf. on Web search and data min. – WSDM '12, Seattle, Washington, USA, pp.603–612, 2012.

[9] G. Liang, X. Yao, Y. Gu, H. Huang, and C. Gu, "Multi-Batches Revenue Maximization for competitive products over online social network," J. Netw. Comput. Appl., vol.201, 103357, May 2022.

[10] Y.-W. Teng, Y. Shi, C.-H. Tai, D.-N. Yang, W.-C. Lee, and M.-S. Chen, "Influence Maximization Based on Dynamic Personal Perception in Knowledge Graph," 2021 IEEE 37th Inter. Conf. on Data Eng. (ICDE), Chania, Greece, pp.1488–1499, April 2021.

[11] T. Vogel and M. Wanke, Attitudes and Attitude Change, 2nd ed., Psychology Press, London, 2016.

[12] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," Proc. ninth ACM SIGKDD Inter. Conf. on Knowl. Discov. and Data Min., New York, NY, USA, pp.137–146, Aug. 2003.

[13] N. Hudson and H. Khamfroush, "Behavioral Information Diffusion for Opinion Maximization in Online Social Networks," IEEE Trans. Netw. Sci. Eng., vol.8, no.2, pp.1259–1268, April 2021.

[14] N. Barbieri, F. Bonchi, and G. Manco, "Topic-aware social influence propagation models," Knowl. Inf. Syst., vol.37, no.3, pp.555–584, Dec. 2013.

[15] N.H. Anderson, "Integration theory and attitude change," Psychol. Rev., vol.78, no.3, pp.171–206, 1971.

[16] N.H. Anderson, Ed., Contributions to information integration theory, vol.1, L. Erlbaum Associates, Hillsdale, N.J, 1991.

[17] J. Jacoby, M. Morrin, J. Jaccard, Z. Gurhan, A. Kuss, and D. Maheswaran, "Mapping Attitude Formation as a Function of Information Input: Online Processing Models of Attitude Formation," Journal of Consumer Psychology, vol.12, no.1, pp.21–34, Jan. 2002.

[18] L. Muchnik, S. Pei, L.C. Parra, S.D.S. Reis, J.S. Andrade Jr, S. Havlin, and H.A. Makse, "Origins of power-law degree distribution in the heterogeneity of human activity in social networks," Sci. Rep., vol.3, no.1, 1783, May 2013.

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

[20] Hamsterster full network dataset – KONECT, http://konect.cc/networks/petster-hamster, Oct. 2017.

[21] B. Rozemberczki and R. Sarkar, "Characteristic functions on graphs: Birds of a feather, from statistical descriptors to parametric models," Proc. 29th ACM International Conference on Information & Knowledge Management, pp.1325–1334, 2020.

Masaaki Miyashita received the B.E. degree and the M.E. degree in information systems science from Soka University, Japan in 2013 and 2015. He is currently working towards the Ph.D. degree in system modeling at Soka University. His research interests include graph theory, social networks. He is a member of IEICE.

Norihiko Shinomiya received the B.E. degree, the M.E. degree and the Ph.D. degree in information systems science in 1995, 1997 and 2001, respectively, from Soka University in Tokyo, Japan. In 2000, he joined the Network Systems Laboratories, Fujitsu Laboratories Ltd. as a research engineer, where he engaged in development of a planning and design tool for wavelength division multiplexing (WDM) networks and a control method of IP networks. Since 2005, he has been in department of information systems science, faculty of engineering, Soka University. Currently, he has been a professor and engaged in the research and development of design method, control architecture and management system based on the graph theoretical algorithms. He received the five Best Paper Awards at international conferences; IEEE ICUMT in 2010, IARIA ICONS (two papers) in 2014, and IARIA ICN in 2015 and 2016. Prof. Shinomiya is a member of IEEE CAS, ComSoc and Computer societies. He has also received the Contribution Award on Steering Committee, IEICE Engineering Sciences Society in 2015 for serving as a secretary of IEEE CAS Society Japan Chapter from 2012 to 2013.

Daisuke Kasamatsu is currently an Associate Professor with the Department of Information Systems Science, Faculty of Science and Engineering, Soka University, Tokyo, Japan and engaged in the research and development of data analytics platform including machine learning and distributed systems based on data engineering. He received the B.E., M.E. and Ph.D. degrees in Information Systems Science from Soka University in 2004, 2006, and 2009, respectively. From 2009 to 2012, he was an Assistant Professor in the Department of Information Systems Science, Faculty of Engineering, Soka University, Tokyo, Japan. Since 2012, he had been a Postdoctoral Research Scientist in the Department of Computer Science at Western Michigan University, Purdue University, and Rochester Institute of Technology in the US. In 2017, he joined the IoT Systems Laboratory, Fujitsu Laboratories Ltd. as a research engineer, where he is engaged in development of services with distributed stream processing systems. He is a member of IEEE, ACM, IEICE, and IPSJ. He received the Best Paper Award at the ICSEA in 2006.

Genya Ishigaki received the B.S. degree in Engineering and the M.S. degree in Engineering from Soka University, Japan in 2014 and 2016, respectively, and the M.S. degree in Computer Science and the Ph.D. degree in Computer Science from the University of Texas at Dallas in 2021. He is currently an Assistant Professor at San Jose State University. His current research interests include network slicing and resource optimization based on learning algorithms.