Journal of King Saud University – Computer and Information Sciences 34 (2022) 7570–7603

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Influence maximization frameworks, performance, challenges and directions on social network: A theoretical study

Shashank Sheshar Singh *, Divya Srivastva, Madhushi Verma, Jagendra Singh

Department of Computer Science and Engineering, Bennett University, Greater Noida, India

ARTICLE INFO

Article history:
Received 5 July 2021
Revised 8 August 2021
Accepted 9 August 2021
Available online 18 August 2021

Keywords:
Influence maximization
Social influence
Information diffusion
Influence evaluation
Social networks

ABSTRACT

The influence maximization (IM) problem identifies the subset of influential users in the network to provide solutions for real-world problems like outbreak detection, viral marketing, etc. Therefore, IM is an essential problem to tackle some real-life problems and activities. Accordingly, many reviews and surveys are presented, and most of them mainly focused on classical IM frameworks for single networks and avoided other IM frameworks. In this context, the IM problem still has some important design aspects along with some new challenges of the problem. Inspired by these facts, a comparative survey of the state-of-art approaches for IM algorithms is presented in this paper. To build the foundation of IM problem, firstly, the well-accepted information diffusion models are discussed. Secondly, a comprehensive study of IM algorithms along with a comparative review is presented based on algorithmic frameworks of IM algorithms. A relative analysis of IM approaches regarding performance metrics is discussed next. At last, the upcoming challenges and future prospects of the research in this field are discussed.

Contents

* Corresponding author.
E-mail addresses: shashankss.rs.cse16@itbhu.ac.in (S.S. Singh), divyalknw@gmail.com (D. Srivastva), madhushi.verma@bennett.edu.in (M. Verma), Jagendra.Singh@bennett.edu.in (J. Singh).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2021.08.009
1319-1578/

1. Introduction

Nowadays, people spend a lot of time over the Internet and 60–70% of the time spent over social networking platforms, such as Facebook, Instagram, Twitter, etc. These social networks have become an essential part of almost everyone's life. The exponential growth of usage of these networks has attracted researcher's attention towards information diffusion (Richardson et al., 2002; Kempe et al., 2005; Kempe et al., 2003), community detection (Biswas and Biswas, 2018; Mishra et al., 2021; Abd Al-Azim et al., 2020), link prediction (Kumar et al., 2020; Kumar et al., 2019; Kumar et al., 2019; Singh et al., 2020a; Berahmand et al., 2021), etc. The information diffusion models spread the information or innovation quickly through word-of-mouth (Brown and Reingen, 1987; Goldenberg et al., 2001a) spreading. This has a wide application potential like viral marketing (Domingos and Richardson, 2001; Singh et al., 2019a; Singh et al., 2019b; Singh et al., 2019c; Tang and Yuan, 2020; Saxena and Saxena, 2019), revenue maximization (Teng et al., 2018), rumor control (Peng and Pan, 2017; Chen et al., 2010d), social recommendation (Ye et al., 2012; Jindal and Sardana, 2020), etc. Inspired by viral marketing, Pedro and Matt (Domingos and Richardson, 2001) introduced influence maximization (IM) as an optimization problem. The IM problem identifies a set of most influential users in the social network in order to maximize the expected adoption of a product.

Kempe et al. (2003) presented a scenario of viral marketing where a social network is given the influence weight that approximates the extent of influence one another. The social networking platform provides a channel of interaction for advertising and marketing. Advertisers aim to maximize product adoption by considering some of the individuals from the network as seed users. These seed users are responsible for product adoption by spreading the product information. The advertiser used to provide a free sample of a product to these seed users by the cost and budget of the advertising. The advertiser hopes that word-of-mouth influence will attract others users to go for the product and will help in maximizing the spread of influence.

Challenges and Issues. Systems are often represented as networks in various areas like sociology, computer science, biology, and physics, where influence maximization plays a vital role. Numerous techniques have been developed for influence maximization, yet the problem is not solved satisfactorily (Li et al., 2018d). Several challenges emerged along with the influence maximization problem, some of which are as follows (Li et al., 2018d).

- The first is how to model diffusion process to propagate the information from seed users in a social network due to stochastic nature, which would heavily affect the adaption of a product of non-seed users in influence maximization problem (Kempe et al., 2003; Guille et al., 2013).
- In general, the influence maximization problem is theoretically complex. Kempe et al. evinced that the identification of seed users is NP-hard under classical diffusion models. Furthermore, it is also observed that the spreading of influence is computationally complex under any diffusion model for a given seed

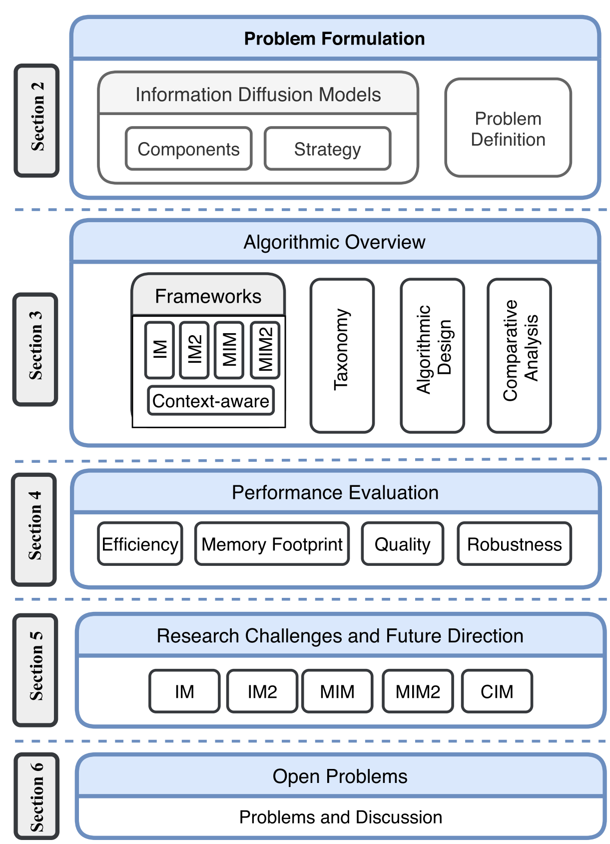


Fig. 1. The survey overview.

set (Barbieri et al., 2012; Li et al., 2018d). Therefore, it is pretty tough to achieve a near-optimal solution for the influence maximization problem and scale for large-scale social networks.
- Mostly information diffusion models and influence maximization algorithms are random (Guille et al., 2013; Li et al., 2018d). A different set of users are activated or influenced in different executions of algorithms for the same network. Accumulation of outputs obtained in different executions of an algorithm during performance evaluation is another challenging task.

Several issues also emerged along with the influence maximization problem, some of which are as follows:

- Identification of effective seed is a major issue in the influence maximization problem. To reduce this issue, contextual features like, location (Guo et al., 2017; Li et al., 2014d, time Gomez-Rodriguez et al., 2016, 2012, topic Li et al., 2015a; Shuo Chen et al., 2015a, and competitive Zhu et al., 2016; Lin et al., 2015), etc were used to propose several context-aware influence maximization. These context-aware methods produce effective seeds having less efficiency and scalability.
- In case of large networks, most of the influence maximization algorithms are time inefficient and non-scalable. Therefore, algorithms need to reduce iteration, improve iteration complexity, reduce search space, and identify seed sets efficiently. To reduce this issue, many algorithms have been proposed such as heuristic metrics based Chen et al., 2009; Wang and Feng, 2009; Kundu et al., 2011), influence path based (Kimura and Saito, 2006; Chen et al., 2010e; Kim et al., 2013), community based (Wang et al., 2010; Li et al., 2015b), sub-modularity based (Sviridenko, 2004; Leskovec et al., 2007; Goyal et al., 2011), etc. However, quality is still an issue.
- Assurance of both effectiveness and efficiency is a significant issue during selecting seed nodes (Arora et al., 2017). Measuring effectiveness incorporates topological and contextual information, whereas measuring efficiency involves the number and complexity of iterations. The fundamental difference between the two measures has led to the trade-off between efficiency and effectiveness. This trade-off is a significant issue during the performance evaluation of influence maximization algorithms.
- In real-world, a company might intend to encourage several competitive and non-competitive products simultaneously in the same social network (Carnes et al., 2007; Bharathi et al., 2007; Sun et al., 2016). Generally in the social networks, interests for different products varies for various users and thus, the acceptance probabilities of promotions from their social network friends varies accordingly. Thus, compared to the traditional influence maximization problem, the main issue is to form the seed users, decide the number of each product among the given m items, and then further identify k individuals who influence the most.
- In the real-world, there have been quite a several users who maintain several accounts simultaneously, which allow them to propagate information across different networks (Zhang et al., 2016; Wang et al., 2016b; Erlandsson et al., 2017). Also, real-world systems are complex as those cover a more comprehensive range of aspects such as multiple relationships, organizational hierarchy, directional associations, etc., (Zhang, 2015). Therefore, identifying seed nodes in those diverse and multiple featured networks with a single algorithm is challenging.

Applications. IM has a wide range of real time contextual based features applications for instance, recruitment, rumor control, social recommendation, social media, population screening, tread analysis and sales prediction, election campaign, epidemiology, blogosphere, revenue maximization, and network monitoring, etc.

Contribution. The present paper provides an extensive review of Influence maximization algorithms concerning their networks. The survey overview is presented in Fig. 1 by focusing extensively on influence maximization algorithms based on the respective framework. The present survey paper majorly focuses on the following.

1. A brief description of traditional diffusion models along with the characteristic comparison is presented.
2. Condensed taxonomy of IM algorithms based on the algorithmic framework, and their comparative analysis is presented.
3. Comparative analysis of existing IM algorithms and their various performance metrics used.
4. Challenges and future scope of research in IM problem concerning their framework.
5. Some open problems in the influence maximization problem are also discussed.

1.1. Differences from existing surveys

Although many surveys (Guille et al., 2013; Sun and Tang, 2011; Zhang et al., 2014a; Chen et al., 2013; Arora et al., 2017; Tejaswi et al., 2016; Li et al., 2018d; Li et al., 2017d; Razaque et al., 2019a; Chang et al., 2018; Sumith et al., 2018) exist on information diffusion analysis. This survey is different in various aspects, as follows.

- The authors of (Guille et al., 2013; Razaque et al., 2019a; Razaque et al., 2019b) present a discussion on information diffusion process, components, and models for social network analysis. This study is not bound to a specific social network domain such as influence propagation and IM problem. Chang et al. (2018) shed light on information diffusion models with their usefulness on three social network applications: information source detection, influence evaluation, and influence maximization.
- Sun and Tang (2011) discuses the approaches to measure social influence in the network. They also discuss the models to analyze the social influence of user-generated data. Although some work (Zhang et al., 2014a; Chen et al., 2013; Tejaswi et al., 2016; Li et al., 2017d; Singh et al., 2019d) describe the models of influence propagation and approximation algorithms for IM. However, these studies are incomplete as they only focus on traditional influence maximization. These studies do not provide insight about recent advances like IM across multiplex (Zhang et al., 2016; Wang et al., 2016b; Erlandsson et al., 2017; Li et al., 2012; Zhan et al., 2015a), context-aware (Singh et al., 2019; Lee and Chung, 2015; Tejaswi et al., 2017; Barbieri et al., 2012; Zhuang et al., 2013; Wang et al., 2017a; Wang, 2016; Tong et al., 2017; Bozorgi et al., 2017), soft computing approaches (Gong et al., 2016; Wang et al., 2017b; Ge et al., 2017; Singh et al., 2019d; Wang et al., 2020), etc.
- Arora et al. (2017) provides an in-depth benchmarking study on some well-known IM approaches. This study is based on experimental analysis rather than theoretical analysis.
- Li et al. (2018d) focuses on theoretical analysis of IM algorithms for a single network based on algorithmic design. The authors also provide a detailed description of context-aware IM algorithms. However, the review of the IM algorithm is incomplete, as they do not cover performance measures of IM approaches. Our work focuses on the theoretical analysis of IM algorithms for a single network and multiplex networks based on the algorithmic framework. Also, the comparison of characteristics covers much broader aspects. This survey discusses research

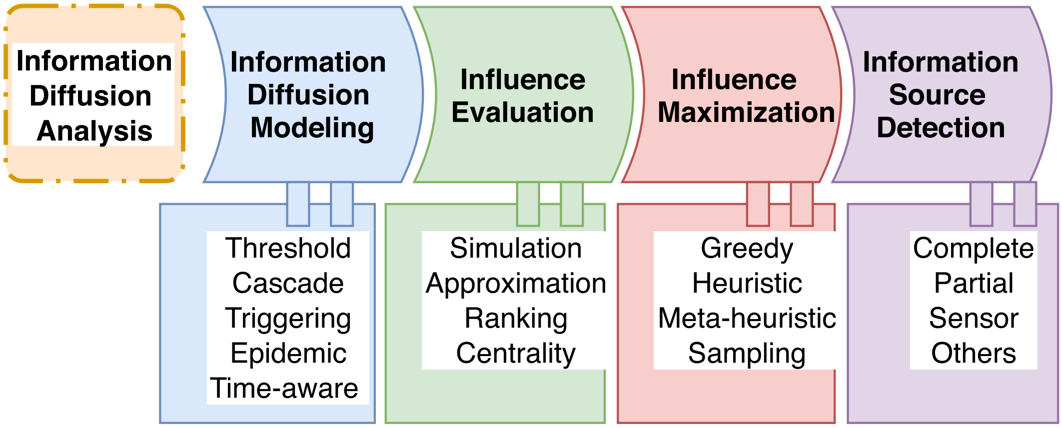


Fig. 2. The information diffusion analysis overview.

challenges and directions respective to each framework instead of only classical IM like (Li et al., 2018d), along with open problems.

Compared to existing surveys, this work presents a study of information diffusion models and the existing IM algorithms for both single and multiple networks. This paper focuses on the rigorous theoretical analysis of the existing influence maximization techniques corresponding to their framework. With this, it also focuses on context-aware IM approaches. Moreover, as per our knowledge, this is the first survey of its kind, which covers IM algorithms comparison concerning performance metrics along with research challenges and future works. This survey also listed some open problems in IM problems need to explore.

2. Preliminary

The information diffusion analysis study can be categorized in four classes: 1) modeling of the information diffusion process, 2) evaluation of influence, 3) influence maximization algorithm design, and 4) influential users identification as shown in Fig. 2 (Chang et al., 2018). This study focuses on the formulation, framework, and algorithms of influence maximization along with information diffusion. The concept of IM as an optimization problem was coined by Pedro and Matt (Domingos and Richardson, 2001) in 2001 by taking inspiration from viral marketing. Later in 2003, Kempe et al. (2003) presented the formal definition of IM problem. They assumed that a social network could be denoted as a graph $G(V,E,W)$, which combines users and their relationships in the network with specific tie strength. IM problem aims to determine the initial seed users denoted by k who create the maximum influence. Further, the influence propagation of an active node can be computed based on the diffusion model. Before formalizing the IM problem definition, some definitions are presented to understand the problem better. The list of abbreviation's are listed in Appendix A. Basic definitions and IM problem formalization are given as follows.

2.1. Definitions

Definition 1 (*Social Network*). A social network which is also referred to as an influence graph, can be represented by a weighted-directed graph $G(V,E,W)$ having N users and M edges. Here, V denotes the set of users in the network, and E signifies a set of relationships such as friendship, follower, re-tweet, mention, co-authorship, etc. The type of relationship is network-dependent. The edge-weight W denotes the relationship strength between peers.

Definition 2 (*Neighbors*). Given a node a, its neighbors $N(a)$ can be presented as the set of users b such that $b \in N(a)$ iff $\exists(a,b) \in E$ and $b \in V$. The in and out neighbors of $a$can be specified as $N_{in}(a)$ and $N_{out}(a)$, respectively. The edge set is the collection of ordered pair of nodes, i.e., $(a,b) \neq (b,a)$ so they may or may not exist in both directions. An edge (a,b) states that node a has some influence on node b but not the vice versa.

Definition 3 (*Degree Centrality*). The count of links incident upon a given node is termed as Degree Centrality and is denoted as $D(a) = |N(a)|$. The number of outgoing links are considered as the degree of centrality of node a in directed social networks i.e. $D(a) = |N_{out}(a)|$.

Definition 4 (*Seed Nodes*). A set of nodes that participate in the information spreading process of social networks and act as the source for propagation are defined as seed nodes or seed set (S) such that $S \in V$ and $|S| = k$.

Definition 5 (*Active Node*). An active node $b \in V$ is one which either belongs to seed set S or has picked up the information transferred by an already activated node $a \in V_A$ using the diffusion model. Once their neighbors activate the node b, it will be added to a set of active nodes V_A, i.e., $V_A \leftarrow \{V_A \cup b\}$.

Definition 6 (*Influence Spread*). The number of active users obtained after implementing the diffusion process using a diffusion model denotes the influence spread $I_S(S)$ of the seed set S and can be represented as $I_S(S) = |V_A(S)|$.

Definition 7 (*Information Diffusion Model (IDM/DM)*). The stochastic process which records the influence spreading of S on G is described as the information diffusion model where $G = (V,E,W)$ is the influence graph and $S \subseteq V$ is the seed set.

2.2. What we need?

In order to address the IM problem, three steps need to be performed.

1. Form or adopt models of influence to propagate information, ideas, innovation, etc., in social networks.

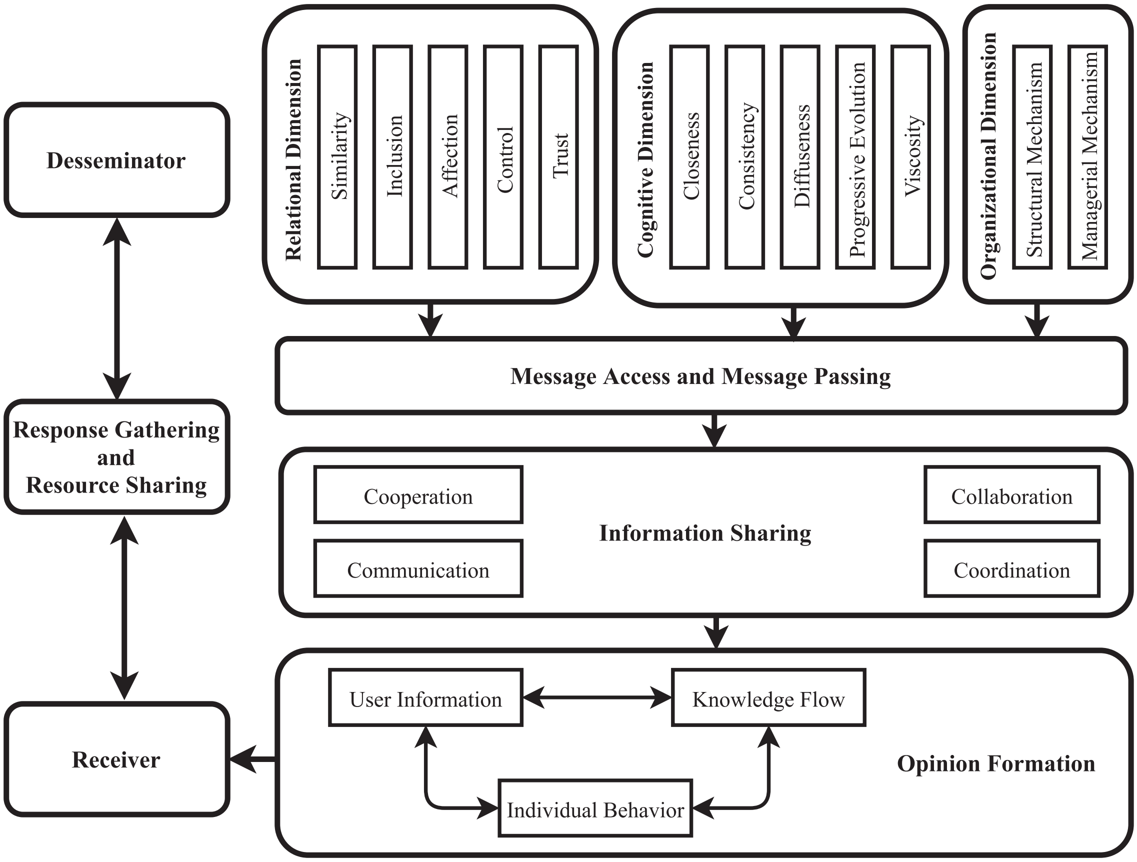


Fig. 3. The component of information diffusion model.

2. Obtain topological information about a particular network to estimate interpersonal influence along with influence and activation probabilities.
3. Finally, devise an algorithm for seed selection to maximize influence spread.

2.3. Diffusion model

Enormous literature exists in social networks, data mining, epidemiology, database, etc. Thus, it is crucial to survey the diffusion models that play a vital role in surveying various aspects of IM-based algorithms. Fig. 3 demonstrates the components of diffusion model (Razaque et al., 2019a). In this work, the classical frameworks of diffusion models for IM in the network have been discussed. Every node $u \in V$ in the framework has a timestamp t associated with either of the two states, i.e., active or inactive. An inactive node is one which is not influenced by its neighbors. In the initial state, except for the seed node, all other nodes are inactive. At $t = 0$, seed nodes are active ($S \subseteq V$) and this starts the diffusion process. The process begins at the seed nodes, and further, the influence is propagated to make their neighbors active. These active nodes then influence their neighbors, and the process continues till no new node can be activated. In general, every model has its own mechanism to capture the active state from inactive based on the neighbor's behavior. Some popular diffusion models are (Guille et al., 2013; Li et al., 2018d; Qiang et al., 2019) as follows.

- **Threshold Model (TM).** Threshold Model was firstly introduced by Schelling (2006) and Granovetter (1978). TM is any model where a single threshold value, or set of threshold values, is used to differentiate ranges of values for the behavior predicted by the model. Linear threshold model (LTM) is the most popular threshold diffusion model. The idea behind this model that a node y becomes active iff $\sum_{x \in N^a(y)} W_{xy} \geqslant T_y$, i.e. the incoming neighbors influence must be greater than the threshold T_y. In linear TM, the value of the user's threshold follows uniform distribution over [0,1]. There are some other threshold models are exist in the literature, distinguish by their threshold values like majority TM ($T_y = \frac{1}{2}D(y)$) (Richardson et al., 2002; Bhagat et al., 2012), small TM (T_y is a small constant) (Eiselt and Laporte, 1989), Separated TM (linear TM with separate competitive cascade) (Kermack and McKendrick, 1927; Borodin et al., 2010b), and unanimous TM ($T_y = D(y)$) (Richardson et al., 2002), where $D(y)$denotes degree centrality of node y. There are some generalized threshold models are presented by replacing the activation function of linear TM with an arbitrary function (Pathak et al., 2010; Bhagat et al., 2012). Bhagat et al. (2012) presented a diffusion model that considers user's experience with a product and captures product adoption rather than influence and named it as linear threshold with color. Banerjee et al. (Pathak et al., 2010) further extend the linear TM to handle the opinion change of users and allow the users to switch back between active and inactive states.
- **Cascading Model (CM).** Inspired by probability theory and interacting particle system (Lecture notes on particle systems and percolation, 1989), dynamic cascade models are introduced for diffusion. The authors of Goldenberg et al. (2001b) and Goldenberg and Muller (2003) were first to introduce cascade models in the field of marketing. The Independent cascade model (ICM) is well-studied and most popular in viral marketing (Goldenberg et al., 2001b). If a node $x$becomes active at time

Table 1
Comparison of the diffusion models characteristics – I.

Diffusion model	Activation condition	Properties	Applications
Linear TM (Kermack and McKendrick, 1927)	$\sum_{x\in N^a(y)} W_{xy} \geqslant T_y$	The objective function $\sigma(S)$ is submodular and IM problem is NP-hard under LTM	Rumors and diseases control
Majority TM (Richardson et al., 2002; Bhagat et al., 2012)	$T_y = \frac{1}{2}D(y)$; T_y is threshold value	IM problem is NP-hard under MTM	Distributed computing, Voting system
Small TM (Eiselt and Laporte, 1989)	$\sum_{x\in N^a(y)} W_{xy} \geqslant T_y$; T_y is a small constant	For $T_y = 1$, IM is P-hard and For $T_y \geqslant 2$ IM is NP-hard under STM	
Unanimous TM (Richardson et al., 2002)	$\sum_{x\in N^a(y)} W_{xy} \geqslant T_y$; $T_y = d(v)$	2-Approximation algorithm,IM is NP-hard	Network security and vulnerability
Separated TM (Kermack and McKendrick, 1927; Borodin et al., 2010b)	$\sum_{x\in N^a(y)\cap\varphi_A^{t-1}} w_{x,y}^A \geqslant T_y^A$	The objective function is monotonic but not submodular and IM is NP-hard	Network with competitive sources
Weighted-proportional TM (Kermack and McKendrick, 1927)	$P[y \in \varphi_A^t \mid y \in \varphi^t \setminus \varphi^{t-1}] = \frac{\sum_{x\in\varphi_A^t} w_{x,y}}{\sum_{x\in\varphi^t} w_{x,y}}$	$\sigma(S)$ is neither submodular nor monotonic and IM is NP-hard.	Deal with two competitive influence
Independent CM (Kempe et al., 2003)	$\prod_{i=1}^{k}(1 - P_y(v_i \mid S\cup M_i)) = \prod_{i=1}^{k}(1 - P_y(v_i^{,} \mid S\cup M_i^{,}))$	$\sigma(S)$ is submodular and IM is NP-hard.	Collective behavior, Viral Marketing
Opinion CM (Zhang et al., 2013)	$\sum_{x\in N^a(y)} W_{xy} \geqslant T_y$	$\sigma(S)$ function is neither submodular nor monotonic and IM is NP-hard	Political campaign and Incorporate user opinions
ICM-NO (Chen et al., 2011)	$\prod_{i=1}^{k}(1 - P_y(v_i \mid S\cup M_i)) = \prod_{i=1}^{k}(1 - P_y(v_i^{,} \mid S\cup M_i^{,}))$	With probability P, each newly active node become positive and with probability 1–P.	Political campaign and Incorporate negative opinions
Decreasing CM (Kempe et al., 2005)	$P_y(x \mid S) \leqslant P_y(x \mid M)$	The objective function is submodular and IM is NP-hard under DCM	Collective behavior, Information spreading
SIR (Kermack and McKendrick, 1927; Kermack and McKendrick, 1991)	–	$\frac{dS}{dt} = -\beta SI$, $\frac{dI}{dt} = \beta SI - \gamma I$,$\frac{dR}{dt} = \gamma I$	Epidemiology
SIS (Kermack and McKendrick, 1927; Kermack and McKendrick, 1991)	–	$\frac{dS}{dt} = -\beta SI + \gamma I$, $\frac{dI}{dt} = \beta SI - \gamma I$	Epidemiology
SIRS (Kermack and McKendrick, 1927; Kermack and McKendrick, 1991)	–	$\frac{dS}{dt} = -\beta SI + fI$, $\frac{dI}{dt} = \beta SI - \gamma I$,$\frac{dR}{dt} = \gamma I - fR$	Epidemiology

Table 2
Comparison of the Diffusion Models Characteristics – II.

Diffusion model	Network		Submodularity		Activation	
	Directed	Weighted	Diminishing Return	Monotone	Multiple Activation	Time Aspect
Linear TM (Kermack and McKendrick, 1927)	✓	✓	✓	✓	✗	✗
Majority TM (Richardson et al., 2002; Bhagat et al., 2012)	✓	✓	✓	✓	✗	✓
Small TM (Eiselt and Laporte, 1989)	✓	✓	✓	✓	✗	✗
Unanimous TM (Richardson et al., 2002)	✓	✓	✗	✗	✗	✗
Separated TM (Kermack and McKendrick, 1927; Borodin et al., 2010b)	✓	✓	✗	✗	✗	✗
Weighted-proportional TM (Kermack and McKendrick, 1927)	✓	✓	✗	✗	✗	✗
Independent CM (Kempe et al., 2003)	✓	✓	✓	✓	✗	✗
Opinion CM (Zhang et al., 2013)	✓	✓	✗	✗	✓	✓
ICM-NO (Chen et al., 2011)	✓	✓	✓	✓	✗	✗
Decreasing CM (Kempe et al., 2005)	✓	✓	✓	✓	✗	✗
SIR (Kermack and McKendrick, 1927; Kermack and McKendrick, 1991)	✓	✓	✗	✓	✗	✗
SIS (Kermack and McKendrick, 1927; Kermack and McKendrick, 1991)	✓	✓	✗	✓	✓	✗
SIRS (Kermack and McKendrick, 1927; Kermack and McKendrick, 1991)	✓	✓	✗	✓	✓	✗

t, it has a only chance to activate its inactive neighbors y with activation probability p_{xy} at stamp $t+1$. The activation process of a node is similar to flipping a coin. If node y becomes active at $t+1$ then it will never be inactive in the future. There are some extensions of the independent cascade model that exist in literature like ICM with negative opinion (Chen et al., 2011), ICM with positive and negative opinion (Nazemian and Taghiyareh, 2012).

- **Triggering Model (TRM).** The triggering model is the generalized form of TM and CM, presented by Kempe et al. (2003). The authors also proved that the triggering model in TM and CM are equivalent. In the triggering model, each node x is associated with a threshold value T_x and a distribution function f_x that maps to a subset of its neighbors S_x with a probability (the likelihood that subset can influence node x). This model independently selects a random subset of neighbors in each instance of the diffusion process for user x. There are two generalized triggering models based on the diffusion behavior of TM and CM. Kempe et al. (2005) presented a more general model than the triggering model, named decreasing cascade model. This model redefines the influence probability of a node x from y as $p_y(x, S_x)$, where S_x represents a subset of active neighbors of x. Decreasing CM enforces $p_y(x, S) \geqslant p_y(x, T), S \subseteq T$ to capture diminishing return property.
- **Epidemic Model (EM).** The epidemic process had a vital impact on transmission of contagious disease, political campaign, and information propagation such as rumors and news (Kumar and Sinha, 2021). In epidemic model, the fixed populace can be divided into three classes as, susceptible (S), infectious (I), and recovered (R). Kermack and McKendrick (1991) presented three epidemic model based on nature of the model cascades: susceptible infectious recovered (SIR), susceptible infectious susceptible (SIS), and susceptible infectious recovered susceptible (SIRS).

Fig. 4. The frequency distribution of the existing diffusion models over various characteristics.

- **Time-aware Model (TAM).** The above discussed diffusion models are time-unaware. These models terminate the diffusion process when no more nodes are activated. However, some propagation campaigns need to maximize the spread of social influence specific to a fixed time- period, i.e., the convergence of propagation depends on the time period rather than the number of iterations. To achieve time-critical demand, time-aware information diffusion models are introduced. The TM is divided into two classes: discrete-time model and continuous-time model. The authors of Chen et al. (2011), Chen et al. (2012) and Lee et al. (2012) presented DTAM models by extending independent CM. These models follow discrete random variables over distinct time-stamps. To handle the scenario, when a node x influencing other in continuous time, CTAM (Lee et al., 2012; Gomez-Rodriguez et al., 2011) were introduced.

Table 1 (Guille et al., 2013; Singh et al., 2019d) and 2 (Sumith et al., 2018) summarize each diffusion model characteristics and properties. Column Diffusion Model gives the algorithm name with reference. Column Activation Condition describes the condition to change the state from inactive to active. Columns Properties and Applications provide the information of properties and their applications respectively. Column Network and Submodularity indicates network type and diffusion function property. The ✓ and ✗ denote the presence and absence of the property. The frequency distribution of the existing diffusion models over different characteristics such as submodularity, network applicability, and activation process is shown in Fig. 4. This also helps to summarize the popularity of different characteristics in IM algorithms.

2.4. Problem Hardness

This section will present the study of the computational hardness of influence spread function of influence maximization problem under TM, CM, TRM, EM, and TAM.

Theorem 1. *The IM problem with viral marketing setting is NP-hard under independent cascade model (ICM) (Kempe et al., 2003, Theorem 2.4).*

Theorem 2. *The IM problem with viral marketing setting is NP-hard under under classical linear threshold model (LTM) (Kempe et al., 2003, Theorem 2.7).*

Theorem 3. *The IM problem with viral marketing setting is NP-hard under continuous time-aware (CTAM) diffusion model (Gomez Rodriguez et al., 2012, Theorem 3).*

Theorem 4. *The IM problem over viral marketing application is NP-hard under triggering (TRM) diffusion model (Kempe et al., 2003).*

Theorem 5. *The Computation of the expected influence spread $\sigma(S)$ of a given subset of active users S is #P-hard under the independent ICM propagation model (Chen et al., 2010a, Theorem 1).*

Theorem 6. *The Computation of the expected influence spread $\sigma(S)$ of a given subset of active users Sis #P-hard under the LTM propagation model (Chen et al., 2010b, Theorem 1).*

Based on observation of these theorems, we can summarize that no such algorithm exists in the literature which can identify a subset of influential users in polynomial time unless $P = NP$. The computation process of the spread of social influence $\sigma(S)$ of seed nodes S under any diffusion models is also complex. Therefore, existing research focuses on developing approximate and efficient IM algorithms.

2.5. Problem definition

Definition 8 (*Influence Maximization (IM) (Kempe et al., 2003)*). The influence spread in G can be maximized by using an information diffusion model that considers a positive integer k and a subset of influential users $S \subseteq V, |S| = k$ such that $\sigma(S) = argmax_{S^* \subseteq V \wedge |S^*| = k}\sigma(S^*)$, where $G = (V, E, W)$ is an influence graph.

Though the IM problem is NP-hard, and finding the optimal solution is not feasible. Therefore, an approximated solution can be found iff the expected influence spread function $\sigma(S)$ is submodular. Any arbitrary function is known as submodular iff the function satisfies both the properties of monotone increasing and diminishing return.

Property 1 (*Monotone Increasing (Singh et al., 2019)*). *An objective function say $\sigma(S)$ is said to be monotone increasing iff $\sigma(S) \leqslant \sigma(T), S \subset T$.*

Property 2 (*Diminishing Return (Singh et al., 2019)*). *An objective function say, $\sigma(S)$ is said to be diminishing return iff $\sigma(S \cup u) - \sigma(S) \geqslant \sigma(T \cup u) - \sigma(T), \forall u \in T and\ S \subset T$.*

The monotonicity states that the addition of more nodes in an influential user set does not reduce its overall expected influence rather, it can be increased, while diminishing return means that marginal gain of node u with a subset of seed set is always more or equal to marginal gain with seed set.

Theorem 7. *The expected spread of influence function $\sigma(S)$ is submodular under ICM information propagation model (Kempe et al., 2003, Theorem 2.2).*

Theorem 8. *The expected spread of influence function $\sigma(S)$ is submodular under LTM information propagation model Kempe et al., 2003, Theorem 2.5).*

Theorem 9. *The expected spread of influence function $\sigma(S)$ is submodular under TRM information propagation model (Kempe et al., 2003, Theorem 4.2).*

Theorem 10. *The expected spread of influence function $\sigma(S)$ is submodular under CTAM information propagation model (Gomez Rodriguez et al., 2012, Theorem 4).*

2.6. The Greedy Algorithm

Most of the simulation-based state-of-the-art methods are based on the greedy framework proposed by Kempe et al. (2003), demonstrated in Algorithm 1. The algorithm starts with taking an empty set as a seed set (line 1) and iteratively identifies a node x with maximum marginal gain (line 3). Then algorithm adds node x to the seed set S (line 4). Finally, the algorithm returns the k distinct nodes as a resultant seed set. The theoretical approximation guarantee of the solution generated by the greedy algorithm depends on the objective function $\sigma(S)$ submodular nature, which holds by the classical DMs as stated in theorems 7 to 10.

Algorithm 1. Greedy Algorithm (Kempe et al., 2003)

Input: Influence graph G, Seed size k.
Output: Seed set S.

1 $S \leftarrow \phi$ ▷ Seed set at start
2 **for** $j = 1, 2, \ldots, k$ **do**
3 $\quad x \leftarrow argmax_{x^* \in V \setminus S}(\sigma(S \cup \{x^*\}) - \sigma(S))$
4 $\quad S \leftarrow S \cup \{x\}$
5 **Return** Seed Set S

Theorem 11. *The relation between expected influence spread of optimal seed and seed generated by greedy algorithm satisfies $\sigma(S) \geqslant (1 - (1 - \frac{1}{k})^k)\sigma(S^*)$, where S^* and S denotes optimal and greedy seed respectively (Nemhauser et al., 1978, Theorem 2.2).*

In general, approximation ratio is considered $(1 - 1/e)$ rather than $(1 - (1 - \frac{1}{k})^k)$ in existing works because of $(1 - 1/e) < (1 - (1 - \frac{1}{k})^k)$ for $\lim_{k \to \infty}(1 - (1 - \frac{1}{k})^k)$ and $k > 0$. Moreover, to account the sampling error in sampling algorithms an additional term ϵ is introduced, i.e., the approximation ratio for sampling methods is $(1 - 1/e - \epsilon)$.

3. Influence maximization frameworks

Although greedy framework guaranteed $(1 - 1/e - \epsilon)$ approximation ratio, it is still challenging to estimate the expected influence in the IM Frameworks. Theorems 5 and 6 stated that the computation of social influence $\sigma(S)$ of a given seed is #P-hard even under classical diffusion models. This theoretical hardness creates an opportunity to perform extensive research for developing efficient approaches to solve the IM problem. Existing research work can be divided into different classes based on how an IM algorithm considers networks, products, contextual features, etc. (Arora et al., 2017; Li et al., 2018d; Singh et al., 2019f). Table 3 illustrates the taxonomy of existing IM techniques (Singh et al., 2019f).

3.1. Basic framework

Arora et al. (2017) performed an experimental study on some popular IM algorithms and proposed a basic benchmarking framework as shown in Fig. 5. The basic framework of IM problem consists of four components: empirical setup, seed selection strategy, evaluation, and insights. All the components except seed selection are same for each framework. Seed selection component has been developed differently based on their algorithmic design and framework. Empirical setup configure the settings and collect information like dataset, diffusion model, algorithm, etc., which is required to address the IM problem. The seed selection process identifies influential users based on their algorithmic design. Evaluation process estimates different performance measures such as influence spread, running time, memory footprints, etc. Finally, the insight component analyzes the findings of the algorithm along with effectiveness, efficiency, scalability, and robustness.

Table 3
Taxonomy of the IM frameworks with references.

	Framework	Categories	Problem Solving Perspective	Representative
Influence Maximization	Classical	Simulation-based	Minimization of #MC Simulations	(Leskovec et al., 2007; Goyal et al., 2011; Zhou et al., 2015a)
			Complexity Improvement of MC	(Wang et al., 2010; Jiang et al., 2011; Chen et al., 2014; Sheng and Zhang, 2018; Singh et al., 2019e; Zhang et al., 2016)
		Heuristic-based	Rank Refinement	(Freeman, 1978; Page et al., 1999; Liu et al., 2014; Chen et al., 2009; Jung et al., 2012; Singh et al., 2020b; Tsai and Liu, 2019)
			Model Reduction	(Kimura and Saito, 2006; Chen et al., 2010e; Kim et al., 2013; Chen et al., 2010c; Goyal et al., 2011; Galhotra et al., 2016; Vega-Oliveros et al., 2020; Banerjee et al., 2019)
		Mixed Approaches	Snapshot Sampling	(Chen et al., 2009; Cheng et al., 2013; Cheng et al., 2013; Ohsaka et al., 2014; Cohen et al., 2014)
			Reverse Reachable	(Borgs et al., 2014; Tang et al., 2014; Tang et al., 2015; Wang et al., 2017a; Nguyen et al., 2016)
	IM2	Simulation-based	Minimization of #MC Simulations	(Saito et al., 2012; Zhang, 2015)
			Complexity Improvement of MC	(Zhang et al., 2016)
		Heuristic-based	Rank Refinement	(Erlandsson et al., 2017; Wang et al., 2016b)
			Model Reduction	(Zhang, 2015)
		Mixed Approaches	Reverse Reachable	(Wang et al., 2016b)
	MIM	Simulation-based	Complexity Improvement of MC	(Sun et al., 2016)
	MIM2	Heuristic-based	Rank Refinement	(Singh et al., 2019f)
	Context-aware	Spatial	Location	(Li et al., 2014a, 2015b, 2016, 2017b, 2018, 2017, 2018a,b)
		Topical	Topic Relevant Targets	(Guo et al., 2013; Li et al., 2015c; Lee and Chung, 2015; Nguyen et al., 2016)
			Topic Dependent Diffusion	(Shuo Chen et al., 2015b; Aslay et al., 2014; Chen et al., 2015; Singh et al., 2019; Barbieri et al., 2012)
		Time	Discrete Time-aware IM	(Chen et al., 2012; Lee et al., 2012; Liu et al., 2014)
			Continuous Time-aware IM	(Gomez-Rodriguez et al., 2011; Gomez Rodriguez et al., 2012; Du et al., 2016; Xie et al., 2015; Tang et al., 2015; Mohammadi et al., 2015)
		Competitive	Known Competitor	(Carnes et al., 2007; Bharathi et al., 2007; Borodin et al., 2010b; Ceren Budak et al., 2011; He et al., 2012)
			Unknown Competitor	(Lin et al., 2015; Li et al., 2015d)
			Comparative	(Wei et al., 2015; Tang et al., 2014; Ou et al., 2016)
		Dynamic	Probing-based	(Aggarwal et al., 2012; Zhuang et al., 2013)
			Sampling-based	(Song et al., 2017, 2015, 2016, 2019)
			Others	(Lei et al., 2015; Tong et al., 2017; Tong et al., 2017; Wang et al., 2017c)
		Conformity	Complexity Improvement of MC	(Tang et al., 2013; Li et al., 2014b; Li et al., 2013a; Li et al., 2018c)
		Semantic	Reverse Reachable	(Chen et al., 2020)
	Extensions	Profit Maximization	Advertisers	(Tang et al., 2018; Wei and Lakshmanan, 2012)
			OSN Providers	(Tang et al., 2017)
		Revenue Maximization	Discounting	(Han et al., 2018)
			OSN Providers	(Khan et al., 2016)

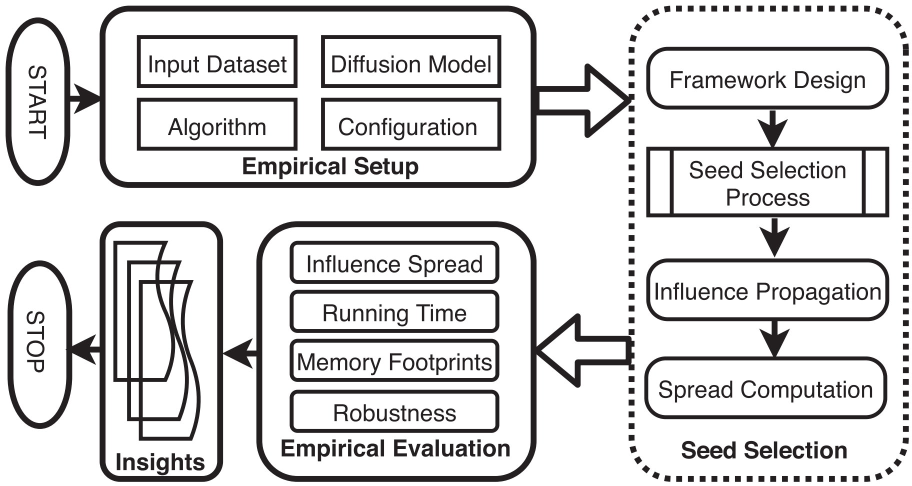


Fig. 5. The basic IM framework.

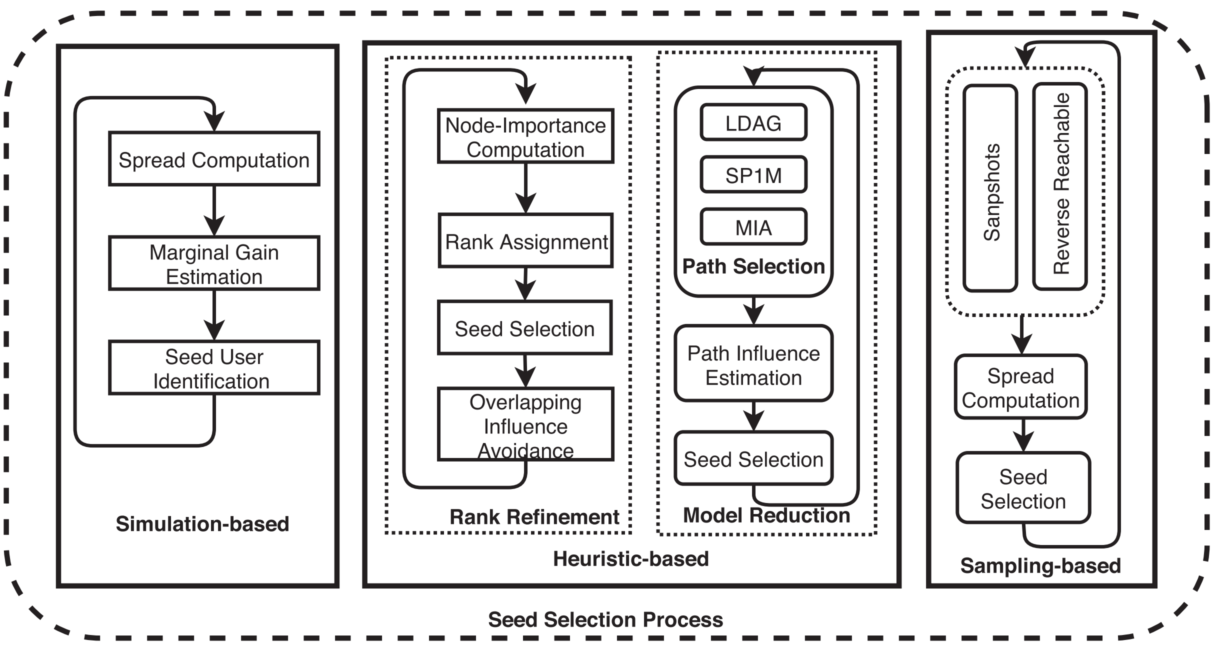


Fig. 6. The seed selection process under classical IM framework.

3.2. Classical IM framework

Most of the work in the field of IM follows the classical IM framework, which identifies the seed users in single networks under single diffusion process. Greedy algorithm (Kempe et al., 2003) was the first work proposed under classical IM framework. The greedy algorithm guarantees an approximation ratio of $(1 - 1/e - \epsilon)$. However, it is still challenging to estimate the expected influence spread using IM. The theoretical hardness of the expected influence computation function $\sigma(S)$ leads to the possibility of conducting extensive research in this area for developing efficient approaches to solve IM problem. Existing research work under this framework can be divided into three classes 1) simulation-based, 2) heuristic-based, and 3) mixed IM approaches. This categorization is done on the basis of how a method overcomes the #P-hardness of influence spread estimation (Arora et al., 2017; Li et al., 2018d). The seed selection process under this framework based on the algorithmic design is shown in Fig. 6. The seed selection process in traditional IM framework is categorized into three classes: simulation-based, heuristic-based, and sampling-based approximation methods. Simulation-based algorithms iteratively computes the marginal gain of each node with current seed set and identifies the most influential users. Heuristic-based algorithms selects the seed by either defining ranking of nodes based on influence or utilizing some topological features like paths, group norms, etc. Sampling-based algorithms utilize the sample of network and reachability of nodes from one another to find the influential users.

1. **Simulation-based Approaches.** These approaches utilize the *Monte-Carlo* (MC) simulations to compute the spread of influence which is time-consuming. It is computed for a given seed set of an information diffusion model in the network. These methods perform a large number of explicit MC simulations

corresponding to every node for computing the average spread of influence which further leads to the estimation of marginal gain of the nodes. The node having highest marginal gain influence corresponding to current seed set is identified as a new seed. Nodes identified as seed are ignored in each subsequent iterations so as to compute the spread of influence. This process is repeated until all the seed nodes are identified. Sviridenko (2004) presents a modified IM problem of Kempe et al., by adding node price constraints which was earlier missing in Kempe et al. (2003). Unlike the traditional unit node price selection process, the model proposed by Sviridenko (2004) selected a subset of seed nodes with different node prices. The greedy approaches are not computationally efficient for large-scale networks due to higher number of MC simulations to be computed. The aforementioned limitation drew the interest of researchers to work on it's optimization, which can be classified in two classes: reducing number of MC simulations (Leskovec et al., 2007; Goyal et al., 2011; Zhou et al., 2015a) and complexity reduction of MC simulations (Wang et al., 2010; Jiang et al., 2011).

- *Minimization of number of MC simulations.* To minimize the number of MC simulations, most of the existing state-of-the-art algorithms compute an upper bound of marginal gains corresponding to node x, i.e., $\sigma(S \cup \{x\})$ and $\sigma(S)$. The non-candidate of seed selection can be eliminated from subsequent rounds. One of the approach in this regard was proposed by Leskovec et al. (2007) named *cost-effective lazy forward* (CELF) which was 700 times efficient when compared to the greedy algorithm. CELF uses Property 2 of a sub-modular function of cascade influence. The marginal gain $\sigma(x|S) = \sigma(x \cup S) - \sigma(S)$ of a node $x$is maintained by the greedy algorithm in each iteration to utilize submodularity. The diminishing returns property states that if a node's marginal gain is greater than other nodes marginal gain at any iteration then the same will also be true for all other subsequent iterations.. Therefore, the nodes with less marginal gain can be avoided in subsequent iterations. Thus, the time efficiency of the greedy method is significantly improved.

  Goyal et al. (2011) presented CELF++ method which is a modification of CELF and estimates an arbitrary node u marginal gain with S and $(S \cup v)$ in each round of iteration and are given as, $\sigma(u|S) = \sigma(u \cup S) - \sigma(S)$ and $\sigma(u|(S \cup v)) = \sigma((u \cup (S \cup v)) - \sigma(S \cup v)$ respectively. By now, v represents maximum marginal gain node in the current iteration. CELF++ simultaneously estimates two marginal gain values in each iteration, and thus, is 30% to 50% times faster than CELF. The algorithm for the estimation of an upper bound to spread of influence of every node by matrix analysis is presented by Zhou et al. (2015a). This method computes an upper bound of expected influence of u by a few multiplications of sparse matrix and avoids most of the initial iteration of CELF++.
- *Improving the complexity of MC simulations.* The other way to improve efficiency of greedy algorithm is to reduce the search space for each node in MC simulation. A *community based greedy algorithm* (CGA) was presented by Wang et al. (2010) to improve the MC complexity by reducing the search space. It utilizes divide and conquer method by first dividing the network as subnetworks and then identifying the influential users within each subnetwork. In this, a cost function was proposed to identify the structure of the community in any mobile networks. Therefore, CGA is not suitable for large-scale networks. Chen et al. (2014) proposed an algorithm viz. *community-based influence maximization* (CIM) that efficiently used the community-based framework to identify seed nodes. It initially detects structure of the community of influence graph and then based on the seed quota, it selects the seed nodes from each community. Though it works efficiently, but influence spread is still a problem. The advantage of CIM and CGA is the reduction of search space by partitioning of the network so that MC simulations run only on subnetworks. The authors of Singh et al. (2019e) adopts similar idea of Chen et al. (2014) to solve the IM problem.

  Sheng and Zhang (2018) presented an algorithm LPIMA based on community detection to identify the seed nodes in the network. The CGA algorithm improves the efficiency by reducing the search space using partitioning but it needs to simulate marginal gain of each node in the community. Therefore, to improve the efficiency of CGA, they proposed LeaderRank method to simulate the influence of community nodes. Then, LPIMA selects candidate nodes based on quantify value of influence. They also incorporate the submodular property to further improve the efficiency of the greedy approach. The experimental results show that LPIMA is more efficient than CGA.

2. **Heuristic-based Approaches.** Unlike simulation based approaches, Heuristic-based IM approaches uses approximate scoring method to evaluate $\sigma(x)$ for each node x. Time-consuming MC simulations are thus avoided in Heuristic-based approaches making them more scalable and efficient. The heuristic approaches are bound to their model most of the time and utilize the network topological features, features of corresponding diffusion model, and quasi-local features. These approaches can be categorized into two groups: 1) rank refinement, and 2) model reduction.
   - *Rank refinement.* Rank refinement methods assign the rank to each individual based on some approximated metrics for estimating the influence spread. Then seed nodes are easily selected based on their ranking. There exists some simple rank refinement methods like, degree, distance centrality (Freeman, 1978), and PageRank (Page et al., 1999) to select the seed users. However, these methods do not provide a good quality solution because these methods avoid influence overlaps and features of diffusion models. To overcome the weaknesses of simple ranking methods, some influence ranking methods like GROUP-PR (Liu et al., 2014), Degree Discount (Chen et al., 2009), IRIE (Jung et al., 2012) etc., are introduced. Chen et al. (2009) proposed Degree Discount algorithm based on the thought that if a node is selected as seed, it cannot be influenced by its neighbors anymore. Therefore, neighbors' degree are reduced by one. Firstly, node having max degree is selected as seed node from the network and it's neighbors' degree are reduced by one. Every time node with the highest degree is selected as seed followed by degree discount. Though this performs better compared to max degree method in terms of accuracy but the improvement is very little and negligible.

     The authors of Liu et al. (2014) extend the idea of PageRank from an individual to a set of users, named as GROUP-PR. This method estimates the ranking score of a node set S by adding the PageRank score of each individual $x \in S$. GROUP-PR is an influence upper bound estimation method and follows the greedy framework. Firstly, it estimate the PageRank value of every individual. Next, it repeatedly finds the maximum marginal gain user x in terms of ranking score and adds x to S. Finally, GROUP-PR is more efficient because it avoids native MC simulations and is more accurate than PageRank. GROUP-PR is more efficient because of it avoids native MC simulations and more accurate than PageRank. Jung et al. (2012) proposed a generalized PageRank approach, IRIE which was based on message passing influence estimation. Utilizing this estimation It designed a system of $|V|$ equations where $|V|$ varies to compute the influence of each individual $x \in V$. To estimate influence spread of x, it combines the influence of itself with its direct influence to neighbors i.e., $(1 + \alpha \sum_{y \in N_{out}(x)} p_{xy}.Inf_y)$, where $\alpha \in [0, 1]$. This method solves the system of linear equations iteratively and after k iterations it returns the seed set S. IRIE matches influence spread with greedy with high efficiency. There are some other rank refinement algorithms such as SPIN (Narayanam and Narahari, 2011), IMRANK (Cheng et al., 2014), etc.

- *Model reduction.* The #P-hardness of expected influence spread $\sigma(S)$ computation of seed S is simplified by the Model reduction method. Model reduction is handled in two ways in IM problems: 1) stochastic models are reduced to deterministic models and estimate the exact spread, and 2) by restricting the influence to local region. Kimura and Saito (2006) utilized shortest influence paths from network and presented two models *shortest path model* (SPM) and SP1M by assuming that only first two shortest paths participate in influence spread process. Thus, the recursive computation of Dijkstra shortest-path method is utilized to compute the spread of influence. SP1M does not need MC simulations. To improve its performance, an objective function is estimated by using approximation strategy. Since, SPM/SP1M both the models only utilize path length among individuals and avoids influence probabilities among them, to achieve a good approximation ratio, SPM/SP1M needs to consider influence probabilities along with path information.

  Chen et al. (2010e) proposed an algorithm named *maximum influence arborescence* (MIA) which was inspired by SP1M. It maintains local arborescence structure to estimate influence speed. It only considers the highest propagation probability paths to evaluate the influence spread and restricts the influence spread to local tree structure. Therefore, it is more time efficient and avoids MC simulations. The authors of Chen et al. (2010e) also present another variant of MIA, named as prefix excluding MIA (PMIA), which is computationally more efficient. Kim et al. (2013) proposed an approach named *independent path algorithm* (IPA) based on reduction of ICM, similar to MIA. IPA assumes that influence paths from node x to y are independent of each other. It considers all the paths from node x to y whose propagation probability is more than threshold unlike MIA/PMIA. Influence spread of paths is computed parallelly, therefore IPA is more time-efficient.

  Using *directed acyclic graph* (DAG) structure, Chen et al. (2010c) introduced an algorithm viz. LDAG. It has the similar idea as PMIA, but tailored for LTM. For each node x, it constructs local DAG structure using Dijkstra shortest-path algorithm and assumes that influence of x is limited to its local region. The construction of LDAG for a node x is both computationally as well as memory intensive for large-scale networks. It uses greedy hill-climbing approach for seed selection which evaluates seed linearly. Therefore, LDAG is more scalable and tractable.

  Goyal et al. (2011) introduced a scoring-based IM approach under LTM, named as SIMPATH. It estimates the influence spread of seed nodes or a set of nodes by enumerating all simple paths starting from seed nodes. SIMPATH uses a pruning strategy to restrict small neighborhood. It uses the vertex cover optimization technique to decrease the running time in the first iteration. It uses look ahead optimization in subsequent iterations. It generates the same level of influence spread as greedy. SIMPATH is more time and space efficient than LDAG. The authors of Galhotra et al. (2016) proposed another proxy method for both ICM and LTM diffusion models, named EASYIM. It enumerate all simple paths within length l to estimates the influence spread of seed nodes. To improve the accuracy of this method, the overlaps between paths are considered. EASYIM incorporates the IRIE method to estimate global influence in an iterative manner and achieves better output.

3. **Mixed Approaches.** Mixed IM approaches improve the theoretical efficiency of simulation based IM approaches with an approximation guarantee. These approaches perform prior computation of a number of samples or sketches under specific diffusion model to avoid rerunning of time-consuming MC simulation. These approaches compute the influence spread by exploiting the sketches. It is also known as sampling-based methods. These IM approaches are categories into two classes based on how samples have been generated: snapshot-based and reverse reachable (RR) set.
   - *Snapshot-based sampling.* The idea of snapshot based sampling approaches is to generate a sketch or sub-graph by creating and selecting an instance of influence diffusion process associated to specific influence propagation model such as ICM and LTM. Next, it computes the expected spread of influence corresponding to seed users S accurately using these snapshots with approximation guarantee. For example, a graph $G = (V, E, W)$ with a diffusion model ICM constructs a snapshot by removing each edge (x, y) with $(1 - p_{x,y})$ probability and form a sub-graph G_i. Let $I_S(G_i)$ represents the influence spread of seed S on G_i and $\{G_i, G_2, \ldots, G_m\}$ are the m-snapshots of G at different instances. Then, the expected influence spread of seed set S is the average of influence spread of S on these snapshots, i.e., $\sigma(S) = \frac{1}{m}\sum_{i=1}^{i=m} I_S(G_i)$. The greedy framework with snapshot sampling can achieve $(1 - 1/e - \epsilon)$ approximate solution. A review of the snapshot sampling based IM approaches has been presented here.

     Chen et al. (2009) proposed a sampling method based on forward influence under ICM diffusion model, named as NewGreedy. It constructs a number of snapshots at different instances to evaluate marginal gain $(\sigma(S \cup x) - \sigma(S))$ of each node $x \in V \setminus S$ at each iteration of NewGreedy. The asymptotic complexity of constructing a snapshot G_i of G is equal to the complexity of running a MC simulation. Therefore, NewGreedy is significantly better than simulation based greedy (Kempe et al., 2003) regarding efficiency with $(1 - 1/e - \epsilon(r))$ approximate influence spread. Cheng et al. (2013) proposed a sampling algorithm viz. StaticGreedy which guarantee the sub-modularity of objective function $\sigma(S)$ in seed selection process. It constructs $(m = (8 + 2\epsilon)N\frac{\log N + \log \binom{N}{k} + \log 2}{\epsilon^2})$ snapshots to achieve $(1 - 1/e - \epsilon)$ approximate solution with $(1 - n^{-1})$ probability. StaticGreedy is much faster than simulation greedy. However, the worst case time-complexity of StaticGreedy is still a concern. To further improve its efficiency, StaticGreedyDU (Cheng et al., 2013) was introduced. It uses a pruning strategy to empirically reduce its efficiency. StaticGreedyDU prunes each node reachable from the seed set Si at iteration i from all the snapshots, and subsequent influence spread estimation performed on pruned snapshots, which would improve its efficiency.

     PRUNEDMC (Ohsaka et al., 2014) was introduced to further improve the time-efficiency of StaticGreedyDU (Cheng et al., 2013) using an index structure. It builds a DAG for each

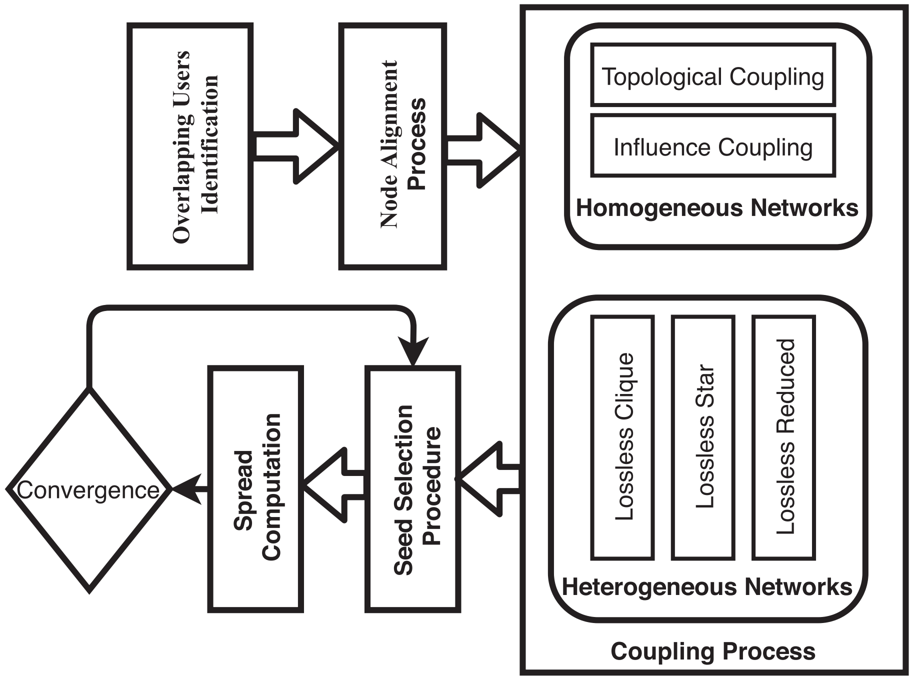


Fig. 7. The seed selection process under IM2 framework.

snapshot G_i, and every node x in DAG is a strong connected component of snapshot G_i. A maximum degree node is selected from each DAG as a hub node. To built an index structure, it marks ancestors and descendants of the hub node on the snapshots. PRUNEDMC speeds up the estimation of influence spread of x by avoiding the traversal of descendants of the hub node, if x is the ancestor of the hub node for the corresponding snapshot. Therefore, the running time of marginal gain estimation of a node x is effectively reduced by combining index structure with StaticGreedyDU pruning technique. Cohen et al. (2014) introduced a snapshot sampling approach SKIM. In order to speed up the computation of influence spread on constructed snapshots, this method uses bottom-K^2 min–Hash. It performs reverse Breadth First Search (BFS) on snapshot and updates bottom-K^2 min–Hash values simultaneously for a number of candidate seed sets. SKIM is faster than simulation based algorithms and some heuristic algorithms, but its worst-case complexity is equal to StaticGreedy.

- *Reverse reachable (RR) sets.* Borgs et al. (2014) are the first to introduce the concept of reverse reachable sets in IM problem. They state that the estimation of influence spread on sketches constructed by operating on the whole graph is not necessary. In the reverse reachable approach, the estimation of influence spread of a seed set S is based on the selection of random nodes and seeing the portion of selected nodes which can be reached by seed set S. Based on RR sets, Borgs et al. introduced a threshold-based method, named as RIS. In this method, they construct RR sets continuously until each edge $(x, y) \in E$ is examined during the RR sets construction process and it reaches a threshold θ.
Tang et al. (2014) proposed TIM, to make RR approach more practical and efficient. It improves the RIS by performing a better analysis on the required number of RR sets to achieve the same theoretical approximation bound. TIM requires m RR sets to ensure theoretical bound, where $(m = O(\frac{\epsilon^{-2}N(\log N + \log \binom{N}{k})}{IS_{optS}}))$ and IS_{optS} denotes the influence spread of the optimal seed set. The authors of Tang et al. (2014) introduced another variant of TIM by improving parameter estimation process, known as TIM+. It gives better empirical performance than TIM with the same worst-case time complexity. To further improve the performance of TIM/TIM+, Tang et al. (2015) proposed a martingale approach, known as IMM. It has a better bootstrap parameter estimation procedure than TIM/TIM+. Therefore, IMM is more time-efficient than TIM/TIM+.
Among the discussed RR sampling based approaches, i.e., RIS (Borgs et al., 2014), TIM/TIM+ (Tang et al., 2014), and IMM (Tang et al., 2015), it is important to notice that these approaches may require a large memory space. This is because of two reasons: a large number of RR sets are generated to preserve theoretical approximation, and every RR set should be stored in the memory for seed selection process by greedy procedure. In order to overcome the memory limitations, the authors of Wang et al. (2017a) proposed BKRIS method based on lazy sampling approach. First, BKRIS computes lower bound on influence spread of optimal seed set IS_{optS}. Next, it estimates the number of RR sets m using lower bound on IS_{optS}. Similar to SKIM (Cohen et al., 2014), BKRIS adopts bottom-K min–Hash strategy. Then, it computes the seed set by fully utilizing every RR set unless necessary. The lazy sampling strategy of BKRIS practically speeds up the IMM by two orders of magnitude. To further improve the performance of IMM, the authors of Nguyen et al. (2016) proposed an orthogonal *stop-and-stare* optimization approach SSA. This approach doubles the number of RR sets iteratively and generate the seed sets using current generated RR sets. It stops the iteration whenever estimated influence $\sigma(S_i)$ at iteration i is close to estimated influence $\sigma(S_{i-1})$ computed at iteration $(i-1)$. They also present an improved variant of SSA named D-SSA. The authors also claim that SSA/D-SSA ensure $(1 - 1/e - \epsilon)$ approximation ratio.

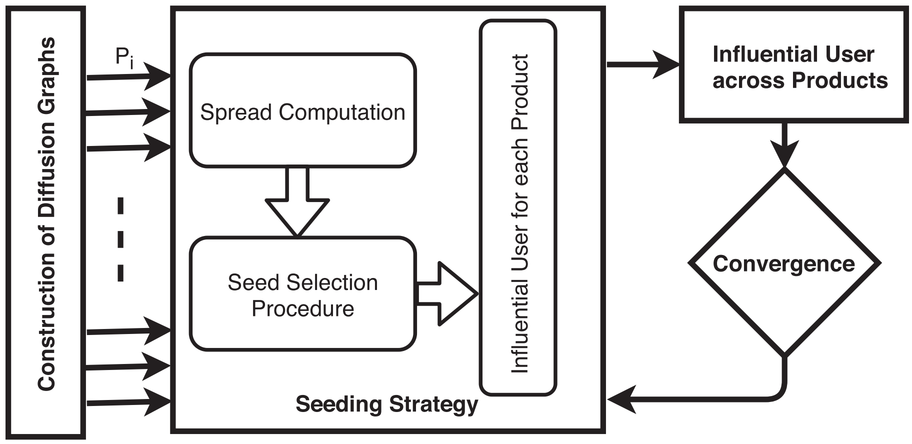


Fig. 8. The seed selection process under MIM framework.

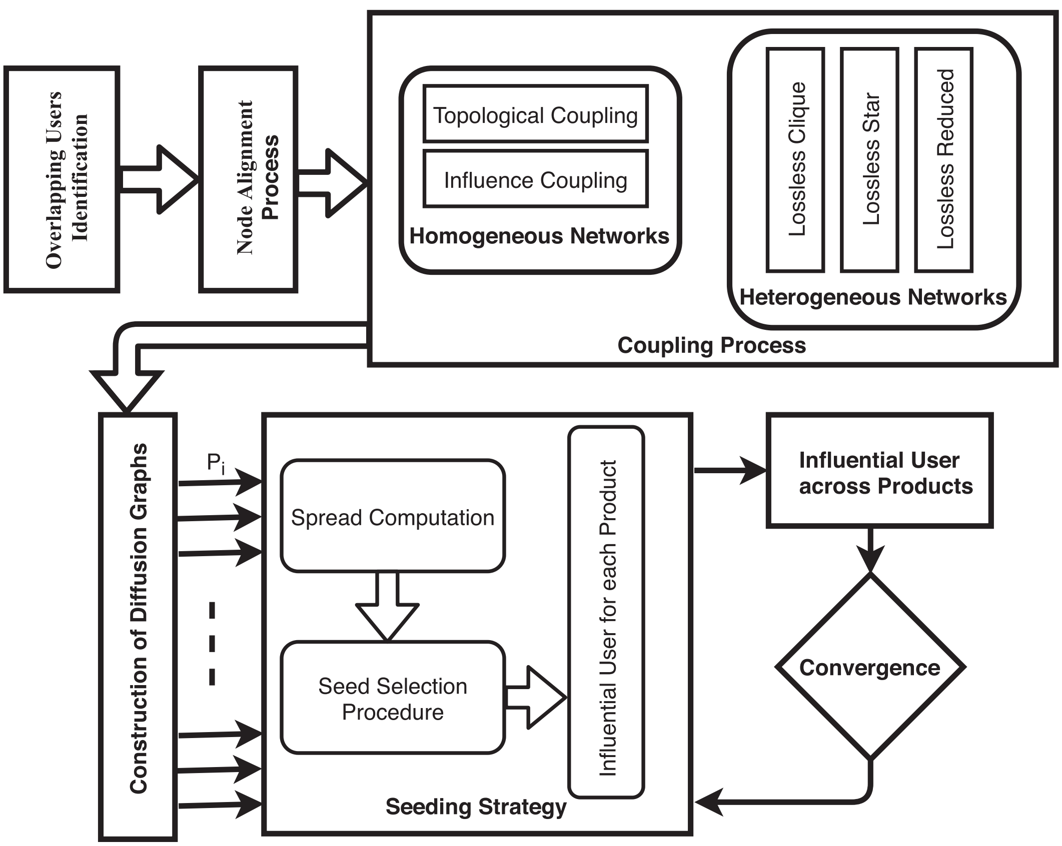


Fig. 9. The seed selection process under MIM2 framework.

3.3. Influence maximization across multiple networks (IM2) Frameworks

Nowadays, people spend a lot of time in online social networks and actively engage in multiple networks simultaneously which allow them to propagate information across the networks. Therefore, to estimate the individual influence accurately, we need to configure multiple networks simultaneously. With this, a new framework is introduced in literature, named as IM across multiple networks (IM2). The seed selection process based on the algorithmic design under this framework is shown in Fig. 7 based on their algorithmic design. Under this framework seed selection process first identifies the overlapping users across different networks then node alignment process needs to be applied for assuring uniformity across networks. Next, coupling of networks have been performed based on homogeneity and heterogeneity nature of the networks. After that a seed selection method can be applied to identify the influential users.

Zhang (2015) studied IM problem on multilayer network. They focused on a specific event on twitter for information diffusion. They adopted CELF++ and SIMPATH (Goyal et al., 2011) to estimate influence spread in the network. The authors used different strategies to assign influence probability in the network over distinct phases. The experimental results show the advantage of IM2 framework over classical IM framework. The authors of Nguyen et al. (2013) and Zhang et al. (2016) the proposed algorithm for

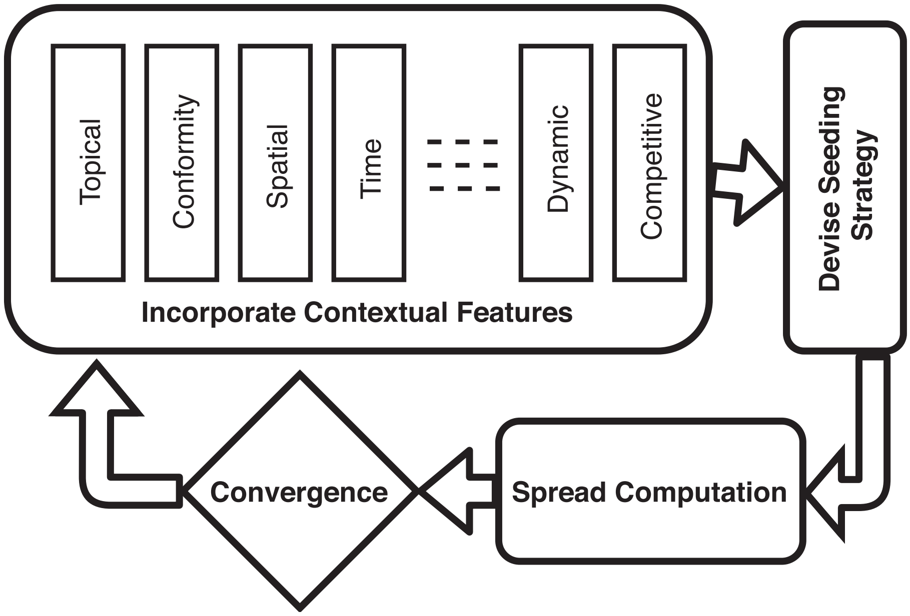


Fig. 10. The seed selection process under context-aware IM framework.

influence maximization across multiple social networks by inducing an improvement over the greedy algorithm. New lossless and lossy coupling schemes were also proposed by them so as to link the multiple social networks effectively. To compute the effective seed, all the network properties of the original networks were preserved by the lossless coupling scheme whereas if the memory consumption and running time were the primary aim then lossy coupling proved to be an attractive alternative. The experimental results show that the presented coupling schemes outperform the real social networks as well as the synthesized networks.

Wang et al. (2016b) presented the IM problem in a distributed framework rather than centralized and named it as an agent selection problem (ASP). They designed a multilayer IM strategy for mobile social network and mobile ad hoc network under IC diffusion model. The authors proved that ASP problem is NP-hard under IC model. The experimental results on synthetic and real datasets show that the proposed distributed cross-layer IM strategy significantly reduced the message overhead compared with the state-of-the-art methods.

3.4. Multiple Influence Maximization (MIM) Frameworks

Previous works do not consider a scenario where an advertiser wants to promote multiple products in the same social network simultaneously. In reality, most of the users have different interest for different products as well as different acceptance probabilities. Hence, an advertiser may consider a promotional strategy where seed users freely recommend multiple products simultaneously and at the same time non-seed users can accept several products. Based on this assumption, the authors of Sun et al. (2016) proposed a framework, named as multiple influence maximization (MIM). The seed selection process under this framework is shown in Fig. 8 based on their algorithmic design. Firstly, MIM framework constructs the product diffusion graph for each product. Then, influential nodes will be selected based on seeding strategy for each product. Next, seed nodes will be compared across product diffusion graph and most influential node across product will be added to the seed set. This process repeatedly computes the remaining seed.

MIM problem in social networks was first introduced by Sun et al. (2016). In this, two things were simultaneously assumed, firstly that a seed user can accept variety of products for free, and secondly non-seed users are equipped with enough purchasing power to accept variety of promotions from their social friends. The greedy algorithm have been utilized in this work to identify seed nodes. MIM guaranteed approximation ratio $(1 - \frac{1}{e})$ only when the expected influence spread function $\sigma(S)$ is submodular. Independent cascade diffusion model have been utilized to propagate influence and stated that MIM is sub-modular under IC model.

3.5. Multiple Influence Maximization across Multiple Networks (MIM2) Frameworks

The MIM2 framework considers both the scenario of IM2 and MIM framework, i.e., it tries to maximize the influence spread of the seed users for several kind of products across the multiple networks simultaneously. The seed selection process under this framework is shown in Fig. 9 based on their algorithmic design. The MIM2 framework first identifies the overlapping users and then performs the node alignment across network to achieve uniformity. Next, a network coupling method is applied to form a new multiplex or multi-layer network for influence propagation. After that, the MIM framework is utilized to identify the influential users. Singh et al. (2019f) were the first one to introduce the MIM2 framework. This work first identifies the overlapping users who propagate information across the networks and assign a unique identification number to each individual. Then the algorithm coupled multiple networks into a single multiplex network using the direct linkage strategy. Next, the algorithm performs backward propagation from non-candidate nodes to compute the diffusion value index for each product. Based on diffusion index value seed nodes are identified.

3.6. Context-aware IM Framework

The studies related to context-aware IM approaches have emerged in the recent years. These approaches are the extensions of the conventional IM problem by further considering some contextual features like location of the users, topic of information or product category, time of information diffusion, competitive marketing, and dynamic nature of online social networks, etc. The seed selection process under this framework is shown in Fig. 10 based on their algorithmic design.

1. **Topic.** The classical IM problem can be extended by considering *topics* of product being propagated, known as Topic-aware influence maximization (TAIM) problem. In order to formalize the definition of TAIM, *topic* denotes user's interests as well as the product's characteristics. The influence spread is dependent on both seed nodes and topics. Then, TAIM is the problem of identifying the optimal seed nodes that maximize the influence over the network for a given topic-aware query. This problem can be classified into two classes based on the users and their relationship characteristics: topic relevant targets and topic dependent diffusion.
   - *Topic relevant targets TAIM.* The topic relevant targets (TAIM) approaches focus on the user's characteristics, i.e., user is topic-aware. There are some efforts such as LGA/ELGA (Guo et al., 2013), KB-TIM (Li et al., 2015c), IMIP/IMAX (Lee and Chung, 2015), CTVM (Nguyen et al., 2016), MFIP (Li et al., 2020), etc., made on maximizing the influence spread over topic-aware users. These approaches differentiate the users in the network and compute the influence spread $\sigma(S)$ of seed nodes over the targeted users.
     Li et al. (2015c) proposed a topic-aware query model to identify the targeted users based on their profile and estimated the seed nodes over the targeted space. The user's profile consists of the information about the preferences of the users regarding specific topics. For example, a user profile $\{< sports, 0.6 >, < tech, 0.5 >, < drama, 0.3 >\}$ denotes the probabilities of user likes to various topics. Then, the algorithm selects targeted users based on a given topic query. The authors incorporate RR sampling strategy and ICM diffusion model to address the targeted IM problem. To find an unbiased estimator for estimating the targeted influence spread, they introduce a weighted sampling strategy with the RR concept. The algorithm select the RR sets based on targeted topic query and merges these sets to calculate the result. Further, to reduce the I/O cost, the authors presented an incremental index structure.
     The authors of Nguyen et al. (2016) generalize TAIM problem by using pre-defined targeted influence function. They adopted RR sampling framework like (Li et al., 2015c). The algorithm avoids generation of too many RR sets by utilizing an early termination rule. Moreover, they also present the cost-aware setting where activation of a user is dependent on the cost function. Guo et al. (2013) presented a special case of TAIM problem, known as personalized IM problem. This problem identifies a set of seed users to maximize the overall influence spread on a given target user which is considered as the topic relevant user among all. They adopted ICM and introduced two algorithms i.e. local greedy (LGA) and efficient local greedy (ELGA). LGA is a simulation-based approach with some pruning rules tailored for the target user local structure. This approach is not applicable for the online query requirement. ELGA is a heuristic approach which considers only the shortest path for information diffusion from each node to the target node. Therefore, the approximation of influence spread can not be guaranteed theoretically.
   - *Topic dependent diffusion TAIM.* The topic dependent diffusion TAIM approaches focus on the user-to-user topic-aware diffusion, i.e., user's relationship with others. Some studies such as AIR-Greedy (Barbieri et al., 2012), INFLEX (Aslay et al., 2014), TIM/MIA (Shuo Chen et al., 2015b), MIS/BTS (Chen et al., 2015), C2IM (Singh et al., 2019), etc., focus on topic dependent diffusion TAIM. These approaches consider the idea that each edge (x, y) between a pair of individuals x and y is topic dependent. The reason behind this is y might be activated by x on some topics (e.g., tech, drama) and still inactive for others (e.g., sports). To formalize the model, let each edge (x, y) be associated with a propagation probability vector $P_{x,y} = \{pp_1, pp_2, \ldots, pp_t\}$ over t topics and a topic dependent query $Q = \{q_1, q_2, \ldots, q_t\}$. To compute the $P_{x,y}$ of an edge (x, y) under ICM, the dot product is calculated, i.e., $P_{x,y} = \sum_{i=1}^{i=t} q_i . pp_i$.
     The authors of Barbieri et al. (2012) and Aslay et al. (2014) were the first to introduce topic dependent TAIM problem. The idea was that the similar topic-aware queries will have approximate influence spread for the queries. The authors introduced an indexing scheme named as INFLEX based on *similarity-search* and *pre-computation of seed set*. This approach cautiously samples reasonable number of topic distribution queries, and pre-computes seed nodes for each query by any IM approach. Similarly, this approach pre-computes level-2 seed nodes under each query distribution, and combines these level-2 seed sets using a *rank-aggregation* method. Finally, it presents maximum-likelihood Dirichlet estimator, Bregman-ball tree, and Kendell scheme for query sampling, similarity search, and seed set aggregation respectively. The authors of Chen et al. (2015) proposed a TAIM approach for some special graph by incorporating similar framework of Aslay et al. (2014). These special graphs follow some properties like sub-additive, typically-separable, etc.
     To provide approximation guarantee, Shuo Chen et al. (2015b) improved the prior works (Barbieri et al., 2012; Aslay et al., 2014; Chen et al., 2015). The authors adopted MIA/PMIA framework under ICM model. The approximation ratio of the algorithm is bounded under MIA/PMIA framework. To estimate an upper bound of influence spread $\sigma(x)$ of a user x, they presented a *best-effort* framework. Then, the algorithm estimates the exact influence spread and prunes the insignificant users based on the upper bound influence. The authors also presented a topical-sample algorithm to pre-compute seed sets for offline-sampled topic distributions. This algorithm outperforms the methods in (Barbieri et al., 2012; Aslay et al., 2014; Chen et al., 2015) regarding influence spread with comparable efficiency.
     Singh et al. (2019) proposed C2IM, which is a community based heuristic approach. They utilized community-based framework that proved out to be better than the existing algorithm. Here, diffusion degree of each node was computed, by detecting the non-desirable nodes for reverse tracing. Diffusion degree of each node play an important role in the selection process of seed nodes based on seed quota. For topic dependent queries, C2IM performs good. However, it might/might not work good for the theoretical approximation of influence spread.
2. **Location.** Location-aware influence maximization (LAIM) problem is the extension of classical IM problem by considering *spatial index* due to commonness of location-based social networks such as Foursquare, Twitter, etc. The objective of LAIM is to maximize the estimated influence spread of location-relevant users unlike generic IM. Some efforts have been made to solve LAIM problem in (Li et al., 2014a; Zhou et al., 2015b; Song et al., 2016; Wang et al., 2017b; Sen et al., 2018; Hosseinpour et al., 2019).
   The authors of Li et al. (2014a) were the first to introduce LAIM framework, focusing on region queries. Given a geographical

region R, LAIM is the problem of identifying k users as seed nodes S to maximize the influence spread over region R. They incorporated the influence estimation model used in MIA/PMIA (Chen et al., 2010e) under ICM diffusion model. The authors proposed an algorithm Expansion which uses the best-first search framework to identify the seed. This search procedure accesses users with large upper bounds on influence and prune the users with insignificant influence. They also focus on developing upper bounds for pruning. *QuadTree* structure is developed for fast allocation of users with *spatial index*. In this work, another algorithm Hint was proposed which performs the pre-computation of seed S_i for every *QuadTree* leaf regions. Then, combines all S_i as hints to compute upper and lower bound of influence. They perform experiments on real-world location-based social networks to evaluate the performance of LAIM algorithms, and reports running time efficiency in milliseconds for different size regions.

Similar to Expansion/Hint, Wang et al. (2017b) introduced a distance-aware pruning-based algorithm MIA-DA, which considers user's distance to a query location as edge weights. The authors adopted MIA/PMIA (Chen et al., 2010e) framework under ICM diffusion model to compute the influence spread. MIA-DA selects a set of locations as *anchor-locations* to estimate the influence bounds. Each anchor-location was selected as a query to compute the influence spread. It utilizes the triangular equality to compute the influence bound for every anchor-location. MIA-DA can also bound with region-based bound estimation in Expansion/Hint. The authors of Song et al. (2016) also adopted distance-aware query and proposed an algorithm Target-IM/Target-IM+. They incorporated RR sets instead of MIA/PMIA proxy in Wang et al. (2017b). The algorithm developed a pool of tree structures based on weighted RR sets. They also highlighted that the approximation ratio of the algorithm is $(1 - 1/e - \epsilon)$. The authors of Zhou et al. (2015b) presented an algorithm TPH based on distance-aware weighting model.

The authors of Guo et al. (2017) proposed IM over trajectory database. They redefined the IM problem as selection of k trajectories on a given advertisement to maximize the influence spread for a large group of audience. The authors used a community-based framework that divides the trajectory database into communities to find the promising trajectories. This work is different from classical IM approaches as it does not incorporate any diffusion models and influence propagation. The authors of Li et al. (2018a) and Li et al. (2018b) proposed a holistic influence diffusion model (HIM) under ICM settings to spread influence over *spatial* social network. They were the first to study how a user x propagates his influence to another user y by spatial interactions together with social influence. They first compute RR sets based on *keyword* query Q. Then they present a baseline method SimRmNN to find seed nodes via their spatial interactions. The authors also proposed two efficient algorithms Index-based and SimRmNN-Upper to answer HIM queries. These improved algorithms are one or two order of magnitude faster than the baseline method.

3. **Time.** The classical IM problem terminates the diffusion process when there are no more nodes that has adopted the product, idea, or innovation. This condition is practically inefficient and unreasonable as influence propagation process may take a long time. For example, discrete time diffusion models may take $O(N)$ steps and continuous time diffusion models may take an arbitrary time length. Therefore, time-aware influence maximization (TimeAIM) problem introduced a time constraint with propagation model to handle the above stated issue.
   - *Discrete Time-aware IM approaches.* Some discrete TimeAIM approaches such as MIA-M/MIA-C (Chen et al., 2012), CT-IPA (Lee et al., 2012), MISP (Liu et al., 2014) has been studied. Chen et al. (2012) introduced the IC-M model, which adopted traditional ICM model.. In IC-M model, a node x contacts y with a meeting probability $m(x,y)$ through edge (x,y). Let node x succeed in meeting with y through edge (x,y), then x has only one chance to activate y with $p_{x,y}$ and y has many chances to contact x. TimeAIM under IC-M model selects seed nodes to maximize expected adoption within τ time-steps over random processes. The authors of Lee et al. (2012) and Liu et al. (2014) presented two identical models CT-IC and LAIC. Under both models, a node x activates y with probability $p(x,y).p_x^{lat}(\Delta_t)$ at $t+\Delta_t$ through edge (x,y), if x is already activated at time-step t. These models are the extension of traditional ICM. ITo find the IM solution, they adopted MIA approach. Therefore, these studies do not have theoretical guarantee due to MIA heuristic nature.
   - *Continuous Time-aware IM approaches.* Gomez-Rodriguez et al. (2011) were the first to introduce *continuous-time independent cascade* (CTIC) model. The objective function $\sigma(S)$ is monotone and submodular under CTIC model. The authors of Gomez Rodriguez et al. (2012) proposed an algorithm INFLUMAX which adopts greedy framework for IM with lazy forward optimization. They also described CTIC as continuous Markov chain model (CTMC). The authors of Du et al. (2016) and Gomez-Rodriguez et al. (2016) proposed a snapshot sampling based approach to efficiently estimate the influence spread with an approximation guarantee. It adopted the sampling strategy from SKIM (Cohen et al., 2014).

     Xie et al. (2015) introduced a new model DynaDiffuse where edge probabilities deteriorate over time. They incorporated CELF greedy (Leskovec et al., 2007) and presented an optimized greedy approach to compute the seed nodes. This approach does not provide an error bound for the stochastic procedure under CTMC diffusion model. Therefore, the algorithm does not provide any theoretical guarantee. Ohsaka et al. (2016) proposed a more general time-aware diffusion model i.e, *time-varying independent cascade* (TV-IC). Let a node x activated at time-step t_1 and contact with y through edge (x,y), then the influence reaches node y at time-step t_2. Then, conditional likelihood of y that gets activated at t_2 is dependent on $(t_2 - t_1)$. The authors also presented the first time-aware extension of LTM and names it as *time-varying linear threshold* (TV-LT). They proposed a RR sampling based approach for IM under both models TV-IC and TV-LT. The algorithm provides a provable theoretical guarantee due to submodularity of both the models.
4. **Competitive.** The competitive influence maximization problem is the extension of generic IM problem by considering competitiveness of the products. This problem considers the scenario where multiple competitors follow the same social network to spread influence simultaneously. Therefore, the competitive IM aims to maximize own influence spread or minimize opponent influence spread under competitive framework. The existing works can be classified into three classes: known, unknown, and comparative.
   - *IM under known competitor.* Carnes et al. (2007) and Bharathi et al. (2007) were the first to introduce competitive IM with known competitor. They considered a scenario where n competitive marketers are present for diffusion over a network and n^{th} marketer want to maximize the spread of influence through diffusion process when seed sets for remaining $(n-1)$ competitors are known. Once node x adopts a product by a marketer then it can not adopt other competitive products. They proved that greedy approach under this framework approximation is same, i.e., $(1 - 1/e - \epsilon)$.

The authors of Borodin et al. (2010b) presented various extensions of LTM under competitive framework. They showed that the objective function $\sigma(S)$ is non-submodular under these models. The finding of seed nodes which have influence spread more than $\sqrt{\sigma(OPT)}$ is NP-hard, where $OPT$is optimal seed set under these models. Another variant of competitive IM problem, known as influence blocking maximization (IBM) was introduced. The idea of IBM was to select a seed set S_2 to minimize the influence spread $\sigma(S_1|S_2)$ of the given seed set S_1 for *misinformation* diffusion. The authors of Ceren Budak et al. (2011) and He et al. (2012) presented greedy algorithms for IBM problem under ICM and LTM diffusion models respectively. Zhu et al. (2016) presented an extension of ICM by considering that a node x can be part of multiple seed sets. They achieved a guaranteed approximation $(1 - 1/e - \epsilon)$ by adopting the greedy framework.

- *IM under unknown competitor.* These approaches consider the scenario that no competitor knows the opponent strategy of seed selection, i.e. does not have pre-knowledge of opponent seed nodes (Lin et al., 2015; Li et al., 2015d; Hong et al., 2020). The authors of Lin et al. (2015) presented competitive IM with unknown competitors as multi-round multi-party game. Li et al. (2015d) proposed a model for competitive IM problem using game theory concept. They model IM problem as Nash equilibrium strategy with n-strategies on a given graph under a diffusion model. Each player can maximize his own spread using a Nash equilibrium strategy. The classical IM approaches were used as building blocks to find the seed solutions of these aforementioned problems.
- *IM under comparative framework.* The diffusion process of comparative IM problem can be classified into two categories: Competitiveness and complementary. In competitive diffusion model, a node x less-likely to adopt product P_2, when it already adopts product P_1. Complementary process is vice versa of competitive diffusion model, i.e., if x adopts P_1 then it has higher chance to adopt P_2. The authors of Wei et al. (2015) presented comparative IC model (Com-IC) by extending ICM to tackle the comparative framework. Then, they introduced two problems self influence maximization (SIM) and complimentary influence maximization (CIM) to address competitive and complementary diffusion process respectively. The authors extended the work in Tang et al. (2014) based on RR sampling to find the seed solutions of these problems. Ou et al. (2016) also studied the competitive IM problem and proposed an interactive LT model by extending LTM. They proposed a heuristic approach TOPBOSS to solve the competitive IM problem.

5. **Dynamic.** The IM approaches discussed so far considers a static scenario, i.e., social graph $G(V, E, W)$ and diffusion probability is fixed. However, social networks continuously keep on evolving in the real-world, i.e. new users may arrive in the network, new relationships may be formed. This continuous network formation may affect the influence process. Therefore, some efforts are made in the *dynamic influence maximization* (DIM) problem such as MaxG (Zhuang et al., 2013), IGA (Wang, 2016), IndexingMethod (Ohsaka et al., 2016), UBI/UBI+ (Song et al., 2017), DIM (Wang et al., 2017a), A-Greedy/H-Greedy (Tong et al., 2017), EFOCS (Huang et al., 2019), CM2D (Talukder and Hong, 2020), etc. DIM approaches can be classified into following categories.
   - *Probing-based DIM strategies.* Aggarwal et al. (2012) proposed a DIM approach by considering a social graph G and evolution samples of graph over period $[t, t + \theta]$. They proposed a proxy method to identify a seed set S at time-step t such that $\sigma(S|(t + \theta))$ is maximized. It is a simple proxy method which is not align with the specific diffusion models. Zhuang et al. (2013) proposed a probing-based strategy MaxG to solve the DIM problem. They considered that network evolution is taken into account periodically by a subset probing. This algorithm follows a two-phase procedure. In the first phase, the algorithm selects a set of nodes S_P for probing at time-step $t_s \in [t, t + \theta]$ and constructs a subgraph G_{t_s} by probing S_P. In the second phase, it selects a seed set S_{t_s} on G_{t_s} using Degree Discount (Chen et al., 2009) algorithm to maximize the influence spread of S_{t_s}. This algorithm is not align with any specific diffusion models.
   - *Sampling-based DIM strategies.* There are some existing works (Song et al., 2017; Nathalie et al., 2015) which focus on modeling the dynamics of the network as snapshots of network $\{G_1, G_2, \ldots, G_T\}$. Sampling-based DIM approaches continuously select the seed sets for each snapshot. Song et al. (2017) adopted SP1M (Kimura and Saito, 2006) and UBLF (Zhou et al., 2015a) to estimate the influence spread and presented a heuristic named upper bound interchange (UBI). Initially, UBI selects seed nodes for G_1 based on offline strategy and estimate influence spread of each node. Then it iteratively updates the expected spread of each node. It also updates the seed set iteratively for each snapshot by interchanging a seed node with a new node which has an influence gain of more than 1% of the total influence. UBI has no approximation assurance due to lack of interchange threshold values and accurate influence computation.
   The authors of Ohsaka et al. (2016) proposed an indexing method for DIM problem under IC model. This approach uses the RR sets instead of snapshots to consider evolving graph. First, it selects the RR sets from G and constructs an indexing structure on these RR sets. Then, it performs re-sampling by deleting and adding nodes from RR sets based on two basic operations SHRINK and EXPAND respectively. Next, it performs set maintenance by sampling any edge or node from a set randomly, and recomputes the sample size to manage these samples. Finally, the algorithm selects the seed sets from these dynamic maintained RR sets like TIM. This algorithm can be aligned with other diffusion models like LTM, TRM, etc. Meng et al. (2019) proposed a snapshot sampling-based approach $T \times$ *oneHope* under DIM settings. This approach considers T snapshots $\{G_1, G_2, \ldots, G_T\}$ over a period $[t_0, t_0 + \Delta t]$ for network evolution, where $T = \Delta t/\omega$. They assume that the nodes in each graph G_i are unchanged but the edges can be added or deleted over a period of time. They incorporated dynamic ICM model for information diffusion. Finally, they proposed a hop-based approach including recursive formula of activation probability for IM problem under dynamic settings. They pointed out that the number of identical users in the seed sets are increases if the consecutive snapshots are more similar.
   - *Other DIM strategies.* In (Lei et al., 2015; Tong et al., 2017) the authors have discussed different dynamics of the networks like uncertainty and incompleteness of the diffusion process. The authors of Lei et al. (2015) considered a scenario wherein the influence probabilities were unknown at the beginning and could be estimated after trails. The authors presented a learning-based algorithm to acquire the propagation probabilities along with diffusion process. They adopted ExploreExploit methods for maximizing the influence spread under DIM settings. Tong et al. (2017) proposed an adoptive greedy approach for DIM problem by considering the propagation probabilities as random variables align with a distribution. The authors of Wang et al. (2017c) stud-

ied the IM problem over stream data. They used the sliding window model to define the users influence. In order to continuously detect the seed set, they proposed stream influence maximization query.

6. **Conformity.** Most of the IM studies consider the individual's capability to influence the others while ignores the individual's propensity to be influenced, known as conformity. Therefore, in order to improve the effectiveness of the seed nodes, some efforts were performed by the authors in (Li et al., 2013a; Li et al., 2014b; Li et al., 2018c; Li et al., 2011a; Tang et al., 2013; Zhang et al., 2014b). Li et al. (2013a) were first to incorporate conformity value to identify the seed users and proposed an algorithm named as CINEMA. This work utilize the community-based framework to reduce the search space of influence spread computation under conformity-aware cascade diffusion process. The seed selection process utilize the greedy procedure to compute the marginal gain of each node in each iteration into respective communities and store it for future references. Li et al. (2014b) further improved the performance of CINEMA by incorporating context-specific conformity using CASINO algorithm. The authors of Li et al. (2018c) considered group norm to adopt conformity based on user profiles. They proposed a group-based influence maximization algorithm using several types of user behaviors based on individual profiles and group profiling, viz., GIM.
7. **Semantic.** Most of previous works ignore user's semantic information like social tags to identify the seed users and lack model generalization. Inspired by user semantic, Chen et al. (2020) introduced a new context-aware problem known as semantic-aware influence maximization (SIM). They proved that under traditional model, SIM is a NP-hard problem. They proposed a sampling-based method GRIS-SIM using reverse influence set to find the seed users. The proposed algorithm GRIS-SIM guarantees $1 - 1/e - \epsilon$ approximate solution. Three different sampling methods were also introduced for GRIS-SIM based on distinguished semantic settings. Furthermore, they had generalized the proposed algorithm to solve the distance-aware IM problem.

3.7. Other extensions

In the recent years, there have been various extensions to the classical IM problems. Some of the extensions are: extension of the the model to consider spreading of more than one information, targeting the influence when the spread is only among the specific nodes, integrating advertisers and network service providers perspective, etc. These extensions lead to new questions, challenges and paves the path as well as provides the perspective of future research.

1. **Profit Maximization** Most of the existing IM approaches neglect the distinction between actual product adoption and social influence. Based on these distinctions, a new extension of profit maximization was introduced (Tang et al., 2018; Wei and Lakshmanan, 2012; Tang et al., 2017; Weersink and Fulton, 2020; Li et al., 2017b). Wei and Lakshmanan (2012) adopted the classical ICM and LTM diffusion models to explicitly distinct between product adoption and social influence by considering influenced and adopting different states. The authors incorporate price and valuation of a product for decision making of the adoption. They proved that the expected profit function is sub-modular under extended diffusion models, but no longer exhibits monotonicity. They utilize greedy framework to compute expected influence spread. They proposed a novel algorithm PAGE to assigns prices based on user's profit potential. The authors of Tang et al. (2017) proposed a profit maximization approach based on online social network providers perspective. They define a profit metric $\phi(S)$ based on the combination of benefit of influence spread $\beta(S)$ and expense of influence propagation $\gamma(S)$, i.e. $\phi(S) = \beta(S) - \gamma(S)$. The profit metric is sub-modular, but no longer monotonic under traditional diffusion models. Tang et al. (2018) adopted the same idea as in Tang et al. (2017) and presented Deterministic Double Greedy algorithm to maximize the profit based on advertisers perspective.
2. **Revenue Maximization.** Recently, viral marketing techniques were influenced by the buyer's behavior, advertisers, network service provider's perspective and introduced a new extension of IM known as revenue maximization in social networks (Han et al., 2018; Khan et al., 2016; Aslay et al., 2016; Babaei et al., 2012). Khan et al. (2016) introduced a novel problem of revenue maximization of online social network provider that provides marketing platform to multiple competitive advertisers. They adopted classical IC and LT models to host's revenue maximization under competitive product campaigners. The authors showed that the objective function $\sigma(S)$ is neither sub-modular nor monotonic under these diffusion models. They proposed a novel scalable and efficient heuristic method to maximize the revenue of a social network host. The experimental results showed that the proposed heuristic method's performance was 5–10% better than the proposed greedy algorithm. The authors of Han et al. (2018) proposed a novel approach to maximize the revenue of a seller by providing discount to the potential users instead of giving free samples of a product. An intelligent seller can take advantage of network diffusion as a powerful means for advertisement to further increase the revenue of marketing. They proposed two revenue maximization approaches based on hill climbing greedy and local search procedures. They showed that the hill climbing greedy algorithm improves the influence spread of local search algorithm by 15% and 11% under monotone and non-monotone diffusion models respectively.

3.8. Discussion

- **Simulation-based Approaches.** These approaches were developed to improve the efficiency of the greedy algorithm. These approaches use MC simulations as black-box, i.e., model generality is preserved but prevents performance improvement by utilizing diffusion model's properties. SA (Jiang et al., 2011) is an exception among all simulation based approaches, as it does not ensure any approximation guarantee. This is because it uses simulated annealing meta-heuristic to explore and search the seed nodes in the network. SA might get stuck in the local optima as it does not provide theoretical guarantee. This algorithm could perform slightly better than other greedy approaches in terms of influence spread with less running time.
- **Heuristic-based Approaches.** To avoid time-consuming MC simulations, these approaches performed scoring procedure based on rank refinement and model reduction. The rank refinement approaches estimates the influence spread efficiently by transforming IM problem to some easier problems, like PageRank (Page et al., 1999), GROUP-PR (Liu et al., 2014, etc. These problems were not closely related to viral marketing scenario, although they are computationally efficient. However, these approaches ignore the diffusion model properties in the ranking process. Therefore, some model reduction heuristics were introduced to account for the properties of the diffusion models. Model reduction heuristics are directly inherited from classical diffusion models and and these model properties were used to compute the influence spread of the seed nodes. These

approaches could achieve the same approximate influence spread as simulation based IM algorithms in most of the cases. However, these approaches cannot maintain a trade-off between influence spread and efficiency when the influence range of nodes and number of influence paths are large. In addition, these approaches were not model generic, i.e., it cannot be generalized to other models.

- **Sampling-based Approaches.** To avoid rerunning of MC simulations, these approaches performed sampling in the graph based on snapshots (forward influence) and reverse reachable sets (backward influence). Most of the snapshot based approaches were presented for ICM, although they can be applied and extended to other models like LTM,TRM, and CATM. This is because these diffusion models are node-independent. These snapshot based methods performed significantly better than the simulation based approaches regarding efficiency with approximation guarantee. However, the theoretical complexity is still an issue in the large-scale network. In general, the backward influence approaches are much faster than forward influence approaches. This is because the snapshots were developed by examining the whole graph while reverse reachable sets were constructed by visiting only those nodes who can activate random sampled nodes. Hence, theoretical complexity of backward influence methods are significantly better than snapshot based approaches.

Tables 4 and 5 compare the characteristics of the existing IM algorithms (Singh et al., 2019f; Li et al., 2018d). Table 4 presents the theoretical analysis such as time-complexity, approximation ratio, problem solving perspective, state-of-the-art and base algorithms, etc., of the existing IM approaches. The state-of-the-art and base algorithm denotes the compared algorithm and core method used in the corresponding algorithm. Table 5 discusses the category, diffusion models and type of the network of the corresponding algorithm. Tables 6 and 7 compare the characteristics of existing the context-aware IM algorithms (Li et al., 2018d; Singh et al., 2019f). Table 6 provides the contextual-category, name of the algorithm, approximation ratio, problem solving perspective, base and the state-of-the-art algorithms. Table 7 provides the algorithmic-category, diffusion models, and type of the network where the corresponding algorithm can be applied. Fig. 11 summarizes the distribution of IM algorithms related to different factors like network applicability, diffusion models, and problem solving perspective from Tables 4–7.

4. Performance analysis

This section discusses the performance metrics used in the evaluation of the IM algorithms. In literature, four primary performance metrics are present: Quality, Efficiency, Scalability, and Robustness.

1. **Quality: Influence Spread.** It is used to equate the number of product adoption in the network by algorithm, with given seed set $S, |S| = k$. In General, the influence spread grows with $k$although there is a possibility of minor fluctuations. The Recent literature survey shows that, some context-aware IM algorithms (Singh et al., 2019; Lee and Chung, 2015; Tejaswi et al., 2017; Barbieri et al., 2012; Zhuang et al., 2013; Wang et al., 2017a; Wang, 2016; Tong et al., 2017; Bozorgi et al., 2017) are introduced to improve the quality of seed nodes.
2. **Efficiency: Running Time.** It is used to measure the running time of the IM algorithm. An algorithm can produce the desired result in a specific time, i.e., seed set S in efficient time. The running time grows with $k$with some exceptions like IMM (Tang et al., 2015) and TIM (Tang et al., 2014) are also presented.
3. **Scalability: Running Time and Memory Consumption.** It is measured in terms of both running time and memory consumption in an IM algorithm. Few algorithms act as a savior like highly efficient self-avoiding random walks generation on GPU(s), algorithms (Li et al., 2019; Nguyen et al., 2019) un several orders of magnitude faster than its CPU's and outperforms the state-of-art methods. Billions of edges can work on bigger networks by stretching the data across multiple GPUs in just a few seconds.
4. **Robustness.** A robust algorithm is one in which, on a slight change in the information propagation process, the optimal solution does not change much. In literature, the robustness evaluation is inadequate because it only focuses on the performance of IM algorithms on different social networks with various structures. Also, it is important to evaluate IM algorithms for various influence probability settings because influence probabilities are key components of influence networks. They can directly affect the performance of IM algorithms. There are some works (Jung et al., 2012; Galhotra et al., 2016; He and Kempe, abs/1602.05240, 2016.; Mehmood et al., 2016) have been done to ensure the robustness of the state-of-the-art methods.

Fig. 12 shows the frequency distribution of the existing IM algorithms over different performance parameters likes efficiency, quality, memory, and robustness. It also helps to identify the research gaps from performance perspective. Table 8 concludes the performance of the existing state-of-the-art IM algorithm and context-aware IM algorithms with respect to performance metrics discussed above and also tabulate some observation based on it.

- None of the existing work can attain all performance measures at the same time. Therefore, a careful selection of algorithms is needed based on the required application.
- Most of the simulation-based algorithms have lower time efficiency and are not scalable to large-scale networks. This is because these algorithms use time-consuming MC simulations. All the simulation-based algorithms have a good quality seed set. The memory consumption of these algorithms is not much high because only the need to store sampled possible world and marginal gain of nodes.
- The time efficiency of heuristics-based algorithms is much higher than simulation-based and much lower than sampling-based algorithms. This is because the score estimation process needs to scan the network several times. The seed quality has no theoretical guarantee on the approximation ratios of the score estimation algorithms, the quality of results is generally high in practice. These algorithms only need to store the information-related propagation paths and scores of nodes. Therefore, memory footprints are the lowest among the existing algorithms. These methods have a low memory footprint and can produce high-quality results when good score estimation functions are used.
- Sampling-based algorithms have much higher time efficiency than simulation-based algorithms. The running time of the sampling-based sample size determines algorithms. These methods are sensitive to influence probabilities and usually have high memory overheads.
- There are very few robust algorithms like IRIE, EASYIM are present in the literature as shown in Fig. 12. Both IRIE and EASYIM are insensitive to influence probabilities. This is because the influence score estimation process only performs arithmetic computations, and the computation time is independent of influence probability values.

Table 4
Comparison of the Existing IM Algorithms and their Characteristics – I.

Algorithm	Time Complexity	Approximation	Problem Solving Perspective	State-of-the-art Algorithms	Base Algorithm
Greedy (Kempe et al., 2003)	$O(kNMI)$	$1-1/e-\epsilon$	Spread Simulation	MaxDegree, Central & Random	–
Knapsack Greedy (Sviridenko, 2004)	$O(N^5)$	$1-1/e-\epsilon$	Spread Simulation	–	Greedy
Diffusion Degree (Kundu et al., 2011)	$O(N+M)$	N.A.	Centrality Based	DD & High Degree	High Degree
CELF (Leskovec et al., 2007)	$O(kNMI)$	$1-1/e-\epsilon$	Sub-modularity	Greedy	Greedy
NewGreedy (Chen et al., 2009)	$O(kIM)$	$1-1/e-\epsilon(r)$	Snapshots	CELF, Greedy & Random	High Degree
MIA/ PMIA (Chen et al., 2010e)	$O(Nt_{i\theta}+kn_{o\theta}n_{i\theta}(n_{i\theta}+\log N))$	$1-1/e$	Influence Path	Greedy, Random, DD & PageRank	SP1M
LDAG (Chen et al., 2010c)	$O(Nt_{\theta}+kn_{\theta}m_{\theta}(m_{\theta}+\log N))$	N.A.	Score Estimation	Greedy,SPIN, DD & PageRank	–
CELF++ (Goyal et al., 2011)	$O(kNMI)$	$1-1/e-\epsilon$	Sub-modularity	CELF	CELF
SA (Jiang et al., 2011)	$O(TIM)$	N.A	Spread Simulation	NewGreedy, CGA, & DD	Greedy
SIMPATH (Goyal et al., 2011)	$O(klNP_{\theta})$	N.A.	Score Estimation	High Degree, CELF & PageRank	LDAG
IPA (Kim et al., 2013)	$O(\frac{NO_v n_{vu}}{c}+k^2(\frac{O_v n_{vu}}{c}+(c-1)))$	N.A.	Influence Path	Greedy, DD & Random	PMIA
StaticGreedy (Cheng et al., 2013)	$O(\frac{kMN^2\log\binom{N}{k}}{\epsilon^2})$	$1-1/e-\epsilon$	Snapshots	CELF, SP1M, DD & High Degree	PMIA
PRUNEDMC (Ohsaka et al., 2014)	$O(\frac{kMN^2\log\binom{N}{k}}{\epsilon^2})$	$1-1/e-\epsilon$	Snapshots	IRIE, Random, PMIA & Degree	Greedy
GROUP-PR (Liu et al., 2014)	$O(kMN)$	N.A.	Influence Ranking	CELF, IRIE,DD & PMIA	PageRank
RIS (Borgs et al., 2014)	$O(\frac{k(N+M)\log^2 N)}{\epsilon^2}$	$1-1/e-\epsilon$	Reverse Reachability	–	–
IMRANK (Cheng et al., 2014)	$O(NTd_{max}\log d_{max})$	N.A.	Rank Refinement	PMIA & IRIE	–
SKIM (Cohen et al., 2014)	$O(\frac{kN^2M\log\binom{N}{k})}{\epsilon^2})$	$1-1/e-\epsilon$	Reverse Reachability	TIM	–
IMM (Tang et al., 2015)	$O(\frac{(k+l)(N+M)\log N)}{\epsilon^2}$	$1-1/e-\epsilon$	Reverse Reachability	TIM, TIM+ & IRIE	–
UBLF (Zhou et al., 2015a)	$O(kINM)$	$1-1/e-\epsilon(I)$	Spread Simulation	Greedy & CELF	Greedy
TIM (Tang et al., 2014)	$O(\frac{k(M+N)\log N}{\epsilon^2})$	$1-1/e-\epsilon$	Reverse Reachability	CELF++, IRIE & SIMPATH	–
Degree Discount (Chen et al., 2009)	$O(k\log N+M)$	N.A.	Heuristic based	CELF, Greedy & Random	High Degree
EASYIM (Galhotra et al., 2016)	$O(kD(N+M))$	N.A.	Influence Ranking	SIMPATH, CELF++ & IRIE	Greedy
SSA/D-SSA (Nguyen et al., 2016)	–	N.A.	Reverse Reachability	IMM & TIM+	RIS
BKRIS (Wang et al., 2017a)	$O(\frac{NM(\log N+\log\binom{N}{k})}{\epsilon^2})$	$1-\frac{1}{e}-\epsilon-\epsilon\prime$	Reverse Reachability	RIS	IMM
SP1M (Kimura and Saito, 2006)	$O(kNM)$	$1-1/e$	Influence Path	Degree, PageRank & Closeness	–
DPSO (Gong et al., 2016; Wang et al., 2017b)	$O(k^2\log kn\bar{D}^2)$	N.A.	Swam Optimization	Degree, CELF++ & SAEDV	–
IRIE (Jung et al., 2012)	$O(k(n_{o\theta}k+M))$	N.A.	Score Estimation	Greedy & PMIA	–
LAIM (Ge et al., 2017)	–	$1-1/e-\epsilon$	Learning Based	Degree, CELF & Random	Greedy
TW Greedy (Wang and Feng, 2009)	$O(kNMI)$	$1-1/e-\epsilon$	Spread Simulation	SCG, KKG & High Degree	Greedy
Cost-Degree (Yang et al., 2018)	$O(M)$	N.A	Score Estimation	Random	–
LAPSO-IM (Singh et al., 2019d)	$O(Ink(\log k+\bar{D}^2)+nkN)$	N.A	Swam Optimization	DD, CELF++ & DPSO	PSO
CGA (Wang et al., 2010)	$O(M+IM_C(N(Z-C)+k(C+N_C)))$	$1-e^{-1/(1+\delta_C)}$	Community Based	DD, MG & Random	–
IM-SSO (Singh et al., 2020b)	$O(n(k\bar{D}^2+n)+n\log n))$	N.A	Swam Optimization	DD, CELF++ & DPSO	SSO
ACO-IM (Singh et al., 2019a)	$O(I_{max}\|R_G\|\|V\|+\|V\|\log\|V\|)$	N.A	Swam Optimization	DD, CELF++ & DPSO	ACO
BP-Greedy (Saito et al., 2012)	–	$1-1/e$	Spread Estimation	Closeness, Betweenness, & Degree	Greedy
LCI (Zhang et al., 2016)	$O((N+M)N.d)$	N.A.	Sub-modularity	Greedy	Greedy
MPMN-CELF++ (Zhang, 2015)	$O(kNMI)$	N.A.	Spread Simulation	CELF++ & SIMPATH	CELF++
MPMN-SIMPATH (Zhang, 2015)	$O(klNP_{\theta})$	N.A.	Influence Ranking	CELF++ & SIMPATH	SIMPATH++
ASMTC (Wang et al., 2016b)	$O(\|V^s\|^2+\|V^s\|)$	N.A.	Reverse Reachability	–	–
SeedSelection-M (Erlandsson et al., 2017)	–	N.A.	Rank Refinement	VoteRank, Degree, & K-Shell	–
MIM-Greedy (Sun et al., 2016)	$O(kmNMI)$	$1-1/e$	Spread Simulation	MaxDegree, Init-First & Random	Greedy
MIM2 (Singh et al., 2019f)	$O((l+m)(N+M)+(k+m)(N\log N+M))$	N.A	Rank Refinement	DD, MaxDegree, CELF & Random	MIM-Greedy

Table 5
Comparison of the Existing IM Algorithms and their Characteristics – II.

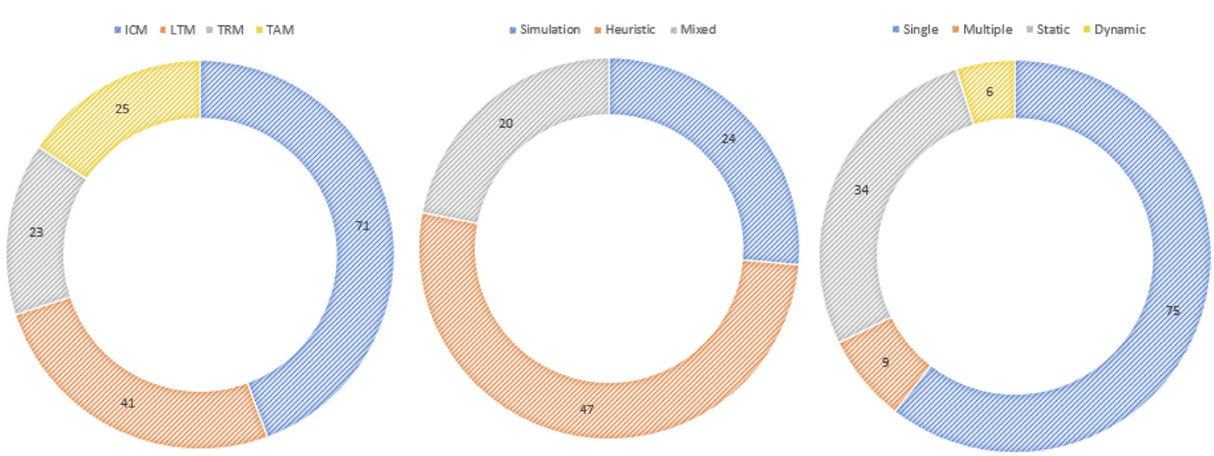
Framework	Algorithm	Diffusion Model				Category				Network	
		Linear Threshold	Independent ascade	Triggering	Continuous Time-aware	Simulation	Heuristic	Meta-heuristic	Mixed	Single	Multiple
IM	Greedy (Kempe et al., 2003)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗
	Knapsack Greedy (Sviridenko, 2004)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗
	SP1M (Kimura and Saito, 2006)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	CELF (Leskovec et al., 2007)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗
	Degree Discount (Chen et al., 2009)	✓	✓	✓	✓	✗	✓	✗	✗	✓	✗
	NewGreedy (Chen et al., 2009)	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗
	TW Greedy (Wang and Feng, 2009)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗
	MIA/ PMIA (Chen et al., 2010e)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	LDAG (Chen et al., 2010c)	✓	✗	✗	✗	✗	✓	✗	✗	✓	✗
	CGA (Wang et al., 2010)	✗	✓	✗	✗	✓	✗	✗	✗	✓	✗
	CELF++ (Goyal et al., 2011)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗
	SA (Jiang et al., 2011)	✗	✓	✗	✗	✓	✗	✗	✗	✓	✗
	Diffusion Degree (Kundu et al., 2011)	✓	✓	✓	✓	✗	✓	✗	✗	✓	✗
	SIMPATH (Goyal et al., 2011)	✓	✗	✗	✗	✗	✓	✗	✗	✓	✗
	IRIE (Jung et al., 2012)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	IPA (Kim et al., 2013)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	StaticGreedy (Cheng et al., 2013)	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗
	PRUNEDMC (Ohsaka et al., 2014)	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗
	TIM (Tang et al., 2014)	✓	✓	✓	✗	✗	✗	✗	✓	✓	✗
	GROUP-PR (Liu et al., 2014)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	RIS (Borgs et al., 2014)	✓	✓	✓	✓	✗	✗	✗	✓	✓	✗
	IMRANK (Cheng et al., 2014)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	SKIM(Cohen et al., 2014)	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗
	IMM (Tang et al., 2015)	✓	✓	✓	✓	✗	✗	✗	✓	✓	✗
	UBLF (Zhou et al., 2015a)	✓	✓	✗	✗	✓	✗	✗	✗	✓	✗
	EASYIM (Galhotra et al., 2016)	✓	✓	✗	✗	✗	✓	✗	✗	✓	✗
	SSA/D-SSA (Nguyen et al., 2016)	✓	✓	✓	✓	✗	✗	✗	✓	✓	✗
	BKRIS (Wang et al., 2017a)	✓	✓	✗	✗	✗	✗	✗	✓	✓	✗
	DPSO (Gong et al., 2016; Wang et al., 2017b)	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗
	LAIM (Ge et al., 2017)	✓	✓	✓	✓	✓	✗	✗	✗	✓	✗
	Cost-Degree (Yang et al., 2018)	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
	LAPSO-IM ((Singh et al., 2019d)	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗
	IM-SSO (Singh et al., 2020b)	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗
	ACO-IM (Singh et al., 2019a)	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗
	DSFLA (Tang et al., 2020)	✓	✓	✓	✓	✗	✗	✓	✗	✓	✗
IM2	BP-Greedy (Saito et al., 2012)	✓	✓	✓	✗	✓	✗	✗	✗	✗	✓
	LCI (Zhang et al., 2016)	✓	✓	✗	✗	✓	✗	✗	✗	✗	✓
	MPMN-CELF++ (Zhang, 2015)	✓	✓	✓	✓	✓	✗	✗	✗	✗	✓
	MPMN-SIMPATH (Zhang, 2015)	✓	✗	✗	✗	✗	✓	✗	✗	✗	✓
	ASMTC Wang et al., 2016b)	✗	✓	✗	✗	✗	✗	✗	✓	✗	✓
	SeedSelection-M (Erlandsson et al., 2017)	✗	✓	✗	✗	✗	✓	✗	✗	✗	✓
	C-IM2 (Singh et al., 2019b)	✓	✓	✗	✗	✗	✓	✗	✗	✗	✓
MIM	MIM-Greedy (Sun et al., 2016)	✓	✓	✗	✗	✓	✗	✗	✗	✓	✗
MIM2	MIM2 (Singh et al., 2019f)	✓	✓	✓	✓	✗	✓	✗	✗	✗	✓

Table 6
Comparison of the Existing Context-aware IM Algorithms and their Characteristics – I.

Categories	Algorithm	Approximation	Problem Solving Perspective	State-of-the-art Algorithms	Base Algorithm
Location	Expansion/Hint (Li et al., 2014a)	$\epsilon(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	PMIA,IRIE,Assembly,Bound	–
	TPH (Zhou et al., 2015b)	N.A	Rank Refinement	MaxDegree,DD,PageRank	–
	Target-IM/Target-IM+ (Song et al., 2016)	$1-1/e-\epsilon$	Reverse Reachable Sets	MIA-L,Expansion,IMM	MIA
	MIA-DA/RIS-DA (Wang et al., 2017b)	$(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	PMIA,MIA/RIS-DA	MIA,RIS
	TOA/TORA (Sen et al., 2018)	$(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	DD,LIA,CELF,PMIA	WRIS,PMIA
Topical	AIR-Greedy (Barbieri et al., 2012)	$1-1/e-\phi$	Spread Simulation	TIC	Greedy
	LGA/ELGA (Guo et al., 2013)	N.A	Model Reduction Heuristic	LDegree,LRandom,LND	–
	KB-TIM (Li et al., 2015c)	$1-1/e-\epsilon$	Reverse Reachable Sets	RR,IRR,WRIS	RIS,TIM
	INFLEX (Aslay et al., 2014)	N.A	Rank Refinement	exactKNN,approxKNN, approxAD	–
	TIM/MIA (Shuo Chen et al., 2015b)	$\epsilon(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	PMIA,INFLEX,MIS	MIA
	IMIP/IMAX (Lee and Chung, 2015)	$1-1/e$	Spread Simulation	CELF,PMIA,IRIE,CD	–
	MIS/BTS (Chen et al., 2015)	N.A	Model Reduction Heuristic	TA-PMIA/Greedy/PageRank	Greedy,PMIA
	CTVM (Nguyen et al., 2016)	$1-1/e-\epsilon$	Reverse Reachable Sets	TIM,CELF++,SIMPATH	BIM
	PITEX (Li et al., 2017a)	$(1-\epsilon)/(1+\epsilon)$	Reverse Reachable Sets	RR,MC,INDEXEST	TIM
	In-out Discounting (Tejaswi et al., 2017)	N.A	Spread Simulation	In-Out, MaxDegee, DD	–
	C2IM (Singh et al., 2019)	N.A	Reverse Reachable Sets	Random, MaxDegree,CIM	CIM,TIM(Shuo Chen et al., 2015a)
	MFIP (Li et al., 2020)	N.A	Approximate Scoring	Greedy,DD,TIM	TIM
Time	MIA-M/MIA-C (Chen et al., 2012)	$(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	Greedy,MaxDegree	MIA
	CT-IPA (Lee et al., 2012)	$(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	Greedy,MaxDegree, Random,IPA	IPA
	MISP (Liu et al., 2014)	$(1-1/e)$ wrt. MIA structure	Model Reduction Heuristic	Random,DC,PMA,ISP	MC
	INFLUMAX (Gomez Rodriguez et al., 2012)	$(1-1/e)$	Markov Chain	Greedy,PMIA,Random, SP1M	–
	FASTMARGIN (Xie et al., 2015)	N.A	Markov Chain	CELF/FastMargin-Static/ Dynamic	CELF
	IMM (Tang et al., 2015)	$1-1/e-\epsilon$	Reverse Reachable Sets	TIM,TIM+,SIMPATH	TIM
	TSDEG/TSREEDY (Mohammadi et al., 2015)	$1-1/e$	Rank Refinement & Simulation	Greedy,Random,BET, MaxDegree	Greedy
Competitive	EIL (Ceren Budak et al., 2011)	$1-1/e-\epsilon$	Spread Simulation	Greedy,Degree, EarlyInfectees	Greedy
	IBM/CLDAG (He et al., 2012)	N.A	Model Reduction Heuristic	Greedy,Degree,Random	LDAG
	RR-SIM/RR-SIM+ (Wei et al., 2015)	$\alpha(1-1/e-\epsilon)$	Reverse Reachable Sets	PageRank,MaxDegree, Random	–
	MinSeed (Zhu et al., 2016)	$1-1/e-\epsilon$	Spread Simulation	LS-Greedy	Greedy
	CI2 (Bozorgi et al., 2017)	N.A	Snapshot (Clustering) Sampling	Greedy,MaxDegree,INCIM, IPA	INCIM
Dynamic	MaxG (Zhuang et al., 2013)	N.A	Rank Refinement	Rand,Deg,Enum,DegRR	Probing
	IGA (Wang, 2016)	N.A	Spread Simulation	Random,MaxDegee,HT	Greedy
	IndexingMethod (Ohsaka et al., 2016)	$1-1/e-\epsilon$	Reverse Reachable Sets	TIM,IMM,PMC,IRIE	RIS
	UBI/UBI+ (Song et al., 2017)	N.A	Rank Refinement	MaxDegree,IMM,IRIE	Greedy
	DIM/Opt-DIM (Wang et al., 2017a)	N.A	Model Reduction	LDAG,SIMPATH	–
	A-Greedy/H-Greedy (Tong et al., 2017)	N.A	Spread Simulation	Random, Greedy	Greedy
Conformity	CASINO (Li et al., 2011a)	N.A	Approximate Scoring	–	–
	CINEMA (Li et al., 2013a)	N.A	Spread Simulation	MIA, LDAG, SIMPATH, MixGreedy	CASINO
	CINEMA-C^3 (Li et al., 2014b)	N.A	Spread Simulation	DD, PMIA, TIC, Greedy	CINEMA
	GIM (Li et al., 2018c)	N.A	Rank Refinement	DD, CELF++	–
Semantic	GRIS-SIM (Chen et al., 2020)	$1-1/e-\epsilon$	Reverse Reachable Sets	SIMPATH, IMM, BWR, LDD	KB-TIM

Table 7
Comparison of the Existing Context-aware IM Algorithms and their Characteristics – II.

Algorithm	Diffusion Model (IDM)			Category			Network			
	ICM	LTM	TM	Simulation	Heuristic	Mixed	Single	Multiple	Static	Dynamic
Expansion/Hint (Li et al., 2014a)	✓	✗	✗)	✗	✓	✗	✓	✗	✓	✗
TPH (Zhou et al., 2015b)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
Target-IM/Target-IM+ (Song et al., 2016)	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗
MIA-DA/RIS-DA (Wang et al., 2017b)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
TOA/TORA (Sen et al., 2018)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
AIR-Greedy (Barbieri et al., 2012)	✓	✓	✗	✓	✗	✗	✓	✗	✓	✗
LGA/ELGA (Guo et al., 2013)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
KB-TIM (Li et al., 2015c)	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗
INFLEX (Aslay et al., 2014)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
TIM/MIA (Shuo Chen et al., 2015b)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
IMIP/IMAX (Lee and Chung, 2015)	Expectation IDM			✓	✗	✗	✓	✗	✓	✗
MIS/BTS (Chen et al., 2015)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
CTVM (Nguyen et al., 2016)	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗
PITEX (Li et al., 2017a)	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗
In-out Discounting (Tejaswi et al., 2017)	Target Adoption IDM			✓	✗	✗	✓	✗	✓	✗
C2IM (Singh et al., 2019)	✓	✓	✗	✗	✗	✓	✓	✗	✓	✗
MFIP (Li et al., 2020)	✓	✓	✗	✗	✓	✗	✓	✗	✓	✗
MIA-M/MIA-C (Chen et al., 2012)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
CT-IPA (Lee et al., 2012)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
MISP (Liu et al., 2014)	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗
INFLUMAX (Gomez Rodriguez et al., 2012)	✗	✗	✓	✗	✓	✗	✓	✗	✓	✗
FASTMARGIN (Xie et al., 2015)	DynaDiffuse IDM			✗	✓	✗	✓	✗	✓	✗
IMM (Tang et al., 2015)	✗	✗	✓	✗	✗	✓	✓	✗	✓	✗
TSDEG/TSDREEDY (Mohammadi et al., 2015)	✓	✓	✗	✓	✗	✗	✓	✗	✓	✗
EIL (Ceren Budak et al., 2011)	✓	✗	✗	✓	✗	✗	✓	✗	✓	✗
IBM/CLDAG (He et al., 2012)	✗	✓	✗	✗	✓	✗	✓	✗	✓	✗
RR-SIM/RR-SIM+ (Wei et al., 2015)	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗
MinSeed (Zhu et al., 2016)	✓	✗	✗	✓	✗	✗	✓	✗	✓	✗
CI2 (Bozorgi et al., 2017)	Decidable Competitive DM			✗	✗	✓	✓	✗	✓	✗
MaxG (Zhuang et al., 2013)	✗	✗	✓	✗	✓	✗	✓	✗	✗	✓
IGA (Wang, 2016)	✓	✓	✗	✓	✗	✗	✓	✗	✗	✓
IndexingMethod (Ohsaka et al., 2016)	✓	✗	✗	✗	✗	✓	✓	✗	✗	✓
UBI/UBI+ (Song et al., 2017)	✓	✗	✗	✗	✓	✗	✓	✗	✗	✓
DIM/Opt-DIM (Wang et al., 2017a)	✗	✓	✗	✗	✓	✗	✓	✗	✗	✓
A-Greedy/H-Greedy (Tong et al., 2017)	✓	✗	✗	✓	✗	✗	✓	✗	✗	✓
CASINO (Li et al., 2011a)	✓	✓	✗	✗	✓	✗	✓	✗	✓	✗
CINEMA (Li et al., 2013a)	✓	✓	✗	✓	✗	✗	✓	✗	✓	✗
CINEMA-C^3 (Li et al., 2014b)	✓	✓	✗	✓	✗	✗	✓	✗	✓	✗
GIM (Li et al., 2018c)	✓	✓	✗	✗	✓	✗	✓	✗	✓	✗
GRIS-SIM (Chen et al., 2020)	✓	✓	✗	✗	✗	✓	✓	✗	✓	✗



(a) Information Diffusion Models (b) Problem Solving Perspective (c) Network Applicability

Fig. 11. The frequency distribution of the existing IM algorithms over various factors.

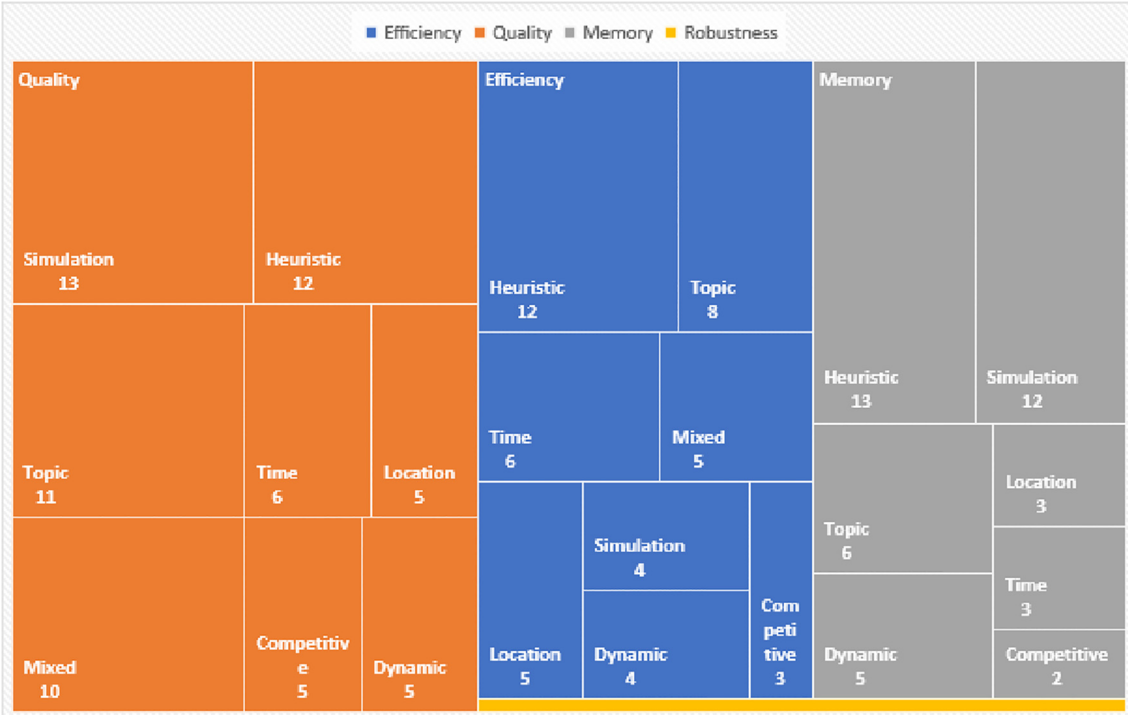


Fig. 12. The performance summarising of the existing IM algorithms over various performance metrics.

5. Research challenges and future directions

In order to discuss research challenges and future directions of the IM problem, we categories these challenges based on the framework as depicted in Table 9. Now, we explain research challenges that arise in each category along with future directions.

5.1. Challenges under classical IM framework

1. **Incorporating Stability.** The instability of IM algorithms is stated in He et al. (2014) when influence probabilities are noisy. Thus, a minor change in the propagation model might lead to a drastic change in optimal seed users. There are some approaches presented to robust IM problems like; IRIE (Jung et al., 2012), EASYIM (Galhotra et al., 2016), Saturate Greedy (He and Kempe, abs/1602.05240, 2016.), and MAX-COVER (Mehmood et al., 2016) keeping the assumption that the structure of the network is fixed. Ideally, the structure of real-world networks changes continuously. Thus, it is a challenging task to identify a robust seed set from given limited graph changes.
2. **Switching from Strict Submodularity to Moderate Submodularity.** To develop theoretical bound IM solutions, the submodularity of the objective function is important. The submodularity property of an objective function is too strict in some scenarios, such as the opinion-aware IM approaches (Galhotra et al., 2016; Li et al., 2011b; Gionis et al., 2013) adopts non-submodularity. This is because nodes can switch their states between negative and positive opinions. Therefore, the greedy algorithm is not applicable in such circumstances. To tackle such circumstances, heuristic solutions can be proposed. A possible direction is to develop an objective function $\sigma(S)$ with a more general submodular function to obtain a better solution than heuristic solutions. The authors of Das and Kempe (2011) introduced a more general function, named as weakly submodular function. The weakly submodular function guarantees the theoretical approximation of the problem.
3. **Incorporating Group Norms.** Most of the existing approaches only focus on the social influence between a pair of individuals x and y with a connection (x, y). However, in the real world, users are influenced by their friends or acquaintances and group norms such as conformity. Users who have similar backgrounds, education, age, etc., in the group conform to each others using conformity behavior. The authors of (Li et al., 2011a; Tang et al., 2013; Zhang et al., 2014b) extract the conformity characteristic from the network and ignore user's profiles regarding the social group. In general, there are two possible future directions: incorporating user profiles with conformity characteristics and consideration of different types of conformity such as identification, compliance, and internalization, etc., in IM problems.

5.2. Challenges under IM2 framework

1. **Identification of Overlapping Users.** Most users in social networking sites create multiple accounts and spread influence to their friends in each network simultaneously. This allows individuals to propagate information across the networks. So, it is required to consider the overall impact of an individual over multiple networks to identify a more accurate and effective seed. Therefore, the identification of overlapping users plays a significant role in information propagation and seed selection. Because these users are responsible for a graph or network coupling, there are some efforts done in this direction (Buccafurri et al., 2012; Suganya et al., 2017; Vosecky et al., 2009). Since network data is unstructured, noisy, incomplete, and big, diversification of user's information over multiple networks leads to challenging tasks for identifying the overlapping users.
2. **Network Coupling and Incorporating Heterogeneous Diffusion Models.** In order to estimate the aggregate influence spread of a user under the IM2 framework, there is a need for a coupling method to form a multiplex network from multiple

Table 8
Comparison of the Performance Metrics of the Existing IM Algorithms.

Framework	Category	Algorithm	Time efficiency	Seed Quality	Memory Footprint	Robustness
		Greedy (Kempe et al., 2003)	✗	✓	✓	✗
		Knapsack Greedy (Sviridenko, 2004)	✗	✓	✓	✗
		CELF (Leskovec et al., 2007)	✗	✓	✓	✗
		CGA (Wang et al., 2010)	✓	✓	✓	✗
	Simulation	CELF++ (Goyal et al., 2011	✗	✓	✓	✗
		SA (Jiang et al., 2011)	✓	✓	✓	✗
		UBLF (Zhou et al., 2015a	✓	✓	✓	✗
		LAIM (Ge et al., 2017	✗	✓	✓	✗
		Cost-Degree (Yang et al., 2018)	✓	✓	✓	✗
		SP1M (Kimura and Saito, 2006	✓	✗	✓	✗
		Degree Discount (Chen et al., 2009)	✓	✓	✓	✗
		TW Greedy (Wang and Feng, 2009)	✗	✓	✓	✗
		LDAG (Chen et al., 2010c)	✓	✓	✗	✗
		Diffusion Degree (Kundu et al., 2011)	✓	✓	✓	✗
		SIMPATH (Goyal et al., 2011)	✓	✓	✓	✗
IM	Heuristic	IRIE (Jung et al., 2012)	✗	✗	✓	✓
		IPA (Kim et al., 2013)	✓	✗	✓	✗
		GROUP-PR (Liu et al., 2014)	✓	✓	✓	✗
		IMRANK (Cheng et al., 2014)	✓	✓	✓	✗
		EASYIM (Galhotra et al., 2016)	✗	✓	✓	✓
		DPSO (Gong et al., 2016; Wang et al., 2017b)	✓	✓	✓	✗
		NewGreedy (Chen et al., 2009)	✓	✓	✗	✗
		IMM (Tang et al., 2015)	✗	✓	✗	✗
		StaticGreedy (Cheng et al., 2013)	✗	✓	✗	✗
		PRUNEDMC (Ohsaka et al., 2014)	✗	✓	✗	✗
	Mixed	TIM (Tang et al., 2014)	✗	✓	✗	✗
		RIS (Borgs et al., 2014)	✗	✓	✗	✗
		SKIM (Cohen et al., 2014)	✓	✓	✗	✗
		SSA/D-SSA (Nguyen et al., 2016)	✓	✓	✗	✗
		BKRIS (Wang et al., 2017a)	✓	✓	✗	✗
		BP-Greedy (Saito et al., 2012)	✗	✓	✓	✗
	Simulation	LCI (Zhang et al., 2016)	✗	✓	✗	✗
		MPMN-CELF++(Zhang, 2015)	✗	✓	✓	✗
IM2	Heuristic	SeedSelection-M (Erlandsson et al., 2017)	✓	✓	✓	✗
		MPMN-SIMPATH (Zhang, 2015)	✓	✓	✓	✗
	Mixed	ASMTC (Wang et al., 2016b)	✓	✓	✗	✗
MIM	Simulation	MIM-Greedy (Singh et al., 2019f)	✗	✓	✓	✗
MIM2	Heuristic	MIM2 (Singh et al., 2019f)	✓	✓	✓	✗
		Expansion/Hint (Li et al., 2014a)	✓	✓	✓	✗
		TPH (Zhou et al., 2015b)	✓	✓	✓	✗
	Location	Target-IM/Target-IM+ (Song et al., 2016)	✓	✓	✗	✗
		MIA-DA/RIS-DA (Wang et al., 2017b)	✓	✓	✗	✗
		TOA/TORA (Sen et al., 2018)	✓	✓	✓	✗
		AIR-Greedy (Barbieri et al., 2012)	✗	✓	✓	✗
		LGA/ELGA (Guo et al., 2013)	✓	✓	✗	✗
		KB-TIM (Li et al., 2015c)	✓	✓	✗	✗
		INFLEX (Aslay et al., 2014)	✓	✓	✓	✗
		TIM/MIA (Shuo Chen et al., 2015b)	✓	✓	✓	✗
	Topical	IMIP/IMAX (Lee and Chung, 2015)	✗	✓	✓	✗
		MIS/BTS (Chen et al., 2015)	✓	✓	✓	✗
		CTVM (Nguyen et al., 2016	✓	✓	✗	✗
		PITEX (Li et al., 2017a)	✓	✓	✗	✗
		In-out Discounting (Tejaswi et al., 2017)	✗	✓	✓	✗
		C2IM (Singh et al., 2019)	✓	✓	✗	✗
Context-aware		MIA-M/MIA-C (Chen et al., 2012)	✓	✓	✗	✗
		CT-IPA (Lee et al., 2012)	✓	✓	✗	✗
		INFLUMAX (Gomez Rodriguez et al., 2012)	✓	✓	✓	✗
	Time	FASTMARGIN (Xie et al., 2015	✓	✓	✓	✗
		IMM (Tang et al., 2015)	✓	✓	✗	✗
		TSDEG/TSDREEDY (Mohammadi et al., 2015)	✓	✓	✓	✗
		EIL (Ceren Budak et al., 2011)	✗	✓	✓	✗
		IBM/CLDAG (He et al., 2012)	✓	✓	✗	✗
	Competitive	RR-SIM/RR-SIM+ (Wei et al., 2015)	✓	✓	✗	✗
		MinSeed (Zhu et al., 2016)	✗	✓	✓	✗
		CI2 (Bozorgi et al., 2017)	✓	✓	✗	✗
		MaxG (Zhuang et al., 2013)	✓	✓	✓	✗
		IGA (Wang, 2016	✗	✓	✓	✗
		IndexingMethod (Ohsaka et al., 2016)	✓	✓	✗	✗
	Dynamic	UBI/UBI+ (Song et al., 2017)	✓	✓	✓	✗
		DIM/Opt-DIM (Wang et al., 2017a	✓	✗	✓	✗
		A-Greedy/H-Greedy (Tong et al., 2017)	✗	✓	✓	✗

Table 9
Illustration of influence maximization challenges.

	Framework	Challenges	Representative
Research Challenges	Classical IM	Incorporating Stability	(Jung et al., 2012; Galhotra et al., 2016; He and Kempe, abs/1602.05240, 2016.; Mehmood et al., 2016)
		Weaker Submodularity	(Galhotra et al., 2016; Li et al., 2011b; Gionis et al., 2013; Das and Kempe, 2011)
		Adopting Conformity	(Li et al., 2011a; Li et al., 2013b; Tang et al., 2013; Zhang et al., 2014b)
	IM across Multiple Networks (IM2)	Overlapping Users Identification	(Buccafurri et al., 2012; Suganya et al., 2017; Vosecky et al., 2009)
		Network Coupling	(Zhang et al., 2016; Wang et al., 2016b; Erlandsson et al., 2017)
		Incorporating Heterogeneous Diffusion Models	(Li et al., 2012; Zhan et al., 2015a)
	Multiple Influence Maximization (MIM)	Budget Fixing	(Sun et al., 2016)
		Incorporating Product Influence	(Sun et al., 2016)
	Multiple IM across Multiple Networks	Budget Fixing	(Singh et al., 2019f)
		Overlapping Users Identification	(Buccafurri et al., 2012; Suganya et al., 2017; Vosecky et al., 2009)
		Network Coupling	(Singh et al., 2019f;Zhang et al., 2016,Wang et al., 2016b; Erlandsson et al., 2017)
		Incorporating Heterogeneous Diffusion Models	(Li et al., 2012; Zhan et al., 2015a)
		Incorporating Product Influence	(Sun et al., 2016; Singh et al., 2019f)
	Contextual IM	Adopting Contextual Features	(Sen et al., 2018; Lee and Chung, 2015; Bozorgi et al., 2017)
		Incorporating Uncertainty and Incompleteness	(Lei et al., 2015; Tong et al., 2017)

single networks. So that the existing single network algorithm can be directly applied to the multiplex system. There are already some work (Zhang et al., 2016; Wang et al., 2016b; Erlandsson et al., 2017) propose in this direction. All these works consider an assumption that each network has the same diffusion model. However, in the real world, these networks may have their diffusion model (Alan Kuhnle et al., 2018; Zhan et al., 2015b; Zhan et al., 2015c). Therefore, one possible future direction is to incorporate heterogeneous diffusion models in network coupling to estimate influence spread accurately.

3. **Incorporating Parallel Architecture.** In order to make IM algorithms computationally efficient, parallel architecture needs to be incorporated. As we studied, most state-of-art algorithms follow iterative execution to identify influential users and maximize overall aggregate influence. So, IM algorithms can be useful and efficient by adopting parallel architecture on GPU and multiple processors. Also, it can utilize distributed architecture by avoiding the serial nature of existing IM algorithms to achieve computational efficiency.

5.3. Challenges under MIM framework

1. **Budget Fixing: Number of Items Per Product.** Budget fixing becomes a vital challenge when in a viral marketing scenario, multiple non-competitive products need to be promoted simultaneously. As different individuals may have other interests in various products in the real-world. Therefore, to maximize the product adoption or profit, the marketing company needs to decide the number of items per product such that $\sum_{i=1}^{i=m}|P_i| = k$, where m and P_i denotes the number of products and product type. The authors of (Sun et al., 2016) provides a strategy to fix the budget and the number of items per product. However, they ignore the profit of individual products and consider that each product generates the same profit. Also, there is a possibility of left-out some products in advertising. So the incorporation of the product characteristics in fixing the budget of each product is one possible future direction.
2. **Incorporating Product Influence and Diffusion over Networks.** Each user has a different interest and influence over distinct products. In order to estimate the seed set under the MIM framework, we need to consider the influence probabilities of users for each product. Although some efforts have been made to incorporate product influence (Sun et al., 2016) by assuming the influence graph structure is fixed and ignore competitive and complementary IM framework. Therefore, one possible direction of research is to incorporate the above-discussed issues.

5.4. Challenges under MIM2 framework

To the best of our knowledge, the only work has been done (Singh et al., 2019f regarding the MIM2 framework. The authors of Singh et al. (2019f) introduced a new framework, MIM2 by handling multiple products and multiple social networks simultaneously. In order to handle the MIM2 problem, several new challenges arise (Singh et al., 2019f), given as follows.

1. Determining the budget of each product is a challenging task under MIM2. To maximize the product adoption or profit, the marketing company needs to decide the number of items per product such that $\sum_{i=1}^{i=m}|P_i| = k$, where m and P_i denotes the number of products and product type.
2. Overlapping users identification along with node-alignment across social networks and coupling these networks into single multiplex one. Accurate measurement of influence spread of non-overlapping and overlapping users in the multiplex network corresponding to every product.
3. Deciding biases of a product corresponding to specific individual and network. Also, which network is more suitable for influence propagation.
4. How to handle heterogeneous diffusion models in different networks.

5.5. Challenges under context-aware IM framework

1. **Considering Contextual Features.** There is a need to incorporate contextual features such as location and topic with classical IM problems for novel applications. Although some works have been done to incorporate contextual features (Sen et al., 2018; Lee and Chung, 2015; Bozorgi et al., 2017). However, context-

aware IM is still a less researched area, and many topics need to be explored. For example, time-aware information diffusion models CTM and TRM are still largely unexplored, with users' utilities varying with time periods, etc.

2. **Incorporating Uncertainty and Incompleteness of the Diffusion Process.** In this framework, there is a concern of uncertainty and incompleteness of the diffusion process. There is some effort already made in this direction (Lei et al., 2015; Tong et al., 2017). They ignore user's history and involvement to learn the influence probability of users and use random variables conforming to a specific distribution. One possible future direction is to consider the user's history to update influence probabilities iteratively.
3. **Incorporating Multiplex Networks and Heterogeneous Diffusion Models.** To the best of our knowledge, there is no work done considering contextual features along with multiplex network and heterogeneous diffusion models (refer column Multiple under the Network in Table 7). Therefore, one possible future direction is to incorporate a multiplex network and contextual features to generate a more effective seed.

6. Open problems

- *How to achieve trade-off between generality and feasibility of diffusion models?* In relation to it, some studies (Borodin et al., 2010a) suggest that models like the competitive threshold diffusion model may not be sub-modular and monotonic, leading to the inapplicability of greedy-based methods. There are some efforts (Gong et al., 2016; Wang et al., 2017b; Singh et al., 2019a) have been made to achieve a trade-off between generality and the feasibility of diffusion models using nature-inspired optimization techniques. A wide gap exists between the generality and feasibility of diffusion models because most of the classical IM approaches focus on the feasibility of diffusion models and avoid abstraction. Besides Zhan et al. (2015d), IM2 methods avoid heterogeneous behavior of diffusion models over different networks.
- *How to combine reality and models?* The matching degree between reality and ideal models needs to be strengthened (Razaque et al., 2019a). The development of diffusion models can be categorized into two classes, i.e., the first is an extension of traditional models. The second uses questionnaires to acquire data from the Internet to validate the rationality and correctness of models. Therefore, it is a difficult task to combine reality and models in order to devise new models. To achieve theoretical and practical aspects of the diffusion model is still challenging and remains much open to exploring.
- *How to acquire complexity of reality in information diffusion models?* The complexity of reality is not consistent with classical diffusion models. The social structure of an individual is replicated with his complex social structure. With this, it is quite possible that the state of nodes may change based on their relationships and social network. Moreover, the state of nodes is also affected by society, the environment, and other ignored factors. Therefore, it is necessary to investigate psychological, social, and political factors to achieve the complexity of reality.
- *How to deal with negative influence in signed social networks?* Some of the social networks also have a negative influence, which is complementary to a positive influence in information diffusion. Therefore, we need to incorporate negative relationships in-addition to positive relationships to compute accurate influence spread. However, most of the previous studies only consider positive influence between a pair of individuals and ignore the negative influence and the corresponding spread of negative opinion. Therefore, IM in signed social networks by considering positive as well as negative influence is still a challenging problem that remains much open to exploring. There are very few studies (Li et al., 2014c; Shen et al., 2015a; Shen et al., 2015b; Li et al., 2017c; Liang et al., 2019; He et al., 2019; Weijia et al., 2020; Wang et al., 2018) which consider positive as well as negative opinion spreading on signed network settings.
- *How to incorporate different competing and complimentary ideas of diffusion process?* As discussed earlier, the viral marketing phenomenon is based on the premise that people influenced each other's purchasing behaviors and decisions. Therefore, to maximize a company's profit, a promotional strategy needs to consider marketing cost, profit from each purchase, purchase pattern, etc. Also, it helps a marketing company in competing and complementary ideas of product advertising based on pattern utility. Nath and Domingos (2010) introduce viral marketing strategy based on utility maximization and repeated inferences. Hence, to decide competing and complementary ideas of product diffusion process in profit maximization is still a challenging task that remains much open to exploring.
- *How to obtains more data about activation process in viral marketing?* The dynamics and processes behind information and opinion spreading gather the attention of researchers within few years. Understanding these dynamics can shed light on the structure of an individual's relationships, impact of promotional strategies, strength of relationship, and activation probability. He et al. (He et al., 2014) addresses the impact of incorrect and noisy parameters like influence probabilities and activation probability. Therefore, estimation of such information is a much open problem to explore in viral marketing.
- *How to incorporate information propagation with other practical applications?* In recent years, researchers attract to the application of information propagation analysis in various practical applications like behavior analysis and prediction using the social influence of users (Wu et al., 2017; Xu et al., 2016; Zhang et al., 2017; Sabot et al., xxxx; Sato and Fujita, 2019; Shen et al., 2019). For example, user consumption behavior and product adaptation are affected by multiple advertisers, user's interests, and group norms at the same time. The authors of Ma et al. (2015) consider the above features to identifying seed users for social marketing. Xu et al. (2016) analyze the behavior of taxi drivers to figure out their future behavior and reaction. Therefore, it is quite interesting to adopt information diffusion with other practical scenarios to tackle real-life problems by analyzing user's social influence.
- *How deep learning techniques utilized for analyzing the information diffusion?* Recently, deep learning techniques are popularly utilized in other social network analysis tasks like link prediction (Liu et al., 2013) and network embedding (Du et al., 2017; Wang et al., 2016a). Therefore, a question naturally arises: Can deep learning methods be utilized to analyze the information diffusion process? For example, when we know an individual's attributes like educational background, gender, age, political view, hometown, etc., and network topology, we can compute an individual's influence using the deep learning model. The author of Bourigault et al. (2016) introduced a deep learning model to identify the source of diffusion without knowing diffusion graph, rather focus on diffusion records.
- *What algorithm should be used for a specific dataset?* Each of the social networks has its properties and features which differentiate it from others. Therefore, it is handy to consider these properties to devise a method for the computation of social influence and seed selection. It is also helpful in selecting an existing process for a specific dataset. To the best of our knowledge, there is no study present in the literature to analyze the algorithms for specific network perspectives, which helps select appropriate methods corresponding to the dataset.

- *How to consider cost, benefit, and time simultaneously?* Whenever a node is identified as a seed, an incentive is required as most of the social networks are formed by rational agents. It is also essential to incorporate the benefit (marginal gain in terms of product adoption, users activation, etc.) of selecting an individual as a seed. Due to the time-sensitive nature of information diffusion, it is also important to consider diffusion time. For example, there is no sense to influence someone after the event is over, such as product advertising for a seasonal product, election campaign before election season, etc. To the best of our knowledge, no study covers all three factors. Therefore, consideration of such elements is a much open problem to explore in viral marketing.
- *How to ensure ethical use of social media?* Nowadays, social media is increasingly becoming an integral part of modern life. An application of the influence maximization problem also needs to address the ethical use of social media and its data from both user's and service provider's perspectives. Therefore, ethical issues regarding the use of social media can also be considered.
- *Others.* Some other possible directions like privacy-preserving AI, Bayesian and deep learning framework, linear programming, etc., can be considered to find elegant and practical solutions to the influence maximization problem. For example, privacy leakage is not considered in existing IM algorithms. These algorithms may results in leakage of private and sensitive information in social and wireless sensor networks (Katti, 2019).

7. Concluding remarks

This study presents a fine-grained classification of existing IM algorithms based on their framework and algorithmic design perspective. First, we briefly discuss information diffusion models. Then, we cover the conventional IM problem and conduct a theoretical study on IM approaches. A rigorous comparative analysis of the state-of-the-art IM techniques is also presented. Furthermore, we show a performance comparison of existing IM algorithms with respect to evaluation metrics. We also categorize the research challenges of the IM problem based on problem formulation and discuss these challenges with some future directions. This study provides researchers new to this problem a better understanding of the IM problem with recent developments and a good starting point to work in this field. The comprehensive survey of IM identifies several research possibilities in viral marketing. However, several pieces had been done in IM, but still many more things to be uncovered. The work done in this area is still in its initial stage. The future work in the IM leads to increase effectiveness, scalability, robustness, and efficiency in the large-scale network. Along with scalability, the context in IM is one of the more challenging issues which must emphasize in future research.

8. Compliance with ethical standards

No conflict of interest is declared by the authors. No human or animal subjects have been been studied in this article. A comprehensive survey on influence maximization approaches is presented.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. List of Abbreviations

IM	≜	Influence Maximization
IM2	≜	Influence Maximization across Multiple Networks
MIM	≜	Multiple Influence Maximization
MIM2	≜	Multiple Influence Maximization cross Multiple Networks
CIM	≜	Context-aware Influence Maximization
TAIM	≜	Topic-aware Influence Maximization
IDM/DM	≜	Information Diffusion Model
TM	≜	Threshold Model
LTM	≜	Linear Threshold Model
CM	≜	Cascade Model
ICM	≜	Independent Cascade Model
ICM-NO	≜	Independent Cascade Model with Negative Opinion
TRM	≜	Triggering Model
TAM	≜	Time-aware Model
DTAM	≜	Discrete Time-aware Model
CTAM	≜	Continuous Time-aware Model
EM	≜	Epidemic Model
OSN	≜	Online Social Network
MC	≜	Monte Carlo Simulations
CELF	≜	Cost Effective Lazy Forward
CGA	≜	Community-based Greedy Algorithm
LPIMA	≜	Label Propagation based Influence Maximization Algorithm
Group-PR	≜	Group Page Rank Algorithm
SPM	≜	Shortest Path Model
MIA	≜	Maximum Influence Arborescence
PMIA	≜	Prefix excluding Maximum Influence Arborescence
IPA	≜	Independent Path Algorithm
DAG	≜	Directed Acyclic Graph
RR	≜	Reverse Reachable Set
BFS	≜	Breadth First Search
SSA	≜	Stop-and-Stare Optimization Approach
ASP	≜	Agent Selection Problem
LGA	≜	Local Greedy
ELGA	≜	Efficient Local Greedy
C2IM	≜	Community-based Context-aware Influence Maximization
LAPSO-IM	≜	Learning Automata based Particle Swarm Optimization Influence Maximization
ACO-IM	≜	Ant Colony based Influence Maximization
LAIM	≜	Location-aware Influence Maximization
HIM	≜	Holistic Influence Diffusion Model
CTMC	≜	Continuous Markov Chain Model
TV-IC	≜	Time-Varying Independent Cascade
TV-LT	≜	Time-Varying Linear Threshold
IBM	≜	Influence Blocking Maximization
SIM	≜	Self Influence Maximization
DIM	≜	Dynamic Influence Maximization
UBI	≜	Upper Bound Interchange
SIM	≜	Semantic-aware Influence Maximization

References

Nouran Ayman R. Abd Al-Azim, Tarek F. Gharib, Mohamed Hamdy, and Yasmine Afify. Influence propagation in social networks: Interest-based community ranking model. Journal of King Saud University - Computer and Information Sciences, 2020.

Charu C. Aggarwal, Shuyang Lin, and Philip S. Yu. On influential node discovery in dynamic social networks. In SDM, 2012.

Alan Kuhnle, Md., Alim, Abdul, Li, Xiang, Zhang, Huiling, Thai, My T., 2018. Multiplex influence maximization in online social networks with heterogeneous diffusion models. IEEE Transactions on Computational Social Systems 5, 418–429.

Akhil Arora, Sainyam Galhotra, and Sayan Ranu. Debunking the myths of influence maximization: An in-depth benchmarking study. In Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD '17, pages 651–666, New York, NY, USA, 2017. ACM.

Çigdem Aslay, Nicola Barbieri, Francesco Bonchi, and Ricardo A. Baeza-Yates. Online topic-aware influence maximization queries. In EDBT, 2014.

Çigdem Aslay, Francesco Bonchi, Laks V.S. Lakshmanan, and Wei Lu. Revenue maximization in incentivized social advertising. CoRR, abs/1612.00531, 2016.

Babaei, Mahmoudreza, Mirzasoleiman, Baharan, Jalili, Mahdi, Safari, Mohammad Ali, 2012. Revenue maximization in social networks through discounting. Social Network Analysis and Mining 3, 1249–1262.

Banerjee, Suman, Jenamani, Mamata, Pratihar, Dilip Kumar, 2019. Combim: A community-based solution approach for the budgeted influence maximization problem. Expert Systems with Applications 125, 1–13.

N. Barbieri, F. Bonchi, and G. Manco. Topic-aware social influence propagation models. In 2012 IEEE 12th International Conference on Data Mining, pages 81–90, Dec 2012.

Berahmand, Kamal, Nasiri, Elahe, Forouzandeh, Saman, Li, Yuefeng, 2021. A preference random walk algorithm for link prediction through mutual influence nodes in complex networks. Journal of King Saud University - Computer and Information Sciences.

Smriti Bhagat, Amit Goyal, and Laks V.S. Lakshmanan. Maximizing product adoption in social networks. In Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, WSDM '12, pages 603–612, New York, NY, USA, 2012. ACM.

Shishir Bharathi, David Kempe, and Mahyar Salek. Competitive influence maximization in social networks. In Xiaotie Deng and Fan Chung Graham, editors, Internet and Network Economics, pages 306–311, Berlin, Heidelberg, 2007. Springer, Berlin Heidelberg.

Biswas, A., Biswas, B., 2018. Fuzag: Fuzzy agglomerative community detection by exploring the notion of self-membership. IEEE Transactions on Fuzzy Systems 26 (5), 2568–2577.

Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. Maximizing social influence in nearly optimal time. In Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '14, pages 946–957, Philadelphia, PA, USA, 2014. Society for Industrial and Applied Mathematics.

Allan Borodin, Yuval Filmus, and Joel Oren. Threshold models for competitive influence in social networks. In Amin Saberi, editor, Internet and Network Economics, pages 539–550, Berlin, Heidelberg, 2010. Springer, Berlin Heidelberg.

Allan Borodin, Yuval Filmus, and Joel Oren. Threshold models for competitive influence in social networks. In Amin Saberi, editor, Internet and Network Economics, pages 539–550, Berlin, Heidelberg, 2010. Springer, Berlin Heidelberg.

Bourigault, Simon, Lamprier, Sylvain, Gallinari, Patrick, 2016. Learning distributed representations of users for source detection in online social networks. In: Frasconi, Paolo, Landwehr, Niels, Manco, Giuseppe, Vreeken, Jilles (Eds.), Machine Learning and Knowledge Discovery in Databases. Springer International Publishing, Cham, pp. 265–281.

Bozorgi, Arastoo, Samet, Saeed, Kwisthout, Johan, Wareham, Todd, 2017. Community-based influence maximization in social networks under a competitive linear threshold model. Knowledge-Based Systems 134, 149–158.

Brown, Jacqueline Johnson, Reingen, Peter H., 1987. Social ties and word-of-mouth referral behavior*. Journal of Consumer Research 14 (3), 350–362.

Francesco Buccafurri, Gianluca Lax, Antonino Nocera, and Domenico Ursino. Discovering links among social networks. In Peter A. Flach, Tijl De Bie, and Nello Cristianini, editors, Machine Learning and Knowledge Discovery in Databases, pages 467–482, Berlin, Heidelberg, 2012. Springer, Berlin Heidelberg.

Tim Carnes, Chandrashekhar Nagarajan, Stefan M. Wild, and Anke van Zuylen. Maximizing influence in a competitive social network: A follower's perspective. In Proceedings of the Ninth International Conference on Electronic Commerce, ICEC '07, pages 351–360, New York, NY, USA, 2007. ACM.

Ceren Budak, Divyakant Agrawal, and Amr El Abbadi. Limiting the spread of misinformation in social networks. In Proceedings of the 20th International Conference on World Wide Web, WWW '11, pages 665–674, New York, NY, USA, 2011. ACM.

Chang, Biao, Tong, Xu., Liu, Qi, Chen, Enhong, 2018. Study on information diffusion analysis in social networks and its applications. International Journal of Automation and Computing 15, 377–401.

Wei Chen, Yajun Wang, and Siyu Yang. Efficient influence maximization in social networks. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '09, pages 199–208, New York, NY, USA, 2009. ACM.

Wei Chen, Chi Wang, and Yajun Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. July 2010.

W. Chen, Y. Yuan, and L. Zhang. Scalable influence maximization in social networks under the linear threshold model. In 2010 IEEE International Conference on Data Mining, pages 88–97, Dec 2010.

W. Chen, Y. Yuan, and L. Zhang. Scalable influence maximization in social networks under the linear threshold model. In 2010 IEEE International Conference on Data Mining, pages 88–97, Dec 2010.

Chen, Wei, Yuan, Yifei, Zhang, Li, 2010d. Scalable influence maximization in social networks under the linear threshold model. Proceedings of the 2010 IEEE International Conference on Data Mining, ICDM '10. IEEE Computer Society, Washington, DC, USA, pp. 88–97.

Wei Chen, Chi Wang, and Yajun Wang. 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, KDD '10, pages 1029–1038, New York, NY, USA, 2010. ACM.

Wei Chen, Alex Collins, Rachel Cummings, Te Ke, Zhenming Liu, David Rincon, Xiaorui Sun, Wei Wei, Yajun Wang, and Yifei Yuan. Influence maximization in social networks when negative opinions may emerge and propagate. In Proceedings of the 2011 SIAM International Conference on Data Mining (SDM'2011), April 2011.

Wei Chen, Alex Collins, Rachel Cummings, Te Ke, Zhenming Liu, David Rincon, Xiaorui Sun, Wei Wei, Yajun Wang, and Yifei Yuan. Influence maximization in social networks when negative opinions may emerge and propagate. In Proceedings of the 2011 SIAM International Conference on Data Mining (SDM'2011), April 2011.

Chen, Wei, Wei, Lu., Zhang, Ning, 2012. Time-critical influence maximization in social networks with time-delayed diffusion process 04.

Chen, Wei, Lakshmanan, Laks V.S., Castillo, Carlos, 2013. Information and Influence Propagation in Social Networks. Morgan & Claypool Publishers.

Chen, Yi-Cheng, Zhu, Wen-Yuan, Peng, Wen-Chih, 2014. Wang-Chien Lee, and Suh-Yin Lee. Cim: Community-based influence maximization in social networks. ACM Trans. Intell. Syst. Technol. 5 (2), 25:1–25:31.

Chen, Wei, Lin, Tian, Yang, Cheng, 2015. Real-time topic-aware influence maximization using preprocessing. Computational Social Networks 3 (1), 8.

Chen, Yipeng, Qiang, Qu., Ying, Yuanxiang, Li, Hongyan, Shen, Jialie, 2020. Semantics-aware influence maximization in social networks. Information Sciences 513, 442–464.

Suqi Cheng, Huawei Shen, Junming Huang, Guoqing Zhang, and Xueqi Cheng. Staticgreedy: Solving the scalability-accuracy dilemma in influence maximization. In Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management, CIKM '13, pages 509–518, New York, NY, USA, 2013. ACM.

Suqi Cheng, Huawei Shen, Junming Huang, Wei Chen, and Xueqi Cheng. Imrank: Influence maximization via finding self-consistent ranking. CoRR, abs/1402.3939, 2014.

Edith Cohen, Daniel Delling, Thomas Pajor, and Renato F. Werneck. Sketch-based influence maximization and computation: Scaling up with guarantees. CoRR, abs/1408.6282, 2014.

Abhimanyu Das and David Kempe. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. Computing Research Repository - CORR, 02 2011.

Pedro Domingos and Matt Richardson. Mining the network value of customers. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '01, pages 57–66, New York, NY, USA, 2001. ACM.

Nan Du, Yingyu Liang, Maria-Florina Balcan, Manuel Gomez-Rodriguez, Hongyuan Zha, and Le Song. Scalable influence maximization for multiple products in continuous-time diffusion networks. CoRR, abs/1612.02712, 2016.

D. Du, H. Wang, T. Xu, Y. Lu, Q. Liu, and E. Chen. Solving link-oriented tasks in signed network via an embedding approach. In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pages 75–80, Oct 2017.

Eiselt, H.A., Laporte, G., 1989. Competitive spatial models. European Journal of Operational Research 39 (3), 231–242.

Fredrik Erlandsson, Piotr Bródka, and Anton Borg. Seed selection for information cascade in multilayer networks. CoRR, abs/1710.04391, 2017.

Freeman, Linton C., 1978. Centrality in social networks conceptual clarification. Social Networks 1 (3), 215–239.

Sainyam Galhotra, Akhil Arora, and Shourya Roy. Holistic influence maximization: Combining scalability and efficiency with opinion-aware models. In Proceedings of the 2016 International Conference on Management of Data, SIGMOD '16, pages 743–758, New York, NY, USA, 2016. ACM.

H. Ge, J. Huang, C. Di, J. Li, and S. Li. Learning automata based approach for influence maximization problem on social networks. In 2017 IEEE Second International Conference on Data Science in Cyberspace (DSC), pages 108–117, June 2017.

Aristides Gionis, Evimaria Terzi, and Panayiotis Tsaparas. Opinion maximization in social networks. CoRR, abs/1301.7455, 2013.

Jacob Goldenberg and Eitan Muller. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. 2003.

Goldenberg, Jacob, Libai, Barak, Muller, Eitan, 2001a. Talk of the network: A complex systems look at the underlying process of word-of-mouth. Marketing Letters 12 (3), 211–223.

Goldenberg, Jacob, Libai, Barak, Muller, Eitan, 2001b. Talk of the network: A complex systems look at the underlying process of word-of-mouth. Marketing Letters 12 (3), 211–223.

M. Gomez Rodriguez and B. Schölkopf. Influence maximization in continuous time diffusion networks. In Proceedings of the 29th International Conference on Machine Learning, pages 313–320, New York, NY, USA, July 2012. Omnipress.

Manuel Gomez-Rodriguez, David Balduzzi, and Bernhard Schölkopf. Uncovering the temporal dynamics of diffusion networks. CoRR, abs/1105.0697, 2011.

M. Gomez Rodriguez and B. Schölkopf. Influence maximization in continuous time diffusion networks. In Proceedings of the 29th International Conference on Machine Learning, pages 313–320, New York, NY, USA, July 2012. Omnipress.

Manuel Gomez-Rodriguez, Le Song, Nan Du, Hongyuan Zha, and Bernhard Schölkopf. Influence estimation and maximization in continuous-time diffusion networks. ACM Trans. Inf. Syst., 34(2):9:1–9:33, February 2016.

Gong, Maoguo, Yan, Jianan, Shen, Bo, Ma, Lijia, Cai, Qing, 2016. Influence maximization in social networks based on discrete particle swarm optimization. Information Sciences 367–368, 600–614.

Amit Goyal, Wei Lu, and Laks V.S. Lakshmanan. Celf++: Optimizing the greedy algorithm for influence maximization in social networks. In Proceedings of the 20th International Conference Companion on World Wide Web, WWW '11, pages 47–48, New York, NY, USA, 2011. ACM.

Goyal, Amit, Wei, Lu., Lakshmanan, Laks V.S., 2011. Simpath: An efficient algorithm for influence maximization under the linear threshold model. Proceedings of the 2011 IEEE 11th International Conference on Data Mining, ICDM '11. IEEE Computer Society, Washington, DC, USA, pp. 211–220.

Granovetter, Mark, 1978. Threshold models of collective behavior. American Journal of Sociology 83 (6), 1420–1443.

Guille, Adrien, Hacid, Hakim, Favre, Cecile, Zighed, Djamel A., 2013. Information diffusion in online social networks: A survey. SIGMOD Rec. 42 (2), 17–28.

Jing Guo, Peng Zhang, Chuan Zhou, Yanan Cao, and Li Guo. Personalized influence maximization on social networks. In Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management, CIKM '13, pages 199–208, New York, NY, USA, 2013. ACM.

Guo, L., Zhang, D., Cong, G., Wu, W., Tan, K.L., 2017. Influence maximization in trajectory databases. IEEE Transactions on Knowledge and Data Engineering 29 (3), 627–641.

Kai Han, Chaoting Xu, Fei Gui, Shaojie Tang, He Huang, and Jun Luo. Discount allocation for revenue maximization in online social networks. In Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, Mobihoc '18, pages 121–130, New York, NY, USA, 2018. ACM.

Xinran He and David Kempe. Robust influence maximization. CoRR, abs/1602.05240, 2016.

Xinran He, Guojie Song, Wei Chen, and Qingye Jiang. Influence blocking maximization in social networks under the competitive linear threshold model. In SDM, pages 463–474. SIAM/ Omnipress, 2012.

Xinran He and David Kempe. Stability of influence maximization. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, pages 1256–1265, New York, NY, USA, 2014. ACM.

He, Qiang, Lei, Zhencheng, Wang, Xingwei, Huang, Min, Cai, Yuliang, 2019. An effective scheme to address influence maximization for opinion formation in social networks. Transactions on Emerging Telecommunications Technologies 30 (6), e3599. e3599 ETT-18-0258.R1.

Hong, W., Qian, C., Tang, K., 2020. Efficient minimum cost seed selection with theoretical guarantees for competitive influence maximization. IEEE Transactions on Cybernetics, 1–14.

Hosseinpour, Mohammad, Malek, Mohammad Reza, Claramunt, Christophe, 2019. Socio-spatial influence maximization in location-based social networks. Future Generation Computer Systems 101, 304–314.

Yongfeng Huang, Haixu Xi, Sayed Chhattan Shah, and Feiyue Ye. The evaluation and application of node influence in dynamic networks based on evolving communities. Transactions on Emerging Telecommunications Technologies, 30(9):e3556, 2019. e3556 ett.3556.

Qingye Jiang, Guojie Song, Gao Cong, Yu Wang, Wenjun Si, and Kunqing Xie. Simulated annealing based influence maximization in social networks. In Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI'11, pages 127–132. AAAI Press, 2011.

Jindal, Honey, Sardana, Neetu, 2020. Web navigation prediction based on dynamic threshold heuristics. Journal of King Saud University - Computer and Information Sciences.

Jung, Kyomin, Heo, Wooram, Chen, Wei, 2012. Irie: Scalable and robust influence maximization in social networks. Proceedings of the 2012 IEEE 12th International Conference on Data Mining, ICDM '12. IEEE Computer Society, Washington, DC, USA, pp. 918–923.

Katti, Anvesha, 2019. Target coverage in random wireless sensor networks using cover sets. Journal of King Saud University - Computer and Information Sciences.

David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03, pages 137–146, New York, NY, USA, 2003. ACM.

David Kempe, Jon Kleinberg, and Éva Tardos. Influential nodes in a diffusion model for social networks. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, Automata, Languages and Programming, pages 1127–1138, Berlin, Heidelberg, 2005. Springer, Berlin Heidelberg.

W.O. Kermack and A.G. McKendrick. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 115(772), 700–721, 1927.

Kermack, W.O., McKendrick, A.G., 1991. Contributions to the mathematical theory of epidemicsùi. Bulletin of Mathematical Biology 53 (1), 33–55.

A. Khan, B. Zehnder, and D. Kossmann. Revenue maximization by viral marketing: A social network host's perspective. In 2016 IEEE 32nd International Conference on Data Engineering (ICDE), pages 37–48, May 2016.

J. Kim, S.K. Kim, and H. Yu. Scalable and parallelizable processing of influence maximization for large-scale social networks? In 2013 IEEE 29th International Conference on Data Engineering (ICDE), pages 266–277, April 2013.

Masahiro Kimura and Kazumi Saito. Tractable models for information diffusion in social networks. In Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou, editors, Knowledge Discovery in Databases: PKDD 2006, pages 259–271, Berlin, Heidelberg, 2006. Springer, Berlin Heidelberg.

Kumar, Pawan, Sinha, Adwitiya, 2021. Information diffusion modeling and analysis for socially interacting networks. Social Network Analysis and Mining 11, 12.

Kumar, Ajay, Singh, Shashank Sheshar, Singh, Kuldeep, Biswas, Bhaskar, 2019. Level-2 node clustering coefficient-based link prediction. Applied Intelligence 49 (7), 2762–2779.

Ajay Kumar, Shivansh Mishra, Shashank Sheshar Singh, Kuldeep Singh, and Bhaskar Biswas. Link prediction in complex networks based on significance of higher-order path index (shopi). Physica A: Statistical Mechanics and its Applications, page 123790, 2019.

Ajay Kumar, Shashank Sheshar Singh, Kuldeep Singh, and Bhaskar Biswas. Link prediction techniques, applications, and performance: A survey. Physica A: Statistical Mechanics and its Applications, page 124289, 2020.

Suman Kundu, C.A. Murthy, and S.K. Pal. A new centrality measure for influence maximization in social networks. In Sergei O. Kuznetsov, Deba P. Mandal, Malay K. Kundu, and Sankar K. Pal, editors, Pattern Recognition and Machine Intelligence, pages 242–247, Berlin, Heidelberg, 2011. Springer, Berlin Heidelberg.

Lecture notes on particle systems and percolation. (book review). SciTech Book News, 13 (1989) 12.

Lee, J., Chung, C., 2015. A query approach for influence maximization on specific users in social networks. IEEE Transactions on Knowledge and Data Engineering 27 (2), 340–353.

W. Lee, J. Kim, and H. Yu. Ct-ic: Continuously activated and time-restricted independent cascade model for viral marketing. In 2012 IEEE 12th International Conference on Data Mining, pages 960–965, Dec 2012.

Siyu Lei, Silviu Maniu, Luyi Mo, Reynold Cheng, and Pierre Senellart. Online influence maximization (extended version). CoRR, abs/1506.01188, 2015.

Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '07, pages 420–429, New York, NY, USA, 2007. ACM.

Liang, W., Shen, C., Li, X., Nishide, R., Piumarta, I., Takada, H., 2019. Influence maximization in signed social networks with opinion formation. IEEE Access 7, 68837–68852.

Hui Li, Sourav S. Bhowmick, and Aixin Sun. Casino: Towards conformity-aware social influence analysis in online social networks. In Proceedings of the 20th ACM International Conference on Information and Knowledge Management, CIKM '11, pages 1007–1012, New York, NY, USA, 2011. ACM.

Yanhua Li, Wei Chen, Yajun Wang, and Zhi-Li Zhang. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. CoRR, abs/1111.4729, 2011.

Cheng-Te Li, Shou-De Lin, and Man-Kwan Shan. Influence propagation and maximization for heterogeneous social networks. In Proceedings of the 21st International Conference on World Wide Web, WWW '12 Companion, pages 559–560, New York, NY, USA, 2012. ACM.

Hui Li, Sourav S. Bhowmick, and Aixin Sun. Cinema: conformity-aware greedy algorithm for influence maximization in online social networks. In EDBT '13, 2013.

Hui Li, Sourav S. Bhowmick, and Aixin Sun. Cinema: Conformity-aware greedy algorithm for influence maximization in online social networks. In Proceedings of the 16th International Conference on Extending Database Technology, EDBT '13, pages 323–334, New York, NY, USA, 2013. ACM.

Guoliang Li, Shuo Chen, Jianhua Feng, Kian-lee Tan, and Wen-syan Li. Efficient location-aware influence maximization. In Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, SIGMOD '14, pages 87–98, New York, NY, USA, 2014. ACM.

Li, Hui, Bhowmick, Sourav S., Sun, Aixin, Cui, Jiangtao, 2014b. Conformity-aware influence maximization in online social networks. The VLDB Journal 24, 117–141.

Dong Li, Zhi-Ming Xu, Nilanjan Chakraborty, Anika Gupta, Katia Sycara, and Sheng Li. Polarity related influence maximization in signed social networks. PLOS ONE, 9:1–12, 07 2014.

Guoliang Li, Shuo Chen, Jianhua Feng, Kian-lee Tan, and Wen-syan Li. Efficient location-aware influence maximization. In Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, SIGMOD '14, pages 87–98, New York, NY, USA, 2014. ACM.

Li, Yuchen, Zhang, Dongxiang, Tan, Kian-Lee, 2015a. Real-time targeted influence maximization for online advertisements. Proc. VLDB Endow. 8 (10), 1070–1081.

Li, Hui, Bhowmick, Sourav S., Sun, Aixin, Cui, Jiangtao, 2015b. Conformity-aware influence maximization in online social networks. The VLDB Journal 24 (1), 117–141.

Li, Yuchen, Zhang, Dongxiang, Tan, Kian-Lee, 2015c. Real-time targeted influence maximization for online advertisements. Proc. VLDB Endow. 8 (10), 1070–1081.

Li, Hui, Bhowmick, Sourav S., Cui, Jiangtao, 2015d. Yunjun Gao, and Jianfeng Ma. Getreal: Towards realistic selection of influence maximization strategies in competitive networks. SIGMOD Conference.

Yuchen Li, Ju Fan, Dongxiang Zhang, and Kian-Lee Tan. Discovering your selling points: Personalized social influential tags exploration. In Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD '17, pages 619–634, New York, NY, USA, 2017. ACM.

Li, Weian, Liu, Wenjing, Chen, Tiantian, Xiaoying, Qu., Fang, Qizhi, Ko, Ker-I, 2017b. Competitive profit maximization in social networks. Theoretical Computer Science 694, 1–9.

Li, Dong, Wang, Cuihua, Zhang, Shengping, Zhou, Guanglu, Chu, Dianhui, Chong, Wu., 2017c. Positive influence maximization in signed social networks based on simulated annealing. Neurocomputing 260, 69–78.

Li, Mei, Wang, Xiang, Gao, Kai, Zhang, Shanshan, 2017d. A survey on information diffusion in online social networks: Models and methods. Information 8 (4).

J. Li, T. Sellis, J.S. Culpepper, Z. He, C. Liu, and J. Wang. Geo-social influence spanning maximization. In 2018 IEEE 34th International Conference on Data Engineering (ICDE), pages 1775–1776, April 2018.

J. Li, T. Cai, A. Mian, R. Li, T. Sellis, and J.X. Yu. Holistic influence maximization for targeted advertisements in spatial social networks. In 2018 IEEE 34th International Conference on Data Engineering (ICDE), pages 1340–1343, April 2018.

Y. Li, X. Gan, L. Fu, X. Tian, Z. Qin, and Y. Zhou. Conformity-aware influence maximization with user profiles. In 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), pages 1–6, Oct 2018.

Li, Y., Fan, J., Wang, Y., Tan, K., 2018d. Influence maximization on social graphs: A survey. IEEE Transactions on Knowledge and Data Engineering 30 (10), 1852–1872.

Li, X., Smith, J.D., Dinh, T.N., Thai, M.T., 2019. Tiptop: (almost) exact solutions for influence maximization in billion-scale networks. IEEE/ACM Transactions on Networking 27 (2), 649–661.

Li, Lingfei, Liu, Yezheng, Zhou, Qing, Yang, Wei, Yuan, Jiahang, 2020. Targeted influence maximization under a multifactor-based information propagation model. Information Sciences 519, 124–140.

Su-Chen Lin, Shou-De Lin, and Ming-Syan Chen. A learning-based framework to handle multi-round multi-party influence maximization on social networks. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '15, pages 695–704, New York, NY, USA, 2015. ACM.

Feng Liu, Bingquan Liu, Chengjie Sun, Ming Liu, and Xiaolong Wang. Deep learning approaches for link prediction in social network services. In Proceedings, Part II, of the 20th International Conference on Neural Information Processing - Volume 8227, ICONIP 2013, page 425–432, Berlin, Heidelberg, 2013. Springer-Verlag.

Qi Liu, Biao Xiang, Enhong Chen, Hui Xiong, Fangshuang Tang, and Jeffrey Xu Yu. Influence maximization over large-scale social networks: A bounded linear approach. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, CIKM '14, pages 171–180, New York, NY, USA, 2014. ACM.

Liu, B., Cong, G., Zeng, Y., Xu, D., Chee, Y.M., 2014. Influence spreading path and its application to the time constrained social influence maximization problem and beyond. IEEE Transactions on Knowledge and Data Engineering 26 (8), 1904–1917.

Ma, Guowei, Liu, Qi, Le, Wu., Chen, Enhong, 2015. Identifying hesitant and interested customers for targeted social marketing. In: Cao, Tru, Lim, Ee-Peng, Zhou, Zhi-Hua, Ho, Tu-Bao, Cheung, David, Motoda, Hiroshi (Eds.), Advances in Knowledge Discovery and Data Mining. Springer International Publishing, Cham, pp. 576–590.

Yasir Mehmood, Francesco Bonchi, and David García-Soriano. Spheres of influence for more effective viral marketing. In Proceedings of the 2016 International Conference on Management of Data, SIGMOD '16, pages 711–726, New York, NY, USA, 2016. ACM.

Meng, Yanhong, Yi, Yunhui, Xiong, Fei, Pei, Changxing, 2019. T+onehop approach for dynamic influence maximization problem. Physica A: Statistical Mechanics and its Applications 515, 575–586.

Mishra, Sneha, Singh, Shashank Sheshar, Mishra, Shivansh, Biswas, Bhaskar, 2021. Tcd2: Tree-based community detection in dynamic social networks. Expert Systems with Applications 169, 114493.

Mohammadi, Azadeh, Saraee, Mohamad, Mirzaei, Abdolreza, 2015. Time-sensitive influence maximization in social networks. Journal of Information Science 41 (6), 765–778.

Narayanam, R., Narahari, Y., 2011. A shapley value-based approach to discover influential nodes in social networks. IEEE Transactions on Automation Science and Engineering 8 (1), 130–147.

Nathalie T.H. Gayraud, Evaggelia Pitoura, and Panayiotis Tsaparas. Diffusion maximization in evolving social networks. In Proceedings of the 2015 ACM on Conference on Online Social Networks, COSN '15, pages 125–135, New York, NY, USA, 2015. ACM.

Nath, Aniruddh, Domingos, Pedro, 2010. Efficient belief propagation for utility maximization and repeated inference. In: Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence. AAAI Press.

A. Nazemian and F. Taghiyareh. Influence maximization in independent cascade model with positive and negative word of mouth. In 6th International Symposium on Telecommunications (IST), pages 854–860, Nov 2012.

Nemhauser, G.L., Wolsey, L.A., Fisher, M.L., 1978. An analysis of approximations for maximizing submodular set functions—i. Mathematical Programming 14 (1), 265–294.

D.T. Nguyen, S. Das, and M.T. Thai. Influence maximization in multiple online social networks. In 2013 IEEE Global Communications Conference (GLOBECOM), pages 3060–3065, Dec 2013.

H.T. Nguyen, T.N. Dinh, and M.T. Thai. Cost-aware targeted viral marketing in billion-scale networks. In IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications, pages 1–9, April 2016.

Hung T. Nguyen, My T. Thai, and Thang N. Dinh. 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 '16, pages 695–710, New York, NY, USA, 2016. ACM.

Nguyen, H.T., Cano, A., Tam, V., Dinh, T.N., 2019. Blocking self-avoiding walks stops cyber-epidemics: A scalable gpu-based approach. In: IEEE Transactions on Knowledge and Data Engineering. 1–1.

Ohsaka, Naoto, Akiba, Takuya, Yoshida, Yuichi, Kawarabayashi, Ken-Ichi, 2014. Fast and accurate influence maximization on large networks with pruned monte-carlo simulations. In: Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, AAAI'14. AAAI Press, pp. 138–144.

Ohsaka, Naoto, Akiba, Takuya, Yoshida, Yuichi, Kawarabayashi, Ken-ichi, 2016. Dynamic influence analysis in evolving networks. Proc. VLDB Endow. 9 (12), 1077–1088.

Ohsaka, Naoto, Yamaguchi, Yutaro, Kakimura, Naonori, Kawarabayashi, Ken-ichi, 2016. Maximizing time-decaying influence in social networks. In: Frasconi, Paolo, Landwehr, Niels, Manco, Giuseppe, Vreeken, Jilles (Eds.), Machine Learning and Knowledge Discovery in Databases. Springer International Publishing, Cham, pp. 132–147.

Han-Ching Ou, Chung-Kuang Chou, and Ming-Syan Chen. Influence maximization for complementary goods: Why parties fail to cooperate? In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, CIKM '16, pages 1713–1722, New York, NY, USA, 2016. ACM.

Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999–66, Stanford InfoLab, November 1999. Previous number = SIDL-WP-1999-0120.

N. Pathak, A. Banerjee, and J. Srivastava. A generalized linear threshold model for multiple cascades. In 2010 IEEE International Conference on Data Mining, pages 965–970, Dec 2010.

Peng, Wu., Pan, Li, 2017. Scalable influence blocking maximization in social networks under competitive independent cascade models. Computer Networks 123, 38–50.

Qiang, Zhecheng, Pasiliao, Eduardo, Zheng, Qipeng, 2019. Model-based learning of information diffusion in social media networks. Applied Network Science 4, 11.

Abdul Razaque, Syed Rizvi, Meer Jaro khan, Muder Almiani, and Amer Al Rahayfeh. State-of-art review of information diffusion models and their impact on social network vulnerabilities. Journal of King Saud University - Computer and Information Sciences, 2019.

Abdul Razaque, Syed Rizvi, Meer Jaro khan, Muder Almiani, and Amer Al Rahayfeh. State-of-art review of information diffusion models and their impact on social network vulnerabilities. Journal of King Saud University - Computer and Information Sciences, 2019.

Matthew Richardson and Pedro Domingos, Mining knowledge-sharing sites for viral marketing. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '02, pages 61–70, New York, NY, USA, 2002. ACM.

Manon E.B. Sabot, Martin G. De Kauwe, Andy J. Pitman, Belinda E. Medlyn, Anne Verhoef, Anna M. Ukkola, and Gab Abramowitz. Plant profit maximization improves predictions of european forest responses to drought. New Phytologist, n/a(n/a).

Saito, Kazumi, Kimura, Masahiro, Ohara, Kouzou, Motoda, Hiroshi, 2012. Efficient discovery of influential nodes for sis models in social networks. Knowledge and Information Systems 30 (3), 613–635.

Sato, Takahiro, Fujita, Masafumi, 2019. A topology optimization method for electric machines and devices through submodular maximization. Electronics and Communications in Japan 102 (6), 3–11.

Saxena, Bhawna, Saxena, Vikas, 2019. Hurst exponent based approach for influence maximization in social networks. Journal of King Saud University - Computer and Information Sciences.

Schelling, Thomas C., 2006. Micromotives and Macrobehavior.

Sen Su, Xiao Li, Xiang Cheng, and Chenna Sun. Location-aware targeted influence maximization in social networks. Journal of the Association for Information Science and Technology, 69(2):229–241, 2 2018.

Chengguang Shen, Ryo Nishide, Ian Piumarta, Hideyuki Takada, and Wenxin Liang. Influence maximization in signed social networks. In Jianyong Wang, Wojciech Cellary, Dingding Wang, Hua Wang, Shu-Ching Chen, Tao Li, and Yanchun Zhang, editors, Web Information Systems Engineering – WISE 2015, pages 399–414, Cham, 2015. Springer International Publishing.

Chengguang Shen, Ryo Nishide, Ian Piumarta, Hideyuki Takada, and Wenxin Liang. Influence maximization in signed social networks. In Proceedings, Part I, of the 16th International Conference on Web Information Systems Engineering — WISE 2015 - Volume 9418, page 399–414, Berlin, Heidelberg, 2015. Springer-Verlag.

Bilu Shen and Changjun Wang. Optimization of key opinion consumer selection strategy for enterprise in the internet environment. Management Science and Engineering, 08:368–375, 01 2019.

K. Sheng and Z. Zhang. Research on the influence maximization based on community detection. In 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), pages 2797–2801, May 2018.

Shuo Chen, Ju., Fan, Guoliang Li, Feng, Jianhua, Tan, Kian-lee, Tang, Jinhui, 2015a. Online topic-aware influence maximization. Proc. VLDB Endow. 8 (6), 666–677.

Shuo Chen, Ju., Fan, Guoliang Li, Feng, Jianhua, Tan, Kian-lee, Tang, Jinhui, 2015b. Online topic-aware influence maximization. Proc. VLDB Endow. 8 (6), 666–677.

Shashank Sheshar Singh, Ajay Kumar, Kuldeep Singh, and Bhaskar Biswas. C2im: Community based context-aware influence maximization in social networks. Physica A: Statistical Mechanics and its Applications, 514:796 – 818, 2019.

Singh, Shashank Sheshar, Singh, Kuldeep, Kumar, Ajay, Biswas, Bhaskar, 2019a. Aco-im: maximizing influence in social networks using ant colony optimization. Soft Computing, 1–23.

Shashank Sheshar Singh, Ajay Kumar, Shivansh Mishra, Kuldeep Singh, and Bhaskar Biswas. A centrality measure for influence maximization across multiple social networks. In Advanced Informatics for Computing Research. 2019.

Shashank Sheshar Singh, Ajay Kumar, Shivansh Mishra, Kuldeep Singh, and Bhaskar Biswas. Influence maximization in social networks. Optimization in Large Scale Problems, Springer Optimization and Its Applications, 152, 2019.

Shashank Sheshar Singh, Ajay Kumar, Kuldeep Singh, and Bhaskar Biswas. Lapso-im: A learning-based influence maximization approach for social networks. Applied Soft Computing, page 105554, 2019.

Shashank Sheshar Singh, Kuldeep Singh, Ajay Kumar, Harish Kumar Shakya, and Bhaskar Biswas. A survey on information diffusion models in social networks. In Ashish Kumar Luhach, Dharm Singh, Pao-Ann Hsiung, Kamarul Bin Ghazali Hawari, Pawan Lingras, and Pradeep Kumar Singh, editors, Advanced Informatics for Computing Research, pages 426–439, Singapore, 2019. Springer Singapore.

Shashank Sheshar Singh, Kuldeep Singh, Ajay Kumar, and Bhaskar Biswas. Coim: Community-based influence maximization in social networks. In Ashish Kumar Luhach, Dharm Singh, Pao-Ann Hsiung, Kamarul Bin Ghazali Hawari, Pawan Lingras, and Pradeep Kumar Singh, editors, Advanced Informatics for Computing Research, pages 440–453, Singapore, 2019. Springer Singapore.

Singh, Shashank Sheshar, Singh, Kuldeep, Kumar, Ajay, Biswas, Bhaskar, 2019f. Mim2: Multiple influence maximization across multiple social networks. Physica A: Statistical Mechanics and its Applications 526, 120902.

Shashank Sheshar Singh, Shivansh Mishra, Ajay Kumar, and Bhaskar Biswas. Clp-id: Community-based link prediction using information diffusion. Information Sciences, 514:402 – 433, 2020.

Shashank Sheshar Singh, Ajay Kumar, Kuldeep Singh, and Bhaskar Biswas. Im-sso: Maximizing influence in social networks using social spider optimization. Concurrency and Computation: Practice and Experience, 32(2):e5421, 2020. e5421 cpe.5421.

Chonggang Song, Wynne Hsu, and Mong Li Lee. Targeted influence maximization in social networks. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, CIKM '16, pages 1683–1692, New York, NY, USA, 2016. ACM.

Song, G., Li, Y., Chen, X., He, X., Tang, J., 2017. Influential node tracking on dynamic social network: An interchange greedy approach. IEEE Transactions on Knowledge and Data Engineering 29 (2), 359–372.

T Suganya, S Thennammai, and Revathi Velusamy. Unique user identification across multiple social network. International Journal of Research in Marketing, 8:137–142, 08 2017.

Sumith, N., Annappa, B., Bhattacharya, Swapan, 2018. Influence maximization in large social networks: Heuristics, models and parameters. Future Generation Computer Systems 89, 777–790.

Jimeng Sun and Jie Tang. A Survey of Models and Algorithms for Social Influence Analysis, pages 177–214. Springer US, Boston, MA, 2011.

Honglin Sun, Xiaofeng Gao, Guihai Chen, Jian Gu, and Yongjian Wang. Multiple influence maximization in social networks. In Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication, IMCOM '16, pages 44:1–44:8, New York, NY, USA, 2016. ACM.

Sviridenko, Maxim, 2004. A note on maximizing a submodular set function subject to a knapsack constraint. Operations Research Letters 32 (1), 41–43.

Talukder, A., Hong, C.S., 2020. Cm2d: Cost minimization under the competitive market in dynamic social networks. IEEE Access 8, 4428–4441.

Tang, Shaojie, Yuan, Jing, 2020. Influence maximization with partial feedback. Operations Research Letters 48 (1), 24–28.

Tang, Jie, Sen, Wu., Sun, J., 2013. Confluence: Conformity influence in large social networks.

Youze Tang, Xiaokui Xiao, and Yanchen Shi. Influence maximization: near-optimal time complexity meets practical efficiency. In SIGMOD Conference, 2014.

Youze Tang, Yanchen Shi, and Xiaokui Xiao. Influence maximization in near-linear time: A martingale approach. In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, SIGMOD '15, pages 1539–1554, New York, NY, USA, 2015. ACM.

Jing Tang, Xueyan Tang, and Junsong Yuan. Towards profit maximization for online social network providers. CoRR, abs/1712.08963, 2017.

Tang, J., Tang, X., Yuan, J., 2018. Profit maximization for viral marketing in online social networks: Algorithms and analysis. IEEE Transactions on Knowledge and Data Engineering 30 (6), 1095–1108.

Tang, Jianxin, Zhang, Ruisheng, Wang, Ping, Zhao, Zhili, Fan, Li, Liu, Xin, 2020. A discrete shuffled frog-leaping algorithm to identify influential nodes for influence maximization in social networks. Knowledge-Based Systems 187, 104833.

V. Tejaswi, P.V. Bindu, and P.S. Thilagam. Diffusion models and approaches for influence maximization in social networks. In 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pages 1345–1351, Sept 2016.

V. Tejaswi, P.V. Bindu, and P.S. Thilagam. Target specific influence maximization: An approach to maximize adoption in labeled social networks. In 2017 9th International Conference on Communication Systems and Networks (COMSNETS), pages 542–547, Jan 2017.

Teng, Ya-Wen, Tai, Chih-Hua, Yu, Philip S., Chen, Ming-Syan, 2018. Revenue Maximization on the Multi-grade Product, pp. 576–584.

Tong, G., Wu, W., Tang, S., Du, D., 2017. Adaptive influence maximization in dynamic social networks. IEEE/ACM Transactions on Networking 25 (1), 112–125.

Tsai, Chun-Wei, Liu, Shih-Jui, 2019. Seim: Search economics for influence maximization in online social networks. Future Generation Computer Systems 93, 1055–1064.

Didier A. Vega-Oliveros, Luciano da Fontoura Costa, and Francisco Aparecido Rodrigues. Influence maximization by rumor spreading on correlated networks through community identification. Communications in Nonlinear Science and Numerical Simulation, 83:105094, 2020.

Vosecky, Jan, Hong, Dan, Shen, Vincent Y., 2009. User identification across multiple social networks.

Wang, Xue-Guang, 2016. A new algorithm for the influence maximization problem in dynamic networks or traffic sensor networks. Multimedia Tools and Applications 75 (8), 4833–4844.

Wang, Yu., Cong, Gao, 2010. Guojie Song, and Kunqing Xie. Community-based greedy algorithm for mining top-k influential nodes in mobile social networks. In: KDD.

Yitong Wang and Xiaojun Feng. A potential-based node selection strategy for influence maximization in a social network. In Ronghuai Huang, Qiang Yang, Jian Pei, João Gama, Xiaofeng Meng, and Xue Li, editors, Advanced Data Mining and Applications, pages 350–361, Berlin, Heidelberg, 2009. Springer, Berlin Heidelberg.

Wang, Daixin, Cui, Peng, Zhu, Wenwu, 2016a. Structural deep network embedding. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Association for Computing Machinery, New York, NY, USA.

Chih-Hang Wang, Po-Shun Huang, De-Nian Yang, and Wen-Tsuen Chen. Cross-layer design of influence maximization in mobile social networks. CoRR, abs/1604.02796, 2016.

Yake Wang, Jinghua Zhu, and Qian Ming. Incremental influence maximization for dynamic social networks. In Beiji Zou, Qilong Han, Guanglu Sun, Weipeng Jing, Xiaoning Peng, and Zeguang Lu, editors, Data Science, pages 13–27, Singapore, 2017. Springer Singapore.

Wang, X., Zhang, Y., Zhang, W., Lin, X., Chen, C., 2017a. Bring order into the samples: A novel scalable method for influence maximization. IEEE Transactions on Knowledge and Data Engineering 29 (2), 243–256.

Qixiang Wang, M. Gong, Chao Song, and Shanfeng Wang. Discrete particle swarm optimization based influence maximization in complex networks. In 2017 IEEE Congress on Evolutionary Computation (CEC), pages 488–494, June 2017.

Wang, X., Zhang, Y., Zhang, W., Lin, X., 2017b. Efficient distance-aware influence maximization in geo-social networks. IEEE Transactions on Knowledge and Data Engineering 29 (3), 599–612.

Yanhao Wang, Qi Fan, Yuchen Li, and Kian-Lee Tan. Real-time influence maximization on dynamic social streams. CoRR, abs/1702.01586, 2017.

Wang, Feng, Jiang, Wenjun, Li, Xiaolin, Wang, Guojun, 2018. Maximizing positive influence spread in online social networks via fluid dynamics. Future Generation Computer Systems 86, 1491–1502.

Wang, Shanfeng, Gong, Maoguo, Liu, Wenfeng, Yue, Wu., 2020. Preventing epidemic spreading in networks by community detection and memetic algorithm. Applied Soft Computing 89, 106118.

Weersink, Alfons, Fulton, Murray, 2020. Limits to profit maximization as a guide to behavior change. Applied Economic Perspectives and Policy 42 (1), 67–79.

Wei Lu and Laks V.S. Lakshmanan. Profit maximization over social networks. CoRR, abs/1210.4211, 2012.

Wei, Lu., Chen, Wei, Lakshmanan, Laks V.S., 2015. From competition to complementarity: Comparative influence diffusion and maximization. Proc. VLDB Endow. 9 (2), 60–71.

Weijia, Ju., Chen, Ling, Li, Bin, Liu, Wei, Sheng, Jun, Wang, Yuwei, 2020. A new algorithm for positive influence maximization in signed networks. Information Sciences 512, 1571–1591.

Wu, L., Ge, Y., Liu, Q., Chen, E., Hong, R., Du, J., Wang, M., 2017. Modeling the evolution of users? preferences and social links in social networking services. IEEE Transactions on Knowledge and Data Engineering 29 (6), 1240–1253.

Xie, Miao, Yang, Qiusong, Wang, Qing, Cong, Gao, de Melo, Gerard, 2015. Dynadiffuse: A dynamic diffusion model for continuous time constrained influence maximization. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI'15. AAAI Press, pp. 346–352.

Tong Xu, Hengshu Zhu, Xiangyu Zhao, Qi Liu, Hao Zhong, Enhong Chen, and Hui Xiong. Taxi driving behavior analysis in latent vehicle-to-vehicle networks: A social influence perspective. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'16, page 1285–1294, New York, NY, USA, 2016. Association for Computing Machinery.

W. Yang, L. Brenner, and A. Giua. Influence maximization by link activation in social networks. In 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), volume 1, pages 1248–1251, Sept 2018.

Mao Ye, Xingjie Liu, and Wang-Chien Lee. Exploring social influence for recommendation: A generative model approach. In Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '12, pages 671–680, New York, NY, USA, 2012. ACM.

Zhan, Qianyi, Zhang, Jiawei, Wang, Senzhang, Yu, Philip S., Xie, Junyuan, 2015a. Influence maximization across partially aligned heterogenous social networks. In: Cao, Tru, Lim, Ee-Peng, Zhou, Zhi-Hua, Ho, Tu-Bao, Cheung, David, Motoda, Hiroshi (Eds.), Advances in Knowledge Discovery and Data Mining. Springer International Publishing, Cham, pp. 58–69.

Zhan, Qianyi, Zhang, Jiawei, Wang, Senzhang, Yu, Philip S., Xie, Junyuan, 2015b. Influence maximization across partially aligned heterogenous social networks. In: Cao, Tru, Lim, Ee-Peng, Zhou, Zhi-Hua, Ho, Tu-Bao, Cheung, David, Motoda, Hiroshi (Eds.), Advances in Knowledge Discovery and Data Mining. Springer International Publishing, Cham, pp. 58–69.

Qianyi Zhan, Zhang Jiawei, Senzhang Wang, Philip Yu, and Junyuan Xie. Influence maximization across partially aligned heterogenous social networks. pages 58–69, 05 2015.

Qianyi Zhan, Zhang Jiawei, Senzhang Wang, Philip Yu, and Junyuan Xie. Influence maximization across partially aligned heterogenous social networks. pages 58–69, 05 2015.

Yue Zhang. Influence maximization on multi-phased multi-layered network. 2015.

H. Zhang, T.N. Dinh, and M.T. Thai. Maximizing the spread of positive influence in online social networks. In 2013 IEEE 33rd International Conference on Distributed Computing Systems, pages 317–326, July 2013.

Zhang, Huiyuan, Mishra, Subhankar, Thai, My, 2014a. Recent advances in information diffusion and influence maximization of complex social networks.

Jing Zhang, Jie Tang, Honglei Zhuang, Cane Wing-ki Leung, and Juanzi Li. Role-aware conformity influence modeling and analysis in social networks, 07 2014.

Zhang, H., Nguyen, D.T., Zhang, H., Thai, M.T., 2016. Least cost influence maximization across multiple social networks. IEEE/ACM Transactions on Networking 24 (2), 929–939.

Zhang, Ting, Qin, Rizhen, Dong, Qiulei, Gao, Wei, Huarong, Xu., Zhanyi, Hu., 2017. Physiognomy: Personality traits prediction by learning. International Journal of Automation and Computing 14, 386–395.

Zhou, C., Zhang, P., Zang, W., Guo, L., 2015a. On the upper bounds of spread for greedy algorithms in social network influence maximization. IEEE Transactions on Knowledge and Data Engineering 27 (10), 2770–2783.

Tao Zhou, Jiuxin Cao, Bo Liu, Shuai Xu, Ziqing Zhu, and Junzhou Luo. Location-based influence maximization in social networks. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, CIKM '15, pages 1211–1220, New York, NY, USA, 2015. ACM.

H. Zhuang, Y. Sun, J. Tang, J. Zhang, and X. Sun. Influence maximization in dynamic social networks. In 2013 IEEE 13th International Conference on Data Mining, pages 1313–1318, Dec 2013.

Y. Zhu, D. Li, and Z. Zhang. Minimum cost seed set for competitive social influence. In IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications, pages 1–9, April 2016.

Y. Zhu, D. Li, and Z. Zhang. Minimum cost seed set for competitive social influence. In IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications, pages 1–9, April 2016.