Expert Systems With Applications 248 (2024) 123429

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review

A survey on influence maximization models

Myriam Jaouadi *, Lotfi Ben Romdhane

MARS Research Lab LR17ES05 Higher Institute of Computer Science and Telecom (ISITCom), University of Sousse, Sousse, Tunisia

ARTICLE INFO

Keywords:
Influence maximization
Social networks
Static networks
Dynamic networks
Influential nodes

ABSTRACT

Influence maximization is an important research area in social network analysis where researchers are concerned with detecting influential nodes. The detection of influential nodes is of great interest in several disciplines including computer science, opinion propagation, political movements, or economics, where systems are often modeled as graphs. The Influence Maximization problem is proved NP-hard. This computational complexity is justified by two main factors. The first factor is about the important size of social networks. Modern social networks like TikTok and Facebook have reached an unprecedented number of users. Dynamic social networks, whose topology or/and informational content is able to evolve, represent the second factor. Maximizing influence in such networks remains a significant task. In this light, several methods have been proposed. Being motivated by this fact, we provide in this paper a detailed survey of influence maximization approaches. Our main concern is to provide a taxonomy of existing models in both static and dynamic networks. In addition, we provide a comparison of the state-of-the-art approaches according to a clear categorization. New trends for detecting influential nodes are also discussed. We provide then some challenges as well as future directions.

1. Introduction

Social networks have become a paramount communication tool in our life due to the explosive use of smartphones. The information's availability in such networks, which concerns the relationship between users as well as data characterizing them, has contributed to a great interest in their analysis. Social network analysis may affect the diffusion arrangements, the appearance of influencers, or the social actors characteristics. Modern social networks such as Twitter, Facebook, and TikTok have become the mainstream networking platforms for people to propagate new ideas or exchange information about new products. During the recent years, the explosive use of such applications has led to a complete industrial chain in social media marketing, including marketing agencies and influencers (Wang et al., 2023). For example, influencers on TikTok have reached tens of millions of followers (Wang et al., 2023). This phenomenon is known as the Influence Maximization problem (IM) in social networks. IM problem appears in a series of practical applications such as political campaigns and epidemic analysis. In politics, American President Trump used the Twitter social network to diffuse tweets dealing with the presidential election thus influencing people to vote for him (Azaouzi et al., 2021). In viral marketing, IM was widely used to promote new products. For example, a company that creates a new product and would like to sell it to a large number of customers starts by choosing an initial subset of users to have free samples of the product to recommend it to their families and friends under the word-of-mouth effect. Starting from this idea, the influence maximization problem appeared in social networks.

The above IM problem intends to choose a minimal set of individuals, called seed set, that are susceptible to maximize the spread of influence under an information diffusion process. Two main diffusion models were proposed by Kempe et al. (2003) that are the IC (Independant Cascade) and the LT (Linear Threshold) models. IM problem was proved NP-hard by Ahmed et al. (2011) under such diffusion models (Jaouadi & Ben Romdhane, 2022).

There have been many efforts on IM problem in recent years. According to the nature of social networks used in existing models, we propose in this paper to divide them into two categories: static networks and dynamic networks. As for static networks, both whole network and reduced network-based models are discussed. The whole static network-based detection models mainly use topological features of nodes such as centrality measures or the community structure. The dynamic network-based models, generally rely on the change of interactions between the network elements. When we talk about the dynamics in this wealth of data, the problem has become even more complicated for the simple reason that both the structure and the information of the network can change over time.

* Corresponding author.
E-mail addresses: jaouadimaryem@gmail.com (M. Jaouadi), lotfi.BenRomdhane@isitc.u-sousse.tn (L. Ben Romdhane).

https://doi.org/10.1016/j.eswa.2024.123429
Received 22 September 2023; Received in revised form 26 January 2024; Accepted 4 February 2024
Available online 15 February 2024
0957-4174/

1.1. Discussion of existing survey works

Survey papers on the IM problem tried to categorize existing models into broad classes (Taha, 2022). Li, Fan et al. (2018) presented in their paper a theoretical analysis of existing IM models. Their survey is the first to tackle context-aware IM problem works. However, they considered only the greedy algorithm to categorize existing models which is time-consuming. To tackle this inefficiency, centrality measures are used. Al-Garadi et al. (2018) have investigated such measures for identifying influential nodes. They classified existing models into local measures, shortest-path measures, and iterative-calculation based measures. A key advantage of this work is to consider machine learning approaches. However, the survey ignored dynamic networks, it is based only on static networks. In the same direction, Jaouadi and Ben Romdhane (2019) provided a detailed survey for IM models. The mentioned work offered a clear categorization for solving the IM problem while mentioning the advantages and drawbacks of the state-of-the-art approaches. However, the authors have considered only static networks. Having the same drawback, the review of Aghaee et al. (2021) considered only topological features of social networks to outline meta-heuristic works. The main idea of this paper is to tackle models that are based on heuristic algorithms such as the bee colony algorithm, the genetic algorithm (Jabari et al., 2022), and the particle swarm. The IM problem was considered as an optimization problem. Bian et al. (2020) reviewed IM approaches while considering dynamic networks, networks' content and topics. In this paper, existing models for IM problem were categorized depending on the network structure, the network content and also based on centrality measures. However, for dynamic networks, the authors have considered only topological changes. In the same context, Hafiane et al. (2020) presented a detailed survey for IM models in dynamic networks. The paper discussed several aspects of dynamics such as incremental approaches, interval approaches and, contact sequence ones. A clear comparison is presented in the paper. However, informational based dynamic models were ignored. The survey of Razis et al. (2021) presented recent approaches for identifying influential users based on social semantics. In addition, they considered community structure and social matching as social semantics roles. Azaouzi et al. (2021) introduced a new concept for the IM problem which is group-based influence. They provided two classes for existing models, node-based models and group-based models. Moreover, the paper tackles the influence maximization problem under privacy protection. The main concern of such a survey is to present a clear categorization of existing models and to describe new trends for the IM problem. However, only static network-based models were considered. Another recent paper discussed the context-aware influence maximization (Ye et al., 2022) and mentioned classic IM algorithms with a clear classification. In this work, the authors presented the IM problem while considering classical diffusion models and a new formulation of diffusion models based on a deep learning approach. The key advantage of a such paper is to consider new methods for solving IM problem such as the context-aware IM. However, their survey was limited to the study of static networks. Existing survey papers on the IM topic classified models into broad classes. Unfortunately, they did not draw distinguishable boundaries among the specific techniques adopted by the existing works. Some works considered only static networks while ignoring dynamic ones. Besides, if dynamic networks are considered, only the structural-based models are taken into account. For modern social networks, that are huge, most models had difficulties for treating such mega-scale networks as they are based on the whole network. However, a reduction phase can be used to tackle the IM problem in a sampled version. Such classification was ignored by all survey papers. In this work, existing models are divided according to the nature of the network into two main categories as seen in Fig. 1, static networks and dynamic networks. As for static networks, both the whole network and reduced network-based models are discussed. The whole static network-based detection models mainly use topological features of nodes such as centrality measures or the community structure. For the dynamic network-based models, we provide works that rely not only on the change of the network structure but also on the change of social content.

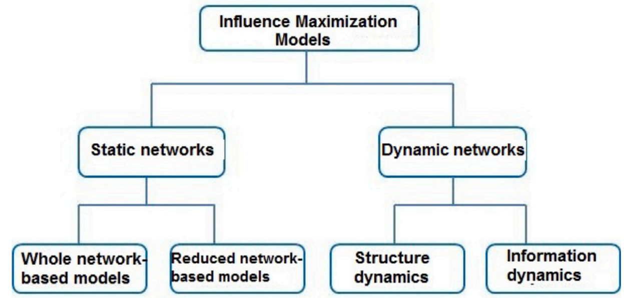


Fig. 1. Taxonomy of models for Influence Maximization in social networks.

1.2. Goals and contributions of the survey

The principal aim of this survey is to present, categorize, and analyze the existing models proposed for the Influence Maximization problem in social networks. The main goals of the paper are as follows:

- To categorize and compare the state-of-the-art IM models
- To provide measures used to characterize, identify, and extract influential users from static networks taking into account whole networks and reduced networks.
- To put worth IM problem according to the dynamic aspect of social networks: the structural-based dynamic models or the informational-based dynamic models.
- To present new trends for detecting influential nodes.

To achieve the goals mentioned above, an overview of traditional works is provided in this paper to give an idea about classical methodologies for solving the IM problem. Before conducting the survey, many research questions have been considered. For example, for a chosen paper, is the proposed model able to scale over large networks? Can the proposed model detect a minimal set of seeds that maximizes influence? The model is parameterless or not? Can the proposed model treat static and dynamic networks? Is the model able to consider more than one diffusion model?

Taking into account such questions, we start to choose papers based on keywords mentioned in Table 1. Then, many most cited papers included a reduction step of a large-scale network to solve the IM problem. From that, the idea of considering works in both the whole social network and a reduced version of a network was born. We tried to have an overview of how to create a reduced graph and can the reduced version maintain the diffusion and key properties of the original network. The review presents a new direction for the study of influence maximization in social networks which is missing in most reviews. To the best of our knowledge, there has not been a survey conducted on IM models in sampled and coarsened networks which is termed as reduced networks. The main objective of this paper is to illustrate such models. Consequently, a theoretical study was carried out on the ability to preserve the diffusion properties in the reduced network (Jaouadi & Ben Romdhane, 2022). For such analysis, real datasets are considered (Leskovec & Krevl, 2014). Moreover, we tried to compare several relevant works based on the reduction of the social network's size. On the other hand, for dynamic networks, we tried to study the structural and semantic aspects of the network. Existing surveys dealing with dynamic networks considered only one aspect, either the structure or the semantics, and not both. Besides, many works are conducted on small networks. For this reason, papers dealing with

Table 1
A summary on IM related work discussed in the paper.

Family	Network criteria	Keywords	Approach	Example of selected papers	Selection criteria
IM in static networks	Large-scale networks, structural aspect, the entire network	- IM problem - IM in social networks - influential nodes - Graph sampling-based IM - IM in reduced networks - Deep learning approach for IM	Greedy algorithm-based models	Goyal et al. (2011) and Leskovec et al. (2007)	Most cited
			Centrality measures-based models	Qiu et al. (2020) and Wu et al. (2018)	Most cited
			Community structure-based models	Li et al. (2020) and Umrawal et al. (2023)	Recent
			Machine learning-based models	Ling et al. (2023) and Ni et al. (2023)	Recent
			Hybrid models	Jaouadi and Ben Romdhane (2016)	Most cited
	Large-scale networks, structural aspect, a reduced network		Sampled networks-based models	Hong et al. (2020) and Li, Xu et al. (2023)	Most cited
			Coarsened networks-based models	Ohsaka et al. (2017)	Most cited
IM in dynamic networks	Large-scale networks, semantic and, structural aspect	- Dynamic IM - IM problem in dynamic networks - Tracking influential nodes	Structure dynamics	Liu et al. (2017), Tong et al. (2017) and Wang, Cuomo et al. (2019)	Most cited
			Information dynamics	Jaouadi and Ben Romdhane (2022), Lotf et al. (2022) and Min et al. (2020)	Recent

such networks are not considered in this review since modern social networks have reached an important size. Several works that could treat large-scale networks are selected and we present in this paper the most cited ones. In summary, a brief overview of the considered influence maximization works is presented in Table 1.

1.3. Structure of the survey

The paper is divided into two main parts. The taxonomy of models for IM problem available in the literature is provided in both static and dynamic networks as illustrated in Fig. 1. New trends for influential node detection will be discussed in the last part to give an idea about the recent formulation of the IM problem. The entire paper is structured as follows. Section 2 is devoted to presenting the preliminary material for the influence maximization problem in social networks as well as diffusion models. Section 3 discusses the taxonomy of the proposed models for the IM problem in static social networks. This section is dedicated to providing an overview of the structural techniques used to choose the seed set from the entire social networks as well as from reduced versions. Proposed models for influential node detection in dynamic social networks are discussed in Section 4. The nature of dynamics: structural/informational, is considered in this section. The new applications of the IM problem are presented in Section 5 to emphasize the new trends in this area. A comparative analysis is discussed in Section 6. Section 7 is devoted to mentioning some challenges and future directions of the IM problem. The last Section 8 concludes the review by providing open issues.

2. Preliminary knowledge for influence maximization in social networks

Social networks are omnipresent in our day-to-day life. Such networks are the subject of study by biologists, economists, sociologists, etc. Graphs are the most obvious way for modeling social networks where vertices are social users and edges are interactions between them (Jaouadi & Ben Romdhane, 2022). Independently of social networks' content, they have common properties that characterize them. For example, a node's neighboring is often distributed in a heterogeneous law, this property is known as node degree distribution (Jaouadi & Ben Romdhane, 2021) which is approximated usually by a power-low distribution. Another structural property is that of the community structure which suggests to detect groups of nodes such that nodes in the same group (community) have more interactions inside their community and fewer interactions with the outside (Rhouma & Ben Romdhane, 2018).

2.1. Problem formulation

A social network can be modeled by an undirected graph $G = (V, E)$ where V is the set of nodes representing the network users and E is the set of edges modeling interactions between them. The influence maximization problem can be defined as follows: Given a graph $G = (V, E)$ modelizing the social network, a diffusion model M that controls the spread of influence among users, and a budget $k < |V|$ as the seed set size, IM aims to detect a seed set e that maximizes the influence spread in G. The following equation can describe more formally the IM problem:

$$IM_M(G, k) = argmax_{e \subseteq V, |e|=k} \sigma_M(e, G) \tag{1}$$

where $\sigma(e, G)$ is the influence spread of e. It illustrates active nodes when the diffusion is stopped. Two states for nodes are to be distinguished in the social graph G: active nodes and inactive nodes. An active node u accepts a new idea and diffuses it to its inactive neighbor v to activate it. If v is influenced by the diffused content, it rediffuses it again and it is said activated. Otherwise, it is still inactive, and thus the node u can never tend to activate it another time. The process of interaction between nodes expresses the information diffusion (Liqing et al., 2020). The diffusion model is used here to describe the interaction process and behavior pattern of nodes. The influence spread of the seed set e represents the number of active nodes in G after an influence diffusion process concludes.

2.2. Diffusion models

Diffusion models are used to control influence spread in social networks. At the end of this process, active nodes are identified. In fact, when a user receives new social content (information, idea) and tends to diffuse it to his/her friends, the latter tend to re-diffuse it among their friends and the information circulates in the network. This phenomenon is termed as the influence spread and it is used to distinguish active nodes in the social network. To recall, a node is said to be active if it receives an idea and re-diffuses it again. However, if a node has not reached the influence or if it refuses the new content propagated from his/her friends he is still inactive. Node states change from inactive to active dependenly on the used diffusion model. Thus, an active node cannot become inactive, it is stills active during the diffusion process. Initially, nodes are inactive and at each iteration, active ones try to activate their neighboring according to a specified diffusion model. Two main diffusion models were proposed by Kempe et al. (2003) are

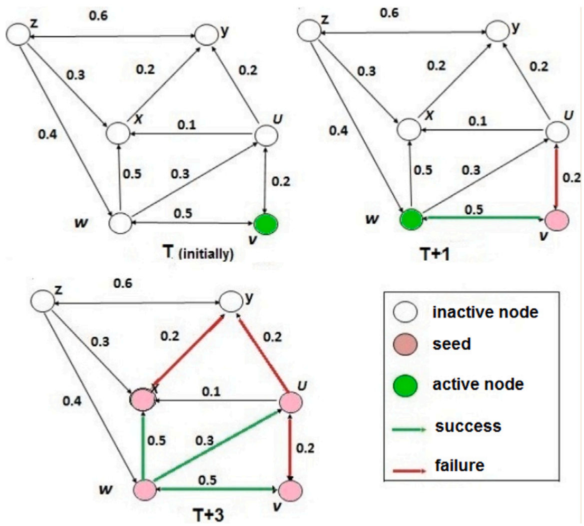


Fig. 2. Principle of the IC model.

the IC (Independent Cascade) and the LT (Linear Threshold) models. In both models, information passes from one node to another according to some probabilistic rule (Zhu et al., 2014). Thereby, in such models, each node tries to activate its neighbor with an influence probability. Consequently, to apply a diffusion model, the social graph should be weighted with influence probabilities as follows. A probability value $p_{u,v} \in [0,1]$ is assigned to the edge (u,v) (connecting nodes u and v)to represent the probability of u activating its neighbor node v. Figs. 3 and 2 show the general principle of the two well-used models where each edge is weighted by an influence probability. Initially, all nodes are inactive, then some nodes said seeds are selected to be activated and propagation starts from them. During the propagation process, an inactive node can transit to active if it adopts the new information and propagates it to its neighbors. However, an active node cannot go into the inactive state. In what follows we will describe formally the principle of the two models.

- Independent Cascade model IC:
  The principle of influence under the IC model is: an active node u has a single chance to activate its inactive neighbor v with an influence probability $\omega_{u,v}$ (Jaouadi & Ben Romdhane, 2018). For influence probabilities, most works rely on the assignment of heuristic probabilities (Li, Fan et al., 2018). For example, a probability $\omega_{uv} = \frac{1}{d_v^{in}}$ (where d_v^{in} is the in-degree of node v) (Li, Fan et al., 2018; Wang et al., 2016) is assigned to the link (u,v). Thus, all active neighbors of v have the same chance of influencing it (an equal probability). The probability that u influences v depends on the neighborhood of v. It is easier to influence low degree nodes than high degree ones. Recently, there have been works that determine influence probabilities from network data e.g. social action spread (Li, Fan et al., 2018).
  For example, as described in Fig. 2, initially at $T = 0$, the input graph is labeled by influence probabilities, and the node v is chosen as active to initiate the diffusion process. At $T = 1$, v tries to activate each of its inactive neighbors (u and w) with the specified influence probabilities that are 0.2 and 0.5 respectively. Consequently, the node v succeeds in activating w and fails to activate the node u. At the end of this iteration, a new active node appears (that is w), and the node v should be added to the seed set as it is managed to activate it. For the next iteration, the new active node w can activate its inactive neighbors and the process of diffusion continues until no inactive node can be activated. For the figure example, after three iterations the diffusion process is stopped.
  The main advantage of the IC model is that the activation process of a node is independent of its neighbors' influence. However, the chance of a node activating its inactive neighbor is only and if it does not succeed, it will no longer be able to activate it (Chen et al., 2014).
- Linear Threshold model LT:
  For the threshold model, a node v is activated if and only if the number or proportion of its already activated neighbors NB_v exceeds its threshold value θ_v (Zhu et al., 2014). For the LT model, the social pressure expresses activation. The condition of activation under this model can be described as:

$$\sum_{u \in NB_v} \omega_{u,v} \geq \theta_v \tag{2}$$

  where $\theta_v \in [0,1]$ is a random threshold and $\omega_{u,v}$ is the weight of the edge (u,v).
  θ_v controls the diffusion of information from node v. Indeed, it is easier to influence a node with a low threshold than one with a high threshold value. In this model, an inactive node can transit to an active state if it admits a sufficient number of active neighbors. For influence probabilities (weight of links), most algorithms use heuristics, for example assigning a probability to the edge (u,v) from the set $\{0.1, 0.01, 0.001\}$ randomly (Li, Fan et al., 2018).
  For example, as illustrated in Fig. 3, initially at $T = 0$, the input graph is labeled by influence probabilities and nodes threshold values. At the first iteration, the node v is chosen as active to initiate the diffusion process. Nodes u and w are inactive neighbors of v. Starting with node w which has a single active neighbor u, and as seen the influence probability of u on w surpasses its threshold value, thus node w moves into the active state and may participate in activating its inactive neighbors in the next iteration. After three iterations there are no inactive nodes that can be activated and the diffusion process is stopped. For the LT model, the chance of an inactive node becoming active increases with the number of its active neighbors. However, the random choice of threshold values can affect the diffusion process.

Several other models have been derived from LT and IC models (Jaouadi & Ben Romdhane, 2019). An improved recent version of the IC model was proposed by Ahmadi Beni et al. (2023). The classical

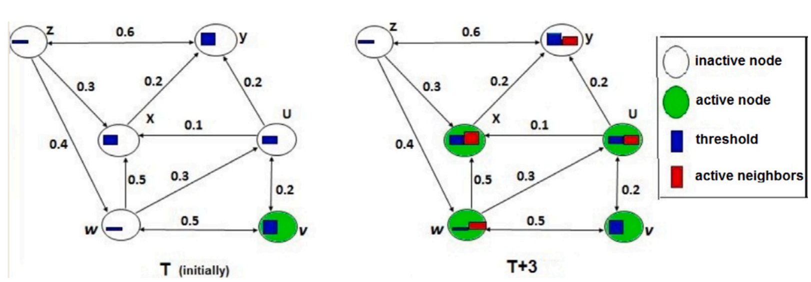


Fig. 3. Principle of the LT model.

IC model only defines the number of activated nodes. However, according to Ahmadi Beni et al. (2023), we cannot rely only on this number under the assumption that in a social network, two nodes can activate the same number of nodes i.e they have the same number of activated nodes but they do have not the same influence expansion. Consequently, the authors proposed to use the network layers for the influence spread process. Based on the K-shell algorithm, they proposed a new measure combining the number of activated nodes with the shell number. The shell number is computed under the assumption that influential nodes generally reside in the innermost layers such that the core layer. WC (Weighted Cascade model) (Wang et al., 2016) is a weighted version of the IC model. Thus, each node admits a non-negative weight which is calculated from its attributes independently of the structure of the network, and using that weight, influence probabilities are computed. Another version of IC is proposed by Tong et al. (2017) which takes into consideration dynamic social networks. This model follows the same principle as the IC model, with the difference of assuming a discrete distribution for influence probabilities instead of assigning fixed values. LT-C (Linear Threshold with colors) (Bhagat et al., 2012) is a model that includes adoption propagation in the diffusion process. Indeed, it considers that a user who has not accepted an idea may diffuse his decision. Moreover, if an individual adopts an idea or information, it includes the adherence degree of the individual to that adopted content. Therefore, influence probabilities under this model depend on two parameters which are adopting social content and liking the adopted content. PSI (Probabilistic Social Influence model) (Myungcheol & Ling, 2014) combine the best aspects of the two standard models IC and LT while ignoring certain drawbacks to propose a probabilistic distribution of social influence.

3. Influence maximization in static networks

Several methods in the literature have been developed to detect influential nodes often called seeds. However, solving this problem remains a challenging and non-trivial task. Most of the existing algorithms are parameterized, with a budget k to select exactly k seeds. Also, most methods are applied to static networks. However, real social networks' data is dynamic: networks evolve. For example, new users are registering on Facebook every day and at the same, time many users are deleting or have deactivating their accounts. In addition, users find new friends and it is also possible that they delete some of their friend list(s). Similarly, such real-world networks are distributed because of their sizes which led to the birth of huge networks and even the notion of Big Data Graphs. The state of the social network is considered in this work, as static or dynamic, whole or reduced. Therefore, existing works are presented based on network models.

To solve the IM problem in static networks, several models have been developed. These models can be categorized into two categories. The first category is based on the totality of the social network to detect influencers, and the second ones are based on a reduced version of the social network to be able to maximize influence in mega-scale networks. The following figures illustrate an example of a social graph, Email-EuAll[1] (having $n = 1005$ nodes and $m = 25\,571$ edges). Figure (a) represents the entire network and as seen, even the small size of this network, it is difficult to distinguish relations and to have a comprehensive study of the graph. Figure (b) describes the community structure of the same network. As seen, it becomes easier to distinguish the network nodes and their relations. For the last figure, a reduced version of the same network is created where some nodes and edges of the original network are kept in the new version. This reduced network is more easier to study and to distinguish influential elements. Moreover, modern social networks are huge, thus having a reduced version of the network, that retains the key properties of the original network, can help to treat such network data.

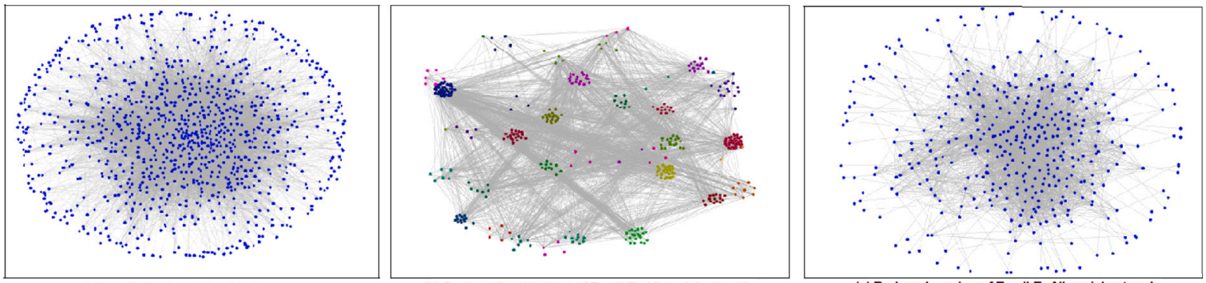
[1] https://snap.stanford.edu/data/email-Eu-core.html.

3.1. Whole network-based models

Much related research considers the IM problem on the entire network. The whole network-based solutions can be divided into four categories: (1) Greedy algorithm-based models; (2) Centrality measures-based models; (3) Community structure-based models; (4) Machine learning-based models and (5) Hybrid models. The present section is dedicated to examining the trend and progress of such models with a comparison of different approaches.

3.1.1. Greedy algorithm-based models

The greedy algorithm iterates k times to select a local optimum at each iteration to converge when choosing k seeds into a global one. In other words, based on an objective function for influence, the greedy algorithm starts by choosing the first seed node that increases the influence gain and this node is added to the seed set S. Then, it parses the rest of nodes to choose the second seed maximizing the influence when added to S and so on, the algorithm iterates until choosing k nodes. The greedy algorithm is based on Monte-Carlo simulations in order to parse the network and choose at each iteration the best node (Shang et al., 2018). Such simulations have been proven time-consuming (Taha, 2022). Several approaches have been proposed in this light. Their main purpose is to reduce Monte-Carlo simulation costs. Leskovec et al. (2007) used the submodularity property of the influence propagation function to improve the time cost gap of the greedy algorithm and they proposed the CELF (Cost-Effective Lazy Forward) algorithm. The main inefficiency of the greedy algorithm lies in the number of nodes evaluated which is quadratic using Monte-Carlo simulations. CELF solved this problem by exploiting the property of submodularity. This exploitation ensures that a node's marginal gain is reduced as the set of seeds increases. CELF++ (Goyal et al., 2011) is an improved version of CELF. It follows the same principle and gives a more appropriate and faster estimate than that of CELF by avoiding unnecessary recalculations of marginal gain. When computing the marginal gain of a node v if added to the seed set S, the CELF++ simultaneously applies the calculation of v while considering $S \cup \{u\}$ where u is the node with the maximum marginal gain. Both gain values can be computed simultaneously over a Monte-Carlo simulation. Experiments have shown that CELF++ is around 30% or even 50% faster than CELF. Although this approach avoids unnecessary calculations and decreases the time complexity of the greedy algorithm, it consumes a lot of memory (Goyal et al., 2011). To solve the memory consumption problem for large-scale social networks, Rostamnia and Kianian (2019) proposed to perform a network pre-processing step before applying the greedy algorithm. They aim to eliminate inefficient nodes from the original network using the node coverage principle. Indeed, a cover D is a subset of nodes of a given graph G where each edge has at least one of its two end nodes being part of D. After deleting inefficient nodes, the authors apply the greedy algorithm under the IC diffusion model to detect influential nodes. In the same light, to solve the memory consumption problem, another framework is proposed by Li and Liu (2019). This work aims to maximize influence based on cliques. Knowing that a clique corresponds to a complete subgraph (i.e. a subgraph in which each node is linked to all the rest of the nodes), the idea is to propose an efficient algorithm for the identification of top k influential nodes based on clique identification. The proposed algorithm is based on three steps starting with the extraction of cliques, then the selection of candidates, and finally the generation of influential nodes. The first step is based on the idea of searching for maximal size cliques. As for the second stage, candidate generation, it is based on the fact that nodes belonging to the same clique can interact easily. Therefore, the node with the highest degree in a clique is chosen as representative and the other nodes are deleted. Finally, the top k influential nodes are chosen starting with the node having the best gain. Li and Liu (2019) introduced a new concept which is influence maximization in signed networks where we distinguish a negative and a positive influence. The



(a) Email-EuAll social network (b) Community-structure of Email-EuAll social network (c) Reduced version of Email-EuAll social network

authors have defined a new diffusion model $PLID$ (Polarity-related Linear Influence Diffusion) which aims to calculate the influence of users without simulations. The main idea of this approach is, instead of having a single influence probability, unlike other models each node will have two probabilities, a probability of being positively influenced and a second of being negatively influenced. Given the input graph G and an initial set of seeds S, the calculation of the negative and positive influence depends on whether the node belongs to S.

3.1.2. Centrality measures-based models

Centrality measures are commonly used in social network analysis tasks. Indeed, for the influence maximization problem, members having high centrality scores are more likely to be adopters of content circulating in the network (Ahsan et al., 2015). Manifold measures have been proposed. For example, the CC measure (*Closeness Centrality*) (Azaouzi et al., 2021) specifies how close an individual is to others in the network. The BC measure (*Betweenness Centrality*) (Ahsan et al., 2015) determines the shortest path number between two network users. The DC measure (Jaouadi & Ben Romdhane, 2021)(*Degree Centrality*) was widely used to detect the top k highest degree nodes as influentials. The k-shell decomposition measure (Zareie & Sheikhahmadi, 2018) is another feature used to select nodes based on their topological positions. k-shell method was widely used. Another measure that has proven effective in calculating influence is the *PageRank* indicator. In the case of the Web, documents are linked by hyperlinks. The structure of the collection is therefore that of a directed graph. The PageRank indicator (PR) of a p_i page is the probability that a user following the links randomly arrives at p_i (Brin & Page, 1998). Adineh and Nouri-Baygi (2018) have proposed two heuristics to reduce the temporal complexity when solving IM problem. These heuristics improve the influence reached by the degree centrality while considering some characteristics of the network. The authors tried to demonstrate that when the node admitting the highest degree is chosen as influential, the probability of propagating the influence to its neighbors will be high. *IgnoringNeighbors*, the first heuristic, starts by selecting the high-degree node. Under the assumption that any neighboring node that is reached by the initially chosen one is going to be influenced, the idea is to remove all reached neighbors and move on to choose the next influential node from the remaining until reaching the budget k. The idea of the second heuristic *DescendingDegreeDecrease* is to decrease the degree of the neighbors reached by the high degree node. Therefore, the next influential element will be selected from the neighbors that have updated their degrees. Yang et al. (2018) have defined influence propagation in two different ways which are either to maximize influence by identifying an initial set of influencers or to minimize influence by link blocking under a certain diffusion model. Indeed, they proposed a model which considers the activation of links under the Independent Cascade model (IC). The proposed solution is essentially based on the calculation of a cost associated with each link. Therefore, based on a budget, only a certain set of links are activated. Referring to the notion of neighborhood, a two-step framework called $LAIM$ (Wu et al., 2018) is proposed. The first step is dedicated to approximate the influence and the second step is for the selection of the seed set. The proposed algorithm begins with the choice of the node admitting the highest local influence as the first seed and then it aims to eliminate this node from the graph and recalculates the local influence of the new graph nodes. Once k nodes are identified the algorithm is stopped. Based on local influence, Qiu et al. (2020) proposed a three-step approach. Initially, the nodes are ordered according to their local influence in a decreasing order and some of them are selected as sources. For the second step, the $LGIM$ (Global Selection based on Local Influence) algorithm is proposed to choose candidates from source nodes' antecedents. Based on a new objective function, the global influence is determined at a two-hop region. Finally, the last step is devoted to select influential nodes from candidates while choosing nodes with the highest marginal gain as seeds. *Bet-clus* (Saxena et al., 2023) is a recent method that combined the betweenness centrality measure and clustering coefficient to identify influential nodes. The basic idea of *Bet-clus* is to rank nodes based on their centrality values. Thus, top-k nodes are chosen as seeds. The proposed method proved the importance of graph measures for social networks' analysis. Centrality measures mentioned above identified important nodes as spreaders and generally seeds are likely to be in the same region. To overcome such issue, the VoteRank model was proposed in order to use a voting mechanism that helped to detect seeds from different areas of the network (Liu et al., 2021). For the voting mechanism, VoteRank considered only direct neighbors (one-hop neighbors) and solved the IM problem in unweighted networks. A weighted version of the voteRank model, called WVoteRank (Sun et al., 2019), was proposed to address the IM problem in weighted networks. Many extensions of such a model were proposed to consider two-hop neighbors (Kumar et al., 2022) and proved efficient and fast.

3.1.3. Community structure-based models

A community refers to a subset of individuals who interact with each other more frequently inside the community than with other individuals outside (Jaouadi & Ben Romdhane, 2021) as seen in Fig. 4. The detection of the community structure of a graph facilitates the search for influential nodes, it thus offers a means of reducing the search space. A framework proposed by Li, Gan et al. (2018) incorporated user profile during influence maximization to define different types of compliance between individuals in the network. The first phase of this framework is dedicated to selecting groups of nodes while generating a new graph in which each node corresponds to a user and can participate in one or many groups at a time. As for the second phase, it consists of choosing influential nodes from each group starting with the largest and under the assumption of choosing a single node from each group. The originality of this model lies in the fact of incorporating two types of conformity between users: friendship conformity and group conformity. Friendship conformity is considered between pairs of nodes having similar interests and residing in the same group. As for group conformity, group profiles gather information about topics that a group might follow. Therefore, the probability that a user is influenced depends on friendship conformity and group conformity. Ye et al. (2018) proposed two effective models for maximizing influence while considering the

Fig. 4. Community structure.

community structure of the network. First, the authors tried to preserve the neighborhood relations of the network and represented the graph by a vector. Then, a predetermined number of communities is detected by applying the k-means algorithm. Finally, two influence maximization models are applied to the detected communities based on the idea that if a node and its neighbors belong to the same community, its influence will be limited within its community, otherwise, the node is considered as a hub node. The notion of community diversity is introduced in a recent work (Li et al., 2020). The purpose of this proposal is to maximize the number of active nodes' communities while choosing the seed set. Thus, to compute influence probabilities, shortest paths are determined. Then, for choosing the seed set the greedy algorithm was adopted. A recent work is proposed by Umrawal et al. (2023) based on community structure. The proposed framework started by partitioning the network into disjoint communities using the well-known methods Louvain and Girvan-Newman as a first step. The second step is devoted to generate candidates for each community separately. Finally, redundant elements are removed to select the final seed set using a progressive budgeting process. Based on users' susceptibility to influence, the CIP (Murugappan et al., 2023) model was proposed. The authors proposed a four-step framework to maximize influence using the community structure of the network. First of all, communities are identified based on users' similarity which is translated by homophily and other graph measures. Then, candidates are identified in each community using the K-means algorithm. As for the third step, the authors designed a prediction model to estimate susceptibility to influence. The final step is devoted to learning susceptibility and predicting the final influence. The major idea of the proposed framework is to track influence under a few number of communities rather than track it across all ones. FIP (Bouyer et al., 2023) is a recent model designed to maximize influence while reducing the computational cost of the IM problem. The basic idea of this model is to reduce the research space of seeds by discovering the community structure of the network. The authors used the LPANNI algorithm to detect overlapping communities as a first step. They emphasized the importance of overlapping nodes to disseminate information to the different communities to which they belong. Then, candidate nodes are identified by combining the best overlapping nodes and the best non-overlapping nodes. In other words, some overlapping nodes are chosen as candidates depending on a probability coefficient of global diffusion in communities. Consequently, nodes were divided into two subsets to define such probability, a subset of nodes having connections outside the community and a subset of nodes that do have not connections outside the community. Thus, the best overlapping nodes are chosen according to their connections to diverse communities. For non-overlapping nodes, the authors combined the number of neighbors of a node, its degree, and the number of communities to which it belongs and proposed a new measure to select the best nodes. Candidate nodes are thus the union of the previously described best nodes. Candidate generation is the second step of the FIP model. As for the last step, top-k best candidate nodes are chosen as seeds.

3.1.4. Machine learning-based models

Machine learning is of great importance nowadays. Researchers include such technique in manyfold disciplines, including social networks analysis, where systems are generally represented as graphs. Graph neural networks (GNN) represent a recent powerful tool to address the IM problem especially when the problem is solved based on machine learning technique. Most machine learning-based models are focused on GNN to estimate influence among nodes and their neighbors across a multi-layer network. A recent work called GLIE (Graph Learning-based Influence Estimation) (Panagopoulos et al., 2023) was proposed to solve the influence maximization problem under the IC model. GLIE was trained on small-size data, then larger networks. The first step of this model is dedicated to using GNN and proposing a CELF version to estimate influence. As for the second step, authors proposed to improve the greedy algorithm for influence spread, and a new algorithm was proposed. The key advantage of this work is to scale over large networks. Kumar et al. (2022) employed the struc2vec node embedding algorithm to extract node features. Then, using the extracted features, a multi-layer network (GNN) was used to estimate influence under both the IC and the LT models. The main idea of the prediction phase is to exploit structural features to extract nodes' feature vectors. The top k nodes with the highest predicted influence values are chosen as seeds. In the same direction, based on GNN, DeepIM model was proposed (Ling et al., 2023). DeepIM is a two-step framework, a learning step and a selection step. The first step is dedicated to controlling the distribution of propagation under the designed GNN. Then, seeds are chosen while optimizing a proposed objective function. The major advantage of this model is its ability to be tested under multiple diffusion models including IC, LT, and SIR models. Fastcover (Ni et al., 2023) reduced the IM problem as a d-hop dominating set problem. In other words, a diffusion step is marked as a hop and the influence can be estimated by k hops neighbors number across a multi-layer network. Fastcover was proposed to solve a k budget-constrained d-hop dominating set problem and authors proposed a new formulation of the classical IM problem. A widely used and well-known application of deep learning was used by Kumar et al. (2023) to tackle IM problem. This technique is known as transfer learning which consists in transferring and using the marginal gain of a previous pre-trained algorithm on a different dataset. Consequently, the authors proposed a recent model based on transfer learning to maximize influence in large-scale networks. They used the LSTM model to label nodes in the new network with influence values obtained on the smallest network, this step was used as a prediction step. Under the idea that this pre-trained model can help to predict influence, nodes in the largest network have been ranked based on predicted influence values. To estimate such values, authors proposed to create a huge Barabasi Albert network, in which they applied the SIR as a diffusion model. Finally, the obtained results for influence spread have been used in a more larger network. Using a machine learning approach, Ahmad et al. (2023) tried to identify influential nodes that they called vital nodes. With a small size sample, the authors trained the network nodes. Then, the train data have been used under the SIR diffusion model to predict vital nodes. For node features, the proposed model incorporated structural properties such as the coreness centrality, node degree, and network connectivity. As for the training step, the authors adopted the RBF kernel as a support vector regression machine and they achieved good results for several real networks. The above machine learning-based models consider only static networks. Temporal networks were considered in Yanchenko et al. (2023) where authors claimed that the future change of a network can be obtained based on a GNN prediction phase. The purpose of this model is to select seed nodes before observing the future evolution of the network. The proposed framework started by predicting the future change of the network based on GNN for links prediction and consequently a set of temporal predicted networks are generated. After that, a greedy IM algorithm was applied in the predicted networks and the initial network was allowed to continue evolving to estimate influence while aggregating snapshots.

3.1.5. Hybrid models

Some approaches suggest a hybridization of the categories mentioned above. ComPath (Rahimnkhani et al., 2014) started by identifying community structure. Considering each community as a node, the most influential communities are chosen using the betweenness centrality. A pruning step is developed to select the highest degree nodes from each community as candidates. Finally, influential nodes are detected based on the shortest path principle. INCIM (Borzogi et al., 2016) combined the closeness centrality with the community structure in order to compute for a given user its local and global influence inside and outside its community respectively. DIN (Jaouadi & Ben Romdhane, 2016) used overlapping communities detection to exploit the structure of social networks. Then, based on the PageRank centrality measure, candidate nodes are chosen by pruning some elements with low values. $TSIM$ (Liqing et al., 2020) is a model that combined two existing approaches, DegreeDiscount (DD) (Chen et al., 2009) and CELF (Leskovec et al., 2007) for solving the IM problem. Using the DD approach, the TSIM model started by choosing candidates. Then, in order to improve the efficiency of the CELF model and to calculate candidates' marginal gain, a new measure was proposed.

3.2. Reduced network-based models

To handle large-scale networks, many works suggested to reduce their size. Depending on the manner of creating the reduced version, two classes of approaches were considered: graph sampling-based models and graph coarsening-based models.

3.2.1. Graph sampling-based models

Graph sampling is a widely used technique for reducing social networks' size. The principle aim of sampling a network is to create a representative pattern from the original one while retaining its key properties. The most simple technique for sampling is to select randomly a subset of vertices or/and edges from the initial network and to form a sampled version (Hu & Lau, 2013). Graph sampling has affected various fields such as community detection and influential nodes detection. It is considered to be a powerful tool helping the analysis of mega-scale social networks (Jaouadi & Ben Romdhane, 2021) notably in the IM problem field where several researchers have proved that using sampling techniques cannot damage the efficiency of the standard IM problem. A node sampling approach was adopted in RIS (Borgs et al., 2014) (Reverse Influence Sampling) for solving IM problem in large-scale social networks. Starting by choosing a random initial node, the sampled network is created by nodes reaching this chosen one. Indeed, the proposed model is based on two steps: selecting the start node and constructing a sample achieving the best influence. To maximize influence, RIS started by selecting R samples. Then, a start node is chosen while considering overlapping size with R. This method has achieved an important influence spread. However, computational time is still important due to the size of the sampled network. Many improved versions of RIS have been developed to tackle its drawbacks (Li, Fan et al., 2018; Tang et al., 2015, 2014; Wang, Zhang et al., 2017). $CRIS$ (Competitive Reverse Influence Sampling) (Hong et al., 2020) is developed as an improvement of the RIS model. The main idea of this model is to avoid unnecessary simulations caused by the greedy algorithm. Firstly, sketches are built using the RIS approach. Then, influence have been estimated while structuring sketches based on seeds competitive situations. At final, the greedy algorithm is used to select influential elements. Several approaches have been developed with the aim of verifying that sampling techniques do not damage diffusion properties and the influence spread of the original network. Based on the RankDegree model, Salamanos et al. (2017) starts by creating a sample network that can achieve an important influence spread that covers 80% of the network users. Experiments have demonstrated that the sampled network retains the key properties of the original one, moreover, it preserves diffusion properties. Another work (Tsugawa & Kimura, 2018) justified that a reduced version of a network is able to maximize influence. In this work, a study on the effects of sampling for the IM problem has been presented. Based on some sampling strategies, small-size sampled networks have been created. Indeed, authors tried to choose nodes having a comparable influence with that achieved in the original network. Authors succeeded in proving that the use of sampling strategies did not damage diffusion properties. In this light, MR-DSIN (MapReduce based Dynamic Selection of Influential Nodes) (Jaouadi & Ben Romdhane, 2022) was proposed recently with the advantage of treating large-scale social networks. The proposed model is based on two phases: a sampling phase and a selection phase. The particularity of such work is to consider information dynamics based on social actions. With a sample not exceeding 30% of the original network, MR-DSIN can reach an influence spread covering 80% of the network users. PIANO (Li, Xu et al., 2023) is proposed recently to exploit deep learning techniques for solving the IM problem. PIANO incorporated the machine learning technique to sample the network. Then, a learning phase is proposed to embed the network and to approximate the influence spread. Consequently, influential nodes are selected based on the learned parameters. For network embedding, the authors used the *Structure2Vec* method.

3.2.2. Graph coarsening-based models

Another research line for reducing the size of a given network is the multilevel approach or graph coarsening. The aim of coarsening is to create smaller graphs by transforming level by level the initial network into smaller ones. In other words, for a given graph G, a series of decreasing graphs' sizes are created (Jaouadi & Ben Romdhane, 2022). At a level $i+1$ the subgraph G_{i+1} is constructed based on the coarsened graph G_i obtained at the previous level i (Rhouma & Ben Romdhane, 2018).

In the context of influence maximization, Purohit et al. (2014) designed a contraction approach in which each pair of nodes is merged while minimizing the change in the first eigenvalue. This method is proven to create a reduced network that maintains the diffusion properties of the original one. MaxInf (Ohsaka et al., 2017) have the same idea of contracting nodes. Based on community detection step, nodes appearing in the same community have been considered as mutually reachable. As for the second step, each community is considered as a single node to weigh the graph. Group-IM (Group-based Influence Maximization) (Li & Liu, 2019) is a work proposed to select groups of collapsed nodes. Based on node proximity, the authors proposed a new diffusion model. Then, nodes' groups are created. Vertices having a large diffusion proximity are collapsed. Influence spread is then calculated using both the size of the created groups and the border nodes. Finally, the seed set is identified based on the greedy algorithm (Goyal et al., 2011).

4. Influence maximization in dynamic networks

Several works tried to solve The IM problem in static networks. However, social networks are dynamic and their topology and/or informational content may change over time. There exist two forms of dynamic networks. Networks in which vertices and links can evolve over time and here we talk about structure dynamics. Networks where information about users such as interests or social interactions can change and this form is about information dynamics (Jaouadi & Ben Romdhane, 2022). This section is dedicated to mentioning existing models for influence maximization in dynamic networks. The nature of dynamics, structural/informational, is considered.

4.1. Structure dynamics

A dynamic structure network is designed as a network in which nodes and/or edges can be added or deleted over time. As seen in

Fig. 5, the network evolution over time design technically an efficient probing strategy. However, it can provide an error bound of influence diffusion since the link between the two seed nodes (n_3 and n_6) may change over time. UBI (Upper Bound Interchange greedy) (Song et al., 2017) is designed to cope with structure change. Based on the *InterchangeHeuristic* measure, UBI proposed to implement node replacement in order to improve the influence. In other words, selecting the seed set at time stamp t is based on seeds chosen at the previous time stamp $t-1$ with a node replacement. The advantage of such work is to track influential nodes over time in order to choose seeds that keep maximizing influence even when the network evolves. A dynamic version of the Independent Cascade model (DIC) is developed by Tong et al. (2017) to capture the dynamics of real-world social networks. In this work, the authors defined an adaptive strategy for detecting influential nodes. They proposed a measure called *H-Greedy* which can reduce Monte-Carlo simulations resulting from the greedy algorithm. At each iteration, the adaptive proposed strategy tried to examine candidate nodes before entering the diffusion process. Then, based on the dynamic diffusion model, influential nodes are selected with an improvement of the hill-climbing model. An incremental approach *IncInf* (Liu et al., 2017) was proposed to detect influential nodes while considering past data. The structure's changes are based on the evolution of the community structure topology. Moreover, high-degree nodes are chosen to analyze the influence spread evolution. Based on nodes localization, an effective approach is designed to analyze influence propagation changes across high degree nodes. An efficient model is developed in Wang, Cuomo et al. (2019) to exploit local detection of influential nodes in dynamic networks. The main idea of this model is to update locally the influence metrics of nodes, without quantifying the influence of all nodes globally. A series of dynamic network evolutions was created by adding and deleting vertices and/or links. Then, the authors combined two metrics that are the degree centrality and the Jaccard similarity to compute influence in dynamic networks. The two chosen metrics can be calculated for each node based on local information. Smani and Megalooikonomou (2022) proposed three algorithms for maximizing influence in dynamic social networks. Including graphs' topological features, a developed version of the MATI existing model was designed in which authors had exploited structure dynamics. Then, a second variant of the same model was developed using the k-core decomposition technique. The last algorithm solved the IM problem using k-truss decomposition. Lotf et al. (2022) developed a recent model for influence maximization based on the genetic algorithm. In their work, strategic nodes and edges are selected while considering changes in social network's structure. The genetic algorithm is adopted to control the diffusion process of the IC model and select influential elements. In the same light, using the genetic algorithm with a hybridization with the ABM model, known as individual-based modeling, Li, Hu et al. (2023) proposed a distributed model that have advantage of tracking the evolution of a social network based on agents. The latter modeled dynamic networks as snapshots in which influence is estimated in a distributed manner. In Yang et al. (2023), the authors illustrated dynamic networks by an edge addition strategy and they proposed an IM model called Dynamic Edge Addition. To maximize influence, the proposed model started by generating a new sampled graph. Then, influence was estimated in the new graph using the IMM model. Finally, to capture dynamics, the last step was designed to add edges, to capture the network structure evolution, based on the shortest paths between seed nodes and other nodes in the network.

4.2. Information dynamics

The previous section is dedicated to mentioning structure dynamics models. However, dynamic networks encompasses both structural and informational change (Jaouadi & Ben Romdhane, 2018). To consider informational content, Min et al. (2020) combined three factors and proposed the Time-Sensitive Influence Maximization model $TSIM$. The state of an individual offline or online is considered as the first factor. To distinguish this state, the authors are based on tweets history for a specific time and only online users may participate in the influence spread process. The second factor is based on the idea of users' favorite topics. Indeed, considering a single subject to measure influence is insufficient. The last factor is the user interaction's delay on the diffused information. In reality, a delay between the reaction of the user and the spread of information may exist. Based on the above three factors, a dynamic version of the IC diffusion model called TTDP (Topic-based Time-sensitive Dynamic Propagation) is proposed. Finally, two algorithms are developed to exploit the local and global influence. Another informational concept is used by Wang, Fan et al. (2017) to define influence probabilities which is social actions (such as 'comment', 'retweets', etc.). Recent social actions are used to compute local influence based on the sliding window model which takes into account most recent actions. Sivaganesan (2021) proposed a parallel framework to consider social behavior while solving IM problem. A new measure combining users' interests and interactive behavior is proposed to maximize influence. A CPU architecture was used for modeling the parallel aspect. In fact, the proposed framework has considered both the network' s structure and informational content.

The study of influence maximization problem in dynamic networks is of great interest and represents a challenging task. In the review of Hafiane et al. (2020), the authors conducted an analysis of the computational cost of IM in temporal networks. Consequently, switching from static to dynamic models increases the complexity of the model as the grain of the available temporal information increases.

5. New trends in influence maximization problem

Existing models for solving the traditional IM problem have been discussed above. All mentioned models focused on a single node influence propagation. However, group-based influence maximization is a recent direction in which a group of activated nodes is specified at each iteration of the diffusion model. Unlike the classical IM problem, where under a diffusion model we try to specify a single node for propagating influence, the purpose of the GIM (Group IM) problem is to maximize the expected number of activated groups. Hence, a person's decision is easily influenced by the majority's opinions, and most of the works or decisions are done by groups (Azaouzi et al., 2021). For example, this can be seen clearly in political movements. When the majority of people support a politician, it can result in maximizing the influence on other persons to support it and thus more efficiently than a single person. Azaouzi et al. (2021) presented a formulation of the GIM problem. They tried to tackle this new trend and they presented some works in this direction. GIN (Group of Influential Nodes) (Aghaee & Kianian, 2020) model specified different groups of nodes with more connections and in order to reduce the search space, some specific nodes are selected from each group. In a recent work (Huang et al., 2022) developed an improved version of the RIS model to solve GIM problem.

The second trend to discuss in this section is the IM problem in multiplex social networks. The traditional IM problem is studied on a single social network. However, an influential node is subject to have many social networks. For example, a person who reads a book on Twitter and finds it important can share his opinion on Facebook social network. Some works (Hosni et al., 2019; Meng et al., 2022; Wang, Liu et al., 2019) studied the problem of detecting influential users in multiplex social networks.

6. Comparison

In this survey paper a detailed review of existing models for influence maximization in social networks is outlined. A technical comparison of the different models is highlighted in Table 2. Such models present some disadvantages in terms of efficiency as well as when dealing with current real social networks. For example in the Facebook

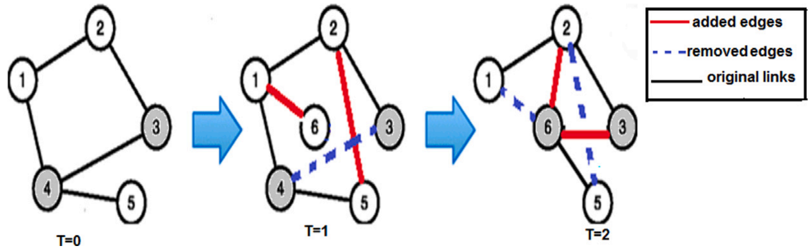


Fig. 5. Principle of probing diffusion influence in a dynamic network.

social network, information about the structure cannot be completed that is designed by distributed networks. Moreover, such a structure is able to change regularly. Another disadvantage is that the majority of models are parameterized and ignore the dynamic aspect of social networks.

7. Challenges and future directions

The influence maximization problem has been widely analyzed and several models are developed to tackle IM problem from numerous perspectives. When studied this topic, we are faced to many challenges. In fact, many aspects have been ignored although many researches have been done in this area of the analysis of social networks. In this paper, some challenges and future directions for IM problem have to be considered.

7.1. Challenges

When studied influence maximization problem, we have proved that some issues have not been covered by researchers.

- *Mutiplex social networks*: An individual is subject to have many social networks. It belongs to many social media. For example, an active person on TikTok can propagate her/his videos on Facebook or Instagram social networks. However, the existing models are applied to a specific social network.
- *Distributed social networks*: Social networks are distributed and thereby the network can be available in a continuous stream form and not in its entire form at one time. Consequently, only "partial" information about the network is available. Thus, the IM problem should be solved adaptively.
- *Influential node tracking*: Social networks evolve quickly. The rapid changes in social network topology; e.g., "friend"/"unfriend" actions are occurring instantly thereby changing dynamically the "authority" of nodes for example important connectivity changes may affect the "authority" of a node; or some influential elements may leave the social network leaving the "authority" to other ones. This challenge is termed influential node tracking and thus predicting the variations in the influence of nodes over short or long time scales is a challenging task. Unfortunately, most existing approaches cannot deal with these characteristics and either assume static networks or networks of moderate size available as a whole to the considered algorithm.
- *Privacy protection*: It is of great importance to protect users' personal information when maximizing influence (Zhang et al., 2022). Indeed, users no longer give confidence easily to diffused content even if they are influenced by that content.

7.2. Future directions

We propose in this section some research directions that should be taken into account in future works.

- *Parallel models*: In some cases, due to the distribution of networks, their huge sizes, or decentralized controls, it is not possible to access the entire network. To tackle this issue, parallel models are used. Depending on the distribution paradigm used, two forms of distribution are distinguished that are data distribution and control distribution. The data distribution allows access to several data segments for calculation. As for the distribution of control, it allows the parallel execution of the algorithm on several nodes (the case of multi-threading for example) and we can talk about parallel models. Most of the existing approaches tried to detect influential nodes in an iterative way. The issue of scalability can be tackled by developing parallel and distributed models.
- *Group-based influence maximization:* The purpose of the GIM (Group IM) problem is to maximize the expected number of activated groups of nodes, unlike the classical IM problem, where under a diffusion model a single node is chosen for propagating influence. Hence, a person's decision is easily influenced by the majority's opinions and most of the works or decisions are done by groups (Azaouzi et al., 2021). For example, this can be seen clearly in political movements. When the majority of people support a politician, it can result in maximizing the influence on other persons to support it and thus more efficiently than a single person.
- *Semantic and structure-based models*: Existing models consider either the topology or the semantics of social networks. However, in reality, and with modern social networks both the structure and link semantics describing the relationship between users are able to change over time. Such hybrid approaches should be considered in future works.
- *Machine learning-based models*: Existing survey works mention traditional techniques used to solve the IM problem. However, the machine learning technique has gained great attention in several domains including social networks analysis. For example, graph neural networks are a powerful tool used in most machine learning works to deal with graphs and to address the IM problem (Li, Gao et al., 2023). In fact, variants of machine learning and deep reinforcement learning-based models should be considered with a more exhaustive analysis in future works.

8. Conclusion

In this paper, we have presented an overview of the influence maximization problem models. Existing models were reviewed and categorized according to the status of social networks: static networks and dynamic networks. A detailed classification for both cases was presented. For static networks, five types of models were distinguished: greedy algorithm-based models, centrality measures-based models, community structure-based models, machine learning-based models, and hybrid models. For dynamic networks both structural and informational dynamic models are reviewed. This article proves that several research works have investigated the IM problem. The strengths and weaknesses of such existing models are outlined in this paper. Finally, some challenges and future directions have been addressed.

Table 2
Summary on models for influence maximization problem in social networks.

Family	Approach	Structure dynamics	Information dynamics	Parameterless	Advantages	Limits
Greedy algorithm-based models	CELF (Leskovec et al., 2007)	✗	✗	✗	- Improve the time inefficiency caused by the greedy algorithm. - Reduce the number of nodes evaluated.	- Redundancy of calculation of the marginal gain of the evaluated nodes. - Cost of Monte-Carlo simulations.
-	CELF++ (Goyal et al., 2011)	✗	✗	✗	- Avoid unnecessary recalculations of the influence. - Generate a gain in terms of execution time as well as the number of evaluated nodes.	- Memory consumption.
	Rostamnia and Kianian (2019)	✗	✗	✗	- Less memory consumption compared to the greedy algorithm. - Applicable on large-scale networks.	- No consensus between selected nodes.
	Li and Liu (2019)	✗	✗	✗	- Inclusion of a new concept: positive influence and negative influence. - New diffusion model ignoring Monte-Carlo simulations	- Sensitivity to the used parameters.
	Li and Liu (2019)	✗	✗	✗	- The inclusion of the notion of cliques brings efficiency to the resolution of the IM problem. - Reduction of memory consumption caused by the greedy algorithm.	- Significant temporal complexity generated by Monte-Carlo simulations.
Centrality measures-based models	Adineh and Nouri-Baygi (2018)	✗	✗	✗	- Reduce execution time. - Simple and intuitive.	- Possibility of loss of information when eliminating nodes.
	Yang et al. (2018)	✗	✗	✗	- Considering link activation which is a new concept. - Simple and intuitive.	- No guarantee on obtained active nodes.
	LAIM (Wu et al., 2018)	✗	✗	✗	- Linear, fast approach with large-scale social networks.	- Low influence propagation, particularly in artificial networks (small or large).
	LGIM (Qiu et al., 2020)	✗	✗	✗	- Two-step filtering approach that reduces computation time.	- Global influence is limited to a two-hop neighborhood.
	Bet-clus (Saxena et al., 2023)	✗	✗	✗	- Achieve a high influence spread based on two centrality measures.	- Limited to static topology networks.
Community structure-based models	Li, Fan et al. (2018)	✗	✗	✗	- Introduction of a new concept: user profile. - Simple and intuitive.	- Sensitivity to the used parameters.
	Ye et al. (2018)	✗	✗	✗	- Fast and efficient. - Highlighting nodes within their communities.	- Identification of influential nodes restricted to the level of the communities to which they belong.
	CDIM (Li et al., 2020)	✗	✗	✗	- Take into account the diversity of communities when searching for influential nodes. - Maintain the diversity of communities even when the community structure is unknown.	
	Umrawal et al. (2023)	✗	✗	✗	- Exploit structural properties to maximize influence. - Prove the importance of community structure for social networks' analysis.	
	CIP (Murugappan et al., 2023)	✗	✗	✗	- Introduce a new concept for influence spread which is the susceptibility. - To address the scalability issue with a parallel framework.	
	FIP (Bouyer et al., 2023)	✗	✗	✗	- Reduce computational cost of classical IM. - Emphasize topological features for detecting seeds.	
Machine learning-based models	GLIE (Panagopoulos et al., 2023)	✗	✗	✗	- GNN tool improves the network visibility.	- Limited to small size networks.
	Kumar et al. (2022)	✗	✗	✗	- Efficient for solving IM problem under IC and LT models	

(*continued on next page*)

Table 2 (*continued*).

Family	Approach	Structure dynamics	Information dynamics	Parameterless	Advantages	Limits
	DeepIM (Ling et al., 2023)	✗	✗	✗	- Scale over huge networks. - The ability to be tested under many diffusion models	
	Fastcover (Ni et al., 2023)	✗	✗	✗	- Fast machine learning solution.	
	Kumar et al. (2023)	✗	✗	✗	- Able to scale over huge networks - Reduce the complexity of the prediction phase.	
	Ahmad et al. (2023)	✗	✗	✗	- Machine learning approach applied in small size samples.	
	Yanchenko et al. (2023)	✗	✓	✗	- Machine learning approach applied in dynamic networks - A distributed model suitable for huge networks	
Hybrid models	ComPath (Rahimnkhani et al., 2014)	✗	✗	✗	- Improve the linear threshold model by a new threshold measure.	- Significant computation time due to all pair shortest paths compute.
	INCIM (Borzogi et al., 2016)	✗	✗	✗	- Demonstrate the effect of community structure detection for IM.	
	DIN (Jaouadi & Ben Romdhane, 2016)	✗	✗	✗	- Combine network structure and semantics.	- High temporal complexity.
	TSIM (Liqing et al., 2020)	✗	✗	✗	- Improvement of both DegreeDiscount and CELF approaches.	- Ignoring some weaknesses of the combined approaches.
Graph sampling-based models	RIS (Borgs et al., 2014)	✗	✗	✗	- Reduce the network size using a node sampling strategy. - The most known approach to maximize influence with a graph sampling strategy.	- Sample size's gap.
	CRIS (Hong et al., 2020)	✗	✗	✗	- Fast version of the RIS model. - Conserve diffusion properties of the initial network.	- Unable to treat large-scale social networks.
	PIANO (Li, Xu et al., 2023)	✗	✗	✗	- Using a new concept for maximizing influence: Deep Learning. - To deal with large-scale networks and to achieve good values for influence spread.	
Graph coarsening-based models	Purohit et al. (2014)	✗	✗	✗	- Linear time complexity for solving IM problem. - Conserve diffusion properties while contracting nodes.	- More suitable for community structure detection.
	MaxInf (Ohsaka et al., 2017)	✗	✗	✗	- Consider the network topology for solving the IM problem. - Maintain diffusion properties.	- Scalability problem.
	group-IM (Li & Liu, 2019)	✗	✗	✗	- Consider a group of nodes in the diffusion process instead of a single node.	- A greedy approach is used to detect seeds.
Structure dynamics-based models	UBI (Song et al., 2017)	✓	✗	✗	- Address IM problem in dynamic networks. - Improve influence convergence.	- Admit a complexity of $O(a^2m)$ in a graph of m nodes each having an average degree a which makes it important for modern large-scale networks.
	DIC (Tong et al., 2017)	✓	✗	✗	- Adaptive selection of influential nodes. - Control detected nodes to maintain their influence. - Take into account dynamic networks	- Greedy approach that ignores the distribution of current networks.
	IncInf (Liu et al., 2017)	✓	✗	✗	- Detect seeds while using past data. - Consider community structure topology's change for modeling dynamic aspect.	- Only structure dynamics is considered. - Analyze influence propagation changes using only high degree nodes.

(*continued on next page*)

Table 2 (*continued*).

Family	Approach	Structure dynamics	Information dynamics	Parameterless	Advantages	Limits
	Wang, Cuomo et al. (2019)	✓	✗	✗	- Update influence probabilities based on local detection of influential nodes. - Simple for computing nodes influence in dynamic networks.	- Even its simplicity, only local influence is considered.
	Lotf et al. (2022)	✗	✓	✗	- Reduce the cost of solving IM problem in large-scale networks by considering snapshot graphs.	- Consider only structure dynamics.
	Yang et al. (2023)	✗	✓	✗	- Consider a new aspect for the network change which is edge addition.	- Only topological features are used.
Information dynamics-based models	Min et al. (2020)	✗	✓	✗	- Propose a new dynamic diffusion model based on information flow. - Take into account the dynamics of information. - Include three factors modeling dynamics: user state, preferred topics, and interaction delay.	
	Wang, Fan et al. (2017)	✗	✓	✗	- Dynamic selection of influential nodes based on social actions. - Using social actions has reduced the computational cost of the greedy approach. - The notion of the checkpoint has improved the result by taking into account the expiration of certain actions and the appearance of others over time.	
	Sivaganesan (2021)	✗	✓	✗	- Parallel framework designed to cope with informational aspect. - Combine users' interests and interactive behavior has a great influence on IM.	
	MR-DSIN (Jaouadi & Ben Romdhane, 2022)	✗	✓	✓	- Detect a minimal seed set that covers more than 80% of the initial network. - Able to cope with mega-scale social networks.	- Need to improve computational time.

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.

Data availability

No data was used for the research described in the article.

References

Adineh, M., & Nouri-Baygi, M. (2018). Maximum degree based heuristics for influence maximization. In *2018 8th international conference on computer and knowledge engineering* ICCKE, (pp. 256–261).

Aghaee, Z., Ghasemi, M., Beni, H., Bouyer, A., & Fatemi, A. (2021). A survey on meta-heuristic algorithms for the influence maximization problem in the social networks. *Computing 103*, 2437–2477.

Aghaee, Z., & Kianian, S. (2020). Influence maximization algorithm based on reducing search space in the social networks. *SN Applied Sciences*, *2*(12), 1–14.

Ahmad, A. R., Munoz, J., Jalili, M., & Khayyam, H. (2023). A machine learning-based approach for vital node identification in complex networks. *Expert Systems with Applications*, *214*, Article 119086.

Ahmadi Beni, H., Bouyer, A., Azimi, S., Rouhi, A., & Arasteh, B. (2023). A fast module identification and filtering approach for influence maximization problem in social networks. *Information Sciences*, *640*, Article 119105.

Ahmed, N., Neville, J., & Kompella, R. (2011). *Network sampling via edge-based node selection with graph induction: Department of computer science technical reports. Paper 1747*.

Ahsan, M., Singh, T., & Kumari, M. (2015). Influential node detection in social network during community detection. In *2015 international conference on cognitive computing and information processing* CCIP, (pp. 1–6).

Al-Garadi, M. A., Varathan, K. D., Ravana, S. D., Ahmed, E., Mujtaba, G., & Khan, M. U. S. (2018). Analysis of online social network connections for identification of influential users: Survey and open research issues. *ACM Computing Surveys*, *51*(1), 1–37.

Azaouzi, M., Mnasri, W., & Romdhane, L. (2021). New trends in influence maximization models. *Computer Science Review*, *40*, Article 100393.

Bhagat, S., Goyal, A., & Lakshmanan Laks, V. S. (2012). Maximizing product adoption in social networks. In *Proceedings of the fifth ACM international conference on web search and data mining* (pp. 603–612). ACM.

Bian, R., Koh, G., & Divoli, A. (2020). Identifying top-k nodes in social networks: A survey. *ACM Computing Surveys*, *52*(1), 1–33.

Borgs, C., Brautbar, M., Chayes, J., & Lucier, B. (2014). Maximizing social influence in nearly optimal time. In *Proceedings of the 25nd annual ACM-SIAM symposium on discrete algorithm*. SODA.

Borzogi, A., Haghighi, H., & Zahedi, M. S. (2016). INCIM: A community-based algorithm for influence maximization problem under the linear threshold model. *Information Processing and Management*, *52*(6), 1188–1199.

Bouyer, A., Ahmadi Beni, H., Arasteh, B., Aghaee, Z., & Ghanbarzadeh, R. (2023). FIP: A fast overlapping community-based influence maximization algorithm using probability coefficient of global diffusion in social networks. *Expert Systems with Applications*, *213*, Article 118869.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, *30*(1), 107–117, Proceedings of the Seventh International World Wide Web Conference.

Chen, W., Wang, Y., & Yang, S. (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD conference on knowledge discovery and data mining*. KDD'2009.

Chen, Y., Zhu, W., Peng, W., Lee, W., & Lee, S. (2014). CIM: Community-based influence maximization in social networks. *ACM Transactions on Intelligent Systems and Technology*, *5*(2), 1–31.

Goyal, A., Lu, W., & Lakshmanan, L. (2011). CELF++: Optimizing the greedy algorithm for influence maximization in social networks. In *Proceedings of the 20th international conference companion on world wide web* (pp. 47–48).

Hafiane, N., Karoui, W., & Ben Romdhane, L. (2020). Influential nodes detection in dynamic social networks: A survey. *Expert Systems with Applications (2020)*, *159*.

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.

Hosni, A., Li, K., & Ahmad, S. (2019). Minimizing rumor influence in multiplex online social networks based on human individual and social behaviors. *IEEE Transactions on Cybernetics*.

Hu, P., & Lau, W. C. (2013). A survey and taxonomy of graph sampling. CoRR.

Huang, P., Longkun Guo, L., & Zhong, Y. (2022). Efficient algorithms for maximizing group influence in social networks. *Tsinghua Science and Technology*, *27*(5).

Jabari, J., Abdollahi Azgomi, M., & Ebrahimi Dishabi, M. R. (2022). An improved influence maximization method for social networks based on genetic algorithm. *Physica A. Statistical Mechanics and its Applications*, *586*, Article 126480.

Jaouadi, M., & Ben Romdhane, L. (2016). DIN: An efficient algorithm for detecting influential nodes in social graphs using network structure and attributes. In *2016 IEEE/ACS 13th international conference of computer systems and applications* AICCSA, (pp. 1–8).

Jaouadi, M., & Ben Romdhane, L. (2018). DSIN : Dynamic Selection of Influential Nodes in social networks using network attributes. In *The 15th international conference on applied computing 2018* AC2018, (pp. 261–268).

Jaouadi, M., & Ben Romdhane, L. (2019). Influence maximization problem in social networks: An overview. In *2019 IEEE/ACS 13th international conference of computer systems and applications*. AICCSA.

Jaouadi, M., & Ben Romdhane, L. (2021). A distributed model for samplig large scale social networks. *Expert Systems with Applications*, *186*.

Jaouadi, M., & Ben Romdhane, L. (2022). A graph sampling based model for Influence Maximization in large-scale social networks. *IEEE Transactions on Computational Social Systems*, *186*.

Kempe, D., Kleinberg, J., & Tardos, E. (2003). Maximizing the spread of influence through a social network. In *In proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining*. KDD 03.

Kumar, S., Mallik, A., Khetarpal, A., & Panda, B. (2022). Influence maximization in social networks using graph embedding and graph neural network. *Information Sciences*, *607*, 1617–1636.

Kumar, S., Mallik, A., & Panda, B. (2023). Influence maximization in social networks using transfer learning via graph-based LSTM. *Expert Systems with Applications*, *212*, Article 118770.

Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., & Glance, N. (2007). Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining* KDD '07, (pp. 420–429).

Leskovec, J., & Krevl, A. (2014). SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data.

Li, J., Cai, T., Deng, K., Wang, X., Sellis, T., & Xia, F. (2020). Community-diversified influence maximization in social networks. *Information Systems*, *92*, 1–12.

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

Li, Y., Gan, X., Fu, L., Tian, X., Qin, Z., & Zhou, Y. (2018). Conformity-aware influence maximization with user profiles. In *2018 10th international conference on wireless communications and signal processing* WCSP, (pp. 1–6).

Li, Y., Gao, H., Gao, Y., Guo, J., & Wu, W. (2023). A survey on influence maximization: From an ML-based combinatorial optimization. *Association for Computing Machinery*, *17*(9).

Li, W., Hu, Y., Jiang, C., Wu, S., Bai, Q., & Lai, E. (2023). ABEM: An adaptive agent-based evolutionary approach for influence maximization in dynamic social networks. *Applied Soft Computing*, *136*, Article 110062.

Li, D., & Liu, J. (2019). Modeling influence diffusion over signed social networks. *IEEE Transactions on Knowledge and Data Engineering*, 1–9.

Li, H., Xu, M., Bhowmick, S., Shafik Rayhan, J., Sun, C., & Cui, J. (2023). PIANO: Influence maximization meets deep reinforcement learning. *IEEE Transactions on Computational Social Systems*.

Ling, C., Jiang, J., Wang, J., Thai, M., Xue, L., Song, J., Qiu, M., & Zhao, L. (2023). Deep graph representation learning and optimization for influence maximization.

Liqing, Q., Chunmei, G., Shuang, Z., Xiangbo, T., & Mingjv, Z. (2020). TSIM: A two-stage selection algorithm for influence maximization in social networks. *IEEE Access*, *8*, 12084–12095.

Liu, P., Li, L., Fang, S., & Yao, Y. (2021). Identifying influential nodes in social networks: A voting approach. *Chaos, Solitons & Fractals*, *152*, Article 111309.

Liu, X., Liao, X., Li, S., Zheng, S., Lin, B., Zhang, J., Shao, L., Huang, C., & Xiao, L. (2017). On the shoulders of giants: incremental influence maximization in evolving social networks. *Complexity*.

Lotf, J., Azgomi, M. A., & Ebrahimi Dishabi, M. R. (2022). An improved influence maximization method for social networks based on genetic algorithm. *Physica A. Statistical Mechanics and its Applications*, *586*.

Meng, Y., Chen, N., Yi, Y., Wang, S., & Pei, C. (2022). Research on the dynamic multisocial networks influence maximization problem based on common users. *IEEE Access*, *9*.

Min, H., Jiuxin Cao, J., Yuan, T., & Liu, B. (2020). Topic based time-sensitive influence maximization in online social networks. *World Wide Web 23*.

Murugappan, V., Pamidighantan, P., Subramanian, S., & Santos, E. (2023). CIP: Community-based influence spread prediction for large-scale networks. In *IEEE international and distributed processing symposium workshops* (pp. 858–867).

Myungcheol, D., & Ling, L. (2014). Probabilistic diffusion of social influence with incentives. *IEEE Transactions on Services Computing*, *7*(3), 387–400.

Ni, R., Li, X., Li, F., Gao, X., & Chen, G. (2023). FastCover: An unsupervised learning framework for multi-hop influence maximization in social networks. CoRR.

Ohsaka, N., Sonobe, T., Fujita, S., & Kawarabayashi, K. (2017). Coarsening massive influence networks for scalable diffusion analysis. In *Proceedings of the 2017 ACM international conference on management of data* (pp. 635–650).

Panagopoulos, G., Tziortziotis, N., Vazirgiannis, M., & Fragkiskos, D. M. (2023). Maximizing influence with graph neural networks.

Purohit, M., Prakash, B. A., Kang, C., Zhang, Y., & Subrahmanian, V. (2014). Fast influence-based coarsening for large networks. In *Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining* KDD '14, (pp. 1296–1305).

Qiu, L., Tian, X., Sai, S., & Gu, C. (2020). LGIM: A global selection algorithm based on local influence for influence maximization in social networks. *IEEE Access*, *8*, 4318–4328.

Rahimnkhani, K., Aleahmed, A., Rahgozar, M., & Moeini, M. (2014). A fast algorithm for finding most influential people based on the linear threshold model. *Expert Systems with Applications*, *42*(3), 1353–1361.

Razis, G., Anagnostopoulos, I., & Zeadally, S. (2021). Modeling influence with semantics in social networks: A survey. *ACM Computing Surveys*, *53*(1), 1–38.

Rhouma, D., & Ben Romdhane, L. (2018). An efficient multilevel scheme for coarsening large scale social networks. *Applied Intelligence*.

Rostamnia, M., & Kianian, S. (2019). Vertex cover preprocessing for influence maximization algorithms. In *2019 5th conference on knowledge based engineering and innovation* KBEI, (pp. 338–342).

Salamanos, N., Voudigari, E., & Yannakoudakis, E. (2017). Deterministic graph exploration for efficient graph sampling. *Social Network Analysis and Mining*, *7*.

Saxena, R., Jadeja, M., & Vyas, P. (2023). An influence maximization technique based on betweeness centrality measure and clustering coefficient (Bet-clus). In *The 15th international conference on computer and automation engineering* (pp. 565–569).

Shang, J., Wu, H., Zhou, S., Zhong, J., Feng, Y., & Qiang, B. (2018). IMPC: Influence maximization based on multi-neighbor potential in community networks. *Physica A. Statistical Mechanics and its Applications*, *512*, 1085–1103.

Sivaganesan, D. (2021). Novel influence maximization algorithm for social network behavior management. *Journal of ISMAC*, *3*, 60–68.

Smani, G. I., & Megalooikonomou, V. (2022). Maximization influence in dynamic social networks and graphs. *Array*, *15*, Article 100226.

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.

Sun, H., Chen, D., He, J., & Chng, E. (2019). A voting approach to uncover multiple influential spreaders on weighted networks. *Physica A. Statistical Mechanics and its Applications*, *519*, 303–312.

Taha, K. (2022). Identifying the top-k influential spreaders in social networks: A survey and experimental evaluation. *IEEE Access*, *10*.

Tang, Y., Shi, Y., & Xiao, X. (2015). Influence maximization in near-linear time: A martingale approach. In *Proceedings of the 2015 ACM SIGMOD international conference on management of data* (pp. 1539–1554).

Tang, Y., Xiao, X., & Shi, Y. (2014). Influence maximization: Near-optimal time complexity meets practical efficiency. CoRR abs/1404.0900.

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

Tsugawa, S., & Kimura, K. (2018). Identifying influencers from sampled social networks. *Physica A. Statistical Mechanics and its Applications*, *507*, 294–303.

Umrawal, K., Christopher, J., & Aggarwal, V. (2023). A community-aware framework for social influence maximization. *IEEE Transactions on Emerging Topics in Computational Intelligence*.

Wang, S., Cuomo, S., Mei, G., Cheng, W., & Xu, N. (2019). Efficient method for identifying influential vertices in dynamic networks using the strategy of local detection and updating. *Future Generation Computer Systems*, *91*, 10–24.

Wang, Y., Fan, Q., Li, Y., & Tan, K.-L. (2017). Real-time influence maximization on dynamic social streams. *Proceedings of the VLDB Endowment*, *10*, 805–816.

Wang, S., Liu, J., & Jin, Y. Y. (2019). Finding influential nodes in multiplex networks using a memetic algorithm. *IEEE Transactions on Cybernetics*.

Wang, Y., Wang, H., Li, J., & Gao, H. (2016). Efficient influence maximization in weighted independent cascade model. In S. Navathe, W. Wu, S. Shekhar, X. Du, S. Wang, & H. Xiong (Eds.), *Lecture notes in computer science*: *vol. 9643*, *Database systems for advanced applications. DASFAA 2016*.

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

Wang, C., Zhao, J., Li, L., Jiao, L., Liu, J., & Wu, K. (2023). A multi-transformation evolutionary framework for influence maximization in social networks. *IEEE Computational Intelligence Magazine*.

Wu, H., Shang, J., Zhou, S., Feng, Y., Qiang, B., & Xie, W. (2018). LAIM: A linear time iterative approach for efficient influence maximization in large-scale networks. *IEEE Access*, *6*, 44221–44234.

Yanchenko, E., Murata, T., & Holme, P. (2023). Link prediction for ex ante influence maximization on temporal networks. *Applied Network Science*, *8*(70).

Yang, W., Brenner, L., & Giua, A. (2018). Influence maximization by link activation in social networks. *vol. 1*, In *2018 IEEE 23rd international conference on emerging technologies and factory automation* ETFA, (pp. 1248–1251). IEEE.

Yang, S., Song, J., Tong, S., Chen, Y., Zhu, G., Wu, J., & Liang, W. (2023). Extending influence maximization by optimizing the network topology. *Expert Systems with Applications*, *215*, Article 119349.

Ye, Y., Chen, Y., & Han, W. (2022). Influence maximization in social networks: Theories, methods and challenges. *Array*, *16*, Article 100264.

Ye, F., Liu, J., Chen, C., Ling, G., Zheng, Z., & Zhou, Y. (2018). Identifying influential individuals on large-scale social networks: A community based approach. *IEEE Access*, *6*, 47240–47257.

Zareie, A., & Sheikhahmadi, A. (2018). A hierarchical approach for influential node ranking in complex social networks. *Expert Systems with Applications*, *93*, 200–211.

Zhang, X., Wang, J., Ma, X., Ma, C., Kan, J., & Zhang, H. (2022). Influence maximization in social networks with privacy protection. *Physica A. Statistical Mechanics and its Applications*, *607*, Article 128179.

Zhu, T., Wang, B., Wu, B., & Zhu, C. (2014). Maximizing the spread of influence ranking in social networks. *Information Sciences*, *278*, 535–544.