Expert Systems With Applications 266 (2025) 125990

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Review

A theoretical review on multiplex influence maximization models: Theories, methods, challenges, and future directions

Oumaima Achour *, Lotfi Ben Romdhane

MARS Research Lab, LR 17ES05, Higher Institute of Computer Science and Communication Techniques of Hammam Sousse (ISITCom), University of Sousse, Street khalifa karoui sahloul 4, Sousse, 526-4002, Tunisia

ARTICLE INFO

Keywords:
Influence maximization
Seed users
Propagation models
Multiplex social networks

ABSTRACT

Online social networks (OSNs) have become an integral part of our daily lives, shaping the way social relationships evolve. Influence maximization (IM) in OSNs has been widely studied by various researchers during the last two decades due to its wide range of applications including viral marketing, public health, recommendation systems, disease spread and prevention, etc. It is known as the problem of selecting a small set of users with size $\leq k$ known as *seed users* from a social network that can maximize the spread of information, influence, or behavior within the network. Recently, online users have joined multiple OSNs simultaneously such as Facebook, Twitter, and Instagram, thereby creating a new complex environment for the IM problem called *multiplex network*. Therefore, studying the IM problem in multiplex social networks, where the same set of users engage in various social networks simultaneously represents a new research direction that researchers have started to explore. While numerous surveys have explored IM algorithms from different perspectives, they primarily focus on traditional methods designed for single networks, overlooking multiplex networks. To fill this gap, in this paper, we present a theoretical review of recent IM algorithms and solutions proposed to solve the IM problem in multilayer networks, with a particular focus on multiplex networks. To build the foundation of the IM problem in multiplex networks, we start by presenting the IM problem in its basic form and reviewing the well-used propagation models and their extensions in single-layer networks to investigate then how they are adapted to cope with the diffusion process within multiplex networks. Next, we present the multiplex IM problem and provide a comprehensive taxonomy of the different existing algorithms proposed to solve it. Afterward, a comparative analysis illustrated by a comprehensive table outlining the strengths and weaknesses of each model is presented. Finally, diverse applications of multiplex IM, highlighting emerging trends and future directions that researchers can consider within this topic are provided.

Contents

* Corresponding author.
E-mail addresses: achour.oumayma@isitc.u-sousse.tn (O. Achour), Lotfi.BenRomdhane@isitc.u-sousse.tn (L.B. Romdhane).

https://doi.org/10.1016/j.eswa.2024.125990
Received 20 June 2024; Received in revised form 15 November 2024; Accepted 28 November 2024
Available online 5 December 2024

1. Introduction

Online social networks (OSNs) have witnessed a rapid evolution over the years as shown in Fig. 1.[1] This evolution has had a significant impact on society, shaping our lives at both personal and professional levels. In 2021, over 4.26 billion people were using social media worldwide, a number projected to increase to almost six billion in 2027.[2] This tremendous usage has offered many great opportunities for researchers to solve many real-world existing problems such as information diffusion (Xu & Liu, 2010), community detection (Moosavi & Jalali, 2014), influence maximization (Peng et al., 2018), recommendation systems (El Kouni, Karoui, & Romdhane, 2020), etc. Recently, influence maximization (IM) (Kempe, Kleinberg, & Tardos, 2003) has gained research attention due to its wide range of applications such as viral marketing, rumor control, epidemic spreading, etc. It aims to identify a small set of individuals known as *seed users* with size $\leq K$ whose actions or opinions can maximize the spread of information, ideas, or behaviors within the network. In fact, the IM problem presents the backbone of the information diffusion process, where viral trends, news stories, and rumors rely on influential nodes to propagate.

In recent years and coinciding with the level of maturity reached by the information and communication technology, OSNs have experienced a radical shift, where the early days of social media were characterized by relatively simple uniplex[3] networks, in which individuals were connected mainly based on a single relationship such as friendships or follows, and information diffusion was largely driven by user interactions within a single-layer network (Amara, Taieb, & Aouicha, 2022). However, as technology has advanced and user behavior has evolved, real-world social interactions have transcended the confines of single-layer networks. For instance, nowadays, most individuals maintain a diverse array of online social accounts such as Facebook, Twitter, Instagram, and even more, allowing them to share the same information across different platforms simultaneously. This phenomenon creates a profoundly interconnected network structure known as *multiplex network*. Multiplex social networks are complex social networks, where the same set of individuals interact with each other across several social networks as illustrated in Fig. 2. These complex networks represent a significant shift in the way people interact and share information, where users engage with each other across multiple social contexts denoted as layers, each of which represents a different aspect of their online presence. These contexts might include social connections (Facebook, Instagram, etc.), interests (music, movies, etc.), professional affiliations (university, company, etc.), or even geographical locations. While the IM problem within traditional single-layer social networks has been a subject of interest for extensive researchers such as in Aghaee and Kianian (2020), Fan, Wang, Zhang, Zhao, and Rui (2024), Singh and Kailasam (2021), Wu and Pan (2017), Yang et al. (2024) and Zareie and Sakellariou (2024), it becomes more challenging and interesting when we shift it to multiplex networks due to their multilayer structure. The IM problem in multiplex social networks has emerged recently as a new and important research direction known as multiplex influence maximization (MxIM), where there has been a growing endeavor to develop solutions (Fig. 4), yet it is still in its infancy stage.

1.1. Challenges and issues

Various challenges and issues have emerged along with the MxIM problem at both theoretical and practical levels, setting it apart from its counterpart which has been proposed for single-layer networks. Some main and fundamental challenges faced when solving the MxIM problem include the increased complexity in modeling the propagation process due to the complex structure of networks. To this end, adapting existing models of information propagation, such as the linear threshold (LT) or independent cascade (IC), to account for multiplex influence propagation, or devising an entirely new model, is not a trivial task. Additionally, the computational complexity nature of the problem is another challenge, where Kempe et al. (2003) proved that finding seed users is NP-hard under classical diffusion models (Singh, Srivastva, Verma, & Singh, 2021). Similar to the traditional IM problem, the MxIM problem remains an NP-hard problem (Keikha, Rahgozar, Asadpour, & Abdollahi, 2020). Beyond these above two challenges, the large size of social networks that reached billions of users, and the inherent link between network size and computational complexity is a pivotal challenge encountered with the MxIM problem. Therefore, achieving a near-optimal solution for the MxIM problem under large-scale social networks is pretty tough. Also, various issues have emerged along with the MxIM problem. For instance, most of the existing solutions for identifying influential users in multiplex networks with large scales struggle with time inefficiency and scalability. These limitations significantly influence the quality of the selected seed nodes. Besides, social networks are dynamic by nature, where new users join and others dissolve, and interactions among them evolve with time (Han et al., 2017). As a result, the topology of social networks undergoes rapid changes, such as *"friend"* or *"unfriend"* actions, which dynamically affect the influence ability of users and subsequently impact the propagation process.

[1] https://www.broadbandsearch.net/blog/complete-history-social-media.
[2] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/.
[3] In this paper, we will use the terms uniplex, monopolex, and, single-layer network interchangeably.

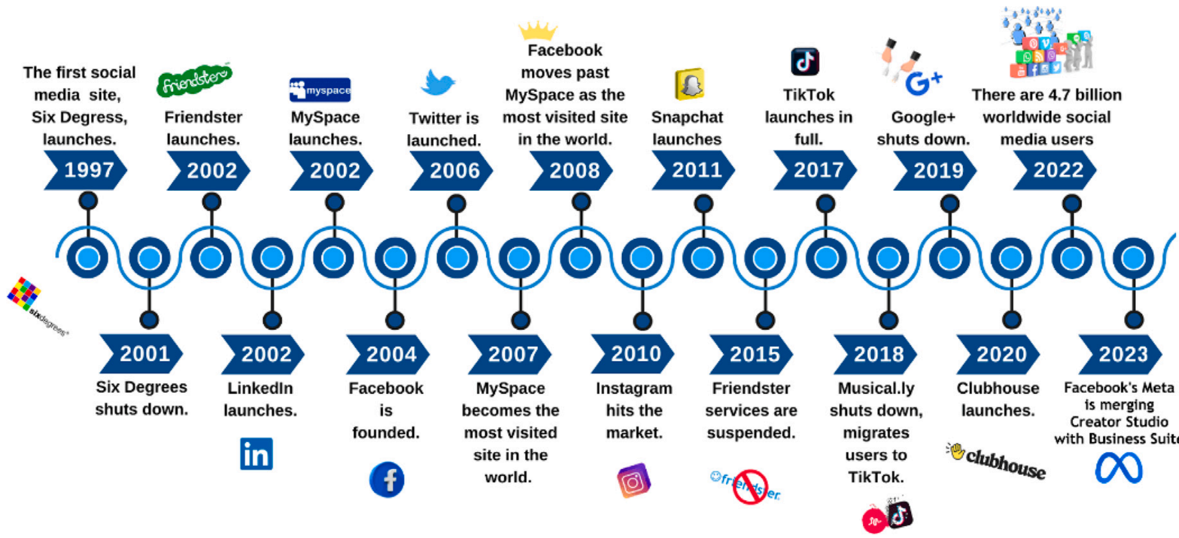

Fig. 1. OSNs evolution over 1997–2023[1].

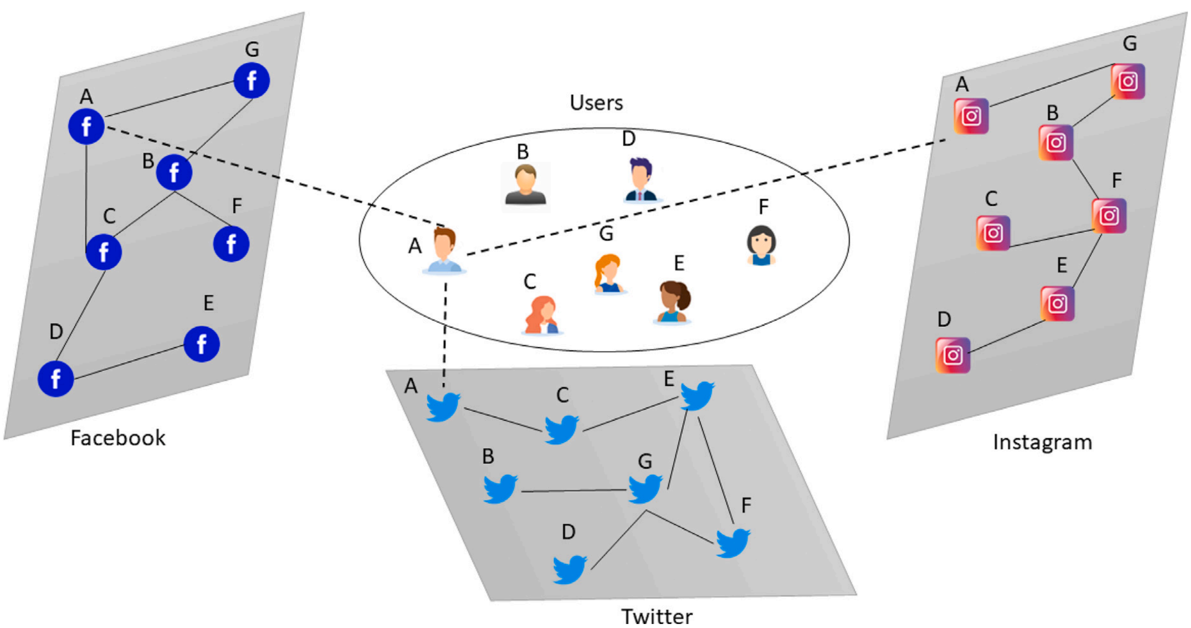

Fig. 2. An illustration of a multiplex network with 7 users and 3 social networks: Facebook, Twitter, and Instagram (Hosni, Li, & Ahmad, 2020).

1.2. Existing reviews

Reviews on the IM problem aim to group existing models into broad categories (Jaouadi & Romdhane, 2024). Many reviews such as in Azaouzi et al. (2021), Banerjee et al. (2020), Jaouadi and Romdhane (2024), Nesrine et al. (2020) and Singh et al. (2021) have been conducted on classical IM frameworks for single social networks taking into account various issues including user privacy protection, large-scale social networks, network dynamic topology, etc. However, these reviews often consider the multiplicity aspect of OSNs. For instance, recently, Jaouadi and Romdhane (2024) introduced a detailed survey on IM approaches, with a particular focus on static and dynamic models. They classified both models into two main families: whole network-based and reduced network-based models for static models and structure dynamics and information dynamics for dynamic models. While the paper offers a valuable discussion on dynamic models from both topological and informational perspectives, it would benefit from exploring different forms of the problem, such as competitive IM and MxIM. In Nesrine et al. (2020), Hafiene et al. presented a detailed review of IM models in dynamic social networks. They provided a comprehensive classification of various IM methods across three distinct network models such as static, snapshot, and dynamic. With a major focus on dynamic models, they discussed several aspects of dynamic approaches such as incremental, interval, and contact sequence. Although the paper provides valuable insights into the diverse methodologies employed for IM models across different types of dynamic networks, it ignores the informational-based dynamic models along with the multiplex models. The paper of Azaouzi et al. (2021), presented a detailed review of node-based and group-based IM models. Besides, they discussed the IM problem under the privacy protection issue. The main advantage of this survey is the introduction and discussion of a new concept for the IM problem: group-based influence. However, it is limited to static network models, overlooking other IM-based models such as dynamic-based and multiplex network-based models. Banerjee et al. (2020) introduced an extensive study of several methodologies of the IM problem. Moreover, they discussed different variants of

Table 1
A summary of IM-related surveys discussed in our paper.

Survey	Year	Keywords	IM aspect	Citation score	Selection criteria
Jaouadi and Romdhane (2024)	2024	– IM in static social networks – IM in dynamic social networks – IM in reduced networks	– Static IM models – Structural-based dynamic IM models – Informational-based dynamic IM models	2	– Recent
Nesrine, Wafa, and RomdhaneLotfi (2020)	2020	– IM in dynamic social networks	– Topological-based dynamic IM models	61	– Most cited – Coverage IM dynamic models
Azaouzi, Mnasri, and Romdhane (2021)	2021	– IM problem – Influential group detection – Privacy protection issue	– Node-based IM models – Group-based IM models – IM models under privacy protection – Parallel IM-based models	51	– Most cited. – Coverage various IM aspects
Banerjee, Jenamani, and Pratihar (2020)	2020	– Social IM problem – Approximation algorithm – Greedy strategy	– Various variants of IM-based models	216	– Most cited – Coverage various IM aspects
Li, Fan, Wang and Tan (2018)	2018	– Context-aware IM problem – IM in social networks	– Greedy-based IM models – Context-aware-based IM models	606	– Most cited
Sumith, Annappa, and Bhattacharya (2018)	2018	– IM in large social networks	– Approximation algorithms – Heuristic-based IM models	46	– Discuss heuristics and guarantee-based models
Bródka, Musial, and Jankowski (2020)	2020	– Spreading process in multilayer networks – Influence diffusion in multilayer networks		51	– Discuss the diffusion process in multilayer networks
Singh et al. (2021)	2021	– IM in social networks – Multiplex IM – Multiple IM	– Classical IM models – Multiplex IM models – Multiple IM models – Multiple IM across multiplex networks models – Context-aware IM models	40	– A detailed survey about the different forms of the IM problem – Discuss the MxIM problem

this problem along with its hardness results in both traditional and parameterized complexity frameworks. Li, Fan et al. (2018) focused on greedy algorithms of IM, where they classified greedy methods into three main families: simulation-based, proxy-based, and sketch-based. More precisely, they provided an overview of existing studies that exploit how the solution methodologies are used to solve the context-aware IM. The main objective of such a review is to provide a clear classification of existing methods and to illustrate new trends for the IM problem. However, this paper only limits their study to the greedy solutions, neglecting other existing categories such as meta-heuristic, game theory, etc. In the same line, Sumith et al. (2018) discussed this problem from various approaches such as models, heuristics, and parameters, where they classified heuristic approaches into four main families: structural features, constraint-specific, multiple competitors, and network and embedding methods. Nevertheless, this paper only provides a superficial overview, lacking any discussion on broader IM-related topics. As we can see from the surveys discussed above, research on IM approaches has primarily focused on single-layer networks. However, real-world interactions are seldom that simple (Mishra, Singh, Kumar, & Biswas, 2022). Recently, a few studies have begun to present surveys on IM in multilayer networks. For example, Bródka et al. (2020) provided an overview of the different types of spreading processes over networks, where they classified it into main four groups: single spread in single-layer networks, single spread in multilayer networks, multi-spread in single layer networks, and multi-spread in multilayer networks. Besides, they introduced a comprehensive study of existing approaches in the area of multiple spreading processes in multilayer networks. In the same line, Singh et al. (2021) offered a detailed review of recent methods for the IM problem and its variants, including multiplex IM models. However, the paper gives less attention to multiplex models, with only a limited number of related works being considered. In nutshell, most of the existing surveys considered only single-layer networks while multilayer ones are neglected, and even multilayer networks are considered, the multiplex networks are ignored (Bródka et al., 2020). To the best of our knowledge, our survey represents a pioneering effort within the area of IM problems under multiplex networks. We aim to provide a detailed review of the existing methods presented so far to solve the MxIM problem. In Table 1, we provide a brief overview of the discussed surveys, focusing on well-recognized and impactful works as indicated by citation scores, and prioritizing recent surveys to reflect the latest advancements and trends in the IM field.

1.3. Motivation and contributions of the review

Although in-depth reviews of the state-of-the-art methods in several directions of the IM problem are conducted, they do not show the IM problem taking into consideration the multiplicity feature of the real-world social networks. As far as we know, this is the first attempt to analyze the existing multiplex-based IM models. Therefore, the key goals of our survey are four-fold:

1. Introduce an overview of the IM problem in social networks.
2. Elaborate the difference between the IM problem in traditional single-layer social models and multiplex models.
3. Conduct a systematic review of multiplex IM-based models.
4. To expand perspectives and offer a forward-thinking outlook in IM research, we identify new trends and future directions in this research area.

To achieve the above goals, this paper introduces, categorizes, and analyzes the existing models for IM in multilayer networks, with a

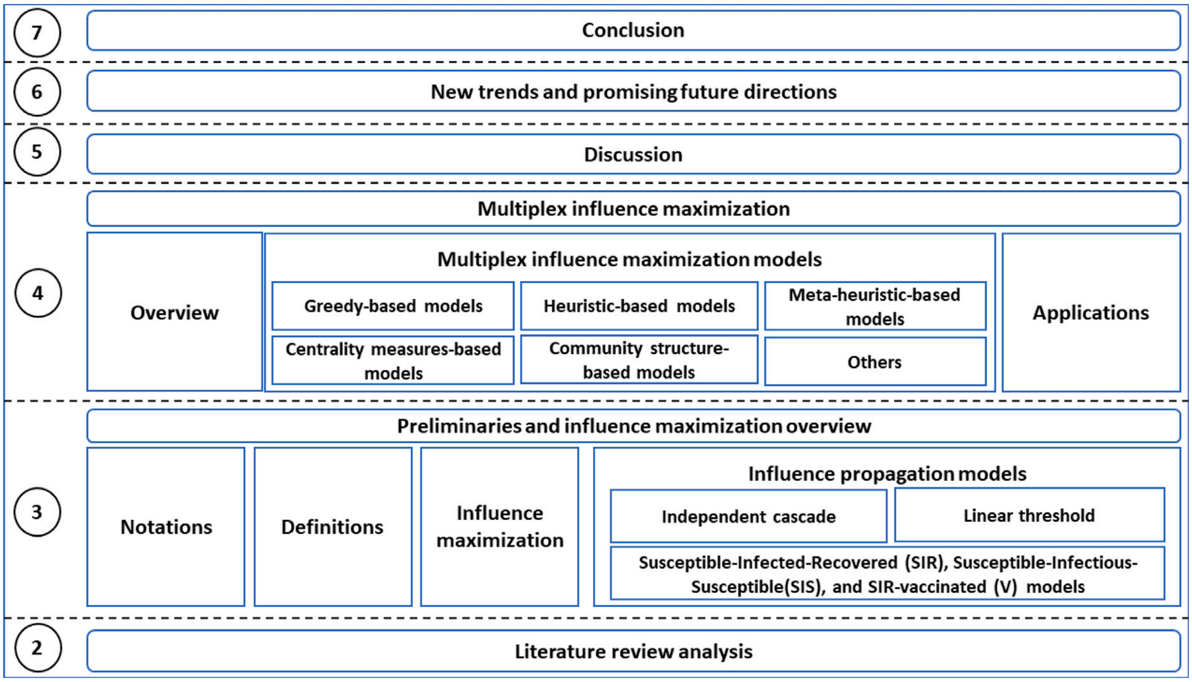

Fig. 3. Survey structure.

particular focus on multiplex networks. To this end, the main contributions of our review are as follows. First, we introduce the IM problem and discuss its main properties and concepts within its basic form. Second, we formally present the MxIM problem. Then, we conduct a comprehensive and systematic review of the existing models for this problem followed by its different applications. Third, we discuss the advantages and limits of the analyzed papers. Finally, we present future directions on this topic.

1.4. Review structure

The present paper is organized into several sections as depicted in Fig. 3 to comprehensively explain the radical transformation of IM problem to multilayer networks. Section 2 is dedicated to introducing the research methodology adopted to extract the analyzed papers. Section 3 presents the preliminary knowledge including related concepts and definitions of the IM problem in social networks along with its basic form and diffusion models. Section 4 introduces the MxIM problem where our contribution takes place. A discussion of the analyzed papers in terms of advantages and limitations and other key features is presented in Section 5. In Section 6, new trends and promising future directions are presented. Section 7 concludes the review.

2. Literature review analysis

To ensure a high-quality and comprehensive review, we present a detailed selection phase for the papers used to study influence maximization within multiplex social networks in this paper. To find relevant papers, we used the most popular academic search engines: IEEE Xplore, Scopus database, ScienceDirect, ACM Digital Library, arXiv, and Google Scholar, and started with the search for influence maximization in multiplex social networks. Our research used a set of explicit and straightforward keywords to make sure that we did not skip any related paper on this topic. Initially, we used broad keywords like “social influence”, “spreading process”, “multiplex networks”, “network of networks”, “heterogeneous networks”, “multidimensional networks”, “information diffusion”, “diffusion process”, “influence maximization”, and “multiplex influence maximization”. Then, as our research progressed, we used more specific and informative terms like “influence maximization in multiplex social networks”, “influence spread in multilayer networks”, “information diffusion in multidimensional networks”, “diffusion process in multiplex networks”, “epidemic spreading in multilayer networks”, “multiplex social influence analysis”, “influential nodes detection in multi-relation networks”, and “spreading process over complex networks”. Our search strategy combined keywords related to the influence spreading process and multiplex network structure to ensure the retrieval of relevant papers. The search phase gave us over 216 papers. After a preliminary screening, this number was reduced to 106 papers that met the eligibility criteria. The preliminary screening excluded papers were based on the title alone we were able to say that they do not fall in the IM in multiplex networks (only consider single-layer models). Applying our eligibility criteria, we further excluded 55 papers that did not directly address multiplex influence maximization. This assessment was based on a review of the abstracts, introductions, and methodologies of these papers. After that, we reached 51 publications that we used in this survey. Fig. 4 illustrated the number of relevant selected papers published each year since 2015.

3. Preliminaries and influence maximization overview

3.1. Notations

Notations and abbreviations used in this paper are listed in Tables 2 and 3, respectively.

3.2. Definitions

Definition 3.1 (*Social Networks*)**.** Social networks are online platforms that allow users to create and maintain personal profiles, connect with other users, and share various types of content, such as text, photos, videos, etc. Kim and Hastak (2018). A social network can be conceptualized as an abstract network of interconnected nodes, often represented by a graph $G = (V, E)$ which can be directed, undirected,

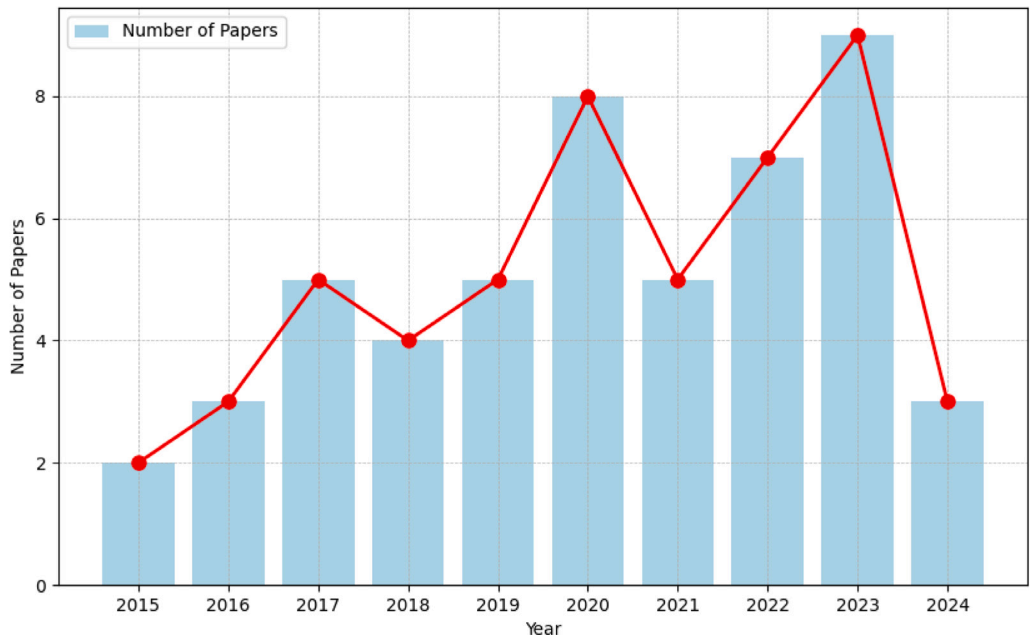

Fig. 4. Research on multiplex influence maximization problem, 2015–2024.

Table 2
List of notations.

Notation	Description
$G = (V, E)$	A social graph G with a set of nodes V and a set of edges E.
$N(x)$	Set of direct neighbors of node x.
$N_{in}(x)$	Set of in-neighbors of node x.
$N_{out}(x)$	Set of out-neighbors of node x.
$w(u, v)$	Influence weight of edge(u, v).
C	A community within the graph G.
V_C	A subset of nodes of community C.
S	Seed set.
$state(x)$	The states of node x;i.e active or inactive.
M	Influence propagation model.
$\sigma(S)$	Influence spread of S.
G_L	A multiplex graph of L-layers.
E_i	Set of edges in layer i.
A	Adjacency matrix.
$dist(x, b)$	Shortest distance between x and b.
$\sigma(a, b\|x)$	The number of the shortest paths between a and b in which x exists.
$\sigma(a, b)$	The number of the shortest paths between a and b.

Table 3
List of abbreviations.

Abbreviation	Description
OSNs	Online social networks
IM	Influence maximization
MxIM	Multiplex influence maximization
MIM	Multiple influence maximization
ST	Stochastic threshold model
RIM	Rumor influence minimization
SNs	Social networks

weighted, or unweighted, where V indicates the set of users within the network and $E \subseteq V \times V$ indicates the relationship connecting them such as friendship, follower, family, colleagues, authorship, etc.

Definition 3.2 (*Node Neighbor*)**.** In a graph, a neighbor of a given node x is any other node y that is directly connected to it by an edge (x, y). We denote by $N(x)$ the set of direct neighbors of x which is defined as $N(x) = \{y \in V; (x, y) \in E\}$. In a directed graph, a node has two types of neighbors: in-neighbors (N_{in}) and out-neighbors (N_{out}). In-neighbors are nodes that have outgoing edges directed towards the given node, while out-neighbors are nodes that have incoming edges directed from the given node.

Definition 3.3 (*Influence Graph* Singh, Kumar, Singh and Biswas, 2019)**.** An influence graph is a directed-weighted graph $G = (V, E, W)$ that model the influence relationships between nodes. Here, an edge (u, v) indicates the influence of node u on node v, and the edge-weight W represents the influence weight between nodes.

Definition 3.4 (*Community* Bedi & Sharma, 2016)**.** A community denoted as C, is a subgroup of nodes V_C in graph G, where $V_C \subseteq V$, that are densely connected within C than outside.

Definition 3.5 (*Seed Nodes*)**.** Seed nodes are defined as a small subset of nodes $S \subseteq V$, such that $|S| \leq k$, that takes part in the information diffusion process within the network and serves as the source of the dissemination (Singh et al., 2021). Ideally, selecting those nodes leads to a cascade effect, where information progressively spreads from them to other nodes in the network, ultimately reaching a large outbreak size.

Table 4
A summary of multilayer networks discussed in this paper (Kivelä et al., 2014).

Network type	Description	Example of applications
Multilayer network (Cardillo et al., 2013)	General structure with multiple layers, where each layer can have different nodes and edges, but typically some interconnections between layers.	Financial systems, transportation networks, ecological networks, etc.
Multiplex networks (De Domenico et al., 2013)	A type of multilayer network where each layer represents a distinct type of relationship between the same set of nodes.	Social networks with various social interaction types between the same individuals
Networks of networks (Zhou, Zemanová, Zamora, Hilgetag, & Kurths, 2006)	A collection of networks, often with distinct node sets, connected through inter-network links.	Infrastructure networks (e.g., transport, power), biological systems (e.g., cellular, genetic), etc.
Heterogeneous multilayer networks (Liu et al., 2024)	Consists of multiple types of nodes and edges, with each layer representing different node and edge types or attributes.	Knowledge graphs, semantic networks, Healthcare Systems, etc.

Definition 3.6 (*Node States*)**.** In the information diffusion process, each node $x \in V$ can belong in two states: *active* or *inactive*. The former indicates that the node either belongs to seed set S or has been influenced by an already activated node, while the latter indicates that the node is not influenced by its active neighbors. Here, we modeled the node states of node x as a binary variable, denoted as $state(x) \in \{0, 1\}$ such that,

$$state(x) = \begin{cases} 1\ if\ x\ in\ active\ state \\ 0\ if\ x\ in\ inactive\ state \end{cases} \tag{1}$$

Definition 3.7 (*Influence Spreading Model*)**.** A mathematical model, denoted as *M*, models the random process of information spreading within the network. It is used to simulate and understand the spread of influence, ideas, information, or behaviors in a network. The influence spreading model describes how one individual's actions or decisions can influence the decisions or behaviors of others, often in the context of social networks, marketing, and epidemiology.

Definition 3.8 (*Influence Spread*)**.** Influence spread, denoted as $\sigma(S)$ represents the number of activated (or influenced) nodes by S at the end of the diffusion process. It is simulated based on an information diffusion model (Azaouzi et al., 2021).

Definition 3.9 (*Multiplex Network Bródka et al., 2020*)**.** Multiplex networks are a special case of multilayer networks, composed of multiple layers, each representing a distinct type of interaction between the same set of nodes (Basaras, Iosifidis, Katsaros, & Tassiulas, 2017). A multiplex graph can be modeled as a collection of L-graphs denoted by $G_L = \{(V, E_i) | i \in \{1, \dots, L\}\}$, where V is the set of nodes representing the users of the multiplex social network, which is the same across all the layers ($V_i = V_j = V$, where $i \neq j$ *and* $(i, j) \in \{1, \dots, L\}$), and $E_i \subset V \times V$ represents the set of links in layer i. In particular, when $L = 1$, multiplex networks will degenerate into single-layer networks. In Table 4, we present the different multilayer networks discussed in this paper.

3.3. Influence maximization

In viral marketing applications, companies frequently aim to maximize product adoption through online platforms. To achieve this goal, they leverage the potential of influential users, by strategically providing them with free samples, companies hope these users will not only enjoy the product but also share their positive experiences with their network and their network share their positive feedback with their network of networks, and so on. This can trigger a cascade of influence, ultimately leading to a significant sales boost (Zhang, Nguyen, Zhang, & Thai, 2015). Selecting the best influencers for distributing free items is not a trivial task. This problem is known as the *influence maximization*. In its basic form, it aims to select *top-k* influential users from a social network to maximize the influence spread within it. Mathematically, given a social network $G = (V, E)$, a seed set, denoted as S $\sqsubseteq$ V, $\sigma(S)$ is the number of nodes expected to be activated by S under a certain influence propagation model M. Here, we formally define the basic form of the IM problem.

Definition 3.10 (*Influence Maximization (IM) Kempe et al., 2003*)**.** Given a social network $G = (V, E)$, an information spreading model M, and a positive integer $k < |V|$, the IM problem looks to select an optimal seed set $S \sqsubseteq V$ with $\leq k$ nodes such that the expected influence of S under M, is maximized.

$$\sigma_{G,M}(S) = argmax_{|S| \leq K} \sigma_{G,M}(S) \tag{2}$$

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

Property 3.1 (*Monotone Increasing Singh et al., 2021*)**.** *The expected influence spread function $\sigma(S)$ is said to be monotone increasing iff $\sigma(S) \leq \sigma(A)$, such that $S \subset A$.*

Property 3.2 (*Diminishing Return Singh et al., 2021*)**.** *The expected influence spread function $\sigma(S)$ is said to be diminishing return iff $\sigma(S \cup v) - \sigma(S) \geq \sigma(A \cup v) - \sigma(A), \forall v \in A$ and $S \subset A$*

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

3.4. Influence spreading models

The influence of any seed set is defined based on the information diffusion process among users. Intuitively, spreading models are widely used to study the spread of information, behaviors, and diseases in social networks. They are defined as discrete processes since they occur in discrete time steps. Generally, in any spreading model, if an individual adopts a new idea or product, it becomes active; otherwise, it is inactive. Besides, each individual in the process can only alter from an inactive to an active state and remain active in the rest of the diffusion process. The spreading process can be analyzed using various metrics, such as the expected number of nodes activated by a given seed set or the expected time for the entire network to become activated. So far, several influence propagation models have been proposed to simulate the underlying influence propagation process, including the two most popular models: the linear threshold (LT) model and the independent cascade (IC), as well as others such as the susceptible–infected–recovered (SIR) model, etc. In the subsequent section, we present the widely used diffusion models in the IM problem, along with their extensions.

3.4.1. Linear threshold model

The linear threshold (LT) model was first proposed by Granovetter (1978) in 1978, in which each individual has an associated activation threshold value λ that reflects its susceptibility to accept the information. In the LT model, an individual is supposed to be either in an active or inactive state. It works as follows: Initially, each node u selects a threshold λ_u in the interval $[0,1]$ uniformly at random, which describes the stochastic process of diffusion. Then, at each time step t, each inactive node v becomes active if the sum of weights of its incoming active neighbors u (activated nodes at time step $t-1$) exceeds its activation threshold λ_v i.e.

$$\sum W(u,v) \geq \lambda_v \tag{3}$$

and the process runs until no more activations are possible. With the LT model, an inactive node is influenced by all of its active neighbors at each time step. An active node influences its inactive neighbors according to the associated weights. The basic idea behind the LT is that a user can switch its status from inactive to active if a *"sufficient"* number of its incoming neighbors are active.

Based on the LT model, a wide range of extensions have been created so far such as the majority LT model (Bhagat, Goyal, & Lakshmanan, 2012) where instead of a simple linear sum as in the LT model, it considers whether a majority of a node's neighbors have adopted the information or behavior. If the majority threshold is met, the node adopts the information or behavior; otherwise, it remains unaffected. In Saito, Kimura, Ohara, and Motoda (2012), the authors introduced the Asynchronous LT (AsLT) model that incorporates asynchronous time delay. It involves defining time-delay parameters for each link in a network. The propagation process unfolds in continuous time, where active nodes influence their neighbors with specific time delays drawn from exponential distributions. When the accumulated influence from active nodes reaches a certain threshold for a node, that node becomes active and subsequently influences its neighbors with associated delays. In He, Song, Chen, and Jiang (2012), a Competitive LT (CLT) model is introduced to capture competitive influence. In this model, nodes can be activated positively or negatively. CLT uses three states: inactive, positively active (+active), and negatively active (-active). Transitions occur only from inactive to one of the active states. In Zhong, Srivastava, and Leonard (2017), a Weighted LT (WLT) model is created which allows for varying strengths of influence along network links. Unlike the LT model, in which all edges are assumed to have equal weight, the WLT model accommodates different influence levels for each link. In Gursoy and Gunnec (2018), the deterministic LT model is introduced. In this model, threshold values are predefined as inputs to the model instead of assuming a random threshold function. In Chan, Ning, and Zhang (2020), a non-progressive LT model is presented which considers the users' non-progressive behavior. Unlike the LT model, it enables active nodes to become inactive in subsequent time steps. In Yang et al. (2023), a Comparative LT (Com-LT) model is developed to describe the spread properties for multiple non-competitive products. Each node is associated with a set of thresholds for each product. Then, when a node adopts one of the products, its threshold for the other product is reduced, representing the complementary.

Theorem 3.1. *The IM problem is NP-hard under the LT model (Kempe et al., 2003).*

Theorem 3.2. *The expected influence spread function $\sigma(S)$ is submodular under LT model (Kempe et al., 2003).*

3.4.2. Independent cascading model

A classical stochastic diffusion model introduced by Goldenberg, Libai, and Muller (2001) in 2001 is called the independent cascading (IC) model, where the randomness resides on the node activation process, which depends on the influence probabilities on edges. In this model, each edge(u,v) is associated with a probability $p(u,v)$, which reflects the influence probability of node u on node v. As in the LT model, in the IC model, there are only two types of node states: active or inactive. The IC model works as follows. Initially, a seed set S_0 is in the active state. At time step t, each node u in S_0 independently attempts to activate its every inactive neighbor v with a probability $p(u,v)$ at time $t+1$. The diffusion process stops when no more activation is possible. The idea behind the IC model is that an active node in time t has only a single chance to activate its inactive neighbors at probability p in time $t+1$. Some extensions of the IC model have been made in the literature, such as the dynamic IC (DIC) model (Tong, Wu, Tang, & Du, 2016), which captures the uncertainty of the diffusion process and the dynamic aspects of real-world social networks. In the DIC model, nodes can be activated more than once. In Ding, Sun, Wu, and Guo (2020), a Realistic IC (RIC) model is proposed where different nodes have different acceptance probabilities that follow a certain distribution, and propagation probabilities between two nodes are different, which also follow a certain distribution. In Wu and Pan (2017), Multi-Campaign IC (MCIC) and Campaign-Oblivious IC (COIC) models are introduced. These models are derived from the basic IC model, describing competitive propagation processes when good information propagates competitively with bad information. In Kim, Lee, and Yu (2014), the Continuously activated and Time-restricted IC (CT-IC) model is presented, incorporating time constraints in the activation process. In this model, the active node activates its neighbors repeatedly, and activation continues until a given time. In Wang, Jin, Yang and Cheng (2017), an emotion-based IC model is proposed to study the process of sentiment spreading. Indeed, numerous improved models have been proposed based on the IC model.

Theorem 3.3. *The IM problem is NP-hard under the IC model (Kempe et al., 2003).*

Theorem 3.4. *The expected influence spread function $\sigma(S)$ is submodular under IC model (Kempe et al., 2003).*

3.4.3. Susceptible–Infected–Recovered (SIR), Susceptible–Infectious–Susceptible (SIS), and SIR-Vaccinated (SIRV) models

These models, known as epidemiological models, are used to study the spread of diseases and infections within populations. The first model is the Susceptible–Infected–Recovered (SIR) model, introduced by Kermack and McKendrick in 1927, where the set of individuals is divided into three compartments: susceptible (S), infected (I), and recovered (R). The susceptible individuals are those who have never been informed of the information and can propagate it; the infected individuals are those who receive and share the information; and the recovered individuals are those who have immunity to the information (i.e., stop sharing the message). In the SIR model, a node can be in one of the three states: S, I, or R. This model works as follows: At each time step, an I-state node informs all its susceptible neighbors with a rate λ and then recovers with β rate. The process stops until there is no $I-state$ node in the network. The second model is the Susceptible–Infectious–Susceptible (SIS) model, where an individual has two possible states: susceptible (S) and infectious (I), which are similar to the S- and I-states of the SIR model. In the SIS model, susceptible individuals will be infected and remain infected for a period of time called the individual's infectious period. Then, at the end of this period, the individual returns to a susceptible state. The third model is the SIRV model, which is the same as the SIR model with an additional state called vaccinated (V). Nodes in the V-state will neither be infected nor transmit disease.

4. Multiplex influence maximization

4.1. Overview

Most of the early existing studies have primarily focused on identifying influential users within single-layer social networks, overlooking the multiplex dimension of these networks. Multiplex models offer a clear and comprehensive understanding of real-world social interactions and relationships by accounting for the diverse ties that individuals maintain in their real lives. Therefore, shifting the IM problem to multiplex models is a challenging task due to the existence of multiple layers. Let us reconsider the viral marketing application, as users engage across multiple social networks simultaneously, companies have adapted their marketing strategies by promoting their products across these various platforms rather than focusing on just one. To achieve this, companies must select a set of influential users who have the most influence over all networks. In this regard, identifying the influential nodes within multiplex social networks is a new extension of the IM problem, and the multiplex influence maximization (MxIM) problem is presented. Singh et al. (2021) proposed in their survey a general framework of the MxIM problem as depicted in Fig. 5. Unlike the classical IM problem (Fig. 3), the MxIM problem consists of four main elements: Overlapping users identification, node alignment process, coupling strategy, and seed selection strategy. The overlapping users identification step is used to identify users who are actively involved in multiple networks. These users are known as *overlapping users*. The node alignment process is the process of reassigning a universal identification number (Uid) to every node in each network so that every overlapping user has the same Uid across networks. In general, every network has its own node-naming system. Therefore, it might be possible that a user has different identification in the different networks (Singh, Singh, Kumar and Biswas, 2019). Next, to propagate information across social networks, there is a need for a network coupling strategy based on the homogeneity and heterogeneity nature of the networks. After that, a seed selection method can be applied to identify the influential users. Mathematically, the MxIM problem is defined as follows:

Definition 4.1 (*Multiplex Influence Maximization (MxIM)*)**.** Given a multiplex social network $G_L = \{(V, E_i) | i \in 1, \ldots, L\}$ with L layers, an information spreading model M, and a budget $k < |V|$, MxIM problem looks to select a subset $S \sqsubseteq V$ of size at most k, such that the expected spread of influence of S in G_L ($\sigma(G_L)$) is maximized (i.e. maximizes the influence across all layers).

$$\sigma(G_L) = \bigcup_{i=1}^{L} \sigma(G_{i,M}, S) \tag{4}$$

where, $\sigma(G_{i,M}, S)$ is defined in Eq. (2), which determines the expected number of nodes influenced in layer i by the seed set S after applying the diffusion model M, and the objective function $\sigma(G_L)$ is the total expected influence spread across all layers, which is sub-modular.

Lemma 4.1. *$\sigma(G_L)$ is sub-modular iff: $\forall S \subseteq A \subseteq V$ and $\forall u \in V$,*

$$\Delta(S, u) \geq \Delta(A, u) \tag{5}$$

where,

$$\Delta(S, u) = \bigcup_{i=1}^{L} (\sigma(G_{i,M}, S \cup \{u\}) - \sigma(G_{i,M}, S)) \tag{6}$$

and

$$\Delta(A, u) = \bigcup_{i=1}^{L} (\sigma(G_{i,M}, A \cup \{u\}) - \sigma(G_{i,M}, A)) \tag{7}$$

Proof. Let S be an arbitrary seed set and $\sigma(G_L)$ is the total expected influence spread defined as the union of influence spreads across L-layers,

$$\sigma(G_L, S) = \bigcup_{i=1}^{L} \sigma(G_{i,M}, S) \tag{8}$$

where $\sigma(G_{i,M}, S)$ indicates the expected influence spread in layer G_i given the seed set S, which is submodular by the influence maximization results in single-layer networks, as proved by Kempe et al. (2003). Consequently, because the function $\sigma(G_L)$ is defined as the union of submodular functions $\sigma(G_{i,M}, S)$ across layers. Submodularity is preserved under the pointwise union function when each function satisfies the diminishing returns property.

Specifically, the submodularity of $\sigma(G_L)$ implies that, for any $S \subseteq A \subseteq V$ and any $u \in V$,

$\Delta(S, u) \geq \Delta(A, u)$,

The union function maintains submodularity as each layer's contribution satisfies diminishing returns, and the combination of submodular functions in this context remains submodular. Thus, $\sigma(G_L)$ is submodular as it is a combination of submodular influence spread functions across the multiplex network layers.

4.2. Multiplex influence maximization models

Several methods in the literature have been developed to find influential nodes (Jaouadi & Romdhane, 2024). Solving this problem remains a challenging and non-trivial task. While most of the existing solutions are applied to single-layer networks, real-world social networks are often multiplex, where users participate in multiple social networks simultaneously. This section discusses the existing works on identifying influential nodes in multilayer social networks, with a primary focus on multiplex models. In this direction, we review various approaches proposed for this purpose based on their adopted seed selection strategy. In our review, we classified the multiplex IM-based models into five families (Fig. 6): (1) Greedy algorithm-based models, (2) Heuristic based-models, (3) Meta-heuristic-based models, (4) Centrality measures-based models, (4) Community structure-based models, and (5) others.

4.2.1. Greedy algorithm-based models

Greedy algorithms, which have been extensively used in IM for single-layer networks due to their simplicity and effectiveness, have also been extended to address the MxIM problem. These extensions aim to handle the challenges of influence propagation across multiple interconnected layers, while retaining the core advantages of greedy methods. The greedy algorithm typically goes through k iterations, selecting a local optimum each time, aiming to converge towards a global optimum by choosing k seeds (Jaouadi & Romdhane, 2024). It starts with an empty set of influential nodes and iteratively adds nodes to the set based on their potential to maximize the influence spread. At each iteration, the algorithm selects the node that can maximize the increase of influence gain, given the current set of influential nodes, and so on. This process continues until k nodes have been selected. The greedy algorithm relies on Monte Carlo simulations to analyze the network and select the best node at each iteration (Heidari, Asadpour, & Faili, 2015). In practice, it is time-consuming and is not scalable to large networks, especially for multiplex networks due to the increased complexity of simulating influence across multiple layers. In this respect, several approaches have been proposed to reduce Monte-Carlo simulation costs. This section thoroughly reviews the different greedy solutions proposed so far to solve the multiplex IM problem. Early solutions, such as the study by Zhang et al. (2015), presented the least cost influence (LCI) problem by incorporating budgetary constraints into the MxIM problem. LCI aims to select influential users at minimal costs. In this regard, an improved greedy algorithm is proposed to maximize the influence in multiplex networks while ensuring minimal costs. Here, the influence propagates independently within each layer under the LT model, where a user becomes active if the total influence from its active in-neighbors exceeds its threshold in *some layer networks*, and can only propagate to other networks via overlapping users. Additionally, to represent the multilayer structure, the authors mapped

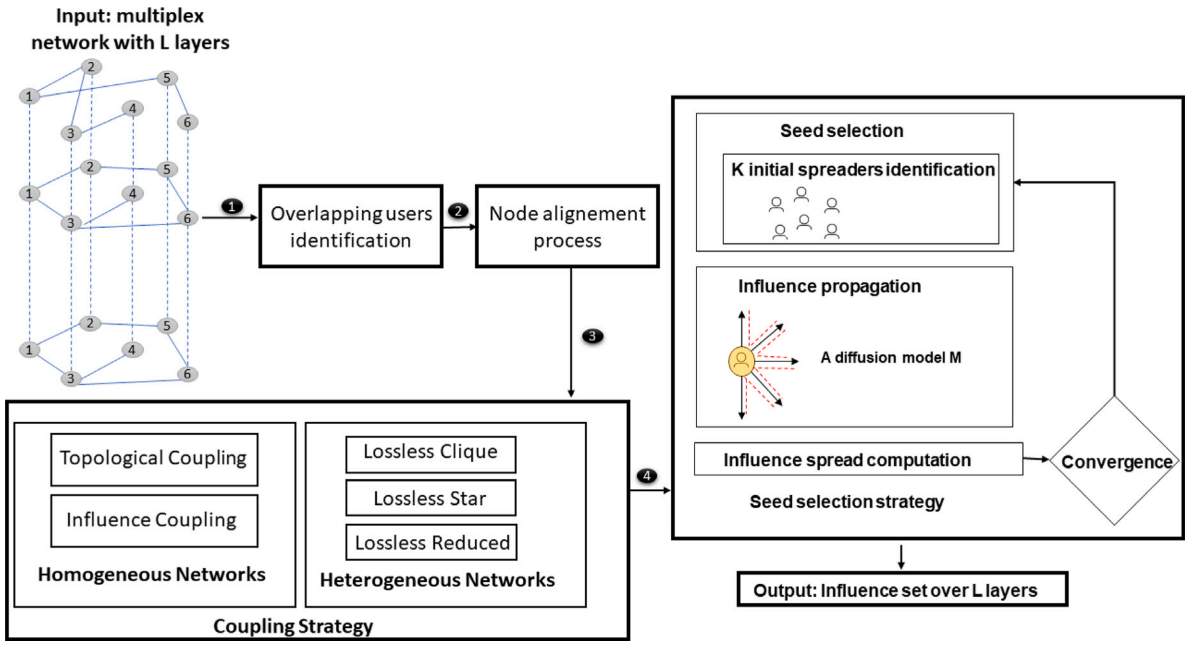

Fig. 5. The general framework of the MxIM problem (Singh et al., 2021).

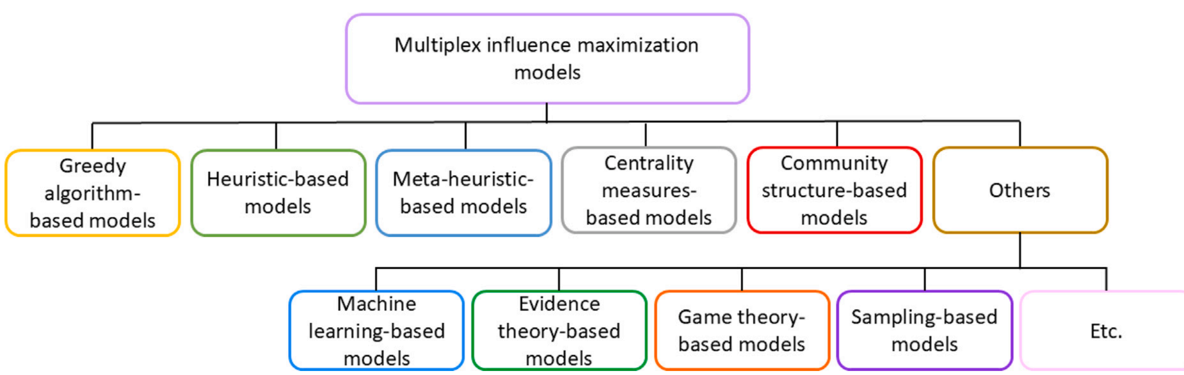

Fig. 6. Taxonomy of multiplex influence maximization models.

the multiplex network into a single-layer network using two proposed coupling methods: lossless and loss. These methods present a trade-off between information loss and memory consumption. The lossless method preserves the network's original properties by adding extra edges and nodes, though it incurs high memory usage. In contrast, the loss coupling method reduces memory and computational demands at the expense of some information loss. Subsequently, building upon the concept of heterogeneous diffusion models, Kuhnle, Alim, Li, Zhang, and Thai (2018) studied the MxIM problem by allowing each network layer to follow distinct diffusion models, such as the LT and IC models, with propagation across layers occurring through overlapping users, as in Zhang et al. (2015). Motivated by this setup, they defined a new diffusion property called generalized deterministic submodular (GDS), which ensures submodularity in the overall multiplex propagation process. In addition, a two-phase approximation algorithm named knapsack seeding of networks (KSN) is introduced. The first phase runs in parallel across layers, storing the set of activated nodes for each seed node, while the second phase selects seed nodes based on the multiple-choice knapsack (MCK) problem. KSN relies on the influential seed finder (ISF) greedy algorithm, which iteratively selects nodes that maximize marginal gain. Although KSN yields high-quality seed selections, its time complexity limits its scalability for large-scale multiplex networks. Thus, to bridge the time complexity gap of the greedy algorithm without compromising its accuracy, Chen and Tan (2020) presented the MNStaticGreedy algorithm for detecting influential nodes in multiplex networks. It starts by transforming the multiplex network into a single-layer network, capturing the varying strengths and interactions between nodes across different layers. Then, by generating static snapshots and leveraging the reverse reachable (RR) set method, MNStaticGreedy selects seed nodes based on the maximum coverage method. Moreover, to model the diffusion process, a multiplex network threshold diffusion model (MNTM), where a node becomes active when the combined influence (cummulative influence) from its neighbors in different layers exceeds its threshold is developed. By utilizing static snapshots rather than running multiple Monte Carlo simulations in each iteration, MNStaticGreedy improves the time complexity of the greedy algorithm. In another research direction, and by leveraging the concept of multiple influences in a competitive environment, Chen, Wang, Feng, Lu, and Gong (2020) introduced the idea of fair seeds allocation with the IM problem in multilayer networks. In this light, a multiple-influence diffusion model MMIC based on the IC model is introduced, in which a node can be activated several times. MMIC works in two steps. The first step finds the influential seeds using the reverse reachable set for entity (RRE) algorithm. While the second step fairly allocating these selected seeds to different customers using the live path method. Referring to the rumor-spreading process within multiplex networks, Hosni, Li, Ding, and Ahmed (2018) presented the Least Cost Anti-rumor Campaign (LCAC) problem, which aims to minimize rumor spread while maximizing the diffusion of true information. LCAC selects a small set of users to initiate the anti-rumor campaign, targeting overlapping users across layers to enhance awareness. The strategy focuses on containing the rumor within one network while minimizing its influence on others. Additionally, a rumor propagation model based on the IC model is proposed, where the rumor transmission probability is defined by two parameters: sending probability p_u^{send}, which estimates the likelihood of a user transmitting the rumor to neighbors, and acceptance probability $p_{v,u}^{acc}$, which evaluates the chance of an individual accepting the rumor. The rumor spreads across layers via overlapping users. In another work (Hosni et al., 2020), Hosni et al. studied how individual behavior influences information dissemination in multilayer networks, modeling individual behaviors as damping harmonic motion. They incorporated users' background knowledge, a hesitation mechanism, and a forgetting–remembering factor into the rumor propagation process to model how individuals are influenced by misinformation. Consequently, based on these factors, they proposed a human individual and social behaviors (HISB) diffusion model to represent the spread of rumors in multiplex OSNs. Then, using this model, a *rumor popularity* metric is presented to establish the truth campaign strategy (TCS). In Hosni et al. (2020, 2018), the truth campaign is developed based on a greedy algorithm with k iteration, where a node with maximum marginal containment influence is added to the truth campaign in each iteration. Li, Zhang, Zhao, Yi, and Li (2020) introduced the concept of global propagation probability to describe the nodes' propagation influence in multilayer networks using the IC model and the theory of bond percolation. By calculating the global propagation probability and expected propagation range of nodes, the study evaluates the influence of nodes across layers. To this end, influential nodes are detected using a greedy algorithm and a de-overlapping method. The greedy algorithm prioritizes nodes with high propagation potential, while the de-overlapping method ensures the selection of non-redundant and diverse propagation sources near the average degree of the network for efficient information dissemination. To summarize, these advancements discussed in greedy algorithms reveal a clear trend towards developing more scalable and nuanced IM models that balance computational efficiency with the complexity inherent in multiplex network structures. However, despite these strides, challenges persist in adapting IM strategies to handle multilayer diffusion efficiently. Consequently, heuristic approaches have emerged as promising alternatives to address these limitations, as discussed in the following sub-section.

4.2.2. Heuristic based-models

Due to the high computational complexity of greedy-based models, various heuristic-based solutions have been proposed to solve the multiplex IM problem. These methods use an approximate scoring system to assess the influence spread for each node (Singh et al., 2021). An earlier solution in this family was proposed by Gaye, Mendy, Ouya, Diop, and Seck (2016), who developed a heuristic method called multi-diffusion degree centrality (C_{dd}^{MLN}) to select seed nodes in multilayer networks using the IC spreading model. C_{dd}^{MLN} uses second-order neighbors along with the diffusion probability $P_{u_k^i}$ of each node u in each kth layer. Here, the authors modeled the multilayer network using equivalence classes and generated mapping matrices between all layers. Consequently, equivalent classes are used to group nodes that have similar influence and connectivity patterns across different layers. Li, Chu, Feng, and Xu (2016) studied the multiplex IM problem using the maximum propagation paths. They used the idea of the upper bound to estimate the influence of nodes in multilayer networks. Using the concept of overlapping influence, which occurs when spreaders affect a group of the same users (Fig. 7), Chen, Deng et al. (2019) proposed an overlapping influence-based (OI-based) method. It aims to select influential users while minimizing overlapping influence among them under the SIR diffusion model. OI algorithm combines the centrality with the relative location of selected users, ensuring that the selected users are central and relatively scattered to maintain low overlapping influence. Then, based on the concept of shortest paths, a multiplex shortest path ESSP (effective spreading shortest path) is introduced. ESSP describes the relative distance between users in multiplex networks, using the spreading rate in the propagation process to measure the collective influence of multiple users. Similar to the previous work, Zhang, Xie, Liu, and Zhan (2023) recently presented a heuristic algorithm called adaptive coupling degree (ACD) for IM in multilayer networks taking into consideration the overlapping influence issue. ACD asses nodes' influence considering neighbor information across different layers and then iteratively selects seed nodes with high spread influence and low overlapping influence. In addition, the authors extended the IC diffusion model, allowing the propagation to occur across layers simultaneously. In this model, at each time step t, the propagation occurs within a particular layer l. Once the propagation process within layer l is complete at time step t, the propagation process moves to the next layer ($l+1$) at time step $t+1$, and so on. Shifting the focus to multiple influence maximization (MIM), Singh, Singh et al. (2019) introduced the MIM2 (multiple influence maximization across multiple social networks) problem under the IC diffusion model. It aims to find influential nodes that can maximize the spread of multiple influences (items) across multiple social networks (or multiplex networks). They first proposed a multi-relationship network aggregation method to fuse the multiplex network into a single-layer network via a direct linkage strategy. Then, based on overlapping users, they developed a heuristic method to find the most influential user over multiple product diffusion multiplex networks. This heuristic starts by finding non-seeding nodes (those who are less likely to be influential in the network). Next, from the selected non-seeding nodes, it identifies the influential nodes per product using the backward propagation method, where it traverses the network in reverse order from these nodes using Breadth First Search (BFS) to accumulate the expected influence spread of each node and selects the node with the maximum expected influence as the most influential user for the corresponding product. Finally, it identifies the most influential user among all products within multiplex networks. While MIMI emphasizes the spread of influence across various products, a different yet related research direction examines aligned heterogeneous network influence maximization (AHI). Here, the challenge lies in maximizing influence in networks that are heterogeneous in nature, involving nodes of different types (e.g., users, locations, etc.) across multiple layers, as explored by Zhan, Zhang, Wang, Yu, and Xie (2015). In this paper, the authors presented the AHI problem which studies the IM problem across two partially aligned heterogeneous OSNs under the LT diffusion model. Using meta paths, they presented the multi-relational network influence maximizer model (M&M). M&M extracts multi-aligned multi-relational networks (MMNs) from aligned heterogeneous OSNs based on a set of inter and intra-network social meta paths. Although the above-discussed works focus primarily on static multiplex networks, overlooking the dynamic nature of many real-world systems, there is a growing need to address the temporal aspect of network evolution. To fill this gap, Meng, Chen, Yi, Wang, and Pei (2021) presented the IM problem in dynamic multi-social networks (DMNIMP) based on common friends. In this paper, influential nodes are identified by building a dynamic multisocial network information propagation model based on the DIC model. It merges multiple dynamic networks into one dynamic network and adds self-propagation edges between common users in each network snapshot of the aggregated network. Additionally, the authors used the T×one hop algorithm to find influential nodes.

4.2.3. Meta-heuristic-based models

Meta-heuristic algorithms have been created in recent years to minimize solution-solving time and pursue improved algorithm efficiency due to the high computational cost of simulation-based algorithms and

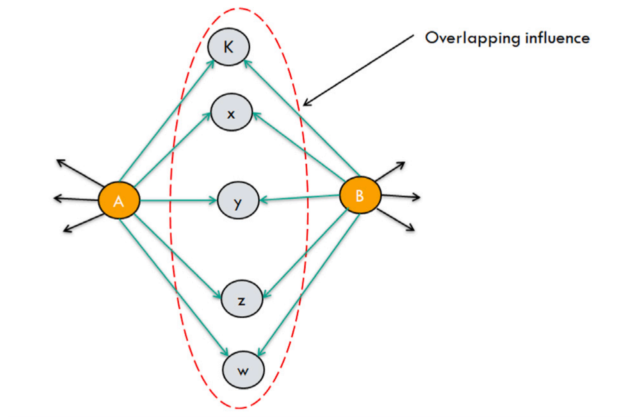

Fig. 7. An example of overlapping influence, where orange nodes represent seed nodes and gray nodes represent ordinary users.

the high efficiency of meta-heuristic algorithms. The MxIM problem is described as an optimization problem in these algorithms by defining a fitness function, and methods such as evolutionary optimization algorithms are used to solve it. For instance, recently Lu, Bu, and Wang (2020) formulated IM in multilayer networks as a multiobjective optimization problem. They proposed a new multi-objective evolutionary approach based on the classic non-dominated Sorting Genetic Algorithm II (NSGA-II) called IMA-NSGA-II to find the influential nodes in multilayer networks under both IC and LT diffusion modes. Unlike traditional greedy algorithms that provide a single solution, IMA-NSGA-II identifies multiple Pareto-optimal solutions, offering decision-makers a wider range of options for seed node selection. It considers both node's centrality and information propagation ability to optimize influence spread in multilayer networks. In the same light, using genetic algorithms, Hu, Zhao, and Yang (2023) introduced a node grouping genetic algorithm (NGGA) to detect influential nodes in multiplex SNs. In this paper, the node grouping strategy is based on node's influence and selection cost. Besides, to avoid Monte Carlo simulations, the authors proposed an evaluation function for accelerating the simulation of information distribution, while minimizing memory consumption. This function measures the influence using the expected propagation value of the seeds in the set within two hops. In another research line, Wang, Liu, and Jin (2019) formulated the MxIM problem as an optimization of multiplex 2-hop influence spread. Using memetic algorithms, a new method called MA-IMmulti to find the best set of influential nodes within multiplex networks is presented. In this paper, the authors extended the IC model by adjusting the activation probability based on the degree discrepancy between activated seeds and non-activated nodes, handling the issue where a node may be highly influential in one layer but negligible in another. Moreover, they demonstrated that activation in a single layer is insufficient to determine a node's activation across all layers. To better address this scenario, they introduced a minimum number of layers, $Lmin$, in which a node must be activated in at least $Lmin$ layers to become active in next steps. In MA-IMmulti, a population of initial seed sets is generated based on an integration of random, roulette-based degree and distance selection strategies. The major limits of this model are twofold: (1) its inability to handle large-scale networks due to the large number of computationally expensive iterations, and (2) the determination of $Lmin$, which is a problem-dependent, making the performance of the diffusion model sensitive to its optimal value. Another work in this family was proposed by Zhang, Zhong, Gao, and Li (2018), in which epidemics spreading across multilayer networks is considered. In this paper, the authors studied the modes of Human Immunodeficiency Virus (HIV) transmissions in multiplex social networks. In this light, a multiplex social network framework is first constructed to capture the multi-mode propagation across two key populations: FSWs and PWID. Then, to identify the most influential individuals for intervention purposes, a novel random search method PRS-NMP (partition-based random search with network and memory prioritization) is developed. PRS-NMP builds upon the classic nested partitions method and incorporates two key innovations. First, it prioritizes nodes with higher network properties like degree and betweenness centralities during the partitioning process. Second, PRS-NMP keeps track of promising regions of the search space in each iteration. This information allows the algorithm to backtrack and focus on areas with a higher likelihood of containing the optimal solution.

4.2.4. Centrality measures-based models

The concept of centrality was first introduced by Freeman et al. in 1979 (Freeman, Roeder, & Mulholland, 1979), in which they show that the centrality measure is a vital indicator for understanding a node's position within a social network. For the IM problem, centrality measures are used to understand how influential nodes are detected. In other words, the greater a node's centrality value, the more influential it is. So far, several centrality measures have been introduced such as degree centrality (DC) (Srinivas & Velusamy, 2015), Closeness Centrality (CC) (Okamoto, Chen, & Li, 2008), Betweenness centrality (BC) (Kourtellis, Alahakoon, Simha, Iamnitchi, & Tripathi, 2013), Eigenvector centrality (EC) (Bihari & Pandia, 2015), K-shell (KS) (Brown & Feng, 2011), etc. However, most of them are applied to single-layer networks. In Table 5 we present typical centrality measures developed in the literature.

In light of the multilayer networks, many centralities have been used to detect influential nodes. For instance, in an attempt to improve the K-shell centrality, Erlandsson, Bródka, and Borg (2017), proposed a hybrid centrality measure that combines both K-Shell and degree centralities to avoid the K-Shell limit (lacks sufficient granularity). They investigated the effect of the degree centrality, the K-Shell, the VoteRank, and the machine learning method ARL, respectively on the spreading processes within eighteen multilayer social networks under the IC model. Subsequently, based on local knowledge of network topology, Basaras et al. (2017) introduced a family of multilayer PCI (Power Community Index) measures, generalized from the h-index centrality, which consider both inter- and intra- connections within the node's neighborhood to detect influential nodes in multilayer networks. In addition, an improved SIR model with two information propagation probabilities: λ_{ii} for intralayer propagation, and λ_{ij} for interlayer propagation, is presented. Unlike the degree centrality, which looks only at the node's degree, PCI intends to find nodes with dense neighborhoods. Referring to the heterogeneity aspect in complex systems, Wan et al. (2022) proposed a multilayer heterogeneous network important node identification method based on multi-relationship fusion (MLC). MLC divides nodes into two distinct layers, referred to as core and auxiliary layers, and detects key nodes in the core layer. It integrates interlayer information with intralayer information to identify influential nodes across layers. Here, interlayer information determines the interlayer influence based on the number of common nodes connected across layers, while intralayer information evaluates the importance of nodes within each layer using the K-shell method and betweenness centrality. Although the authors presented an interesting approach, it is not within the scope of our interest, as it overlooks the true structure of multiplex networks, where the same set of nodes is replicated across all layers. Referring to entropy-based centrality methods, Wang, Tian, and Wei (2023) investigated a centrality measure for multilayer networks based on local structure entropy. The proposed centrality combines the global centrality (betweenness centrality) inside each network layer with the influence of inter-layer connections. Further exploring centrality methods, based on the concept of layer weighting, Bouyer, Mohammadi, and Arasteh (2024) and Zhou, Bouyer, Maleki, Mohammadi, and Arasteh (2023) recently introduced a three-step approach for detecting influential nodes in multiplex networks under

Table 5
A comparison of the typical centrality measures (Huang, Chen, Wang, & Ren, 2020).

Centrality	Description	Equation	Complexity
DC (Srinivas & Velusamy, 2015)	It is defined by counting the number of links incident upon a node. For a directed network, we have two-degree centrality-based measures: in-degree centrality, which considers the number of edges incoming to the node, and out-degree, which considers the number of outgoing edges of the node.	$\forall x \in V, DC(x) = \lvert N(x)\rvert$, $DC_{in}(x) = \lvert N_{in}(x)\rvert$, and $DC_{out}(x) = \lvert N_{out}(x)\rvert$	$O(n)$
CC (Okamoto et al., 2008)	Closeness centrality (CC) is the reciprocal of the average shortest distance from a particular node to all other reachable nodes of the graph.	$CC(x) = \frac{n-1}{\sum_{b\in V \neq xx} dist(x,b)}$	$o(nm + n^2 log(n))$
BC (Kourtellis et al., 2013)	Betweenness centrality (BC) measures the times a node acts as a bridge along the shortest path of two other nodes.	$BC(x) = \sum_{a,b\in V} \frac{\sigma(a,b\mid x)}{\sigma(a,b)}$	$o(nm + n^2 log(n))$
EC (Bihari & Pandia, 2015)	Eigenvector centrality is an extension of degree centrality. It is based on the idea that a node is important if its neighbors are important.	$EC(x) = \frac{1}{\lambda} \sum_{j\in N(x)}^{n} A_{xj} x_j$	$o(n + m)$
PR (Brin & Page, 1998)	PageRank is the fundamental search engine mechanism of Google. It is used to rank the quality of pages that have hyperlinks to some web pages. The basic idea is to measure the importance of a web page based on the number and quality of links pointing to it. In the context of social network analysis, PageRank can be applied to find the most influential nodes.	$PR(x) = \frac{(1-\epsilon)}{N} + \epsilon \times \sum_{v\in\pi_x} \frac{PR(v)}{d_v^+}$	$o(n + m)$
H-index (Hirsch, 2005)	H-index centrality is an adaptation of the h-index, a metric originally designed to measure the productivity and citation impact of the publications of a scholar.	$H(x) = argmax_{h\in N} \{\forall N(j) \geq h, 1 \leq h \geq \lvert N(x)\rvert, j \in N(x)\}$	$o(n + m)$
KC (Katz, 1953)	Katz centrality (KC) is a metric to measure the relative influence of a node in social networks. This measure is a generalization of both eigenvector centrality and the node degree centrality. It takes into account both the quality of connections (importance of neighbors) and the quantity of connections.	$KC(x) = \sum_{K=1}^{\infty} \sum_{y=1}^{n} \alpha^k (A^k)_{xy} + \beta$	$O(N^3)$
K_sC (Seidman, 1983)	K-shell centrality (K_sC)is a measure of node importance in a network based on its position within the network's core structure. It involves iteratively removing nodes with a degree less than k, starting from $k = 1$ and incrementing K until all nodes are removed. The highest value of k at which a node is removed is its k-shell index.	$K_sC(x) = k, s.t x \in S_{shell}^k, x \notin S_{shell}^{k+1}$	$O(n + m)$

the SIR diffusion model. Initially, a family of layer metrics is presented to evaluate the spreading ability of nodes within each layer. These metrics are defined as active node connectivity to all unique interlayer connections (ANAUIC), normalized uniqueness of intralayer communication compared to other layers(NUICL), and layer coreness alignment (LCA). Second, using these metrics a layer weighting method is developed to assign weights to the different layers describing their importance. Finally, the third step involves two substeps. First, the centrality measure of each node is determined independently in each layer to create a centrality vector. Then, by incorporating the layer weights, a mapping function is applied to convert the centrality vector of each node into a scalar value, denoted as the singular rank value (SRV), which is used to identify the influential nodes. In this paper, closeness, percolation, Page PageRank, Katz, and Laplacian centralities are used. Taking advantage of gravity model, several studies have built upon the concept of gravity centrality, each contributing uniquely to enhance its application in multilayer networks. For example, a weighted gravity centrality method (WGCM) based on both an individual's neighborhood size and the social distance is proposed by Ni, Yang, Pang, and Gong (2022) for identifying influential nodes in multiplex networks. In this regard, to simulate the diffusion process, the IC spreading model is used, in which the activation probability for inactive nodes depends on the activation probabilities of their active neighbors within the same network layer. Interlayer propagation occurs through replica nodes: when one replica node is activated on a layer, all corresponding replica nodes on other layers are activated simultaneously. Based on the proposed WGCM centrality, the authors introduced three seeding strategies: the Single-Stage Seeding Strategy (SSS), Timed Sequential Seeding Strategy (TSSS), and Revived Sequential Seeding Strategy (RSSS). SSS activates seeds immediately, TSSS uses a timed sequence for seed selection, and RSSS revives seeds in phases. In the same direction, Ni and Yang (2022) developed a two-step seeding strategy based on multiplex gravity centrality (MGC). The first step identifies the target layer for the seeding identification process, and the second step selects the seed nodes (target seeds). The target layer is determined based on two centrality measures: multiplex betweenness centrality (MBC) and multi-hop multiplex neighbors (MMNs), where the layer with the largest sum of MBC or MMNs for each node is considered the target. Seed nodes are then chosen using the MGC centrality, which combines degree centrality and shortest path distance. This paper also introduces a redefined IC model (r-ICM), which evaluates the influence spread of selected seeds through two diffusion probabilities: intra-layer (based on common neighbors' information) and inter-layer (defined using inter-layer conversion costs). Using r-ICM, influence propagation begins in the target layer and spreads to other layers through replica nodes. More recently, Lv et al. (2024) presented an improved gravity centrality (PRGC) for identifying key nodes in multilayer networks based on multi-PageRank centrality. PRGC works in two phases. In the first phase, it defined the multi-PageRank centrality by combining the importance of a node's neighbors and the network layers in which they are located. Then, based on that, PRGC proposed a distance calculation method, which incorporates the layer's important influence on the shortest path between nodes. In addition, based on the proposed PRGC centrality, the authors described the propagation process under the

LT model, where influence propagates within each network layer, and when a node is activated in one layer, it becomes active in all layers of the network, reflecting inter-layer propagation. In a different research direction, and leveraging the tensor analysis method, Wang, Wang and Zou (2017) proposed a centrality measure called essential nodes determining based on CANDECOMP/PARAFAC (CP) tensor decomposition centrality (EDCPTD) to find the most influential nodes in multilayer networks. They considered four key roles for each node and layer in the multilayer tensor representation: the authority and hub of nodes, as well as the authority and hub of layers. To this end, EDCPTD begins by modeling the multilayer network as a fourth-order tensor and then applies CP tensor decomposition to extract significant factors, such as principal singular vectors, and produces quadruplet vectors to compute the hub and authority scores for each node across all layers. Unlike gravity-based centralities, which primarily emphasizes node influence, tensor analysis allows for a detailed evaluation of multilayer dependencies through higher-dimensional representations. In Zeng, Liu, Tang, and Gong (2021), Zeng et al. introduced a coupling-sensitive (CS) centrality, which focuses on identifying super-spreaders in the dynamics of information-disease transmission on multiplex networks under the SIR–SIRV spreading model. Here, the multiplex network is modeled as a duplex network of contact and communication layers. The latter reflects a kind of communication channel including OSNs where information on the epidemic is diffused. While the former indicates the physical contact layer where the disease spreads. *CS* centrality incorporates various dynamical and structural coupling factors including (1) the two-layer relative spreading rate, (2) the inter-layer degree correlation, and (3) the inter-layer coupling strength. It extends standard centrality measures, such as degree and eigenvector centrality, by introducing additional parameters, namely transmission rates, immunization rates, and informing rates. While the proposed centrality effectively identifies influential nodes in the multilayer network, its performance relies heavily on the spreading rates of two layers: immunization and informing rates. For privacy and security reasons, Hu et al. (2022) introduced a secure multiparty computation ranking (SMPC) method to collaboratively identify influential nodes in multiple private networks (PNs) under the SIR spreading model. The SMPC-ranking method is used to detect influential nodes in multiple private networks while preserving the privacy of each network. The proposed method starts by computing the influence of nodes based on a centrality function such degree of centrality within each network layer using the secure multiparty computation (SMPC) protocol and then aggregates the results from all the private networks to determine the overall influence of nodes across all networks.

4.2.5. Community structure-based models

To improve the efficiency of the centrality measures and the greedy algorithm-based models, researchers have started investigating the community structure into MxIM-based solutions (Kumar, Singhla, Jindal, Grover, & Panda, 2021). A community is a subset of individuals who interact with each other more frequently inside the community than with other individuals outside (Jaouadi & Romdhane, 2021). It has a significant impact on information diffusion, because of dense connections between the nodes within the community. The identification of the community in the graph facilitates the search for influential nodes, by reducing the search space (Jaouadi & Romdhane, 2024). Katukuri, Jagarapu, et al. (2021) presented a clique-based influence maximization (CIM) framework to find influential nodes in multilayer under the IC spreading model. CIM starts by detecting the communities based on the principle of maximal cliques. Then, influential nodes are detected as the maximal cliques based on their size. Moreover, to improve the efficiency of seed nodes, CIM used a node pruning method, leveraging the fact that once a node is activated in one layer, it is no longer necessary to re-evaluate it in the rest layers. CIM assumes that the community structure of the network is known. However, in many real-world multilayer networks, the community structure is unknown beforehand. Similar to the CIM model, Rao and Chowdary (2022) proposed a community-based influence maximization (CBIM) method to find seed nodes in multilayer networks using LT and IC models. CBIM involves two main phases. In the first phase, CBIM detects small communities in each network layer using the dice neighborhood similarity index. Then, to enhance community qualities, a certain fraction of these small communities is merged into larger ones. In the second phase, seed nodes are selected from each community based on a quota-based selection method using a proposed edge weight sum (EWS) measure. EWS is like Katz's centrality but incorporates both degree and distance in calculating the edge weights. Although CBIM ensures a balanced selection of nodes across different communities, it eliminates interlayer information during the community detection process, treating each layer independently when selecting seed nodes. Further, Lv et al. (2023) proposed a community-based centrality model (CBCM) approach to find influential nodes in multilayer networks. CBCM begins by modeling multilayer networks, including interlayer links, using a fourth-order tensor model. It then calculates interlayer similarity to measure the strength of interaction between different network layers. Using this tensor representation, CBCM computes centrality scores, capturing node importance across layers simultaneously. To this end, a new centrality measure called PR_BIS (PageRank-Based Interlayer Similarity) is introduced, combining elements of traditional metrics like PageRank and gateway ranking. CBCM assesses node importance by integrating three key factors: a node's centrality within its own layer, its role in connecting communities, and the significance of the community itself. By weighting and fusing these factors across network layers, CBCM determines the overall importance of nodes in multilayer networks. Drawing on the concept of structural holes, Mittal and Bhatia (2017) proposed a method to detect the top-k structural hole nodes in multilayer networks by leveraging centrality measures and community structures. The approach begins by identifying multiplex communities within the network. Next, the method calculates the eigenvector centrality of each node within its respective community. Nodes with the highest eigenvector centrality scores are then identified as the top-k structural holes, representing key connectors that bridge different parts of the network. Chowdary et al. (2023) extended the K++ Shell decomposition algorithm to identify influential nodes in multilayer networks by adapting the K-Shell decomposition method and integrating the Label Propagation Algorithm (LPA) to detect communities. This method operates in two phases: first, it strategically prunes nodes based on their degrees, assigning them to specific buckets. During the pruning process, reward points are incrementally assigned to the neighbors of pruned nodes, encouraging influence from a broader range of buckets. In the second phase, the K++ Shell algorithm selects seed nodes for information dissemination. By incorporating reward points, this approach overcomes the traditional limitation of selecting only the highest-bucket nodes, enabling nodes from lower buckets to serve as influential seed nodes as well. More recently, Fujita and Tsugawa (2023) studied the problem of limiting the spread of negative influence in multilayer social networks. In this light, they introduced a method that blocks negative influence by strategically spreading positive influence, targeting nodes based on degree and community structure using the IC model. In comparison to traditional methods that focus on single-layer networks, Huang et al. (2020) proposed a MINE algorithm for identifying influential nodes in both monolayer and multilayer networks based on the folding of community structures hierarchically under the SIR spreading model. Using the maximum influential neighbor's expansion strategy, MINE treats communities as nodes. These nodes undergo an iterative folding process until convergence is achieved. This iterative process aims to identify seed nodes hierarchically by aggregating and refining community structures.

4.2.6. Others

In recent years, many other emerging solutions have been proposed to solve the IM problem in multilayer networks including machine learning-based algorithms, evidence theory-based algorithms, game theory-based algorithms, etc. This sub-section outlines these solutions. Recently, taking advantage of machine learning-based models, Keikha et al. (2020) proposed a deep learning-based influence maximization algorithm (DeepIM) to detect influential nodes in interconnected networks (a special case of multilayer networks) under IC and LT models. DeepIM used deep learning techniques to learn network node feature vectors while preserving local and global structural information. To this end, it combined the community-aware method (CARE (Keikha, Rahgozar, & Asadpour, 2018)) with skip-gram (Mikolov, Chen, Corrado, & Dean, 2013) to learn network node embedding vectors and further use the learned embedding vectors to select seed nodes. The major advantage of this model is its ability to cope with large-scale networks. In the same direction, another recent machine learning-based, MIM-Reasoner model is presented by Do, Chowdhury, Ling, Zhao, and Thai (2024) to address the IM problem in multiplex networks under heterogeneous diffusion models. MIM-Reasoner decomposes the multiplex network into individual layers and leverages deep reinforcement learning to find near-optimal seed nodes for the multiplex network. It begins by applying the Knapsack approach to assign budgets to each layer and then learns a lightweight policy to find feasible solutions for each layer using a decomposition strategy. Through the learning process, MIM-Reasoner models interlayer relationships between layers, which avoids reactivating nodes several times leading to reducing the time complexity. Referring to sampling-based models, Chen, Qian, Wu, Chen and Wang (2019) addressed the adoption maximization problem in multilayer social networks using a sampling-based approach. They proposed a multilayer Reverse Reachable Adoption (MRRA) method, which extends the Reverse Influence Sampling (RIS) technique to identify the top-k influential users across multilayer networks. MRRA generates sampled sets of nodes reachable by any node across layers. Then, a greedy algorithm is applied to select seed nodes that cover the most MRRA sets. To optimize the sampling process, they introduced the MAIMM baseline approach, along with an h-hop-based method that includes a strict theoretical bound to accelerate the search. Furthermore, a Multilayer Adoption IC (MAIC) model to simulate information propagation is presented. A key advantage of this method is its efficiency in handling large-scale networks. Yu, Yang, Ai, Zhu, and Wang (2020) studied the impact of limited contacts and network multiplexity on information spreading. Accordingly, they proposed an information-spreading model with limited contact capacity within a two-layered multiplex network based on the SAR diffusion model, in which individuals can only transmit information to a limited number of their neighbors. Then, a generalized edge-based compartmental theory is introduced to quantitatively analyze the influence spread. This paper demonstrates that information spreading on multilayer networks exhibits a crossover phenomenon between the information outbreak size and the transmission probability. Taking into consideration heterogeneous multilayer networks, Rani and Kumar (2022) presented a heterogeneous degree ranking (HDR) method to find influential nodes in multilayer networks under the SIR spreading model. Consequently, they developed five multiple-criteria decision-making (MCDM)-based methods such as the analytic hierarchy process (AHP), the technique for order of preference by similarity to ideal solution (TOPSIS), fuzzy AHP (F-AHP), fuzzy TOPSIS (F-TOPSIS), and the analytic network process (ANP) to identify and rank the most influential nodes in multilayer networks. Shifting the focus to game theory-based approaches, Yu, Liu, and Han (2021) explored the effects of individuals' heterogeneous properties — such as social position, link degree distributions, and trust levels — on the evolution of cooperation within multilayer networks. They proposed a heterogeneous diffusion model in which nodes are classified into authority and non-authority groups based on their social position and influence. Each node adopts one of two strategies: cooperating (C) or defecting (D), and participates in evolutionary games like the Prisoner's Dilemma and Snowdrift Game. The model also accounts for the memory lengths of nodes, as well as the impacting and reacting forces between nodes and groups. This interplay of properties and strategies drives social interactions and influences the diffusion of cooperative behavior within multilayer networks. Li, Zhang and Deng (2018) presented a key node identification method for multilayer networks based on evidence theory. In this approach, the influence of each node is quantified within a single layer using the relevance value (RV) metric. RV is defined as the sum of the node's shortest paths to all other nodes. Nodes with smaller RVs are considered more influential within their respective layers. These individual RVs are then aggregated using the Dempster–Shafer evidence theory to identify the final set of influential nodes across layers. The major limitation of this method is its computational cost, which makes it unsuitable for large-scale social networks. In the same direction, Lei, Liu, and Xiao (2023) introduced a weighted information fusion (WIF) method for identifying influential nodes within a Network of Networks (NONs) using evidence theory. WIF divides NONs into individual networks, calculates node relevance for each network, and constructs Basic Probability Assignments (BPAs). These BPAs are then weighted using the effective distance approach and fused to determine influential nodes across NONs. While WIF evaluates influential nodes within NONs, it does not consider interlayer information. Benefiting from hybrid-based models, Li, Tang, Du and Li (2021) presented a key node discovery (KND) algorithm based on the shortest path, degree centrality, and random walk features in a multi-relationship network under the SIR diffusion model. Here, the multi-relationship network is represented as a multiplex network with various social interactions such as forwarding, replying, and mentioning. Zhong et al. (2017) explored social influence within duplex networks, a special case of multiplex networks, focusing on the cascade centrality of individual nodes. They generalized the LT model to multiplex networks by introducing OR and AND protocols for combining signals (or influence) from different layers. These protocols represent respectively sensitive and conservative responses to active neighbors. In addition, this paper extends the definition of both Live-Edge Models (LEMs) and reachability to the case of two-layer multiplex networks, in which an equivalence between LEMs and the LT model is demonstrated. Al-Garadi, Varathan, Ravana, Ahmed, and Chang (2016) analyzed the multilayer structure of OSNs using established centrality models. They first transformed the multilayer network into a single-layer network, considering both multilayer interactions and overlaying links as weights. This study explores the impact of different aggregation methods on algorithm accuracy, demonstrating the importance of improving network topology and ranking algorithm efficiency for identifying influential users. Wang and Tan (2022) presented a multi-factorial evolutionary approach (MFEA-RIM$_m$) to solve the robust IM (RIM) problem on multilayer networks under the IC diffusion model. MFEA-RIM$_m$ uses problem-directed operators such as the lattice-based initialization operator, the distance-aware crossover operator for transferring knowledge, and the three-phase local search operator to improve its search-ability and to handle scenarios involving multiple instances of damage. Zhuang and Yagan (2019) investigated the IM problem in multiplex networks using a multi-stage content-dependent LT contagion model. They examined the impact of network connectivity on the expected size of global cascades (i.e. global influence), revealing how complex network structural properties, such as assortativity influence contagion dynamics. As a result, this study uncovers a connection between network structural properties and the influence of hyperactive nodes on cascade size. Lately, Liu, Zeng, Pan, and Tang (2023) analyzed the interaction between information and disease spreading on multiplex networks using an asymmetrically interacting information-disease spreading model (SIR-SIRV model). They investigated how interlayer coupling factors, including relative spreading rate between layers, the strength of interlayer connection, and interlayer degree correlation between layers affect the node's influence ability. The authors then introduced a method for identifying influential nodes in disease spreading on multiplex networks, based on outbreak size calculated using percolation and a message-passing approach. Their findings indicate that structural centralities, such as degree and eigenvector centralities, within the contact layer alone are insufficient for predicting a node's influence due to the suppressive effect from the information layer.

4.3. Applications

The IM problem with multiplex social networks has a wide range of potential real-world applications.

- *Recommender systems*: Influence maximization in recommender systems enhance the reach and effectiveness of recommendations by leveraging the power of social influence (Sun, Cautis, & Maniu, 2023). It involves identifying a subset of users who, if influenced to adopt certain items such as films, musics, movies, etc., would maximize the spread of those items throughout the social network. Taking advantage of multiplex social networks, recommendations then can be enhanced by considering the cross-platform influence of users. For instance, a movie recommendation system aims to recommend a new film to influential users on Twitter, who then share their experiences on Facebook and Instagram, leading to a wider adoption.
- *Politics*: In politics, campaigns use IM algorithms to identify key influencers within social networks, who can effectively propagate political ideas and messages. For instance, in the 2016 U.S. presidential election, Trump used the social network Twitter to influence and attract people to vote for him by posting tweets about his program, activities, etc. Jaouadi and Romdhane (2024). Therefore, by leveraging the multiplex network social networks, political campaigns might synchronize posts and advertisements across multiple social networks simultaneously to amplify the message and reach a broader audience effectively, ultimately influencing voter behavior and enhancing election outcomes.
- *Epidemic analysis*: In epidemics spreading, various studies have been performed to study the coupled epidemic-information spread in multiplex social networks to maximize the spread of awareness and prevent infections. For example, Wang and Xia (2020) worked on awareness of epidemics in two-layer networks. In these two-layer networks, one layer propagates awareness about the disease and the other layer spreads the disease.
- *Public healthcare*: Influence maximization in public healthcare improves the dissemination and adoption of health practices, information, and interventions by harnessing the power of social influence within healthcare networks (Tang & Yang, 2012; Xu et al., 2021; Zhan, Zhuo, & Liu, 2019). It involves identifying a core group of individuals or organizations who, if influenced to adopt or promote certain health behaviors, vaccination, or messages, would maximize the spread of those behaviors throughout a population. By leveraging multiplex networks, public health companies can extend their reach across multiple layers of society. For example, in a vaccination campaign, multiplex IM might focus on identifying trusted healthcare professionals (e.g., doctors, nurses, or health experts) who can share information about healthcare protocols through platforms such as Facebook, forums, and outreach programs simultaneously. Therefore, by strategically engaging these influential figures, public health authorities can increase awareness, foster positive health behaviors, and enhance the overall impact of their campaigns across diverse communities.
- *Scientific collaborations*: Influence maximization in scientific collaborations can enhance the reach and impact of scientific ideas. It involves identifying a subset of key researchers who, if encouraged to collaborate or promote specific ideas, would maximize the spread of those ideas across the scientific community (Li, Chu et al., 2021). By taking advantage of multiplex collaboration networks, where influence spreads across different platforms and types of interaction, scientific contributions can be amplified across multiple channels.
- *Smart cities*: In recent years, multiplex influence maximization has also found applications in smart city networks, particularly in optimizing the spread of information across different urban services such as traffic management, energy usage, and emergency services (Molaei, Rahsepar Fard, & Bouyer, 2024). Here, multiple information flows (e.g., traffic updates, and emergency alerts) need to be managed simultaneously, and the goal is to efficiently influence citizens' behaviors, enhancing service delivery and response times.
- Etc.

5. Discussion

As we can see from Section 4.2, there is an interesting endeavor in addressing the multiplex influence maximization problem. In our review, we categorized the different proposed solutions into five main families commonly used for addressing the IM problem in single-layer networks. Each family brings unique strengths and challenges in managing multiplex network complexity. Greedy algorithms are widely recognized for their high-quality solutions and strong theoretical guarantees. However, they often require extensive Monte Carlo (MC) simulations, making them computationally intensive, particularly in large-scale multiplex networks. This high complexity limits their applicability in real-time scenarios where rapid decisions are critical. Consequently, to overcome the computational demands of greedy approaches, heuristic-based models have been developed as an efficient alternative. These methods prioritize efficiency by approximating solutions more quickly, though often at the cost of some accuracy. They are especially suited for applications where decision speed outweighs the need for optimal solutions, offering a practical balance between performance and resource use. Building on the efficiency of heuristic approaches, meta-heuristic models, including techniques such as genetic algorithms and simulated annealing, bring additional adaptability and robustness to the IM problem by allowing fine-tuning across diverse network structures. These approaches generally strike an effective compromise between solution quality and efficiency, although their results can vary due to the sensitivity to parameter settings. In addition, centrality-based models, are among the most popular approaches for identifying influential nodes in multiplex networks, primarily due to their low computational cost. By leveraging centrality measures, these methods efficiently identify key nodes for maximizing the influence spread, which is particularly beneficial in large-scale networks where computational resources are limited. However, centrality-based models often focus solely on node importance within individual layers, potentially overlooking influence dynamics that span across multiple layers, which can limit their accuracy in modeling multiplex interactions. To address this limitation, Community structure-based models leverage network modularity to identify influential nodes within communities, allowing for efficient scaling in networks with distinct community structures. These models excel in networks where communities are well-defined but may be less effective when influence needs to extend beyond community boundaries. Furthermore, by often ignoring interlayer dependencies, community-based models can fall short of accurately capturing influence propagation in complex, interconnected networks. Recently, as research interest in multiplex IM has grown, alternative methodologies have emerged, including machine learning-based, game theory-based, and evidence theory-based approaches. While these newer methods hold promise for advancing IM in multiplex networks, they remain relatively underexplored, with only a limited number of studies compared to other approaches. To this end, we aggregated these emerging approaches, along with other less prevalent methodologies, into a section named "others" (Section 4.2.6). This subsection contains works from various families with a limited number of papers, typically ranging from one to two. In addition to these method families, the literature on multiplex IM can also be grouped based on how the structure of the multiplex network is handled into three main categories. The first common family is to aggregate multiplex networks into a single-layer representation and then estimate influence within this simplified structure. This approach has the benefit of reducing computational complexity but may risk information loss by oversimplifying inter-layer relationships. The second family treats each layer independently, assessing influence within each layer before combining the results through a fusion function to estimate network-wide influence. This method captures layer-specific influence patterns but can introduce redundancies due to the independent treatment of layers. A third family involves calculating influence directly within the multiplex network, preserving inter-layer dependencies to achieve a more nuanced view of influence spread. Although this method provides a more accurate representation of multiplex interactions, it is typically associated with higher computational costs. While we did not adopt this categorization in our review, each approach has unique implications for balancing accuracy, complexity, and computational efficiency.

In Table 6, we present a theoretical comparison of the different models discussed in Section 4.2. We evaluate each model's effectiveness in capturing both inter- and intra-layer information, the types of multilayer networks used, their time complexity, and key advantages and limitations. Intra- and inter-layer information columns denote whether each model can account for dependencies within and between layers.

6. New trends and future directions

The different papers discussed in this review have made significant contributions to the influence maximization area within multilayer networks. However, many aspects have been ignored which opens the door to several new trends and directions for future research. This section presents some future directions that should be addressed when studying the IM problem in multilayer networks.

- *Dynamic-based multiplex IM models:* Real-world social networks are dynamic by nature and their structure changes over time. They exhibit many practical scenarios where new users are added and others are deleted (Nesrine et al., 2020). Therefore, the study of IM problems in multiplex dynamic and temporary graphs is suggested. Recently, many efforts have been introduced for dynamic IM in the single OSN. However, determining the most influential users in dynamic multiplex networks has not yet been explored, which represents an interesting research line.
- *Privacy protection-based multiplex IM models:* As user privacy protection is a crucial part of OSNs, it recently has gained wide interest among researchers, especially in studying the IM problem. This problem aims to detect top-k influential users while taking into consideration the concerns of confidentiality. Since the public profile of a user has become an important factor of social influence, it is difficult to find things or people that he would be influenced by without public information about a user. Therefore, IM in multiplex OSNs under privacy protection can be a good research axe in which research can be considered in the future.
- *Group-based IM in multiplex OSNs:* Recently, a new research direction has emerged along with the IM problem within single-layer OSNs called the Group-based influence maximization (GIM) (Azaouzi et al., 2021). This problem consists of selecting a k-top group instead of selecting k-top nodes to maximize influence over the network. To study the GIM Problem, a group is abstracted as several users and this group will be activated if a certain number of users in this group are activated. Therefore, extending the GIM problem to multiplex social networks is a very quite interesting research area in which researchers might consider it.
- *Semantic-based multiplex IM models:* So far existing multiplex influence maximization models, focus only on structure-based models which take into consideration the topological structure features of the network. However, the semantic richness of the information associated with users and the relationships between them is not yet considered. Thus, such semantic-based models should be proposed in future works.

7. Conclusion

In this paper, we provided a condensed literature review of the different IM models within multilayer networks, with a major emphasis on multiplex networks. Existing models were reviewed and classified according to the adopted seed selection strategy. Our review starts by presenting the influence maximization problem and its relative concepts in its basic form. Then, focusing on the MxIM problem, a detailed classification of the multiplex IM models was introduced, where we identified five main families: greedy-based models, heuristic-based models, meta-heuristic-based models, centrality measures-based models, and community structure-based models. These families were mainly adopted with the IM problem in single-layer networks. In addition, this paper discussed the strengths and limitations of the different

Table 6
Comparison of the multiplex IM approaches. NONs: Network of networks, HMultilayer: Heterogeneous multilayer networks.

Family	Approach	Intralayer information	Interlayer information	Network type	Time complexity	Advantages	Disadvantages
	Zhang et al. (2015)	✓	✗	Multiplex	$O((m+n)\cdot nd)$	– Efficient coupling schemes for solving MxIM problem under various diffusion models. – Applicable for large networks with millions of nodes. – Improved time complexity by 700 times compared to the greedy algorithm.	– High memory consumption due to the use of extra links and nodes during the coupling process. – Possibility of information loss caused by the coupling process.
	Hosni et al. (2018)	✓	✗	Multiplex	[a]	– Efficient to minimize the spread of rumors. – Present a new rumor propagation model based on the IC model.	– Depend on the overlapping users. – High computational cost.

(*continued on next page*)

Table 6 (*continued*).

	Hosni et al. (2020)	✓	✗	Multiplex	$O(nkN\lvert\bar{E}\rvert)$	– Propose a new rumor diffusion model based on individual and social behaviors – Tested under multiple RIM strategies. – Achieve an average of 69% reduction in the number of people most likely to be infected by the rumor.	– High computational complexity.
	Kuhnle et al. (2018)	✓	✗	Multiplex	$O((k+l)(m+n)logn/\epsilon^2+(kl)^{\lceil 1/\epsilon-1\rceil}logk)$	– Introduce the concept of heterogeneous diffusion models, by allowing each layer to have its own propagation model. – Propose a new diffusion property (GDS). – Parallel approximation algorithm designed using algorithm A and a knapsack-based approach to deal with multiplex network.	– High computational complexity caused by Monte-Carlo simulations. – Depend on the number of overlapping users. – Inefficient for large-scale networks.
	Chen and Tan (2020)	✓	✓	Multiplex	[a]	– Improve the time complexity of the greedy algorithm by ignoring the use of Monte Carlo simulations. – Presented a new diffusion model based on the LT model, which considers the impact of topological relationships between layers on propagation, the overlapping effects of nodes between layers, and the strength of relationships among them.	– Depend on the number of snapshots. – Inefficient for large networks.
	Chen et al. (2020)	✓	✗	Multilayer	$O(\frac{(k+l)(n+m)log(n)}{\epsilon^2})$	– Introduce the concept of multiple influences in multilayer networks. – Use the fair seed allocation problem with the multiple influences. – Develop a new diffusion model based on the IC model to account for multiple influences competing in multilayer networks.	– Scalability issue. – High time complexity.
	Li et al. (2020)	✓	✓	Multilayer	$o(m\bar{d}N)$ and $O(m\bar{d}(d-1))$	– Significantly reduces the time complexity of the greedy algorithm – Present a new metric for evaluating the influence spread of nodes: the global propagation probability. – Propose a new spreading model based on the IC model and the theory of bond percolation. – Linear time complexity for selecting seed nodes in multilayer networks.	– Inefficient for large-scale networks. – Tested only on the synthetic multilayer network. – Require network structure to be known.

(*continued on next page*)

Table 6 (*continued*).

Heuristic-based models	Gaye et al. (2016)	✓	✓	Multilayer	[a]	– A new centrality measure designed based on the diffusion probability and degree centrality.	– Scalability issue.
	Meng et al. (2021)	✓	✓	Multiplex	[a]	– Study the multiplex IM problem in dynamic networks. – Develop a new spreading model based on common users under the DIC model	– Inefficient for large-scale networks. – Common users' number constraint.
	Chen, Deng et al. (2019)	✓	✓	Multiplex	[a]	– Use the concept of overlapping influence. – Detect seeds with a relatively sparse location and a lower overlapping influence. – Propose a multiplex shortest path based on the spreading rate.	– High computational complexity. – Inapplicable for networks with huge size. – Sensitivity to the used parameters.
	Zhan et al. (2015)	✓	✓	Multilayer	[a]	– Study the IM problem in heterogeneous aligned networks. – High Accuracy. – Extend the LT model by two diffusion weights: intra- and inter-network diffusion weights.	– High computational complexity. – Depend on the ratio of anchor users.
	Zhang et al. (2023)	✓	✓	Multiplex	[a]	– Achieve a high influence spread with low influence overlap based on neighborhood information. – Propagation occurs simultaneously across different layers under the IC model. – Simple and intuitive.	– Consider only one-hop neighborhood information.
	Singh, Singh et al. (2019)	✓	✓	Multiplex	$O((l+m)(M+N)+(k+m)(M+N\log N))$	– Propose various coupling strategies to transform the multiplex network into a monoplex network with low information loss. – High influence spread.	– Possibility of loss of information during the aggregation process.
	Li et al. (2016)	✓	✓	Multiplex	[a]	– Simple and intuitive.	
Meta-heuristic-based models	Lu et al. (2020)	✓	✓	Multilayer	$O(n^2 \log n)$	– Define the concept of pairwise reciprocal length and pairwise influence based on information diffusion in multilayer networks. – Cope with large-scale networks. – High influence spread – Evaluated under many spreading models.	
	Hu et al. (2023)	✓	✓	Multiplex	[a]	– Consider the node selection cost in solving the IM problem in multiplex networks.	– Influence spread is limited to a two-hop neighborhood.
	Wang et al. (2019)	✓	✓	Multiplex	[a]	– Fast and efficient. – Evaluated on synthetic and real datasets. – Extended the IC model by introducing the degree centrality into the influence spreading probability. – Define a minimal active layer (L_{min}) into the diffusion process.	– Consider only the influence within two hops. – Sensitive to the value of L_{min}.

(*continued on next page*)

Table 6 (*continued*).

	Zhang et al. (2018)	✓	✓	Multiplex	[a]	– Effective and efficient	– Depend only on the structure properties of networks. – Convergence issue.
	Basaras et al. (2017)	✓	✓	Multilayer	[a]	– Achieve high influence spread. – Improve the SIR model by the inter-layer spreading probability. – Tested within a variety of network structures and sizes.	– Based only on the local information topology of networks.
	Bouyer et al. (2024) and Zhou et al. (2023)	✓	✓	Multiplex	$O(n+NL^2+LE^a+nL)$	– Investigate new layer metrics to quantify the importance of layers based on a layer weighting method. – Consider inter-layer information alongside centrality measures to find influential nodes.	– Low influence propagation. – Tested on small-size networks.
	Ni et al. (2022)	✓	✓	Multiplex	[a]	– Improve the IC model by a new influence diffusion probability taking into account strength and superposition of relationships induced by multiple social connections. – Define three patterns of seeding strategies.	
	Ni and Yang (2022)	✓	✓	Multiplex	[a]	– Redefine the IC model defining two influence probabilities: intra-layer and inter-layer influence probabilities. – Introduce the concept of inter-layer conversion costs of influence diffusion on seeding strategies in multilayer networks. – High influence coverage.	– Inapplicable for large-scale networks. – Evaluated on small-size networks. – Sensitivity to the used parameters. – High time complexity caused by the computation of multiplex betweenness centrality.
Centrality-based models	Lv et al. (2024)	✓	✓	Multiplex	$O(N^2)+O(N\log N)$	– Improvement of the gravity centrality using multi-PageRank centrality. – Use both local and global topological information of networks.	– High computational complexity caused by the calculation of all possible shortest paths between any two nodes across all network layers. – Tested on small-size networks. – Sensitive to the used parameters. – Enable to cope with large-scale multilayer networks.
	Zeng et al. (2021)	✓	✓	Multiplex	[a]	– Propose a new centrality measure based on the dynamical and structural couplings between layers. – Test on synthetic and real datasets. – Robust and simple.	– Limited to the case of a two-layer multiplex network.
	Erlandsson et al. (2017)	✓	✗	Multilayer	[a]	– Improve the k-shell centrality by creating a hybrid centrality. – Tested various seeding strategies under the IC model. – Simple and clear.	– Low influence propagation. – Ignore the interlayer information. – Not suitable for complex multilayer networks.

(*continued on next page*)

Table 6 (*continued*).

	Wang et al. (2023)	✓	✓	Multilayer	$O(2N)$	– Take into account the effect of layer number and inter-layer coupling on node centrality. – Evaluated with real multilayer networks of different sizes. – Low computational complexity – Cope with large-scale networks.	– Not verified with real applications such as detection of influential nodes.
	Wang, Wang et al. (2017)	✓	✓	Multiplex	[a]	– Present a new centrality metric which take into consideration the intra- and inter-layer information. – Cope with networks with huge size.	– Sensitive to the used parameters. Computational Complexity issue. – Depend on the effectiveness of tensor representation. – Lack of validation with diffusion models.
	Hu et al. (2022)	✓	✗	Multiplex	$O(L \times n)$	– Consider privacy protection issue with the IM in multilayer networks. – Study the problem of collaboratively identifying the most influential nodes in multilayer private networks.	– Ignore the interlayer information. – High computational complexity, particularly for large scale-networks.
	Wan et al. (2022)	✓	✓	Multilayer	$O(n^2 + \|N\| \times \|E\|)$	– Simple and intuitive. – Achieve high-quality seeds.	– Not applicable to the case of multiplex networks. – Limited to unweighted and undirected networks.
Community-based models	Katukuri et al. (2021)	✓	✓	Multiplex	$O(Y(n)) + O(rNOMC + NOMClog(NOMC)) + O(K)$	– Leverage the concept of clique structures. – Ignore the information redundancy.	– Supposed that the community structure of the network is known. – Low influence propagation.
	Rao and Chowdary (2022)	✓	✗	Multiplex	$O(V^2 + V logV + 2VM + 1) + O(CS^3V + M logM + V)$	– Fast and efficient. – Reduce the research space by merging small communities into larger ones.	– Ignore the interlayer information. – Seeds are selected independently from each layer.
	Lv et al. (2023)	✓	✓	Multiplex	$O((K_1 + K_2)M + N(k_{max})^2 + L + logN + K_{kL})$	– Integrate the interlayer information into the community structure detection. – Improve the PageRank centrality by the random walks concept.	– Sensitive to the used parameters.
	Huang et al. (2020)	✓	✓	Multiplex	$O(2n(K + log(n)) + K^2 logk(1 + log(n)) - n)$	– Lower computational complexity. – Tested on 32 real-world datasets with different sizes.	
	Mittal and Bhatia (2017), Mittal et al.	✓	✓	Multiplex	[a]	– Introduce the concept of structural holes.	– High computational complexity. – Deal with small multiplex networks.
	Chowdary et al. (2023)	✓	✓	Multiplex	$O(V + V^2 + n \times V + E + V logV + K)$	– Fast and efficient.	
	Fujita and Tsugawa (2023)	✓	✗	Multiplex	[a]	– Propose a solution for Limiting the spread of negative influence.	– Sensitive to the used parameters.

(*continued on next page*)

Table 6 (*continued*).

	Li, Zhang et al. (2018)	✓	✗	NONs	[a]	– Use the evidence theory to fuse the influence of node on different single-layer networks. – Reduce information loss caused by the aggregation of influence from different network layers.	– Tested only in one dataset with a small dataset. – Ignore the interlayer information.
	Lei et al. (2023)	✓	✗	NONs	[a]	– Efficient solution to find influential nodes in multiplex networks.	– Test in small datasets. – Ignore the interlayer information.
	Li, Tang et al. (2021)	✓	✓	Multi-relationship	[a]	– Present a solution to detect overlapping users in multilayer networks.	– High computational complexity issue.
	Zhong et al. (2017)	✓	✓	Multiplex	[a]	– Robust & Consistent – Define a weighted LT model using live-edge models.	– Limited to two-layer networks. – Ignore the influence spread analysis.
	Zhuang and Yagan (2019)	✓	✓	Multiplex	[a]	– Define a new nodes state: Hyper-active.	– Test only on synthesized networks. – Low accuracy. – Limited to random networks.
	Liu et al. (2023)	✓	✓	Multiplex	[a]	– Study the information-disease coupled spreading dynamics on multiplex networks. – Investigate the effect of three interlayer coupling factors on the node's spreading influence on disease spreading.	– Evaluated with synthetic dataset. – Sensitive to the used parameters.
	Rani and Kumar (2022)	✓	✓	HMulti-layer	$O(n^2)$	– Propose various seed selection strategies. – law time complexity.	– Consider heterogeneous networks with different types of nodes and edges.
Others	Yu et al. (2020)	✓	✗	Multiplex	[a]	– Introduce the concept of limited contact, in which an individual can only transmit information to a limited number of neighbors.	– Influence coverage is not analyzed. – Sensitive the several parameters. – Evaluated on a synthetic network.
	Al-Garadi et al. (2016)	✓	✓	Multiplex	[a]	– Analyze the effect of different topological network coupling methods on IM problem. – Simple and clear.	– Information loss caused by the coupling process. – Based only on classical seeding selection methods (No innovation).
	Do et al. (2024)	✓	✓	Multiplex	$O(\max_{h\in k} \left(tc(A,G_h,l)\right)^{\left\lceil \frac{1}{1-\epsilon} \right\rceil} \log(k) + \lvert Y\rvert^2 \times (k-1)+Q)$	– Leverage the use of deep reinforcement learning. – Able to scale over huge networks. – Used heterogeneous propagation models. – Balancing the model size and computational efficiency. – Reduce the time complexity.	
	Keikha et al. (2020)	✓	✓	Interconnected networks	$O(nlogn)$	– Cope with different types of networks and large scale. – Use the network embedding method to learn the local and global structure of each network node.	– Consider networks with different sizes and heterogeneous nodes. – Comparison with very old algorithms.

(*continued on next page*)

Table 6 (*continued*).

Wang and Tan (2022)	✓	✓	Multiplex	$O(MaxGen \times P \times K)$	– Parallel model developed to cope with multiple optimization scenarios.	– Tested on small datasets. – Influence computation is limited to the region's two hops. – Sensitive to the used parameters.
Yu et al. (2021)	✓	✓	Multiplex	[a]	– Investigate an heterogeneous diffusion model based on game-theory.	– Tested on a synthetic network. – Sensitive to the used parameters.
Chen, Qian et al. (2019)	✓	✓	Multiplex	[a]	– Cope with large-scale networks.	– High computational complexity. – Depend on the number of hops.

[a] Not mentioned.

existing multiplex IM models. Finally, new trends and opportunities have been presented and discussed.

CRediT authorship contribution statement

Oumaima Achour: Conceptualization, Writing – review & editing. **Lotfi Ben Romdhane:** Supervision, Validation.

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

Aghaee, Z., & Kianian, S. (2020). Efficient influence spread estimation for influence maximization. *Social Network Analysis and Mining*, *10*(1), 1–21.

Al-Garadi, M. A., Varathan, K. D., Ravana, S. D., Ahmed, E., & Chang, V. (2016). Identifying the influential spreaders in multilayer interactions of online social networks. *Journal of Intelligent & Fuzzy Systems*, *31*(5), 2721–2735.

Amara, A., Taieb, M. A. H., & Aouicha, M. B. (2022). Cross-network representation learning for anchor users on multiplex heterogeneous social network. *Applied Soft Computing*, *118*, Article 108461.

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

Banerjee, S., Jenamani, M., & Pratihar, D. K. (2020). A survey on influence maximization in a social network. *Knowledge and Information Systems*, *62*(9), 3417–3455.

Basaras, P., Iosifidis, G., Katsaros, D., & Tassiulas, L. (2017). Identifying influential spreaders in complex multilayer networks: A centrality perspective. *IEEE Transactions on Network Science and Engineering*, *6*(1), 31–45.

Bedi, P., & Sharma, C. (2016). Community detection in social networks. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *6*(3), 115–135.

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

Bihari, A., & Pandia, M. K. (2015). Eigenvector centrality and its application in research professionals' relationship network. In *2015 international conference on futuristic trends on computational analysis and knowledge management* (pp. 510–514). IEEE.

Bouyer, A., Mohammadi, M., & Arasteh, B. (2024). Identifying influential nodes based on new layer metrics and layer weighting in multiplex networks. *Knowledge and Information Systems*, *66*(2), 1011–1035.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, *30*(1–7), 107–117.

Bródka, P., Musial, K., & Jankowski, J. (2020). Interacting spreading processes in multilayer networks: A systematic review. *IEEE Access*, *8*, 10316–10341.

Brown, P., & Feng, J. (2011). Measuring user influence on twitter using modified k-shell decomposition. *vol. 5*, In *Proceedings of the international AAAI conference on web and social media* (pp. 18–23).

Cardillo, A., Zanin, M., Gómez-Gardenes, J., Romance, M., García del Amo, A. J., & Boccaletti, S. (2013). Modeling the multi-layer nature of the European air transport network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*, *215*, 23–33.

Chan, T.-H. H., Ning, L., & Zhang, Y. (2020). Influence maximization under the non-progressive linear threshold model. In *International workshop on frontiers in algorithmics* (pp. 37–48). Springer.

Chen, J., Deng, Y., Su, Z., Wang, S., Gao, C., & Li, X. (2019). Identifying multiple influential users based on the overlapping influence in multiplex networks. *IEEE Access*, *7*, 156150–156159.

Chen, S., Qian, H., Wu, Y., Chen, C., & Wang, X. (2019). Efficient adoption maximization in multi-layer social networks. In *2019 international conference on data mining workshops* (pp. 56–60). IEEE.

Chen, S., & Tan, W. (2020). Influence diffusion model in multiplex networks. *Computers, Materials & Continua*, *64*(1).

Chen, Y., Wang, W., Feng, J., Lu, Y., & Gong, X. (2020). Maximizing multiple influences and fair seed allocation on multilayer social networks. *Plos One*, *15*(3), Article e0229201.

Chowdary, C. R., et al. (2023). K++ shell: Influence maximization in multilayer networks using community detection. *Computer Networks*, *234*, Article 109916.

De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., et al. (2013). Mathematical formulation of multilayer networks. *Physical Review X*, *3*(4), Article 041022.

Ding, J., Sun, W., Wu, J., & Guo, Y. (2020). Influence maximization based on the realistic independent cascade model. *Knowledge-Based Systems*, *191*, Article 105265.

Do, N., Chowdhury, T., Ling, C., Zhao, L., & Thai, M. T. (2024). MIM-reasoner: Learning with theoretical guarantees for multiplex influence maximization. arXiv preprint arXiv:2402.16898.

El Kouni, I. B., Karoui, W., & Romdhane, L. B. (2020). Prucars: improved association rule-based social recommender systems using overlapping community detection. *Procedia Computer Science*, *176*, 956–965.

Erlandsson, F., Bródka, P., & Borg, A. (2017). Seed selection for information cascade in multilayer networks. In *International conference on complex networks and their applications* (pp. 426–436). Springer.

Fan, C., Wang, Z., Zhang, J., Zhao, J., & Rui, X. (2024). Influence maximization in social networks based on discrete harris hawks optimization algorithm. *Computing*, *106*(2), 327–351.

Freeman, L. C., Roeder, D., & Mulholland, R. R. (1979). Centrality in social networks: II. Experimental results. *Social Networks*, *2*(2), 119–141.

Fujita, Y., & Tsugawa, S. (2023). Limiting the spread of misinformation on multiplex social networks. In *2023 IEEE 47th annual computers, software, and applications conference* (pp. 406–411). IEEE.

Gaye, I., Mendy, G., Ouya, S., Diop, I., & Seck, D. (2016). Multi-diffusion degree centrality measure to maximize the influence spread in the multilayer social networks. In *International conference on e-infrastructure and e-services for developing countries* (pp. 53–65). Springer.

Goldenberg, J., Libai, B., & Muller, E. (2001). Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, *12*, 211–223.

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

Gursoy, F., & Gunnec, D. (2018). Influence maximization in social networks under deterministic linear threshold model. *Knowledge-Based Systems*, *161*, 111–123.

Han, M., Yan, M., Cai, Z., Li, Y., Cai, X., & Yu, J. (2017). Influence maximization by probing partial communities in dynamic online social networks. *Transactions on Emerging Telecommunications Technologies*, *28*(4), Article e3054.

He, X., Song, G., Chen, W., & Jiang, Q. (2012). Influence blocking maximization in social networks under the competitive linear threshold model. In *Proceedings of the 2012 siam international conference on data mining* (pp. 463–474). SIAM.

Heidari, M., Asadpour, M., & Faili, H. (2015). SMG: Fast scalable greedy algorithm for influence maximization in social networks. *Physica A. Statistical Mechanics and its Applications*, *420*, 124–133.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572.

Hosni, A. I. E., Li, K., & Ahmad, S. (2020). Minimizing rumor influence in multiplex online social networks based on human individual and social behaviors. *Information Sciences*, *512*, 1458–1480.

Hosni, A. I. E., Li, K., Ding, C., & Ahmed, S. (2018). Least cost rumor influence minimization in multiplex social networks. In *International conference on neural information processing* (pp. 93–105). Springer.

Hu, W., Xia, X., Ding, X., Zhang, X., Zhong, K., & Zhang, H.-F. (2022). SMPC-ranking: a privacy-preserving method on identifying influential nodes in multiple private networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *53*(5), 2971–2982.

Hu, X.-M., Zhao, Y.-Q., & Yang, Z. (2023). Nodes grouping genetic algorithm for influence maximization in multiplex social networks. In *2023 26th international conference on computer supported cooperative work in design* (pp. 1130–1135). IEEE.

Huang, X., Chen, D., Wang, D., & Ren, T. (2020). Mine: identifying top-k vital nodes in complex networks via maximum influential neighbors expansion. *Mathematics*, *8*(9), 1449.

Jaouadi, M., & Romdhane, L. B. (2021). A distributed model for sampling large scale social networks. *Expert Systems with Applications*, *186*, Article 115773.

Jaouadi, M., & Romdhane, L. B. (2024). A survey on influence maximization models. *Expert Systems with Applications*, *248*, Article 123429.

Katukuri, M., Jagarapu, M., et al. (2021). CIM: clique-based heuristic for finding influential nodes in multilayer networks. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, 1–12.

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika.*, *18*(1), 39–43.

Keikha, M. M., Rahgozar, M., & Asadpour, M. (2018). Community aware random walk for network embedding. *Knowledge-Based Systems*, *148*, 47–54.

Keikha, M. M., Rahgozar, M., Asadpour, M., & Abdollahi, M. F. (2020). Influence maximization across heterogeneous interconnected networks based on deep learning. *Expert Systems with Applications*, *140*, Article 112905.

Kempe, D., Kleinberg, J., & Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 137–146).

Kim, J., & Hastak, M. (2018). Social network analysis: Characteristics of online social networks after a disaster. *International Journal of Information Management*, *38*(1), 86–96.

Kim, J., Lee, W., & Yu, H. (2014). CT-IC: Continuously activated and time-restricted independent cascade model for viral marketing. *Knowledge-Based Systems*, *62*, 57–68.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, *2*(3), 203–271.

Kourtellis, N., Alahakoon, T., Simha, R., Iamnitchi, A., & Tripathi, R. (2013). Identifying high betweenness centrality nodes in large social networks. *Social Network Analysis and Mining*, *3*(4), 899–914.

Kuhnle, A., Alim, M. A., Li, X., Zhang, H., & Thai, M. T. (2018). Multiplex influence maximization in online social networks with heterogeneous diffusion models. *IEEE Transactions on Computational Social Systems*, *5*(2), 418–429.

Kumar, S., Singhla, L., Jindal, K., Grover, K., & Panda, B. (2021). IM-ELPR: Influence maximization in social networks using label propagation based community structure. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, *51*(11), 7647–7665.

Lei, M., Liu, L., & Xiao, F. (2023). Identify influential nodes in network of networks from the view of weighted information fusion. *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies*, *53*(7), 8005–8023.

Li, H., Chu, Z., Feng, P., Lv, H., Wang, T., & Liu, J. (2021). Application of maximum influence in scientific collaboration network. In *2021 36th youth academic annual conference of Chinese association of automation* (pp. 881–885). IEEE.

Li, G., Chu, Y., Feng, J., & Xu, Y. (2016). Influence maximization on multiple social networks. *Chinese Journal of Computers*, *39*(4), 643–656.

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

Li, X., Tang, Y., Du, Y., & Li, Y. (2021). Key node discovery algorithm based on multiple relationships and multiple features in social networks. *Mathematical Problems in Engineering*, *2021*, 1–12.

Li, M., Zhang, Q., & Deng, Y. (2018). Evidential identification of influential nodes in network of networks. *Chaos, Solitons & Fractals*, *117*, 283–296.

Li, X., Zhang, X., Zhao, C., Yi, D., & Li, G. (2020). Identifying highly influential nodes in multilayer networks based on global propagation. *Chaos. An Interdisciplinary Journal of Nonlinear Science*, *30*(6).

Liu, Y., Li, A., Zeng, A., Zhou, J., Fan, Y., & Di, Z. (2024). Motif-based community detection in heterogeneous multilayer networks. *Scientific Reports*, *14*(1), 8769.

Liu, Y., Zeng, Q., Pan, L., & Tang, M. (2023). Identify influential spreaders in asymmetrically interacting multiplex networks. *IEEE Transactions on Network Science and Engineering*.

Lu, Q., Bu, Z., & Wang, Y. (2020). A multiobjective evolutionary approach for influence maximization in multilayer networks. In *Proceedings of the 2020 6th international conference on computing and artificial intelligence* (pp. 431–438).

Lv, L., Hu, P., Zheng, Z., Bardou, D., Zhang, T., Wu, H., et al. (2023). A community-based centrality measure for identifying key nodes in multilayer networks. *IEEE Transactions on Computational Social Systems*.

Lv, L., Zhang, T., Hu, P., Bardou, D., Niu, S., Zheng, Z., et al. (2024). An improved gravity centrality for finding important nodes in multi-layer networks based on multi-PageRank. *Expert Systems with Applications*, *238*, Article 122171.

Meng, Y., Chen, N., Yi, Y., Wang, S., & Pei, C. (2021). Research on the dynamic multisocial networks influence maximization problem based on common users. *IEEE Access*, *9*, 127407–127419.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

Mishra, S., Singh, S. S., Kumar, A., & Biswas, B. (2022). Mnerlp-mul: Merged node and edge relevance based link prediction in multiplex networks. *Journal of Computer Science*, *60*, Article 101606.

Mittal, R., & Bhatia, M. S. (2017). Mining top-k structural holes in multiplex networks. In *2017 8th international conference on computing, communication and networking technologies* (pp. 1–6). IEEE.

Molaei, R., Rahsepar Fard, K., & Bouyer, A. (2024). An improved influence maximization method for online advertising in Social Internet of Things. *Big Data*, *12*(3), 173–190.

Moosavi, S. A., & Jalali, M. (2014). Community detection in online social networks using actions of users. In *2014 Iranian conference on intelligent systems* (pp. 1–7). IEEE.

Nesrine, H., Wafa, K., & RomdhaneLotfi, B. (2020). Influential nodes detection in dynamic social networks: A survey. *Expert Systems with Applications*, *159*, Article 113642.

Ni, C., & Yang, J. (2022). Two-step seeding strategy in multiplex networks with inter-layer conversion cost of influence. *Chaos. An Interdisciplinary Journal of Nonlinear Science*, *32*(8).

Ni, C., Yang, J., Pang, Z., & Gong, Y. (2022). Seeding strategy based on weighted gravity centrality in multiplex networks. *IEEE Transactions on Network Science and Engineering*, *10*(1), 331–345.

Okamoto, K., Chen, W., & Li, X.-Y. (2008). Ranking of closeness centrality for large-scale social networks. In *International workshop on frontiers in algorithmics* (pp. 186–195). Springer.

Peng, S., Zhou, Y., Cao, L., Yu, S., Niu, J., & Jia, W. (2018). Influence analysis in social networks: A survey. *Journal of Network and Computer Applications*, *106*, 17–32.

Rani, S., & Kumar, M. (2022). Influential node detection and ranking with fusion of heterogeneous social media information. *IEEE Transactions on Computational Social Systems*.

Rao, K. V., & Chowdary, C. R. (2022). CBIM: Community-based influence maximization in multilayer networks. *Information Sciences*, *609*, 578–594.

Saito, K., Kimura, M., Ohara, K., & Motoda, H. (2012). Learning asynchronous-time information diffusion models and its application to behavioral data analysis over social networks. arXiv preprint arXiv:1204.4528.

Seidman, S. B. (1983). Network structure and minimum degree. *Social Networks*, *5*(3), 269–287.

Singh, A. K., & Kailasam, L. (2021). Link prediction-based influence maximization in online social networks. *Neurocomputing*, *453*, 151–163.

Singh, S. S., Kumar, A., Singh, K., & Biswas, B. (2019). C2IM: Community based context-aware influence maximization in social networks. *Physica A: Statistical Mechanics and Its Applications*, *514*, 796–818.

Singh, S. S., Singh, K., Kumar, A., & Biswas, B. (2019). MIM2: Multiple influence maximization across multiple social networks. *Physica A. Statistical Mechanics and its Applications*, *526*, Article 120902.

Singh, S. S., Srivastva, D., Verma, M., & Singh, J. (2021). Influence maximization frameworks, performance, challenges and directions on social network: A theoretical study. *Journal of King Saud University-Computer and Information Sciences*.

Srinivas, A., & Velusamy, R. L. (2015). Identification of influential nodes from social networks based on enhanced degree centrality measure. In *2015 IEEE international advance computing conference* (pp. 1179–1184). IEEE.

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

Sun, Y., Cautis, B., & Maniu, S. (2023). Social influence-maximizing group recommendation. *vol. 17*, In *Proceedings of the international AAAI conference on web and social media* (pp. 820–831).

Tang, X., & Yang, C. C. (2012). Ranking user influence in healthcare social media. *ACM Transactions on Intelligent Systems and Technology*, *3*(4), 1–21.

Tong, G., Wu, W., Tang, S., & Du, D.-Z. (2016). Adaptive influence maximization in dynamic social networks. *IEEE/ACM Transactions on Networking*, *25*(1), 112–125.

Wan, L., Zhang, M., Li, X., Sun, L., Wang, X., & Liu, K. (2022). Identification of important nodes in multilayer heterogeneous networks incorporating multirelational information. *IEEE Transactions on Computational Social Systems*, *9*(6), 1715–1724.

Wang, Q., Jin, Y., Yang, T., & Cheng, S. (2017). An emotion-based independent cascade model for sentiment spreading. *Knowledge-Based Systems*, *116*, 86–93.

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

Wang, S., & Tan, X. (2022). Determining seeds with robust influential ability from multi-layer networks: A multi-factorial evolutionary approach. *Knowledge-Based Systems*, *246*, Article 108697.

Wang, D., Tian, F., & Wei, D. (2023). A new centrality ranking method for multilayer networks. *Journal of Computer Science*, *66*, Article 101924.

Wang, D., Wang, H., & Zou, X. (2017). Identifying key nodes in multilayer networks based on tensor decomposition. *Chaos. An Interdisciplinary Journal of Nonlinear Science*, *27*(6).

Wang, Z., & Xia, C. (2020). Co-evolution spreading of multiple information and epidemics on two-layered networks under the influence of mass media. *Nonlinear Dynamics*, *102*, 3039–3052.

Wu, P., & Pan, L. (2017). Scalable influence blocking maximization in social networks under competitive independent cascade models. *Computer Networks*, *123*, 38–50.

Xu, B., & Liu, L. (2010). Information diffusion through online social networks. In *2010 IEEE international conference on emergency management and management sciences* (pp. 53–56). IEEE.

Xu, B., Zhang, Y., Chen, L., Yu, L., Li, L., & Wang, Q. (2021). The influence of social network on COVID-19 vaccine hesitancy among healthcare workers: a cross-sectional survey in Chongqing, China. *Human Vaccines & Immunotherapeutics*, *17*(12), 5048–5062.

Yang, S., Du, Q., Zhu, G., Cao, J., Chen, L., Qin, W., et al. (2024). Balanced influence maximization in social networks based on deep reinforcement learning. *Neural Networks*, *169*, 334–351.

Yang, W., Shi, Q., Yan, J., Wang, C., Song, M., & Wu, M. (2023). Complementary influence maximization under comparative linear threshold model. *Expert Systems with Applications*, Article 121826.

Yu, J., Liu, Z., & Han, X. (2021). Cooperation evolution in multiplex networks with the heterogeneous diffusion model. *IEEE Access*.

Yu, X., Yang, Q., Ai, K., Zhu, X., & Wang, W. (2020). Information spreading on two-layered multiplex networks with limited contact. *IEEE Access*, *8*, 104316–104325.

Zareie, A., & Sakellariou, R. (2024). Fuzzy influence maximization in social networks. *ACM Transactions on the Web*, *18*(3), 1–28.

Zeng, Q., Liu, Y., Tang, M., & Gong, J. (2021). Identifying super-spreaders in information–epidemic coevolving dynamics on multiplex networks. *Knowledge-Based Systems*, *229*, Article 107365.

Zhan, Q., Zhang, J., Wang, S., Yu, P. S., & Xie, J. (2015). Influence maximization across partially aligned heterogenous social networks. In *Pacific-Asia conference on knowledge discovery and data mining* (pp. 58–69). Springer.

Zhan, Q., Zhuo, W., & Liu, Y. (2019). Social influence maximization for public health campaigns. *IEEE Access*, *7*, 151252–151260.

Zhang, H., Nguyen, D. T., Zhang, H., & Thai, M. T. (2015). Least cost influence maximization across multiple social networks. *IEEE/ACM Transactions on Networking*, *24*(2), 929–939.

Zhang, S.-S., Xie, M., Liu, C., & Zhan, X.-X. (2023). Influence maximization in multilayer networks based on adaptive coupling degree. arXiv preprint arXiv:2311.08663.

Zhang, Q., Zhong, L., Gao, S., & Li, X. (2018). Optimizing hiv interventions for multiplex social networks via partition-based random search. *IEEE Transactions on Cybernetics*, *48*(12), 3411–3419.

Zhong, Y. D., Srivastava, V., & Leonard, N. E. (2017). On the linear threshold model for diffusion of innovations in multiplex social networks. In *2017 IEEE 56th annual conference on decision and control* (pp. 2593–2598). IEEE.

Zhou, X., Bouyer, A., Maleki, M., Mohammadi, M., & Arasteh, B. (2023). Identifying top influential spreaders based on the influence weight of layers in multiplex networks. *Chaos, Solitons & Fractals*, *173*, Article 113769.

Zhou, C., Zemanová, L., Zamora, G., Hilgetag, C. C., & Kurths, J. (2006). Hierarchical organization unveiled by functional connectivity in complex brain networks. *Physical Review Letters*, *97*(23), Article 238103.

Zhuang, Y., & Yagan, O. (2019). Influence propagation with multiple stages over random multiplex networks. In *2019 IEEE global communications conference* (pp. 1–6). IEEE.