Knowledge-Based Systems 246 (2022) 108697

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Determining seeds with robust influential ability from multi-layer networks: A multi-factorial evolutionary approach

Shuai Wang, Xiaojun Tan *

School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen 518107, China

ARTICLE INFO

Article history:
Received 2 December 2021
Received in revised form 25 March 2022
Accepted 27 March 2022
Available online 6 April 2022

Keywords:
Complex networks
The influence maximization problem
Robustness
The multi-tasking optimization
Evolutionary algorithm

ABSTRACT

There has been a great stream of literature in studying the dynamics of information diffusion processes attached on networked systems. And the corresponding influential seeds selection task can be modeled as the influence maximization problem. Effective diffusion models and methods are developed to detect powerful seeds from both single- and multi-layer networks. Meanwhile, some recent studies indicate that the structural destruction also perturbs seeds' spreading procedures, and robust seeds are expected in daily applications. Based on a performance measure for the robust influence maximization problem, the existing solution intends to determine a specific damage percentage first; guided by which, solutions are provided. However, the generality of such method is non-guaranteed, and the knowledge in search processes towards multiple scenarios is completely neglected. Focusing on these deficiencies, the multi-tasking optimization theory has been introduced into the seed determination task from multi-layer networks. Multiple optimization scenarios are considered parallelly and the synergy between these tasks is exploited in the search process. Combining informative knowledge from both genetic and fitness domains, a multi-factorial evolutionary algorithm containing problem-directed operators, named MFEA-RIM$_m$, has been designed to solve the robust influence maximization problem on multi-layer networks. Structural characteristics in the inputted network are also emphasized in the local refinement process. Empirical analyzes demonstrate the notable performance of MFEA-RIM$_m$ over existing methods, and diversified results can be obtained simultaneously to cater to challenges in multiple information diffusive scenarios.

1. Introduction

Practical systems always possess complicated structures, which causes difficulties in related design and analysis tasks. Utilizing the implicit connective information in systems, the network theory [1] introduces possible approaches to fetch and analyze the dynamics and crucial information from systems. Increasing attention has been paid to modeling different kinds of networked systems, and a series of network models are constructed to simply modern systems, including the random network [2], the scale free network [3], and the small world network [4]. Plenty of properties such as the assortativity [5] and the community structure [6] have been discovered and summarized to further reveal the in-depth discipline pertaining to networks. These studies indicate the significance of the network theory in theoretical and practical fields [7], and boost related investigations in a wide range of topics.

Meanwhile, the thriving of multimedia technologies leads to a new informative era. The real-time access towards news and personal status is achievable with the help of online social network service platforms. Compared with traditional media like newspaper, broadcast, and TV, the emerging social network service like Tiktok, Facebook, and WeChat allows users to obtain massive information. Under this context, the online virtual marketing [8] has shown non-negligible values in advertising and propagating, and becomes a hotspot in recent decade. The discovery and determination of key systematic members may facilitate the dynamical analysis [9] and predication [10] in other systems. The spreading behavior or mechanism between nodes is fundamental for a valid virtual marketing operation. Several possible models have been proposed in existing studies; for example, the independent cascaded (IC) model [11] and the weighted cascaded (WC) model [12] are validated on plain and weighted networks, respectively. In allusion to determining those influential nodes in a system, the influence maximization (IM) problem [13] has been presented. Several nodes may be selected to constitute the seed set, which is expected to reach the maximal influence range in a certain network. Both heuristic-based [11,14] and population-based [15] optimization methods have been employed to find

* Corresponding author.
E-mail address: tanxj@mail.sysu.edu.cn (X. Tan).

https://doi.org/10.1016/j.knosys.2022.108697
0950-7051/

influential seeds. In addition, some network properties [16,17] are also helpful when solving the influence maximization problem. A solid theoretical basis has been established to model the information diffusion process on networks, and methods that intend to find influential nodes are also provided.

Based on the influence maximization problem, there has been a growing interest on two relevant topics. The first one is the extension on multi-layer networks. As shown in [18], the structure consisting of only single layer may be inadequate to depict modern complicated systems, and networks with multiple layers are optional to describe the functional distribution and interdependence of multiplex networked systems [19]. Focusing on the information diffusion process on multi-layer networks, the structural discrepancy is significant to the seed determination task [20]. Seeds that have steady spreading ability in the whole system are expected [21] to generate a considerable influence across all attached network layers. The population-based search technique has shown effectiveness when solving this problem [22]. Another topic is the robustness of the influence diffusion process. In terms of networks, the robustness is referred to the structural tolerance against attacks and errors [7]. Corresponding evaluation [23] and optimization [24] methods have been greatly emphasized in previous studies. Combined with the spreading behavior of seeds, the effect caused by structural destructions on the diffusion process is worth of investigations. Some recent studies indicated that the changeable diffusive models and parameters are likely to impact the performance of selected seeds [25,26]. A numerical measure has been designed in [27] to assess the robustness of seeds suffering from structural damages, which also provides a criterion for the selection task. Seeds with robust influential ability are expected to achieve stable spreading results even under sabotages. However, little work has touched upon the combination of the above two topics. The diffusion ability of seeds selected from a multi-layer networked system under structural perturbations is unclear; especially when there exist multiple damage scenarios. In other words, the robust influence maximization (RIM) problem on multi-layer networks still remains to be solved.

Several challenges should be addressed to solve the problem. First, the decision variable for the robust influence maximization problem is the selected nodes, which is highly discrete and different from the network structure required by structural robustness optimization like in [23,24]. Efficient strategies tend to be necessary in the search process to explore the network-based solution space. Meanwhile, the application environment of networked systems is complicated, it is thus hard to determine a general damage prototype for diverse networks. One possible solution is to define a relatively universal prototype through trials and errors as shown in [27], which is intuitive but the rationality cannot be guaranteed on common networks. A better approach may be through an overall consideration on different scenarios. Further, diversified information is provided in dealing with multiple optimization scenarios. The method that can efficiently utilize such information and transfer rational knowledge across optimization processes is thus significant. Operators like crossover and local search are required to fit the network-based data and excavate valuable information from current individuals.

In this paper, we manage to solve the RIM problem on multi-layer networks. Based on several related studies, multiple damage scenarios are tackled simultaneously to give an overall consideration of different situations. In order to utilize the valuable optimal information across tasks in the parallel optimization process, the multi-factorial evolutionary algorithm [28] has been introduced into the multi-tasking optimization process on multi-layer networks. Several features of the RIM problem on multi-layer networks are considered when designing the algorithm, including the discrete decision variable for constructing seed sets, the connective information intra-network together with the structural difference inter-networks, and the optimal knowledge implied in the genetic domain and the fitness domain. Problem-oriented genetic operators such as the lattice-based initialization operator, the distance-aware crossover operator for transferring knowledge, and the three-phase local search operator are developed to improve the search ability of the algorithm and address existing challenges. Equipped with which, a multi-factorial evolutionary algorithm (MFEA) has been devised to solve the RIM problem on multi-layer networks, termed MFEA-RIM_m. Experiments have been conducted on plenty of networks to demonstrate the performance of the proposed algorithm compared with existing methods. Meanwhile, the remarkable efficiency improvement induced by multi-tasking optimization has also been shown.

Contributions of this paper can be summarized as follows. First, in terms of the robust influence maximization problem, an extended research directed at multi-layer networks has been shown. It can be concluded the distributed functionality on multiple layers results in further complexity on networks' dynamics. Also, the vulnerability of multi-layer networks under structural failures incurs additional difficulty when finding influential seeds. Experiments reveal that a fixed parameter in R_S^{Multi} may not work well on some instantiations. The multi-tasking optimization theory has been introduced to solve this network-related problem in a parallel manner, which has seldom been touched upon in previous studies. The proposed algorithm represents a remarkable performance and improved efficiency compared with existing methods, and multiple solutions can be generated for decision makers simultaneously.

Second, for the multi-factorial evolutionary algorithm, a strategy that considers information from both genetic domain and fitness domain is designed to solve the robust seed determination task from multi-layer networks. Equipped with which, the proposed MFEA-RIM_m achieves remarkable performance compared with existing methods. We can see that an incisive employment of the obtained knowledge contributes to the promotion of the optimality ability.

Last, the RIM problem is network-based and exhibits discrete decision variables, and only little attention has been paid on such problem. This paper provides a successful case study of utilizing theory and technique in the computational field to tackle the information excavation task from networked systems. More attempts are imperative in the future to provide better solutions on the mining and optimizing challenges of networks.

The rest of this paper is organized as follows: Section 2 presents related work on the multi-layer network, the influence maximization problem together with its robustness, and the multi-factorial evolutionary algorithm. Section 3 depicts the multi-tasking framework for the RIM problem on multi-layer networks. Section 4 gives the details MFEA-RIM_m. Section 5 shows the experimental results on both synthetic and real-world networks. Section 6 reports conclusions and discussions of this work.

2. Related work

2.1. The network model and its robustness

The effectiveness of the complex network theory has been widely validated in engineering and theoretical fields. Plenty of complicated systems like infrastructures [1], social networks [6], the Internet [16] can be modeled as corresponding networks, which dramatically improves the efficiency in analyzing tasks. Generally, the single-layer network can be expressed as a connection matrix G consisting of node and link information. A popular

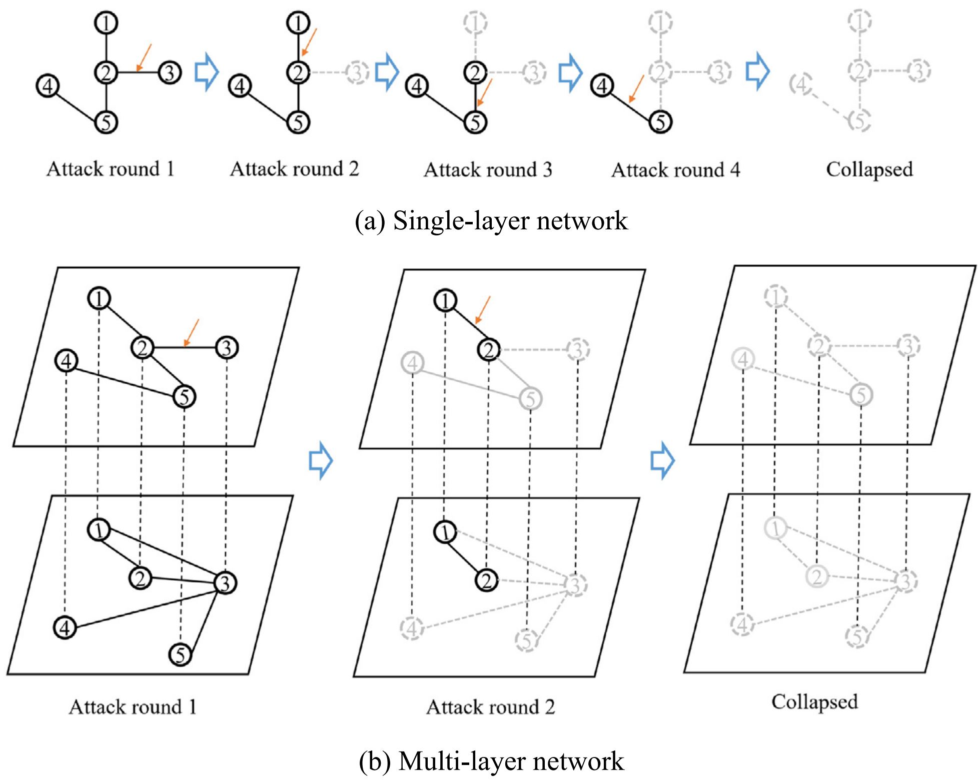


Fig. 1. The destruction process in the link-based removal process on (a) single-layer and (b) multi-layer networks. Failed nodes and links are marked in dashed gray. In (b), nodes are interdependent across layers. Once a node gets failed in one layer, this node is unfunctional in all layers. Four rounds of attacks lead to the collapse of the single-layer network in (a), but only two rounds of attacks result in the collapse of the multi-layer network in (b).

form is $G = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ and $\mathbf{E}$ represent the set of nodes and links, respectively. We assume that there is a total of N nodes $\{1, 2, \ldots, N\}$ in $\mathbf{V}$ and a total of M links $\{e_{ij}$, where $i, j \in \mathbf{V}\}$ in $\mathbf{E}$. In this manner, G gives an objective description on the connection information of a specific network. Based on which, the in-depth functional information is maintained in the topology, and some crucial topics such as the degree distribution [3], the structural property [5], and the community allocation [6] have been discovered and investigated. Meanwhile, the tolerance against attacks and errors, i.e., the robustness of G, is considered as a crux in the complex network field. G provides approaches to define and evaluate the robustness. For example, the structural cost to destroy the connectivity is one feasible measure to define networks' robustness [7]. This measure is intuitive but may omit the situation that the network is not totally destructed, and the measure R was proposed in [23] to tackle universal cases when the network still keeps partly functional suffering from destructions. R is defined as,

$$R = \frac{1}{N} \sum_{q=1}^{N} s(q) \tag{1}$$

In the equation, N represents the scale of G, and $1/N$ works as the normalization factor for networks with different scales. $s(q)$ is the largest connective fraction after q nodes are failed. This measure gives a numerical evaluation on the robustness of networks against node-based attacks, and a larger R indicates that the network tends to be more robust under destructions. Several extensions have been proposed to deal with link-based [19] and community-based [29] attacks.

The research on single-layer networks has been intensively emphasized; however, the deficiency lies in that such studies may omit the interdependence between different networked systems. As indicated by several pilot studies, the functionality of a whole system may rely on multiple networks [18] which are correlated with each other. A specific nodal member tends to possess diverse roles in different network layers. The multi-layer network model shows validity in the modeling and analyzing process of many real systems [30]. Generally, a network with multiple layers can be denoted as $G_m = [G_1, G_2, \ldots, G_L]$ where L is the number of layers. For simplification, we suppose that each layer G_l in G_m consists of N nodes. And the functionality of each node i in all layers is interdependent; i.e., the failure of i in only one layer is enough to cause the malfunction of i in G_m. This feature leads to the fragility of multi-layer networks against structural destructions. An intuitive example of the destruction process on a toy network can be found in Fig. 1. As shown in the figure, link-based attacks are destructive to the connectivity of both single-layer and multi-layer networks. In Fig. 1(a), the connectivity gets damaged gradually; but in Fig. 1(b), a cascading failure can be noticed which causes significant losses to the network. Focusing on node 3 in the multi-layer network, this node is structurally important in the lower layer but just ordinary in the upper layer. When node 3 is disconnected in the upper layer, this node is considered as unfunctional, and all attached links in other layers are to be removed. In this way, nodes 4 and 5 in the lower layer also loss the functionality, and the damage can spread back and forward between layers. Due to the interdependence across layers, attacks tend to be more destructive, and the removal of two links causes the collapse of the whole multiplex system. Whilst the single-layer network can tolerate more attacks. More attention is still required to solve attached optimization problems on multi-layer networks.

2.2. The influence maximization problem and its robustness

The structure of networks gives a direct description of connection between nodes, revealing the underlying interactive information in this system. Based on which, the influence maximization problem [8] focuses on the information diffusion process pertaining to networks, and the related studies can be roughly classified into two categories.

The first one is to simulate the influence spreading process between nodes. Several diffusion models have been proposed to tackle different simulation scenarios. K nodes are expected to be determined from the network to form the seed set S, which spreads the subsequent influence through connections between other nodes. For a specific S, the expected influence range is denoted as $\sigma(S)$, and the crux of diffusion models is to obtain a reliable estimation on $\sigma(S)$. Each node in the network may possess two statuses, i.e., either activated or inactivated. Seeds in S are the initial activated nodes, and those inactivated ones can be transformed into activated following a certain criterion, but not vice versa. The independent cascaded (IC) model is a widely employed one [11] to modulate the diffusion behavior in networks. In detail, activated nodes manage to activate each of the connected inactivated neighbors at time step t, and those successfully activated ones can spread influence in the following procedures. A probability parameter p is required to decide the activation operation is successful or not. The feasibility of this model has been verified in [11,15].

IC model mainly considers network without weight information, the weighted cascaded (WC) model can deal with weighted networks. The activation probability is not fixed but regulative based on the weighted link information [12]. Another approach is the linear threshold (LT) model, where the activation criterion is the accumulative influence received by a node [31]. Considering the IC model exhibits generality in applications, this work mainly concentrates on this model.

With the help of these models, the influence diffusion process on networks can be simplified and presented. Following the spreading mechanism of a specific model, the generated influence $\sigma(S)$ is estimated via the number of activated nodes in the diffusion process. In detail, only K nodes in the original seed set S are activated in the initial time step $t = 0$. Seeds in S attempt to activate those inactive neighbors under the probability restriction p at each following time step t, and s_t is taken to preserve the successfully activated ones. Then S is updated with s_t, as $S = S \cup s_t$. The whole diffusion process terminates if there are no nodes get activated at time step ε,i.e., $s_\varepsilon = \varnothing$. The influence range of S, $\sigma(S)$, can be evaluated through the number of all activated nodes in the diffusion process, as $\sigma(S) = |s_0 + s_1 + \cdots + s_\varepsilon|$. A direct method to evaluate $\sigma(S)$ is via the Monte Carlo process [11,12] which is intuitive but often requires numerous independent realizations. Therefore, this Monte-Carlo-based method tends to be time-consuming and may not be feasible on complicated problems. A fast approximation approach has been proposed based on the information theory [32], where the influence range is estimated inside the 2-hop range of seeds. This method is defined as follows,

$$\hat{\sigma}(S) = \sum_{s\in S} \hat{\sigma}\{s\} - \left(\sum_{s\in S} \sum_{c\in C_s \cap S} p(s,c)(\sigma_c^1 - p(c,s))\right) - \chi \tag{2}$$

where C_s is the connected neighbors of a seed s in S, p is the activation probability. σ_c^1 is the 1-hop range of node c. χ denotes the duplicated influence generated between activated nodes with the initial seeds, which is defined as $\chi = \sum_{s\in S} \sum_{c\in C_s/S} \sum_{d\in C_c \cap S/\{s\}} p(s,c)p(c,d)$. Briefly, $\hat{\sigma}(S)$ calculates the 2-hop influence of the initial seeds and deducts the duplicated influence.

Further, the other category of studies focuses on how to determine those influential seeds from networks. The Monte Carlo process can be conducted to evaluate the performance of candidates and guide the search process as in [11,14], which may not be applicable on large-scale networks. The heuristic-based and population-based optimization methods [15,27–33] have been employed to improve the efficiency guided by the approximation approach in Eq. (2). In addition, the structural property may provide valuable information for the seed determination task, including the centrality measure [17,33], cooperative behaviors [34], and structural differences [35].

Based on which, the influence maximization problem has been extended into multi-layer networks. The diffusion dynamics across different layers was investigated [20,21], and an approximation method on the performance of seeds was given in [22]. The centrality information [21] and the meta-heuristic optimization method [22,33] have been verified to be effective to determine influential spreaders. Based on the spreading behavior in single-layer networks, the diffusion processes of seeds in different layers are independent, but the activation of an inactive node requires more criteria. The activation in just one layer is not enough to completely activate this node; as indicated in [22], the node should be activated in at least L_{min} layers to become an influential one in following steps. L_{min} can be problem-related, and a possible rational configuration is three. Meanwhile, several attempts have been made to model the diffusion process on multi-layer networks [36,37], and the cost for generating crossing-layers influence is shown [38]. Further, methods including multiple adoption [39], heuristic mapping method [40], and deep learning [41] have been developed to find influential seeds from multi-layer networks.

Meanwhile, the influence diffusion process may be threatened by potential perturbances, and its robustness has attracted increasing attention. As shown in [25,26], some systematic factors, including possible changes of the spreading behavior such as the activation probability, the diffusion model, and the discrepancy between the ideal the realistic diffusion environments may hinder the spreading ability of selected seeds. The hyperparameter may also impact the performance of seeds [42]. A theoretical approximation algorithm is given in [43] to detect potential robust seeds. The game theory has been demonstrated to be closely correlated with the diffusion process with dynamic parameters [44]. However, the applicable value of the robustness towards systematic parameters seems to be limited, and the robustness against structural failures should be considered in the diffusion process as indicated in [27]. A measure R_S has been developed based on Eq. (1) to numerically assess the robustness of seeds under structural failures; guided by which, seeds with robust diffusion ability can be selected [27].

These studies reveal the mechanism in the information diffusion on social networks, which contributes to related analysis and knowledge mining tasks. The solution of some dilemmas like the political mobilization [8], the pandemic control [10], and the digitized marketing [13] may also be benefited.

2.3. The multi-factorial evolutionary algorithm

Motivated by the biological inheritance and the evolution theory, evolutionary algorithms (EAs) [45] are a kind of optimization method to solve complicated problems, especially those without gradient information. A population-based search mechanism is employed to explore the solution space. Compared with the heuristic-based optimization method like in [23], the optimal information from diversified individuals can be utilized to improve the optimization ability. Further, the problem-orientated local improvement strategy has been considered in EAs, which is denoted as the Memetic algorithm (MA) [46]. The included individual-based search process provides extra information to exploit the available genetic materials. Traditionally, only one target is considered for EAs, and a specific candidate is expected as the output of the whole search process. Combined with problem-directed exploration and exploitation operators, EAs have shown validity on plenty of optimization problems [45,46].

However, some recent studies indicated that there may exist potential common knowledge between the optimization process on different tasks [28,47], and traditional EAs tend to omit possible genetic complementarities when tackling different problems. Employing the multi-factorial inheritance theory, genetic materials are allowed to be transferred across the optimization on multiple tasks, and the multi-factorial evolutionary algorithm (MFEA) has been developed in [28]. MFEA is aimed at achieving the simultaneous optimization on multiple tasks with similar or even different data shapes. As shown in [28], the gene transfer operation contributes to an efficient solution of both benchmark and realistic optimization problems.

In [28], the knowledge transfer operation is probability-based, and a parameter *rmp* is necessary to balance the inter-task exploitation and intra-task exploration search operations. The detailed determination of *rmp* is closely related to the performance of MFEA and a fixed configuration may not be the best choice in the search process. Consequently, a probabilistic estimation model is considered to reach the online learning of *rmp* in [47]. In this manner, predominantly negative transfers can be avoided and prevent a plethora of inter-task information. Further, the multitasking evolutionary approach has also been investigated within the domain of multiobjective optimization [48]. The concurrent solution of different multi-objective optimization problems may reveal underlying similarities between these tasks. Based on which, the exploitation of latent complementarities between tasks can improve the optimization efficacy.

The boon of multitasking optimization has been demonstrated in a wide range of applications. For instance, the parallel optimization on operational indices contributes to a decreased budget of beneficiation processes [49]. The evolutionary multitasking framework also shows efficacy in multiple sparse reconstruction tasks, and a successful attempt has been made on the hyperspectral image unmixing problem [50].

Yet, little attention has been paid to solve the network-related multitasking dilemmas; e.g., the RIM problem. Compared with existing studies in [28,47,48], several characteristics of networked systems should be considered when designing algorithms. First, the decision variable for RIM is a set of seeds. The problem is with a discrete shape and seeds tend to be independent with each other in the solution space. Comparatively, tasks in [28,47,48] tend to possess numerical decision variables, which cannot be directly applied on the solution of RIM. The design of transfer operator between multiple seed sets remains an open question. Further, the performance evaluation of networks is likely to be costly, and an efficient search procedure is therefore expected to avoid prohibitive computational consumption. Meanwhile, the structure of networks often maintains nonnegligible information for the search process. Several structural properties including the connectivity [23], the structural importance of nodes [22], and the interdependence between different layers [21] have shown significance in previous studies. A rational consideration of structural information is critical when designing the corresponding algorithm.

3. The multifactorial optimization of the RIM problem on multi-layer networks

Two basic conceptions should be clarified for designing the problem-directed MFEA to find influential seeds from multi-layer networks. The first one is how to evaluate the robust information diffusion ability of seed candidates in the whole networked system; and the second one is how to constitute the multitasking optimization problem. Based on related studies, possible solutions are given in this section.

3.1. The performance measure of the RIM problem

Given a specific seed set, the performance measure is expected to give an accurate and reliable evaluation result on its robust information diffusion ability. The measure can work as a criterion to judge the superiority among seed candidates, and thereby guide the optimization process. Considering the frequent invoking of the measure in the search process, a low-cost and high-efficiency measure is desired. As indicated in Section 2.2, the Monte Carlo process seems to exhibit a remarkable evaluation accuracy for its direct simulation procedure; nevertheless, the extensive computational budget makes it unavailable in the optimization process. In allusion to assessing the robustness of seeds under structural perturbances, the measure may need to record the performance change of seeds in the destruction process, as measure R in Eq. (1). Here, R is designed to record the connectivity loss of networks under nodal attacks, whose mechanism has been validated in the related literature [23,24]. In terms of the RIM problem, the influential ability of seeds is key for the evaluation process, and the approximation approach in Eq. (2) has been considered as the criterion in previous studies [27]. A measure R_S has been designed to numerically reflect the robust influential ability of a seed set S, defined as,

$$R_S = \frac{1}{M \times Per} \sum_{q=1}^{M \times Per} \hat{\sigma}(S|q) \tag{3}$$

where M is the number of links and Per is a changeable parameter to decide the impaired fraction, lying in the range of [0, 1]. Two extreme cases are as follows, there is no attack conducted when Per is set as 0, R_S is replaced by $\hat{\sigma}(S)$; and all links are to be removed when Per is set as 1. $\hat{\sigma}(S|q)$ refers to the influence performance of S when q attacks are conducted. R_S mainly considers single-layer networks, and has been verified in a related optimization task in [27]. Different from R in Eq. (1), R_S cannot be directly applied as guidance for optimization algorithms, and a specific Per is to be determined. A possible solution is through empirical study on synthetic networks. A theoretically feasible Per that performs steady on several scenarios is obtained to guide following optimization procedures, as shown in [27].

3.2. The evaluation towards multi-layer networks

The multi-layer network is generally defined as $G_m = [G_1, G_2, \ldots, G_L]$ where L is the number of layers, and each layer G_l consists of N nodes. The multiplexity of the whole system is represented in the network. Several attempts have been made to evaluate the robustness or functionality of multi-layer networks [20,22]. From the structural perspective, each layer G_l is relatively independent, and different layers may exhibit disparate connective features. From the functional perspective, layers are correlated with each other, and the malfunction of a specific node in only one layer leads to the failure of this node in other layers. This notable interdependence should be considered when evaluating influential ability of seeds on multi-layer networks.

Based on R_S, the measure for seeds on multi-layer networks can be developed considering the feature of this kind of networks. Some pilot studies [20,22] indicated that the information diffusion process on multiple network layers is relatively independent with each other, and the influential ability of seeds in S can be evaluated layer by layer. In this way, the influential ability of S can be acquired via an accumulation operation across all layers. Meanwhile these studies indicated that the corresponding seed determination process on multi-layer networks is different from that on single-layer ones. The structural discrepancy among layers is significant and should be intensively emphasized in the search process. Further, the robustness of seeds' influential ability is evaluated under structural perturbances, and the link-based attack, which aims at obstructing connections between nodal members in the system, is considered in this work. In detail, given a seed set S and a multi-layer network G_m, the influential ability of S is estimated within the 2-hop area of seeds in each layer. The structural destruction is achieved by link removal, and links in G_1 are sequentially removed to cause cascaded failures in the whole system like in Fig. 1.

Considering the generality and computational cost, the degree centrality [19] is adopted to evaluate the importance of links, and links with larger degree are removed first in the destruction process. The influence generated by S is evaluated across layers as the structural failures proceeding. A measure $\hat{\sigma}_{multi}(S)$ has been devised in [22] to approximate the influence ability of S on multi-layer networks; based on which, the robustness of S's diffusion process under link-based attacks can be evaluated by the following measure R_S^{Multi}, as,

$$R_S^{Multi} = \frac{1}{Atk}\sum_{q=1}^{Atk}\hat{\sigma}_{multi}(S|q), \text{ where:} \\ \hat{\sigma}_{multi}(S|q) = P \times \sum_{l\in L}\omega_l\ \hat{\sigma}_l(S|q) \tag{4}$$

where Atk is the number of removal attempts, which is caused by losing links between nodes. Atk can be calculated as $Atk = M \times Per$ (M is the total number of links in G_m). Similar with R_S in Eq. (3), Per is the parameter to decide to what extent G_m is destructed. P and ω_l are parameters to estimate the influential ability of S, as defined in [22]. $\hat{\sigma}_l(S|q)$ represents the influence generated by S in layer l after q links are removed. The IC model is considered in [22], with an activation probability of p. p is adjustable based on the degree discrepancy between activated seeds (u) and inactivated nodes (v), as $p = p' \times \text{degree}(u)/\text{degree}(v)$. The basic p' can be set as 0.1.

With the help of R_S^{Multi}, the robustness of the diffusion process originated from S can be evaluated in a numerical manner. The comparison between different seed sets is achievable, which further works as a criterion to guide the seed determination process. And the RIM problem on multi-layer networks can be modeled as an optimization problem with a discrete decision variable.

R_S^{Multi} follows the evaluation mechanism of other measure towards multi-layer networks like in [21,22]. The influential diffusion process is independently proceeded in each layer, the structural independence is thus guaranteed. Further, the activation of new seeds gives an overall consideration across multiple layers, and dynamic activation probability together with conditioned activation is adopted. The functional interdependence is also shown. Therefore, R_S^{Multi} theoretically reflects the diffusive behavior on multi-layer systems, and the rationality is guaranteed.

3.3. The multifactorial optimization model

As shown in Eq. (4), a changeable parameter Per is maintained in R_S^{Multi}, which causes direct impact on the evaluation process. A smaller Per represents that the measure shows preference on cases when the network suffers from trivial structural destructions; whilst a larger Per represents that the measure emphasizes the situation when distinct structural losses arise. We can see that detailed configurations of Per lead to different selection principles, and the obtained seeds tend to exhibit considerable performance on the considered situation but only perform poorly on other situations. A possible and intuitive approach is to determine a specific Per value to work as the performance indicator and guide the optimization process [27]. The deficiency is that the determination on different networks is independent, and the solution based on multiple synthetic networks in [27] may not achieve a considerable influence on untrained networks. For alleviating the unexpected preference induced by the fixed Per, the multi-tasking optimization theory is introduced into this optimization problem in this work. R_S^{Multi} with different Per values can be considered simultaneously as tasks, and possible correlated genetic information (i.e., the seed tactics) is exploited across tasks to leverage optimal knowledge.

Following the definition in [28], θ optimization tasks $\{T_1, T_2, \ldots, T_\theta\}$ are considered simultaneously in MFEA, and each task may be referred to a specific problem. Theoretically, tasks with similar or dissimilar fitness landscape can be tackled through transferring knowledge across different domains. In terms of the RIM problem, seed selection processes guided by R_S^{Multi} with different Per are taken as independent optimization tasks, and these tasks exhibit similarity in the solution space and the coding space. In other words, all possible tasks are operated on the same multi-layer network G_m with the same output form which is a set of nodes in discrete shape. Furthermore, G_m provides potential structural information like the degree, the connection intensity, and other centrality properties. Such information is valuable in finding nodes with distinct properties and importance to work as competitive candidates as influential seeds. In this manner, resemblance may exist when finding feasible seeds guided by different R_S^{Multi}, and this pool of knowledge can be leveraged to yield computational benefits.

θ tasks are tackled simultaneously in MFEA, where two attributes are necessary for a candidate p_i to solve the multi-tasking optimization problem as in [28].

(1) *Skill factor*: The skill factor τ_i of p_i is the one task that p_i is specialized on among all θ tasks.

(2) *Scalar fitness*: The scalar fitness φ_i of p_i is calculated based on the rank of p_i in the fitness space, as $\varphi_i = 1/r_i^{\theta_i}$ where $r_i^{\theta_i}$ is the rank of p_i on task θ_i. Candidates with better performance tend to possess lower rank and larger scalar fitness.

Individuals in the population of MFEA are expected to exhibit a uniform distribution on considered tasks to avoid latent preference during the search process. On the other hand, multiple scenarios can be considered parallelly in one realization of MFEA, which utilizes potential synergies between tasks to improve the efficiency. Compared with the algorithm with fixed parametric configuration proposed in [27], multiple solutions can be obtained in MFEA to cater to extensive applications. Concentrating on the RIM problem on multi-layer networks, the instantiation of the multi-tasking theory still remains an open question.

4. MFEA-RIM$_\mathrm{m}$

Existing MFEAs like in [28,38,48] mainly consider optimization problems with numerical decision variables in continuous fitness landscapes, yet the RIM problem exhibits discrete decision variables and G_m is required when evaluating fitness. Problem-orientated operators are necessary in the algorithm to conduct effective searches. Meanwhile, the knowledge transfer strategy is the crux for MFEAs, and a rational strategy guarantees pertinent information between multiple tasks is properly exploited whilst avoiding predominantly negative transfers. Solving numeric-based optimization problems, the transfer operation in [28,38,48] considers the knowledge underlying in the fitness domain to decide whether an inter-task transfer operation is conducted or not. While on network-related problems like RIM, the network structure is closely correlated to the fitness evaluation and the search process, structural information in the genetic domain is thus nonnegligible in the search process. Based on these considerations, MFEA-RIM$_\mathrm{m}$ has been developed to determine influential seeds from multi-layer networks.

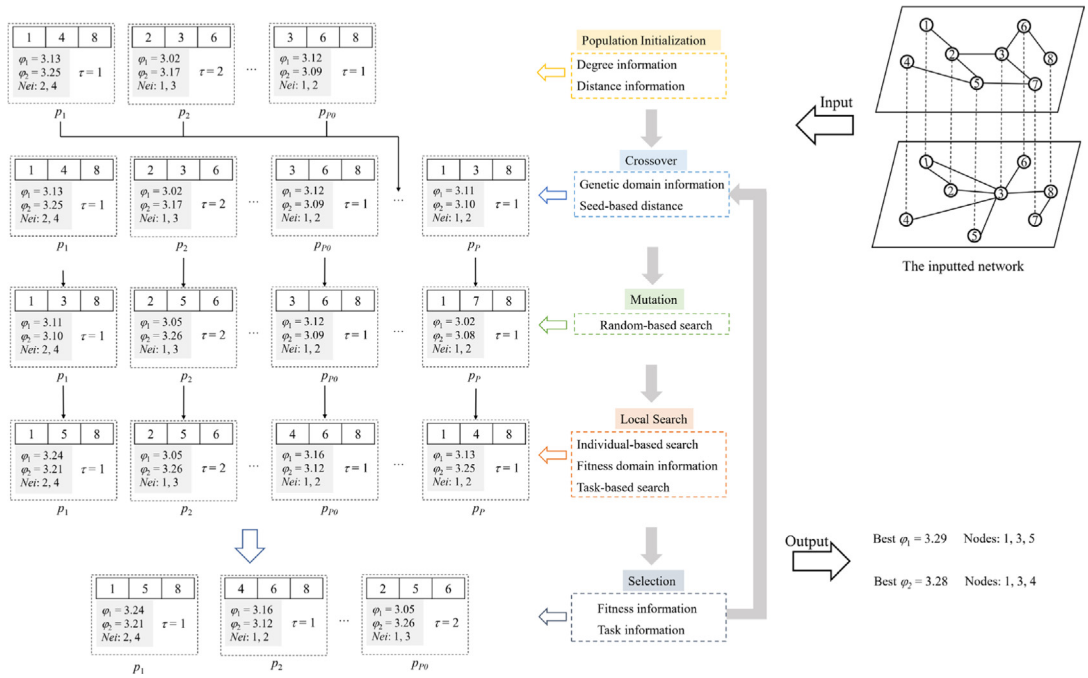


Fig. 2. An intuitive framework of MFEA-RIM$_m$. A network maintaining eight nodes is taken as the input, and two tasks are tackled. Five components are included in the algorithm, including initialization, crossover, mutation, local search, and selection. Nodes are selected to work as seeds to spread influence. Guided by the performance measure in Eq. (4) and optimal information, powerful seeds are expected as the output.

4.1. The framework

In MFEA-RIM$_m$, θ tasks are tackled, and each task is guided by R_S^{Multi} with a specific configuration of *Per*. θ seed sets are to be obtained as the output of the algorithm, where each set maintains K seeds. Two kinds of distance information are considered in MFEA-RIM$_m$ including the structure-based distance and the seed-based distance. The former one is calculated via shortest paths in G_m, and the distance between nodes i and j is equal to the required steps jumping from i to j. This distance metric is determinate as G_m is given, with a limited computational budget. But considering some networks may not be fully connected, only the metric relying on the shortest path may cause difficulties when finding closer seeds. The later one is calculated by the two-hop neighboring area between two seed sets, and the result is obtained through the *xor* operation. All neighboring nodes of a seed set are detected and compared with those of another seed set. If more divergent neighbors can be found, then the two seed sets may possess a longer distance, and vice versa. These two metrics are optional in the following search process.

In terms of the detailed optimization process, a Memetic-algorithm-liked framework is employed. Given the multi-layer network G_m, the original population is generated by the initialization operator to attain P_0 individuals with K seeds, and a wider distribution in the fitness landscape is preferred. This operator includes randomness-based strategy, degree-based strategy, and distance-based strategy to achieve a diversified population. Then, all individuals are evaluated under θ tasks using R_S^{Multi} with corresponding *Per*. The skill factor τ_i and the scalar fitness φ_i of each individual p_i are determined accordingly. Followed by the neighbor initial procedure, which is aimed at finding a number of *Nei* neighbors for each individual p_i based on the distance information. Details of these operations are described in the subsequent section.

Based on the obtained initial population with P_0 individuals, the crossover operator is conducted to extend the population into the size of P. At a probability of p_c, an individual p_i is randomly determined first, and another individual p_j is selected from either the whole population or the neighbors of p_i. The task information (τ_i) of the two individuals is considered here via utilizing intra-task or inter-task crossover strategy. Then, the mutation operator is conducted to generate minor genetic changes on existing individuals at a relatively low probability of p_m. In allusion to improving the overall performance of the whole population, the local search operator is implemented to exploit local area of individuals under the guidance of R_S^{Multi}. In addition, the knowledge from individuals specific on different tasks is explored to leverage synergies between them. To end a genetic iteration, the selection operator is conducted to preserve those better individuals into the child population. The elitism strategy is adopted to avoid the degeneration in the search process. Also, the neighbors of each individuals are also updated. The whole genetic iteration stops when the termination criterion is reached, and the best-found candidates for each task are the output of MFEA-RIM$_m$.

The framework of the algorithm is given in Algorithm 1. An intuitive framework can be found in Fig. 2.

4.2. The genetic operators

The initialization operator

As described in the last subsection, this operator intends to initialize $\boldsymbol{Pop}_0$ using diversified individuals. The whole population is divided into three part equally. Individuals in the first part are equipped with K stochastically selected seeds from G_m, which intends to introduce the randomness into the population. Those in the second part are equipped with K high-degree seeds, and the selection is based on the roulette considering the synthetic degree across all layers. Nodes with higher degree are likely to be selected as seeds in this part. Those in the third part are combined with the high-degree nodes and distance-aware strategy. The first seed is selected via the roulette based on the synthetic degree; for the rest $K-1$ seeds, L candidates are randomly found, and the

Algorithm 1: MFEA-RIM$_m$

Input:
- G_m: The inputted multi-layer network;
- K: The seed set size;
- Nei: The number of neighbors;
- P_0: The initial population size;
- P: The whole population size;
- p_c: Probability of conducting the crossover operator;
- p_m: Probability of conducting the mutation operator;
- p_l: Probability of conducting the local search operator;
- $MaxGen$: The maximum number of genetic iterations;

Output:
- $\boldsymbol{S}^*$: Best-found seed sets for all considered tasks;

Initialize the population ***Pop***$_0$ using the initialization operator (G_m, P_0, K, Nei), set $g = 1$;
while $g < MaxGen$ ***do***:
  Extend the population into ***Pop*** with a size P using the crossover operator (p_c, g, ***Pop***$_0$);
  Update ***Pop*** using the mutation operator (p_m, ***Pop***);
  Update ***Pop*** using the local search operator (p_l, ***P***, g);
  Find better individuals in ***Pop*** and preserve in ***Pop***$_0$ using the selection operator;
  Update $\boldsymbol{S}^*$, $g = g + 1$;
end while;
Output $\boldsymbol{S}^*$;

one with the longest distance towards the first seed is selected into the seed set. Here the structure-based distance metric is employed. In this way, a larger structural distance between seeds is guaranteed to alleviate the effect of overlapping influence. Then, the fitness of all individuals in ***Pop***$_0$ is evaluated on all considered tasks.

Based on which, the task preference is determined. Traditionally, the skill factor τ_i is chosen based on the best φ_i for numerical optimization problems in [28]. And the task preference is implied in τ_i, which incurs different orientations in following search procedures. Considering the complexity of RIM problem, the task preference simply decided by the initial state may misdirect the seed determination process. The random task determination strategy is employed, while making sure that each task has at least one individual.

In addition, a neighbor-based niching strategy is employed, and a number of Nei neighbors with close distance are found for each individual as its neighbors. The seed-based distance is considered here to avoid possible unreachable scenarios between seeds. For each individual p_i, the distance between p_i and other individuals in ***Pop***$_0$ is calculated and the Nei closest neighbors are detected and preserved. Information in the genetic domain is partly reflected here.

The crossover operator

This operator is aimed at extending the population, and both inter-task transfer operation and intra-task crossover operation are considered. An individual p_i is randomly first from ***Pop***$_0$, and another p_j is to be determined to work as the crossing mate. At a probability of $p_c \times (MaxGen$ - $g)/$ $MaxGen$, p_j is selected from the Nei neighbors of p_i; otherwise p_j is randomly selected from ***Pop***$_0$. If p_i and p_j are with the same τ_i, then the uniform crossover operation is conducted. In detail, each seed s_i in the generated individual x_c is chosen either from p_i or p_j at the same probability. When p_i and p_j are with diversified tasks, the transfer operation is conducted. Only one seed s_k is transferred from p_j to p_i to prevent distinct gene changes. The selection of s_k mainly considers information from the genetic domain. The seed-based distance metric is adopted to get $Dis[j]$ for each seed j in p_j to all seeds in p_i, then $Dis[j]$ is normalized as $\mathrm{Dis}[j] = g/\mathrm{Dis}[j]$ to introduce the preference into genetic iteration. Based on the obtained Dis values, s_k is selected based on the roulette. In the early genetic phase, seeds with smaller distance towards p_i are preferred, but not in the later genetic phase. The seed in the same position of p_i is replaced by s_k and saved as x_c, and the fitness of x_c is evaluated. Note that the related seed generation process should keep the validity of the seed set, i.e., no repetitive seeds are allowed. Those duplicated seeds are replaced by randomly-generated ones. Considering the genetic-domain information, the inter-task transfer operation has been achieved in this operator.

The mutation operator

This operator intends to generate minor perturbances on existing individuals in ***Pop***. One seed is randomly replaced in the operating individual p_i at a probability of p_m. The validity of p_i is also guaranteed, and the fitness of p_i is re-evaluated.

The local search operator

This operator is aimed at improving the fitness level of the whole population. Two phases of search procedures are maintained to excavate the optimal information from both local and global areas.

(1) *The individual-based search phase*. At a probability of p_l, for every seed s_i of each individual p_i in ***Pop***, the connected neighboring nodes of s_i are searched. Those valid ones, which can replace s_i in the seed set, are preserved in a temporary set $Neigh$. A number of L candidates in $Neigh$ are selected to replace s_i in p_i and L new seed sets are obtained. The selection is a roulette considering both the synthetic degree information (syn_deg) and the seed-based distance towards s_i of each candidate (dis), as $syn\_deg \times (MaxGen - g) + dis \times g$. In this manner, candidates with larger synthetic degree are preferred in the early generation; but those with longer distance are in priority in the later generation. All L candidates are evaluated under the specific performance metric of p_i. The best one in the obtained L seed sets is compared with p_i, and p_i is updated if an improvement has been achieved. This phase mainly considers the local connection information

Algorithm 2: The local search operator

Input:
Pop: The current population;
p_l: The crossover probability;
g: The current iteration label;
MaxGen: The maximum number of genetic iterations;
Output:
Pop: The operated population;

for each individual p_i in ***Pop***:
/*The first phase*/
if $rand < p_l$ ***do***: * *rand* is a randomly-generated number in the range of [0, 1] *\
Conduct the *individual-based search phase* on p_i;
/*The second phase*/
if $rand < p_l$ ***do***:
Conduct the *fitness-domain learn phase* on p_i;
end for;
/*The optional phase*/
if $rand < p_l \times (MaxGen - g) / MaxGen$ ***do***:
Conduct the *task-based learning procedure* on ***Pop***;

provided by seeds in the population, and manages to explore better seed combinations in the local area guided by R_S^{Multi} with a specific *Per*.

(2) *The fitness-domain learn phase*. The knowledge transfer operation is considered based on the information in the fitness domain, which is collected from neighbors of the operated individual. For each individual p_i, the fitness-domain learn phase is conducted at a probability of p_l. In the learn process, one neighbor n_i of p_i is randomly selected. If n_i reaches a better fitness level over p_i, then p_i is replaced by n_i. Otherwise, the random learn is proceeded. Considering the synthetic degree information, the roulette is conducted to select two seeds i and j from p_i and n_i, respectively. p_i is updated through replacing i by j on the condition that the seed set is valid and a fitness promotion can be achieved.

(3) *The task-based learning procedure*. An optional task-based learning procedure is also given, which is conducted at a decreasing probability $p_l \times (MaxGen - g)/ MaxGen$ to avoid a premature convergence. Individuals in ***Pop*** are classified considering their task information, and θ groups can be obtained. Inside each group, the poorest individual intends to learn from the best individual using the intra-task crossover operation to improve its fitness level. The procedure of the local search operator is depicted in Algorithm 2.

As shown in aforementioned sections, MFEA-RIM$_m$ has two parts: the initialization part and the evolutionary part. The time complexity is dominated by the evolutionary part for its iterative conductions. For the crossover process, suppose there are P_0 individuals in the population and each has K seeds as gene, the exchange between two individuals is required to generate new individuals in each generation, whose time complexity follows $O(P_0/2 \times K)$. The mutation operator employs stochastic searches on each individual, and holds a complexity of $O(P)$. Then, the local search process includes several phases to exploit the optimal information of P individuals, following a complexity of $O(P \times K)$. The selection operator has a complexity of $O(P_0)$. These operators are conducted iteratively for *MaxGen* times, and the complexity of the algorithm is $O(MaxGen \times P \times K)$. For the space complexity, each individual requires storage for the selected seed set, the space complexity thus is $O(P \times K)$.

4.3. Comparison against other optimization techniques

MFEA-RIM$_m$ adopts the framework of multi-factorial evolutionary paradigm, and different optimization tasks are parallelly considered in the optimization process. A knowledge transfer operation is also maintained to exploit the latent complementarities between tasks. Considering the RIM problem on multi-layer network contains several potential optimization scenarios, and similarity can be found between scenarios. Therefore, MFEA has remarkable potential to solve this problem.

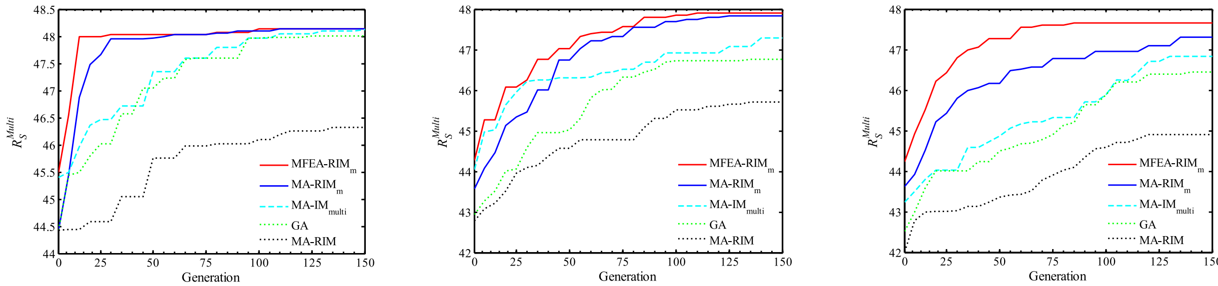
Compared with single-objective optimization approaches like in [15,22,27], only a specific target can be considered and solved in one iteration, and the potential pertinent information between different targets is completely omitted. Focusing on the RIM problem, such methods are of low efficiency. The MFEA-based method is preferred for the multiple considerations on diverse targets, which significantly improves the efficiency and contributes to the promotion of obtained results.

Compared with heuristic-based methods like in [11,14], the optimizing process mainly relies on one candidate, and only partial solution space tends to be explored. The obtained results may not be satisfactory. On the contrary, MFEA-RIM$_m$ follows the population-based search mechanism, and a much wider solution space can be searched by diversified individuals. Also, the transfer operation promotes efficiency, which makes MFEA-RIM$_m$ competitive over traditional heuristic-based methods.

Additionally, compared with those simple structural-property-based methods [16,17], related structural characteristics are included in MFEA-RIM$_m$ to provide potential candidates, and seeds with better robust influential ability can be determined.

5. Empirical results

The performance of MFEA-RIM$_m$ is validated on both human-made and real-world networks. Parameters of the algorithm are set as listed in Table 1. The independent cascaded model is employed to simulate the information diffusion process on multi-layer networks as in [22], and the activation probability p is set as 0.1.



(a) $Per = 0.3$ (b) $Per = 0.5$ (c) $Per = 0.7$

Fig. 3. The convergence curves of five tested population-based optimization methods.

Table 1
Parameters for MFEA-RIM$_m$.

Parameter	Meaning	Configuration
P_0	The initial population size	30
P	The whole population size	50
p_c	The probability of conducting the crossover operator	0.6
p_m	The probability of conducting the mutation operator	0.4
p_l	The probability of conducting the local search operator	0.6
MaxGen	The maximum number of genetic iterations	150
Nei	The number of neighbors of each individual	3

5.1. Benefit of employing the multi-tasking optimization theory

Traditionally, the single-objective optimization technique has been adopted to determine seeds with robust influential ability like in [27]. A specific *Per* configuration is given first through empirical tests on synthetic networks, which is likely to perform reliably on other possible damage scenarios. In MFEA-RIM$_m$, different values of *Per* in R_S^{Multi} can be considered, lying in the range of {0, 0.3, 0.5, 0.7, 0.9} (when θ is set as 5). The multi-tasking optimization theory makes the cross-task knowledge available in the search process, and multiple optimization problems can be tackled simultaneously. The efficiency promotion of MFEA-RIM$_m$ over single-objective optimization methods is validated first. Corresponding experiments have been conducted on human-made networks with distinct structural features, including random networks generated from the Erdős–Rényi model (ER) [2], scale-free networks with a power-law degree distribution (SF) [3], and small-world networks with the local connection structure SW [4]. For each kind of model, networks with $N = 200$ and averaged degree $\langle k \rangle = 4$ are generated, and four layers ($L = 4$) are gathered to constitute multi-layer networks. Several existing single-objective optimization methods are implemented to provide comparisons, including MA-IM$_{multi}$ in [22], MA-RIM in [27], the former one aims at finding influential seeds without considering the robustness and the later one aims at solving the RIM problem on single-layer networks. Guided by the proposed measure R_S^{Multi}, the modified MA-RIM towards multi-layer networks is denoted as MA-RIM$_m$. In addition, the plain genetic algorithm (GA) [27] and the heuristic-based optimization method simulated annealing (SA) [51] are also implemented. Parameter settings for algorithms follow recommendations in related papers. Different values of K are considered in the experiment, which vary in the range of [5,10,20]. Taking $K = 10$ and *Per* in the range of {0.3, 0.5, 0.7} as an example, the convergence curves of four population-based optimization techniques on SF networks are drawn in Fig. 3.

As shown in Fig. 3, the tested population-based optimization methods can gradually approach the optima, but with diversified search ability. The proposed MFEA-RIM$_m$ exhibits a steady performance on the three shown optimization scenarios, and outperforms other existing methods. MA-IM$_{multi}$ was designed to find influential seeds without considering the robustness, and only inferior results can be obtained especially when *Per* gets increased. The difference between the plain IM problem and the RIM problem is marked here. Some seeds may reach a considerable influential ability, but a sharp degeneration can arise when the network suffers from structural failures. And the purpose of studying the robustness of the diffusion process is to detect those candidates with better tolerance against connective losses. Results here indicate that the diffusion process correlates with the structure of networks, and an integration of multiple destructive situations like in Eq. (4) may be necessary to tackle complicated applications of networked systems. GA is the basic genetic algorithm without considering the local search procedure. GA only performs well when the optimization problem is not complicated (*Per* is small), and shows mediocre search ability compared with those population-based optimization algorithms utilizing local information. In terms of MA-RIM, the algorithm was devised to solve the RIM problem on single-layer networks. A distinct discrepancy is likely to exist between networks with only one layer and those with multiple layers as shown in Fig. 1. This discrepancy directly causes the poor performance of MA-RIM in the experiment. But the modified version MA-RIM$_m$ guided by R_S^{Multi} can achieve competitive results especially when *Per* is small.

Note that for each *Per* configuration in Fig. 3, one independent realization of the aforementioned algorithms is necessary to obtain required results, which is time consuming and omits the potential parallelism between different tasks. Yet, only one realization is enough for MFEA-RIM$_m$ to obtain solutions for multiple tasks. The remarkable efficiency of employing the multi-tasking optimization theory is validated here. The latent complementarities between tasks further promote the search ability of the algorithm, and the superiority over existing methods can be noticed when the task is complicated (*Per* is large). The search ability of algorithms is validated on the scenario of $Per = 0.9$, numerical optimization results on three synthetic networks with different K are given in Table 2. The Wilcoxon rank sum tests with a significance level of $a = 0.05$ are adopted to analyze the statistical difference between these results and those obtained by the MFEA-RIM$_m$.

As listed in Table 2, MFEA-RIM$_m$ and MA-RIM$_m$ tend to show competitive performance over other algorithms, and the proposed MFEA-RIM$_m$ outperforms MA-RIM$_m$ in more cases. In detail, the two algorithms reach similar performance on multi-layer SF networks when K is small (5 and 10), but MFEA-RIM$_m$ tends to outperform MA-RIM$_m$ on multi-layer ER and SW networks in most cases. This phenomenon may be caused by the structural differences between these human-made networks. For SF networks, only some key hubs dominate the connectivity for its significant degree over other plain nodes. These hubs possess

Table 2
R_S^{Multi} values of seeds obtained by different algorithms when *Per* is set as 0.9. Results are averaged over five independent realizations. The best result is marked for each test scenario. In terms of the significance analysis result, "−" indicates that the compared algorithm is inferior to MFEA-RIM$_m$, "+" indicates the compared algorithm outperforms MFEA-RIM$_m$, while "≈" means the two algorithms show no significant difference.

Network		MFEA-RIM$_m$	MA-RIM$_m$	MA-IM$_{multi}$	GA	MA-RIM	SA
ER	$K = 5$	**21.661**	**21.661** (≈)	20.741 (−)	20.736 (−)	20.023 (−)	20.551 (−)
	$K = 10$	**42.997**	42.922 (−)	41.255 (−)	41.247 (−)	40.828 (−)	41.027 (−)
	$K = 20$	**85.412**	85.394 (−)	85.033 (−)	82.001 (−)	81.751 (−)	81.994 (−)
SF	$K = 5$	**24.520**	**24.520** (≈)	22.003 (−)	21.971 (−)	21.536 (−)	21.971 (−)
	$K = 10$	**47.506**	**47.506** (≈)	46.881 (−)	46.221 (−)	45.896 (−)	46.129 (−)
	$K = 20$	**91.483**	91.607 (−)	90.231 (−)	88.873 (−)	88.325 (−)	88.556 (−)
SW	$K = 5$	**21.149**	21.118 (−)	20.661 (−)	20.236 (−)	20.178 (−)	20.226 (−)
	$K = 10$	42.152	**42.153** (≈)	41.088 (−)	41.104 (−)	40.881 (−)	41.011 (−)
	$K = 20$	**84.025**	84.019 (−)	83.515 (−)	83.311 (−)	82.951 (−)	83.401 (−)

plenty of neighbors and tend to spread influence in a wide range. Powerful candidates can thus be detected via the inputted structural information in G_m, which contributes to the search process. On the contrary, ER and SW networks do not maintain such evident nodal preference, and the search process can only obtain limited information from the structure of G_m. Effective genetic operators and the employed knowledge transfer operation in MFEA-RIM$_m$ make for the performance promotion as in Table 2. Also, results of the heuristic-based optimization algorithm SA are compared, which reveals the effectiveness of implementing population-based search techniques to solve complicated optimization problems. But the advantage of SA is the low space complexity, which allows it work as a component in the genetic algorithm as in [24]. The optimization ability of MFEA-RIM$_m$ in determining robust and influential seeds is verified in Table 2.

5.2. Comparison against other MFEAs

In MFEA-RIM$_m$, several problem-orientated operators have been devised to tackle the network-related problem. Meanwhile, the knowledge transfer operation considers the information from both genetic domain and fitness domain to leverage pertinent complementarities. Several existing MFEAs have also been implemented to provide comparisons, including the basic MFEA in [28], SREMTO in [52], and MFDE in [53]. These algorithms were designed to solve numeric-variable-based optimization problems, and they cannot tackle the optimization on networks directly. In the experiment, only the framework of these algorithms has been adopted; following which, the network-related operators are considered to finish the required optimization task. Meanwhile, several variants of MFEA-RIM$_m$ are tested as well, including the one without genetic domain information termed MFEA-RIM$_m$-NG, the one without fitness domain information termed MFEA-RIM$_m$-NF, and the one without the local search operator MFEA-RIM$_m$-NL. θ is set as 5 for all mentioned MFEAs, tested on multi-layer SF networks with 200 nodes and $K = 20$. Results are grouped into two categories, the low-cost group including $Per = \{0, 0.3, 0.5\}$ and the high-cost group including $Per = \{0.7, 0.9\}$, the corresponding fitness changes in the genetic process are compared in Figs. 4 and 5, respectively.

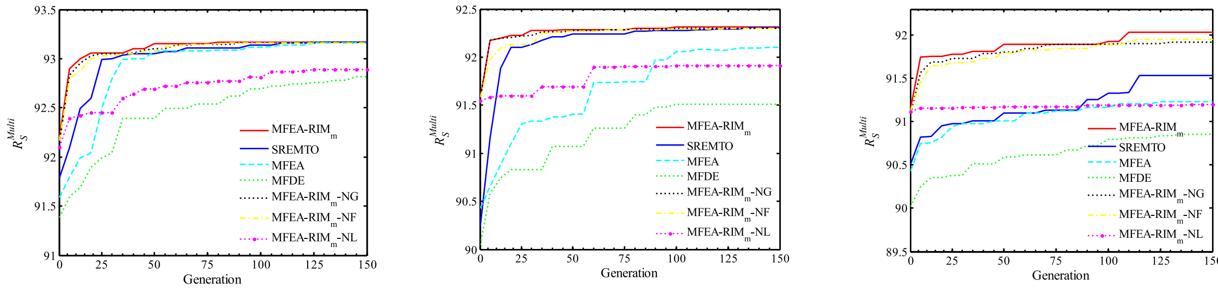
As shown in Fig. 3, MFEA-RIM$_m$ and two of its variants perform competitively compared with other MFEAs in the low-cost group. When the *Per* value in R_S^{Multi} is relatively small, the optimization problem may not maintain much disturbance in the search process, and most algorithms obtain considerable results (such as when *Per* is set as 0). Comparatively, MFDE shows inferiority in the experiment, which is caused by the search mechanism of differential evolution (DE) maintained in MFDE. DE shows efficiency in the solution of numerical-variable-based problems [44], but may not be valid in the discrete problem. SREMTO and MFEA also implement the multi-tasking optimization theory, but do not concentrate on network-related problems. It can be seen the performance of the initial population of the two algorithms tend to be poor, which reveals the stochastic initialization operation is not enough to generate promising candidates for the RIM problem. Also, as *Per* increasing, problem-orientated operators are necessary to guarantee the quality of obtained candidates, which are not included in these two algorithms. For MFEA-RIM$_m$-NL, the obtained candidates are inferior and the promotion extent is quite limited. Such result validates the effectiveness of the local search operator when solving the RIM problem on multi-layer networks.

In the high-cost group, we mainly focus on the comparison of MFEA-RIM$_m$, MFEA-RIM$_m$-NG, and MFEA-RIM$_m$-NF for their competitive performance in the experiment of Fig. 4. Detailed results are given in Fig. 5. As shown in the figure, MFEA-RIM$_m$ tends to outperform the two variants as the complexity of optimization tasks gradually increasing. For MFEA-RIM$_m$-NG, the transfer operation in the genetic domain is omitted and replaced by the plain randomness-based knowledge learn procedure as in MFEA [28]. The algorithm can still harness possible inter-task synergies in the fitness domain, and it keeps a remarkable convergence ability. But the search process may be disturbed by local optima, which makes the final results less competitive. For MFEA-RIM$_m$-NF, individuals cannot directly learn better candidates, which causes decrease in the convergence speed. Yet the inter-task crossover provides potential knowledge for the search process. And the results are also inferior compared with MFEA-RIM$_m$.

Experiments here validate the efficacy of designed multi-tasking transfer operators. Equipped with these operators, the advantage of MFEA-RIM$_m$ over existing MFEAs in determining robust and influential seeds from multi-layer networks has been shown. The proposed algorithm maintains diversified optimization knowledge from both genetic domain and fitness domain, the contribution of these operations is validated. Considering the uncertainty and complexity in multi-layer networked systems, the multi-tasking theory seems to be rational when solving the RIM problem, and an improved efficiency can be achieved in the seed determination process.

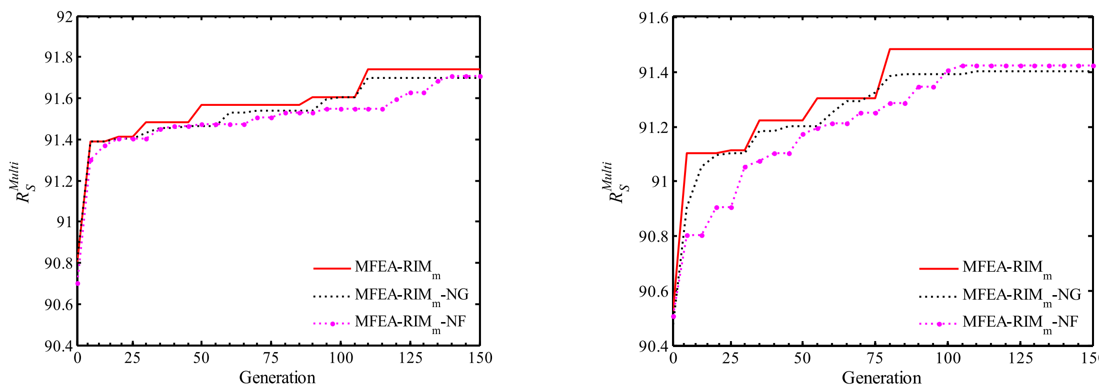
5.3. The parametric sensitivity analysis

Several parameters are included in MFEA-RIM$_m$, and the sensitivity of these parameters is tested. Taking results of multi-layer SF network with 200 node, $K = 20$, $Per = 0.9$ as an example, convergence curves under different configurations of p_c, p_m, p_l, and *Nei* are given in Fig. 6. As shown in the figure, different parameters generate diverse impact on algorithm's performance. p_c controls the frequency of interactions between individuals including both inter-task and intra-task ones. If p_c is undersized, the algorithm may not receive enough information to promote the search process, and a rational setting is encouraged. p_m restricts the intensity of random adjustments on the population. An



(a) $Per = 0$ (b) $Per = 0.3$ (c) $Per = 0.5$

Fig. 4. The convergence curves of MFEAs in the low-cost group.



(a) $Per = 0.7$ (b) $Per = 0.9$

Fig. 5. The convergence curves of MFEAs in the high-cost group.

oversized p_m may disturb existing elitists and cause unexpected degeneration, and an undersized p_m tend to work poorly in resisting local optima. These two parameters do not show distinct influence on the performance. For p_l, the parameter directly decides the intensity of individual-based searches, which is closely correlated to the optimization ability. As given in Fig. 6(c), a performance decrease can be detected when p_l is reset as smaller values. On the other hand, an overlarge p_l is not recommended for the incurred prohibitive computational cost. For *Nei*, the parameter controls the scope range of individuals. Smaller *Nei* leads to less neighbors of individuals, which may hinder the learn efficacy in the fitness domain. Larger *Nei* allows more learn chances in the search process, which results in a fast improvement in the early genetic phase, but may generate excessive similar candidates and interfere the later genetic phase. p_l and *Nei* are closely correlated with algorithm's performance, the detailed configuration is crucial for getting satisfactory results. The rationality of parametric configuration in Table 1 has been verified here.

Further, MFEA-RIM$_\mathrm{m}$ is tested on networks with larger sizes. Synthetic SF networks with 500 and 1000 nodes are considered, and corresponding numerical results are depicted in Table 3. We can see that MFEA-RIM$_\mathrm{m}$ can reach the best result on most scenarios, and it shows superiority when the parameter *Per* enlarges. The scalability of tested algorithms is validated here.

Table 3
The performance of MFEA-RIM$_\mathrm{m}$ and other three algorithms on synthetic multi-layer SF networks with large size. Results are averaged over three independent realizations. The result of MFEA-RIM$_\mathrm{m}$ is highlighted if it reaches the best in each test scenario. Labels are the same with those in Table 2.

Size	K	*Per*	MFEA-RIM$_\mathrm{m}$	MA-RIM$_\mathrm{m}$	MA-IM$_\mathrm{multi}$	SREMTO
500	10	0	**54.175**	54.175 (≈)	54.175 (≈)	54.175 (≈)
		0.3	**53.262**	53.262 (≈)	53.191 (−)	53.003 (−)
		0.9	**52.779**	52.311 (−)	52.021 (−)	52.202 (−)
	20	0	**102.338**	102.338 (−)	102.338 (−)	102.338 (−)
		0.3	**101.041**	101.041 (−)	100.893 (−)	100.997 (−)
		0.9	**99.702**	99.652 (−)	98.761 (−)	99.593 (−)
1000	10	0	**60.597**	60.597 (≈)	60.597 (≈)	60.597 (≈)
		0.3	**58.895**	58.895 (≈)	58.652 (−)	58.699 (−)
		0.9	**58.104**	58.104 (≈)	58.001 (−)	57.795 (−)
	20	0	112.708	**112.717** (+)	112.717 (+)	112.695 (−)
		0.3	**109.600**	109.599 (≈)	109.578 (−)	109.489 (−)
		0.9	**107.651**	107.444 (−)	106.897 (−)	107.233 (−)

5.4. Validations on real-world networks

The above networks are human-made, which are with distinct structural properties including the small-world connections, connective randomness, and the power-law degree distribution. To further validate the performance of MFEA-RIM$_\mathrm{m}$, three real-world networks are also considered. Multiple structural properties can be found in different parts of a realistic networked system. The first network is a social network, gathered from an IT corporation at Aarhus [54] (SN), which contains 5 layers and 61 nodes. Each layer represents relations between employees in different platforms. The second network is a transportation network (TN), gathered from airlines in Europe [55], which contains 6 layers and 450 nodes. For TN, some less dense layers are omitted, and each considered layer represents an airline dispatch for a certain operator. The last one is a multiplex neuron network (NN), gathered from the nematode, which contains 3 layers and 279 nodes [56].

On these three networks, different configurations of θ are tested to show the effect incurred by the intensity of intertask knowledge transfer. Considering the R_S^{Multi} with *Per* setting

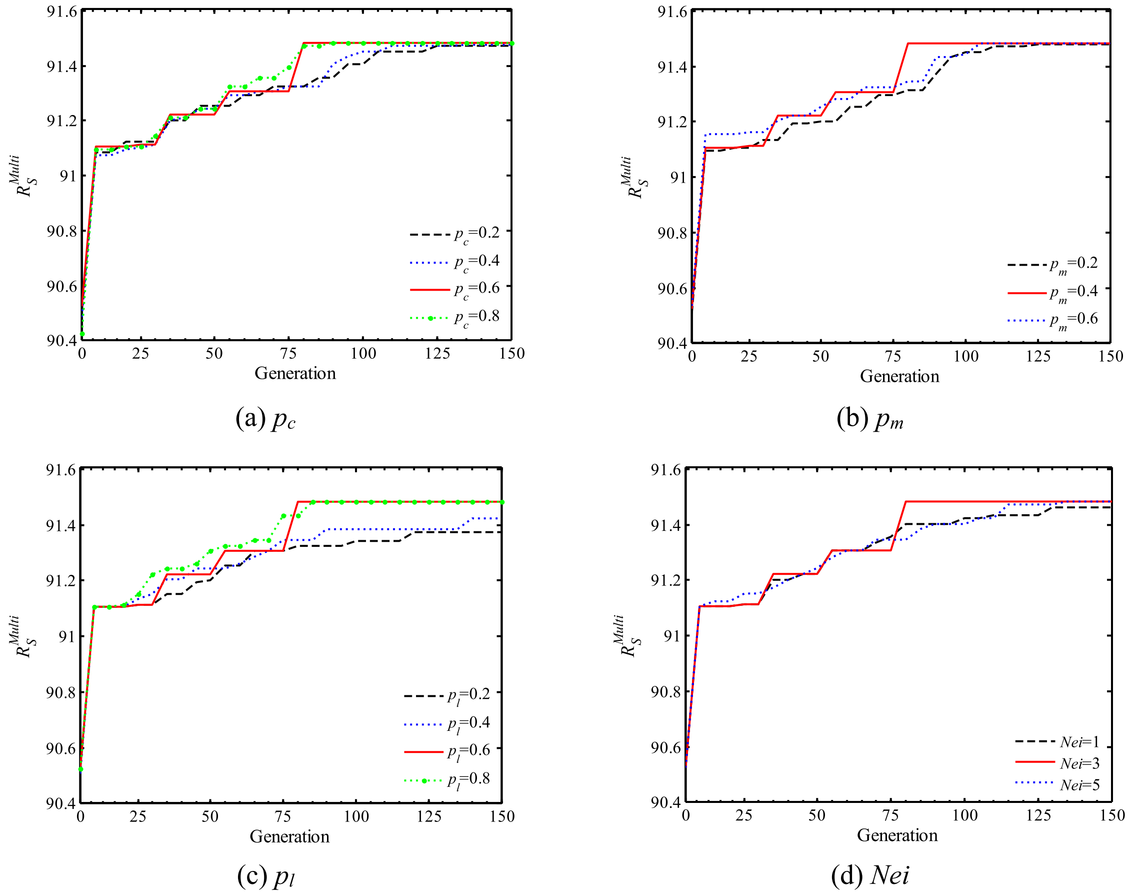


Fig. 6. The parametric sensitivity analysis of MFEA-RIM$_m$.

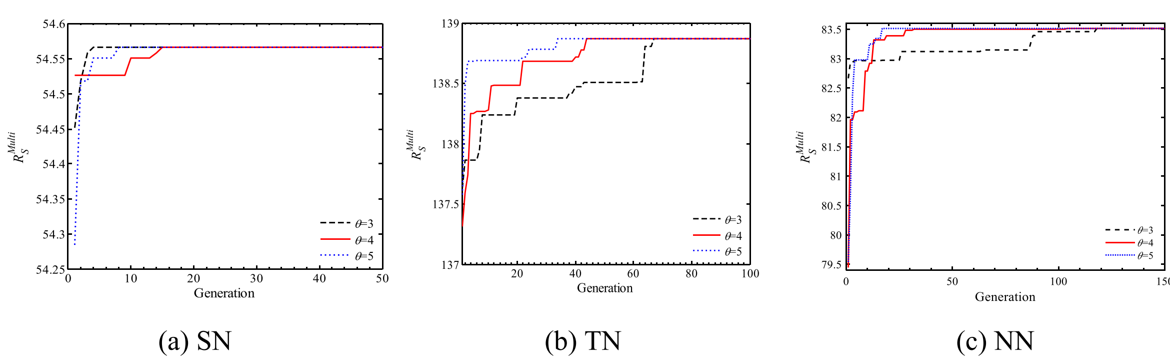


Fig. 7. The convergence curves of MFEA-RIM$_m$ with different θ. Part of genetic process is shown to improve the visuality.

as 0.3 shows potential universality in applications, this specific evaluation scenario with $K = 10$ for SN and 20 for TN and NN is extracted to show the convergence procedure. The curves are drawn in Fig. 7. For SN, the scale of the network is relatively small, the corresponding optimization task is thus not hard. Algorithms get converged in the very early genetic process, the increase of θ contributes little to the result. But for TN and NN, the complicated structure may hinder the seeds determination process, and more tasks provide a richer pool of knowledge to be leveraged in the multi-tasking search procedure. As plotted in Fig. 7(b) and (c), the algorithm with θ setting as 5 exhibits better ability to find optima over other parametric configurations. Analogous to some pilot studies [28,47], it can be concluded that the introduction of multiple tasks contributes to utilizing the in-depth information between optimization processes, and a concurrent promotion on tasks is expected to be achieved.

In [27], $Per = 0.3$ is deemed as a universal solution towards diversified scenarios based on empirical analyzes on synthetic networks. On the real-world networks, a validation is conducted to examine the reliability of this assumption. For the obtained results on θ tasks, the one that performs comprehensively well on all scenarios is chosen, and the result is listed in Table 4 with the result of $Per = 0.3$ as a comparison (labeled as Baseline). Similar with the experiment in Fig. 7, K is set as 10 on SN and 20 on TN and NN. Outlined in the table, R_S^{Multi} with Per setting as 0.3 may not consistently be the best choice to lead the seed determination process. For SN, it can be noticed that a larger Per is expected to reach the overall efficient solution. But for TN and NN, 0.3 or 0.5 seems to be a rational configuration. This phenomenon may be caused by the structural difference between human-made and real-world networks. Notable complexity exists in real-world networks instead of a certain structural property,

Table 4
The overall best solution obtained by MFEA-RIM$_m$ on the three real-world networks. The column labeled "Sum" represent the summation of following tested scenarios. The one labeled "*Per*" indicates corresponding setting of *Per* values in the optimization process.

		Sum	0	0.3	0.5	0.7	0.9	*Per*
SN	MFEA-RIM$_m$	272.231	54.791	54.558	54.412	54.288	54.182	0.5/0.7
	Baseline	272.166	54.705	54.566	54.418	54.292	54.185	0.3
TN	MFEA-RIM$_m$	691.997	139.381	138.867	138.381	137.915	137.453	0.3/0.5
	Baseline	691.997	139.381	138.867	138.381	137.915	137.453	0.3
NN	MFEA-RIM$_m$	416.137	84.718	83.490	82.936	82.615	82.377	0.3/0.5
	Baseline	615.987	84.786	83.514	82.894	82.528	82.265	0.3

which impacts the information diffusion pertaining to networks. Results here reveal that a fixed value like in [27] shows limitation when dealing with untouched networks, and the multi-tasking optimization theory provides potential solutions to this knotty optimization problem. The unexpected preference brought in by the changeable parameter *Per* in R_S^{Multi} can be avoided via the parallel co-optimization on multiple tasks. The remarkable efficiency of the proposed algorithm over single-objective techniques is demonstrated.

The obtained multiple results on the realistic networked systems can work as potential solutions towards dilemmas in daily operations. For example, results on SN indicate key players in the social platform. Being different from the solution in [22], these players are robust against structural losses, while keeping a considerable diffusion ability to accomplish advertising and marketing assignments. Results on TN help to identify influential junctions in the transportation system, which provides information for data excavations and future developing. Influential seeds are detected on NN, which may work as significant trunk nodes to support vital moments.

6. Concluding remarks

The influence maximization problem shows significance on complicated networked systems with multiple layers. Based on related studies, we focus on the robustness of seeds' diffusion process against structural failures. A numerical measure R_S^{Multi} has been developed to evaluate the robustness of seeds, which contains a changeable parameter *Per*. A feasible method is to determine a specific value for *Per* and guide the following search process. Modeled as a multi-tasking optimization problem, MFEA-RIM$_m$ has been devised considering knowledge from both genetic and fitness domains. The superior performance and computational efficiency have been validated on networks.

The robustness of the information diffusion process on multilayer networks has been emphasized in the view of link-based attacks. Yet, nodal members in a system tend to dominate the diffusion process [57], the destruction caused by node-based attacks and its effect on seeds' spreading ability deserves attention. The correlation between the optimization towards node-based and link-based remains to be studied. Further, robust seeds are expected to keep a stable performance against multiple damage scenarios. The weighted-optimization or multi-objective optimization may be considered in following studies. Meanwhile, the computational cost required by of the robustness measure like R_S^{Multi} is non-trivial, which causes delay in the search process. Nevertheless, iterative optimization processes like genetic algorithms require a frequent invoking of such measure to guarantee the search efficacy. Some pilot studies on surrogate-assisted method [58] may be valuable in further improving the efficiency. The representation on potential solution and its explorative strategies is worth of investigations. From a practical perspective, the application of related theories and algorithms on real systems is challenging as well, the adjustment towards detailed networks is to be solved.

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.

Acknowledgments

This work was supported in part by Guangdong Basic and Applied Basic Research Foundation, China under Grant 2021A1515110543 and in part by the Key-Area Research and Development Program of Guangdong Province, China under Grant 2020B090921003.

References

[1] M.E.J. Newman, Networks: An Introduction, Oxford University Press, Oxford, 2010.
[2] P. Erdős, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960) 17–61.
[3] A.L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[4] D.J. Watts, S.H. Strogatz, Collective dynamics of small-world networks, Nature 393 (1998) 440–442.
[5] M.E.J. Newman, Assortative mixing in networks, Phys. Rev. Lett. 89 (20) (2002) 208701.
[6] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. USA 99 (12) (2002) 7821–7826.
[7] R. Albert, H. Jeong, A.L. Barabási, Error and attack tolerance of complex networks, Nature 406 (2000) 378–382.
[8] R. Bond, C. Fariss, J. Jones, A. Kramer, C. Marlow, J. Settle, J. Fowler, A 61-million-person experiment in social influence and political mobilization, Nature 489 (7415) (2010) 295–298.
[9] V. Latora, M. Marchiori, Efficient behavior of small-world networks, Phys. Rev. Lett. 87 (19) (2001) 198701.
[10] L. Yan, H. Zhang, J. Goncalves, et al., An interpretable mortality prediction model for COVID-19 patients, Nat. Mach. Intell. 2 (2020) 283–288.
[11] D. Kempe, J. Kleinberg, É. Tardos, Maximizing the spread of influence through a social network, in: Proc. 9th ACM SIGKDD Int. Conf. Knowledge Discovery and Date Mining, Washington, DC, 2003, pp. 137–146.
[12] W. Chen, Y. Wang, S. Yang, Efficient influence maximization in social networks, in: Proc. 15th ACM SIGKDD Int. Conf. Knowledge Discovery and Date Mining, Paris, 2009 pp. 199–208.
[13] E. Cambria, M. Grassi, A. Hussain, C. Havasi, Sentic computing for social media marketing, Multimedia Tools Appl. 59 (2012) 557–577.
[14] A. Goyal, W. Lu, L. Lakshmanan, CELF++: Optimizing the greedy algorithm for influence maximization in social networks, in: Proc. 20th ACM SIGKDD Int. Conf. Companion on World Wide Web, Hyderabad, India, 2011 pp. 47–48.
[15] M. Gong, C. Song, C. Duan, L. Ma, B. Shen, An efficient memetic algorithm for influence maximization in social networks, IEEE Comput. Intell. Mag. 11 (3) (2016) 22–33.
[16] S. Brin, L. Page, The anatomy of a large-scale hypertextual Web search engine, Comput. Netw. ISDN Syst. 30 (1998) 107–117.
[17] K. Saito, M. Kimura, K. Ohara, H. Motoda, Super mediator-a new centrality measure of node importance for information diffusion over social network, Inform. Sci. 329 (2016) 985–1000.
[18] S. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, Catastrophic cascade of failures in interdependent networks, Nature 464 (2010) 1025–1028.
[19] S. Wang, J. Liu, Robustness of single and interdependent scale-free interaction networks with various parameters, Physica A 460 (2016) 139–151.
[20] S. Gómez, A. Guilera, J. Gardeñes, C. Vicente, Y. Moreno, A. Arenas, Diffusion dynamics on multiplex networks, Phys. Rev. Lett. 110 (2013) 028701.
[21] P. Basaras, G. Iosifidis, D. Katsaros, L. Tassiulas, Identifying influential spreaders in complex multilayer networks: a centrality perspective, IEEE Trans. Netw. Sci. Eng. 6 (1) (2019) 31–45.
[22] S. Wang, J. Liu, Y. Jin, Finding influential nodes in multiplex networks using a memetic algorithm, IEEE Trans. Cybern. 51 (2) (2021) 900–912.
[23] C.M. Schneider, A.A. Moreira, J.S. Andrade, S. Havlin, H.J. Herrmann, Mitigation of malicious attacks on networks, Proc. Natl. Acad. Sci. USA 108 (10) (2011) 3838–3841.
[24] S. Wang, J. Liu, Designing comprehensively robust networks against intentional attacks and cascading failures, Inform. Sci. 478 (2019) 125–140.
[25] X. He, D. Kempe, Stability and robustness in influence maximization, ACM Trans. Knowl. Discovery Data 12 (6) (2018) 66.

[26] W. Chen, T. Lin, Z. Tan, M. Zhao, X. Zhou, Robust Influence Maximization, in: Proc. 15th ACM SIGKDD Int. Conf. Knowledge Discovery and Date Mining, Paris, 2016 pp. 795–804.
[27] S. Wang, J. Liu, A memetic algorithm for solving the robust influence maximization problem towards network structural perturbances, Chinese J. Comput. 44 (6) (2021) 1153–1167.
[28] A. Gupta, Y.S. Ong, L. Feng, Multifactorial evolution: towards evolutionary multitasking, IEEE Trans. Evol. Comput. 20 (3) (2016) 343–357.
[29] L. Ma, M. Gong, Q. Cai, L. Jiao, Enhancing community integrity of networks against multilevel targeted attacks, Phys. Rev. E 88 (2013) 022810.
[30] J. Gao, S.V. Buldyrev, H.E. Stanlay, S. Havlin, Networks formed from interdependent networks, Nat. Phys. 8 (2012) 40–48.
[31] K. Rahimkhani, A. Aleahmad, M. Rahgozar, A. Moeini, A fast algorithm for finding most influential people based on the linear threshold model, Expert Syst. Appl. 42 (3) (2015) 1353–1361.
[32] J. Lee, C. Chung, A fast approximation for influence maximization in large social networks, in: 23rd ACM SIGKDD Int. Conf. Companion on World Wide Web, Seoul, Korea, 2014, pp. 1157–1162.
[33] Y. Zhou, J. Hao, F. Glover, Memetic search for identifying critical nodes in sparse graphs, IEEE Trans. Cybern. 49 (10) (2019) 3699–3712.
[34] D. Xue, S. Hirche, M. Cao, Opinion behavior analysis in social networks under the influence of cooperative media, IEEE Trans. Netw. Sci. Eng. 7 (3) (2020) 961–974.
[35] T. Pan, X. Li, A. Kuhnle, M.T. Thai, Influence diffusion in online social networks with propagation rate changes, IEEE Trans. Netw. Sci. Eng. 7 (4) (2020) 3100–3111.
[36] D. Nguyen, H. Zhang, S. Das, M. Thai, T. Dinh, Least cost influence in multiplex social networks: model representation and analysis, in: IEEE 13th International Conference on Data Mining, 2013.
[37] D. Nguyen, S. Das, M. Thai, Influence maximization in multiple online social networks, in: IEEE Global Communications Conference, 2013.
[38] H. Zhang, D. Nguyen, H. Zhang, M. Thai, Least cost influence maximization across multiple social networks, IEEE/ACM Trans. Netw. 24 (2) (2016) 929–939.
[39] S. Chen, H. Qian, Y. Wu, C. Chen, X. Wang, Efficient adoption maximization in multi-layer social networks, in: International Conference on Data Mining Workshops, 2019.
[40] S. Singh, K. Singh, A. Kumar, B. Biswas, MIM2: Multiple influence maximization across multiple social networks, Physica A 526 (2019) 120902.
[41] M. Keikha, M. Rahgozar, M. Asadpour, M. Abdollahi, Influence maximization across heterogeneous interconnected networks based on deep learning, Expert Syst. Appl. 140 (2020) 112905.
[42] D. Kalimeris, G. Kaplun, Y. Singer, Robust influence maximization for hyperparametric models, in: International Conference on Machine Learning, 2019.
[43] G. Nannicini, G. Sartor, E. Traversi, R. Calvo, An exact algorithm for robust influence maximization, Math. Program. 183 (2020) 419–453.
[44] Y. Gong, S. Liu, Y. Bai, Efficient parallel computing on the game theory-aware robust influence maximization problem, Knowl.-Based Syst. 220 (2021) 106942.
[45] B. Li, J. Li, K. Tang, X. Yao, Many-Objective evolutionary algorithms: A survey, ACM Comput. Surv. 48 (1) (2015) 13.
[46] A. Gupta, Y.S. Ong, Memetic Computation: The Mainspring of Knowledge Transfer in a Data-Driven Optimization Era, Springer, 2019.
[47] K.K. Bali, Y.S. Ong, A. Gupta, P.S. Tan, Multifactorial evolutionary algorithm with online transfer parameter estimation: MFEA-II, IEEE Trans. Evol. Comput. 24 (1) (2020) 69–83.
[48] A. Gupta, Y.S. Ong, L. Feng, K.C. Tan, Multiobjective multifactorial optimization in evolutionary multitasking, IEEE Trans. Cybern. 47 (7) (2017) 1652–1665.
[49] C. Yang, J. Ding, Y. Jin, C. Wang, T. Chai, Multitasking multiobjective evolutionary operational indices optimization of beneficiation processes, IEEE Trans. Autom. Sci. Eng. 16 (3) (2019) 1046–1057.
[50] H. Li, Y.S. Ong, M. Gong, Z. Wang, Evolutionary multitasking sparse reconstruction: framework and case study, IEEE Trans. Evol. Comput. 23 (5) (2019) 733–747.
[51] S. Wang, J. Liu, X. Wang, Mitigation of attacks and errors on community structure in complex networks, J. Stat. Mech. Theory Exp. (2017) 043405.
[52] X. Zheng, A.K. Qin, M. Gong, D. Zhou, Self-regulated evolutionary multitask optimization, IEEE Trans. Evol. Comput. 24 (1) (2020) 16–28.
[53] M. Gong, Z. Tang, H. Li, J. Zhang, Evolutionary multitasking with dynamic resource allocating strategy, IEEE Trans. Evol. Comput. 23 (5) (2019) 858–869.
[54] M. Magnani, B. Micenkova, L. Rossi, Combinatorial analysis of multiple networks, 2013, arXiv:1303.4986.
[55] A. Cardillo, et al., Emergence of network features from multiplexity, Sci. Rep. 3 (2013) 1344.
[56] M. Domenico, M. Porter, A. Arenas, MuxViz: A tool for multilayer analysis and visualization of networks, J. Complex Netw. 3 (2) (2015) 159–176.
[57] F. Morone, H.A. Makse, Influence maximization in complex networks through optimal percolation, Nature 524 (2015) 65–68.
[58] S. Wang, J. Liu, Y. Jin, Surrogate-assisted robust optimization of large-scale networks based on graph embedding, IEEE Trans. Evol. Comput. 24 (4) (2020) 735–749.