Journal of Computing and Security

Research Article, 2026, Volume 13, Number 1 (pp. 1–21)

http://www.jcomsec.org

An Ant-Colony-Optimization-Based Method for Community-Aware Link Prediction in Social Networks

Seyed Mohammad Hosseini [a] Afsaneh Fatemi [a,*] Maryam Hosseini-Pozveh [b]

[a] *Faculty of Computer Engineering, University of Isfahan, Iran.*

[b] *Department of Computer Engineering, Shahreza Campus, University of Isfahan, Iran.*

ARTICLE INFO.

Article history:

Received: 21 April 2025

Revised: 17 October 2025

Accepted: 15 February 2026

Published Online: 26 April 2026

Keywords:

Link Prediction, Ant colony optimization, Community structure, Nodes similarity.

ABSTRACT

Link prediction is one of the primary issues in the social network analysis domain, focusing on predicting the emergence of future links. Among state-of-the-art approaches, metaheuristic-based methods stand out, as they can combine existing heuristic measures for link formation with greedy exploration, thereby improving the effectiveness of future link prediction. Many of these heuristics are designed based on the structural characteristics of social networks. However, only a limited number of methods have incorporated community structure, a fundamental property of social networks. In this article, a novel link prediction method is introduced by integrating the concept of community structure into the ant colony optimization algorithm. Specifically, three common neighbor-based heuristics are extended to consider the community structure and the distinct nature of intra-community and inter-community relationships. These enhanced heuristics are then utilized in the ant colony optimization search mechanism. The effectiveness of the proposed method is evaluated against several benchmark techniques using nine real-world datasets. The results demonstrate the outperformance of the presented approach considering AUC and precision comparison metrics.

1 Introduction

Social networks are composed of individuals and the interactions or relationships between them. These systems are typically modeled as graphs, where nodes represent individuals and edges represent their relationships [1]. A key challenge in social network analysis is the link prediction problem. As social networks evolve over time, through the addition of nodes and links, link prediction serves as a valuable tool for understanding the mechanisms that drive their evolution [2]. The formation of new links depends on both the current structure of the network and the characteristics of its nodes. Community structure is one of the main inherent properties of social networks. Communities are densely connected subgroups of nodes where intra-community links are strong, and inter-community links are relatively weak [3]. Examples of communities include circles of friends in friendship networks and clusters of researchers connected by shared research interests in co-authorship networks [3]. This tendency for connections to form between similar nodes is also explained by the principle of homophily[4], where, analytical assessments reveal individuals with many

* Corresponding author.

Email addresses: Mohammad_hoseini@icloud.com (M. Hosseini), a_fatemi@eng.ui.ac.ir (A. Fatemi), m.hosseini@shr.ui.ac.ir (M. Hosseini-Pozveh)

https://dx.doi.org/10.22108/jcs.2026.145007.1166

ISSN: 2322-4460

common characteristics (age, job, gender) usually interact with each other more densely in comparison with those that do not have common features [5]. Link prediction methods can be classified into several categories, including similarity-based, probabilistic and statistical, meta-heuristic-based, matrix factorization-based, and machine learning technique-based methods [6–10]. Among these, similarity-based approaches, which utilize local link-structural measures to calculate node similarity, are notably run-time efficient. However, their heuristic nature does not guarantee effective prediction outcomes [6–8]. Nature-inspired meta-heuristic methods can address this limitation by combining their powerful greedy exploratory search with various local and non-local link structural heuristic measures. Additionally, their scalable time complexity makes them suitable candidates to be applied for link prediction in social networks [11]. Ant Colony Optimization (ACO), introduced by Dorigo et al. [12], is a prominent nature-inspired algorithm based on swarm intelligence. Modeled on the collective foraging behavior of ants, ACO leverages key mechanisms such as positive feedback, distributed computing, and greedy heuristic search [13]. While several ACO-based methods for link prediction have been proposed [11, 13–16], a significant gap remains: none explicitly incorporate the community structure inherent to social networks. Based on the preceding discussion, in this study, a new ACO-based method considering the community structure property of the social networks is proposed for link prediction. The proposed approach leverages this structure by modeling the distinct effects of existing intra-community and inter-community relationships on future link formation. While intra-community links are critically important, inter-community connections also provide significant signals and must not be neglected. To address this, three new community-aware local similarity-based measures are proposed, which are applied for computing the heuristic values of the edges in the ACO. The contributions of this study are outlined as follows:

- Three existing local measures—Common Neighbors, Resource Allocation, and Jaccard— are extended based on the community structure information and different essences of intra-community and inter-community relationships.
- Three versions of ACO are proposed in which the three mentioned proposed community-based measures are applied as the exploitation ability of ACO, for link prediction in social networks.
- A new fitness function proper for link prediction is proposed considering the community structure of the social networks.

The rest of this article is organized as follows. The related works are reviewed in section 2. In section 3, the proposed method is discussed. The details of the experiments on nine real-world datasets for the proposed method and benchmark ones are presented in section 4. Finally, the study is concluded in section 5.

2 Related Works

In recent years, various methods have been proposed for predicting future links in social networks. These methods can be broadly categorized into similarity-based methods, probabilistic and statistical methods, meta-heuristic-based methods, matrix factorization-based methods, and machine learning technique-based methods. The last category itself includes various subcategories, including classic classification-based methods, classic clustering-based methods, deep learning-based methods, graph neural network-based methods, and reinforcement learning-based methods [6–10]. These categories differ in terms of runtime complexity, the type and amount of information they require, and the mechanisms they employ for link prediction.

- Similarity-based methods: In the similarity-based methods, the heuristic assumption is that nodes tend to form links with similar nodes. Accordingly, a similarity measure is defined to calculate a score representing the similarity value between every two nodes of the social network. After computing these scores, the node pairs with the highest similarity values are recommended as the most likely future links [6–10]. The existing similarity measures are divided into three categories: local, global, and quasi-local [6–10]. In local similarity-based methods, local link structural information about nodes is applied to compute similarity values. Common neighbors [17], Adamic-Adar index [18], Jaccard index [19], and resource allocation index [20, 21] are considered in this category. These methods are more runtime efficient and more parallelizable than non-local methods; nevertheless, they are less effective because they rely solely on local structural information. Global similarity-based methods apply the global characteristics and link structural details of the whole network to compute the scores of every pair of nodes. Due to having a more global view of the network in calculations, the effectiveness of predictions in this category of methods is higher than the local methods. However, they have higher time complexity, making them less scalable for large networks. Additionally, the parallelization of global methods is more challenging. Negated Shortest Path [22], Katz index [23, 24], random walk [25], random walk with restart [26], and flow Propagation [27] are considered in this category.

The quasi-local methods strike a balance between local and global approaches. Their runtime complexity is closer to that of local methods, but they incorporate additional topological information, enhancing their effectiveness. Local path index [28], local random walk [25], FriendLink [29], Mutual Influence Random Walk (MIRW) [30], and asymmetric influence-based superposed random walk [31] are considered in this category. Some of the methods exist in this category are presented based on community structure, offered to improve effectiveness or run time efficiency [10], which are discussed under Section 2-2.

- Probabilistic and statistical methods: In probabilistic and statistical methods, statistical analysis and probability theory are applied for link prediction. First, a model is constructed based on an assumed network structure, where the model parameters can be estimated according to that structure. Next, these computed parameters are applied to the probabilistic prediction of future links, where the possible links are ranked based on the probabilistic values [6–8]. The hierarchical Structure Model [32] and Stochastic Block Model [33, 34] are considered in this category. The main drawback of these methods is their high time complexity [6–8].
- Meta-heuristic-based methods: Nature-inspired meta-heuristics can do greedy search based on any similarity-based measure to compute the similarities between nodes. Accordingly, the effectiveness of the applied similarity-based measures may be increased, while at the same time, the runtime complexity of nature-inspired heuristics methods is proper for running on social networks [6, 35]. Two important subgroups of these methods are evolutionary algorithms and swarm intelligence-based algorithms. The genetic algorithm is one of the most well-known and widely applicable evolutionary algorithms, which is applied in some studies to predict future links [36–40]. Swarm intelligence algorithms are inspired by the behavior of natural populations such as birds, ants, and fish, which are used to solve complex optimization problems. ACO algorithm is one of the most widely applied methods in this category. The advantages of the ACO include positive feedback through the communication between ants, distributed computing, and greedy heuristic search [13]. In this research, the ACO algorithm is considered for link prediction. Given that this study focuses on applying the Ant Colony Optimization (ACO) algorithm, a detailed discussion of related works is provided in Subsection 2.1.
- Matrix factorization-based methods: Dimensionality reduction plays a vital role in link prediction tasks, as it simplifies complex, high-dimensional graph data. This simplification is particularly beneficial for large-scale datasets, where processing and analysis can be computationally challenging. By creating a lower-dimensional representation, the data becomes more manageable, leading to gains in processing speed and potentially enhancing the predictive effectiveness of models. A widely adopted technique in this domain is matrix factorization. This method involves decomposing a matrix representing the graph to generate embedding vectors for nodes, which subsequently facilitate the efficient prediction of missing links [9, 10]. While numerous matrix factorization-based methods have been introduced in recent years [41–45], the approach has notable drawbacks. A primary limitation is its strong reliance on approximating the observed network with a low-rank matrix—an assumption that may not hold for networks with intricate, non-linear structures. Additionally, the factorization process can be computationally intensive, and without proper regularization, it is prone to overfitting, especially in the context of large networks [10].
- Machine learning technique-based methods: Machine learning is applied to link prediction, leveraging its capabilities across diverse domains. A common approach involves feature engineering, where informative characteristics are extracted from network data. These features serve as input for learning algorithms designed to estimate the likelihood of a link forming between two nodes. The existing literature on machine learning-based link prediction can be broadly classified into several categories, including classic classification-based methods [46–49], classic clustering-based approaches [50, 51], deep learning models [52–56], Graph Neural Network (GNN)-based techniques [57–62], and methods utilizing reinforcement learning [63–65].

2.1 Ant Colony-Based Link Prediction Methods

In [13], an ACO-based method is proposed for link prediction, where, first, the complete graph of the social network named the logical graph is constructed. Next, an initial pheromone value and a heuristic value are assigned to every link of the logical graph. Eventually, a set of iterations begins, where in each iteration, ants find their paths considering the pheromone and heuristic values. At the end of every iteration, the pheromones of the links of the path are updated based on the quality of that path. The final pheromone of every link would be equivalent to the similarity score of two endpoint nodes of that link. In this method, the initial value of the pheromone of every link in the logical graph is determined based on the presence or absence of that link in the main social network graph. Also, the initial value of the heuristic of every link is determined based on the number of common neighbors between two endpoint nodes in the main social network graph. In the method presented in [11], ACO is applied in the network subgraphs detection phase. This method is presented based on the fact that membership in the same subgraph structure means the similarity between the nodes of that structure, and this heuristic can be applied to predict the future links between the nodes. These subgraphs are the tringles of the network, and the probability of link formation between their belonging nodes is higher. By detecting the triples of the network using ACO, a pheromone is assigned to the links of the logical graph, which is applied to assign a score to the links for link prediction. In [14], quantum computing and ACO are combined for link prediction. Accordingly, a heuristic value named visibility of link and a pheromone value is assigned to every link in the logical graph. Moreover, a quantum pheromone value is set for every node. In each iteration of the ACO, ants explore their paths considering these three values. At the end of the iterations, based on both mentioned values of every link, the similarity score between each pair of nodes is obtained. The initial value of the pheromones is set equal to the same constant value. Also, the initial value of the visibility of links is determined based on the degree of common neighbors of the two endpoints of every edge. Quantum pheromone values are determined based on the quantum bits in quantum computing. In [15], the chaotic perturbation model and ant colony optimization are combined to predict the missing links. Accordingly, ants detect missing links through quasi-local information of nodes and link weights, while a chaotic perturbation is applied to keep the ant colony from being trapped in local optimums. This strategy leads to higher prediction accuracy as demonstrated by conducted evaluations. In [16], a complete logical graph containing nodes and missing links of the graph are constructed, to be traversed by the ants in the ACO. Pheromone of every link is updated based on the Euclidian distance between nodes, and the heuristic value is determined based on the degree of nodes and their common neighbors. At the end of every iteration, the pheromones of the links of the path are updated based on the quality of that path.

2.2 Community-Based Link Prediction Methods

The similarity scores between pairs of nodes can be calculated based on the different characteristics of the social network structure. One of the main characteristics of the social networks is the community structure where there are strong relationships within communities and weak relationships between communities. This important property has been considered in the process of link prediction in some studies of this domain. In this regard, the method presented in [66] as a community-based statistical method and the method presented in [67] as a supervised learning-based method can be mentioned. Moreover, the methods presented in [68–70] have developed the global similarity-based measures to be community-based. Since the goal of the current study is to combine the local similarity-based measures with the meta-heuristic-based ACO algorithm, in the following, the existing community-based local similarity-based link prediction methods are reviewed in more details. The general idea of all these methods is to extend the local similarity-based measures based on the number of common neighbors between pairs of nodes, considering the community structure. In this regard common neighbors within the same community as a pair of nodes are treated differently from common neighbors belonging to other communities. In some of the presented methods, after obtaining the community structure of the network, the similarity score is calculated only for pairs of nodes inside the communities [71, 72], while in other methods, predicting the links between different communities is also of concern [73–77]. In the methods presented in [71] and [72], first, the community structure of the network is determined. Next, by finding the similarity scores between the pairs of nodes within each community, the links that will be formed in the future within the communities, are predicted. In [71], after community detection, for every node i in every community, a set of candidate nodes for link establishment is selected. These candidate nodes are the ones that are in the same community as node i, their distance from node i is less than or equal to the average shortest path of the social network, and their degree is greater than or equal to the degree of i. Then, based on a new similarity measure, which is defined considering the common neighbors of two nodes, the similarity score of every node i with its candidate nodes is calculated. Eventually, pairs of nodes whose calculated similarity score is higher than a threshold value are recommended as a future link. In [72], nine local similarity-based measures are extended based on the community structure. The extended similarity measures are calculated only for pairs of nodes within the same community. Since all these local measures are defined based on the number of common neighbors, the idea of extending them based on the community structure, is to first divide the common neighbors of a pair of nodes into two groups of: a) members of the two nodes' community and b) none-members, and next, applying the numbers of these two groups in the equations based on the different weight factors. Although with a higher probability, the future links are those that are formed within communities, the prediction of edges between communities is also important, and not considering them is a limitation. This limitation has been addressed in some studies [73–77]. In [75], a similarity measure is proposed to be calculated between each pair of communities based on whether every two nodes and how many of their common neighbors are members of these two communities. This measure is combined with a similarity-based measure for pairs of nodes to determine the probability of link formation between nodes. In [73], similarity-based local measures are extended based on the community structure of the social network. In this regard, after identifying the network's communities, common neighbors between pairs of unconnected nodes and their communities are identified. For every common neighbor of pairs of unconnected nodes that are in the same community as those nodes, their similarity score in the applied measure is increased. Accordingly, the existing nodes and their neighbors in the same community are considered more important. In [76], the common neighbors similarity-based measure is extended so that, for every pair of nodes, the number of their common neighbors is added to the number of their common communities. In [77], two different threshold values are determined for the local similarity score between pairs of nodes a) inside the same community, and b) belonging to two communities. If the similarity score of a pair of nodes is higher than this value, a possible future link is predicted. To determine the threshold values, first, the similarity scores between all pairs of nodes are determined; next, the pairs of nodes are divided into two categories: intra-community and inter-community. Eventually, the threshold values are determined so that the same percentage of every category is selected as future links. In [74], the common neighbors of a pair of nodes are divided into four categories, according to whether the two nodes are in the same community or not, and whether the community of each of their common neighbors is the same as the community of the two nodes or not. To determine the similarity score between pairs of nodes, a function is assigned to each common neighbor. The similarity score is then calculated as the weighted sum of the function values of the common neighbors, with a distinct weight assigned to each category of common neighbors. All the mentioned methods in [71–77] has been compared with their basic versions and the evaluation results demonstrate their superiority. However, no experimental or analytical comparison between the mentioned methods has been presented. In this study, an ACO-based algorithm is presented that considers the community-based local similarity-based measures. In this regard, through presenting various heuristics functions based on extended local similarity measures, local methods have been applied, and through leveraging the interaction, cooperation, and feedback among ants during each iteration of the ACO, the algorithm harnesses the power of meta-heuristic methods for efficient greedy search.

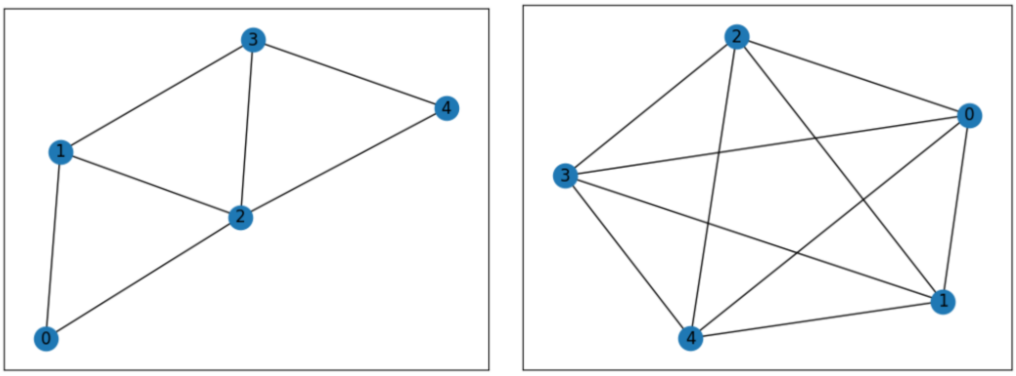
3 Proposed Method

An undirected, unweighted social network is modeled as a simple graph $G = (V, E)$, where V is the set of nodes, and E is the set of existing links of the social network. All the possible links of such a graph form the set U, whose size, for $N = |V|$, is $|U| = \frac{N(N-1)}{2}$. Link prediction is the problem of determining the links that will be formed in the future in the social network. Such links belong to the set $U - E$. In this study, the pheromone, heuristic, and fitting function of the ACO algorithm are redefined to make a community-based model for predicting links in social networks. To apply the ACO on the social network graph, the graph $G = (V, E)$ is converted to a complete graph $G_C = (V, E_C)$, where all the nodes are connected to every other node (Fig.1). Accordingly, all the possible paths for the ants to explore are considered. As every link in the complete network is assigned a heuristic value and an initial pheromone, ants will tend to move in paths with higher pheromone and heuristic values than other paths. The proposed community-based ACO is designed and implemented based on the complete graph G_C. The flowchart of this algorithm is illustrated in Fig. 2, and its details are explained as follows:

Pheromone matrix initialization: Having $G = (V, E)$ as the primary network graph, and $C_G = \{c_1, c_2, \ldots, c_m\}$ as the communities of the network, the initial values of the pheromone matrix $\tau = [\tau_{ij}]$, are assigned based on the adjacency matrix values $A = [a_{ij}]$ (1 if a link exists between nodes, and 0 if no link exists between nodes), and community matrix values $M_C = [c_{ij}]$ (1 if nodes are in the same community, and 0 if nodes are in the different communities), through equation 1:

$$\forall v_i, v_j \in V, v_i \in c_i, v_j \in c_j \Rightarrow$$
$$\tau_{ij} = \begin{cases} 1 & c_i = c_j, a_{ij} = 1 \\ 0.50 & c_i = c_j, a_{ij} = 0 \\ 0.75 & c_i \neq c_j, a_{ij} = 1 \\ 0.25 & c_i \neq c_j, a_{ij} = 0 \end{cases} \tag{1}$$

Heuristics matrix initialization: Heuristic values between nodes, denoted as $\eta = [\eta_{ij}]$, are initialized through one of the proposed community-based similarity measures. These measures take into account



a) b)

Figure 1. a) An Example Graph of a Network, and b) Its Corresponding Complete Graph.

the distinction between intra-community and inter-community links. In following, these measures are discussed:

(1) Community-based resource allocation measure (CRA)
The modified ACO presented based on this measure is named CRA-ACO. To calculate this measure, three distinct scenarios should be considered: a) when two endpoint nodes of the link belong to the same community, $\eta = [\eta_{ij}]$ is initialized through equation 2; b) when two endpoint nodes of the link belong to different communities, $\eta = [\eta_{ij}]$ is initialized through equation 3; c) When two nodes have no common neighbors, regardless of whether they are in the same community or not, $\eta = [\eta_{ij}]$ is initialized through equation 4. The equations are as follows:

$$\begin{aligned}\eta_{ij} &= \acute{\alpha}(\Sigma_{Z\in\{\Gamma_i\cap\Gamma_j\},(C_i=C_j\;and\;C_Z=C_i)}\frac{1}{|\Gamma_Z|}) \\ &+ \acute{\beta}(\Sigma_{\acute{Z}\in\{\Gamma_i\cap\Gamma_j\},(C_i=C_j\;and\;(C_{\acute{Z}}\neq C_i\;or\;C_{\acute{Z}}\neq C_j))}\frac{1}{|\Gamma_{\acute{Z}}|}) \\ &+ \epsilon\end{aligned} \tag{2}$$

$$\begin{aligned}\eta_{ij} =& \acute{\beta}(\Sigma_{\acute{Z}\in\{\Gamma_i\cap\Gamma_j\},(C_i\neq C_j\;and\;(C_{\acute{Z}}\neq C_i\;or\;C_{\acute{Z}}\neq C_j))}\frac{1}{|\Gamma_{\acute{Z}}|}) \\ &+ \epsilon\end{aligned} \tag{3}$$

$$\eta_{ij} = \epsilon \tag{4}$$

where, Γ_i and Γ_j are the set of neighbors of nodes i and j respectively, $|\Gamma_Z|$ is the degree of node Z, and ϵ is a constant value. As shown in equation 2, the first term is based on common neighbors that belong to the same community as the two target nodes, and the second term is based on common neighbors that belong to a different community from two target nodes. To differentiate the weight of common neighbors within and between communities, two coefficients $\acute{\alpha}$ and $\acute{\beta}$ are applied in calculating the final value of Γ_{ij} in both equations 2 and 3. Also, according to equation 4, the role of weak links in forming new links is considered.

(2) Community-based common neighbors measure (CCN)
The modified ACO presented based on this measure is named CCN-ACO. To calculate this measure, three distinct scenarios is considered similar to the previously discussed measure: a) when two endpoint nodes of the link belong to the same community, η_{ij} is initialized through equation 5; b) when two endpoint nodes of the link belong to different communities,η_{ij} is initialized through equation 6; c) When two nodes have no common neighbors regardless of whether they are in the same community or not, η_{ij} is initialized through equation 4. Equations 5 and 6 are as follows:

$$\eta_{ij} = \acute{\alpha}.|Z| + \acute{\beta}.|\acute{Z}| + \epsilon \tag{5}$$

$$\eta_{ij} = \acute{\beta}.|\acute{Z}| + \epsilon \tag{6}$$

where, $|Z|$ represents the number of common neighbors of nodes i and j that belong to the same community as both nodes, and $|\acute{Z}|$ is the number of the common neighbors of nodes i and j that are not at least in the same community with one of them. ϵ is a constant value, and $\acute{\alpha}$

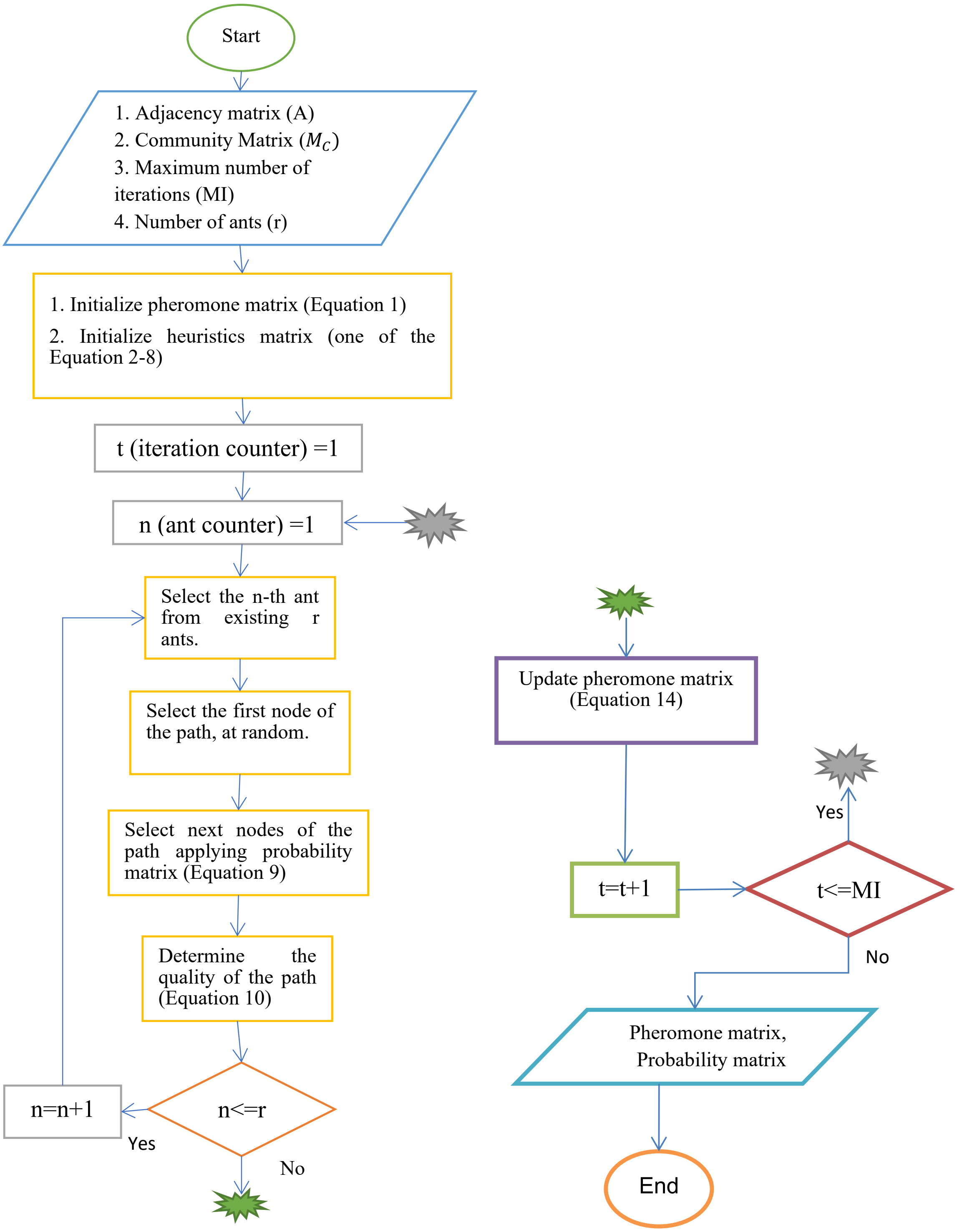


Figure 2. Flowchart of the Proposed ACO for Link Prediction.

and $\acute{\beta}$ are applied coefficients to differentiate between the weight of common neighbors within and between communities in calculating the final value of η_{ij}.

(3) Community-based Jacquard measure (CJ)
The modified ACO presented based on this measure is named CJ-ACO. To calculate this measure, three distinct scenarios are considered here as well: a) when two endpoint nodes of the link belong to the same community, η_{ij} is initialized through equation 7; b) when two endpoint nodes of the link belong to different communities, η_{ij} is initialized through equation 8; c) When two nodes have no common neighbors regardless of whether they are in the same community or not, η_{ij} is initialized through equation 4. Equations 7 and 8 are as follows:

$$\eta_{ij} = \acute{\alpha}.\frac{|Z|}{|\Gamma_i + \Gamma_j|} + \acute{\beta}.\frac{|\acute{Z}|}{|\Gamma_i + \Gamma_j|} + \epsilon \qquad (7)$$

$$\eta_{ij} = \acute{\beta}.\frac{|\acute{Z}|}{|\Gamma_i + \Gamma_j|} + \epsilon \qquad (8)$$

where Γ_i and Γ_j are the set of neighbors of nodes i and j respectively, $|Z|$ represents the number of common neighbors of nodes i and j that belong to the same community as both nodes, and $|\acute{Z}|$ is the number of the common neighbors of nodes i and j that are not at least in the same community with one of them, ϵ is a constant value, and $\acute{\alpha}$ and $\acute{\beta}$ are applied coefficients to differentiate between the weight of common neighbors within and between communities in calculating the final value of η_{ij}.

(4) Resource allocation measure by Soundarajan and Hopcroft [73] (SHRA):
In addition to the previously proposed versions of the ACO algorithm, which are based on the three measures introduced in this research, the two measures proposed by Soundarajan and Hopcroft [73] are also applied to initialize the η_{ij} values in the proposed ACO algorithm. This provides a mechanism for comparison with existing community-based link prediction methods. In the first measure (SHRA) which is based on resource allocation concept, first the common neighbors of two unconnected nodes are identified, then the algorithm examines their communities. If common neighbors with unconnected node pairs are in the same community, a positive score is assigned to them according to the resource allocation measure. The proposed ACO algorithm based on this criterion is named SHRA-ACO.

(5) Common neighbors measure by Soundarajan and Hopcroft [73] (SHCN):
To initialize a η_{ij} value, in this measure, first, the common neighbors of two unconnected nodes are identified, then their communities are examined. If common neighbors with unconnected node pairs are in the same community, a positive score is assigned to them according to the common neighbors measure. The proposed ACO algorithm based on this measure is named SHCN-ACO.

Parameter discussion: In equations (2), (5), and (7), which apply to node pairs within the same community, the parameters $\acute{\alpha}$ and $\acute{\beta}$ weight the contributions of common neighbors from inside and outside the community, respectively. If $\acute{\alpha} = \acute{\beta}$, CRA behaves like RA. For $\acute{\alpha} > \acute{\beta}$, and with a relative increase in $\acute{\alpha}$, nodes within the same community as the two target nodes gain a higher weight in producing results compared to other neighboring nodes. In this research, based on the previously discussed points, the goal is to relatively reduce the influence of common neighbors in other communities rather than completely eliminating their impact. Therefore, the values $\acute{\alpha} = 0.6$ and $\acute{\beta} = 0.4$ are set. Furthermore, in equations (3), (6), and (8), the results are calculated for node pairs that do not belong to the same community. In this case, the term containing $\acute{\alpha}$ is removed because all common neighbors belong to other communities. However, considering them with a weight of $\acute{\beta}$ neither eliminates their impact nor considers it to be as significant as that of common neighbors within the same community.

Calculating the selection probability of the next node in a path: Having the initial pheromone matrix and heuristics matrix, the probability of selecting the node v_j, when the kth ant is in the node v_i, is determined based on equation 9:

$$p_{ij}^k = \frac{(\tau_{ij})^\alpha \times (\eta_{ij})^\beta}{\Sigma_{a \in N(v_i)} (\tau_{ia})^\alpha \times (\eta_{ia})^\beta} \qquad (9)$$

where, the α and β are controller coefficients that are used to increase or decrease the influence of the pheromone and heuristic weight in calculating the probability of selecting a node. $N(v_i)$ represents the set of neighbors of the node v_i. The higher the probability of a link according to equation 9, the more the ants will choose that link in the forward path.

Evaluating the quality of the path traveled by the ants: Based on the previously presented content, every ant explores a path with a pre-defined size, l. According to the flowchart in Fig.2, the number of paths of length l equals m (the number of ants in each iteration). In addition, a measure is needed to evaluate the quality of the paths traveled by each ant considering intra-community and inter-community links. When the traveled path by ant k is shown with $W(k)$ where $W(k) = (w_1, w_2, \dots, w_l)$ and $w_i \in V$, the quality of the path is computed through equation 10:

$$Q(W(k)) = \frac{1}{|l|}\Sigma Score(e_{ij}) \tag{10}$$

where, $Score(e_{ij})$ is obtained through equation 11:

$$\forall \quad e_{ij} \in W(k): \quad Score(e_{ij}) = \begin{cases} 1 & c_i = c_j \\ 0.50 & c_i \neq c_j \end{cases} \tag{11}$$

Based on equations 10 and 11, it is observable that the quality of the path traveled by ant k depends on the type of navigated links. If the navigated link connects two nodes in the same community, then a score of 1 will be assigned to it, and if it connects nodes of two different communities, a score of 0.5 will be assigned to it. Accordingly, the quality of the traveled path is always a number between $[0.5, 1]$. The closer this number is to one, the higher the quality of the path.

Updating the pheromones of links: In every iteration, m ants travel the paths with length l. More than one ant may pass a link between two nodes. Accordingly, at the end of every iteration, the quality of every link is calculated through equation 12:

$$\Delta\tau_{ij}(t) = \Sigma_{k=1}^{m}\Delta\tau_{ij}^{k}(t) \tag{12}$$

where $\Delta\tau_{ij}^{k}(t)$ is the pheromone change of the link between nodes i and j by ant k, and is calculated through equation 13:

$$\Delta\tau_{ij}^{k}(t) = Q(W(k)) \tag{13}$$

It can be seen that the more ants choose a link, the more the pheromone of that link will be, and as a result, its probability of being selected in the next iteration is increased. Based on this, the pheromone of each link is calculated as equation 14:

$$\tau_{ij}(t+1) = (1-\rho) \times \tau_{ij}(t) + \Delta\tau_{ij}(t) \tag{14}$$

where the ρ is the evaporation parameter of the pheromones. In fact, at the end of every iteration, for every link, a percentage of the pheromone value of the previous iteration remains, while some of the pheromones evaporate. Then, the remaining pheromone values for each link are aggregated with the values of the pheromone changes obtained in the current iteration and will update the pheromone matrix. As a result, after each iteration, according to the changes of the pheromone matrix, the probability matrix also changes, and the probability of high-quality links increases.

Checking the final condition of the iterations: Eventually, as the end condition of the iterations, one of the following two conditions must be met:

(1) Reaching the maximum number of iterations of the algorithm, displayed by MI.
(2) $max_{i,j}|\tau_{ij}(t+1) - \tau_{ij}(t) < \lambda|$

At the end of the algorithm, the final pheromone matrix is considered as the similarity matrix. The values in this matrix indicate the similarity scores between nodes. Accordingly, the higher the values of this matrix, the greater the similarity of the nodes in the row and column corresponding to that similarity value.

Time complexity: To discuss the time complexity of the proposed algorithm, different parts of the algorithm are considered separately. First, calculating the community matrix and the initial values of the pheromone matrix is of $O(n^2)$ time complexity, where n is the number of the nodes. To calculate the initial values of the heuristic matrix based on each of the mentioned measures (equations 2 to 8), determining the set of common neighbors of pairs of nodes is needed. In Python, if the size of two sets is p and q, determining the intersection of two sets is conducted in $O(min(p,q))$. Accordingly, having the average degree of the nodes in the network as k, the time complexity of the heuristic matrix initialization is equal to $O(n^2.k)$. Also, the time complexity of choosing a path by an ant is $O(n^2)$. Accordingly, the time complexity for t times of iterations, where in every iteration r ants choose a path, is equal to $O(rtn^2)$. Eventually, the overall time complexity of the algorithm is $O(rtn^2 + n^2.k)$. The provided time complexity assumes that the community structure of the network is already available. If the community structure is not known, the cost of community detection must be added to this time complexity. However, since many community detection methods are scalable, the overall cost of the algorithm remains acceptable for running on social networks. For example, the time complexity of label propagation-based community detection algorithms is $O(n+m)$, where n is the number of nodes and m is the number of edges [78].

4 Evaluations

Experiments are run to address the following main questions:

- Do the proposed community-based similarity measures outperform the existing community-based and none-community-based similarity measures in ranking the future links for link prediction?
- Can the proposed community-based ACO method

improve the effectiveness of the link prediction in social networks by combining its greedy search ability with local similarity-based measures enriched based on the intra-community and inter-community information?

All the implementations are run in Python, applying the networkX package [79]. Experiments are run on a 2.5 GHz Intel core i7 processor with 16 GB memory.

4.1 Evaluation Measures

To run the evaluations, the set of existing links, E, is randomly divided into two sets E^T and E^P, where $E = E^T \cup E^P$ and $E^T \cap E^P = \varnothing$. E^T denotes the set of training links and includes the 0.8 of existing links, and E^P denotes the set of test links. The needed sets are summarized in Table 1. To have a fair evaluation, the cross-validation method is applied here, where, the existing links are randomly divided into k subsets with the same size. The training and testing process is repeated k times, and in each iteration, one set is selected as the test set, and the rest of the sets are selected as the training set [13]. In this research, the value of k is considered equal to 10 based on similar studies [13].

Accordingly, applied evaluation measures in this study are as follows:

Area under curve (AUC) [68]: it is a widely applied measure to evaluate the link prediction algorithms, and is presented through equation 15:

$$AUC = \frac{\acute{n} + 0.5.\acute{\acute{n}}}{n} \tag{15}$$

where n is the number of independent comparisons, $\acute{n}$ represents the number of times that the random selection of a missing link (link within E^P) has a higher similarity score than the random selection of a non-existent link (link within $U - E$), and $\acute{\acute{n}}$ represents the number of times that their similarity scores are equal to each other. The closer the value of the AUC is to 0.5, the higher the random behavior of that method, and the closer it is to 1, the higher the effectiveness of it.

Precision [68]: This measure is another widely applied one to evaluate the effectiveness of the proposed algorithms in the link prediction field. This measure is presented through equation 16:

$$precision = \frac{L_r}{L} \tag{16}$$

where, L represents the number of links with the highest score among the predicted links and L_r represents the number of these links that are inside $E^P(|L| \leq |E^P|)$. In the current study, and in line with existing literature, $|L|$ is considered equal to $|E^P|$.

Table 1. Applied sets in the evaluations

applied sets	Description
E^T	set of training links
E^P	set of test links
$E = E^T \cup E^P$	set of existing links
U	set of all possible links
$U - E$	set of none-existing links
$U - E^T$	set of none-observed links

4.2 Datasets

Applied datasets are presented as follows:

Karate [80]: This social network includes the relationships of friendships between 34 members of a karate club. Due to disagreements between the manager and coach of the club, the nodes that represent the members are divided into two different groups, which has caused intra-community and inter-community links in this network.

Dolphin [81]: This undirected social network comprises 62 bottlenose dolphins. These dolphins live in small groups and communicate with each other using sounds. The behavior of these 62 dolphins has been studied for seven years. This network has two communities.

Football [82]: This network models American football matches in American high schools. The nodes and links in this network represent the teams, and the games played between them, respectively. The teams are divided into 12 different conferences so that almost every conference has between 8 to 12 teams. In this network, the number of games played between teams of the same conference is much more than the number of games played between teams from different conferences.

Books about American politics [83]: This network contains information on the books sold on Amazon during the 2004 presidential election. The nodes represent the sold books and the links represent books purchased by the same buyer. Nodes are in three categories: liberal, neutral, or conservative. The number of links between books from one category is more than the links between books from different categories. These three categories represent the three communities in this network.

Astro Physics collaboration network [84]: This network contains the information of the e-print arXiv considering the collaborations between authors of the papers submitted to Astro Physics category. The information is from January 1993 to April 2003.

Twitter [85]: This dataset consists of the circles from Twitter and it was crawled from public sources. The dataset includes user profiles, circles, and ego networks.

Email data from a European research institution [84]: This dataset includes all incoming and outgoing emails between members of a large European research institution. Members constitute the nodes and the emails between them are the edges.

Facebook [85]: This dataset consists of 'friends lists' from Facebook including profiles, circles, and ego networks.

Twitch Social Network (users of England) [86]: This dataset includes the Twitch user-user relationships of gamers in England. Nodes are the users and the links are the friendships between them. These social networks is collected in May 2018.

Among the applied datasets, the ground truth community structure of Karate, Dolphine, Football, Books, and email-Eu-core datasets are available. For the Astro Physics collaboration network, Twitter, Facebook, and Twitch datasets, one of the most applicable community detection methods, louvain community detection method [87] is applied here.

The information on datasets is summarized in Table 2, and their degree distribution is presented in Fig.3.

4.3 Benchmark Methods

As described earlier, this study focuses on extending the similarity-based local measures. Accordingly, first, three local similarity-based measures, which are Common neighbors (CN) [17], Jaccard index (J) [19], and resource allocation index (RA) [20, 21], are extended considering the information of community structure. These three proposed community-based measures are called CRA, CCN, and CJ. Next, these three proposed measures in addition to two existing community-based measures (SHRA) proposed by Soundarajan and Hopcroft [73], and SHCN proposed by Soundarajan and Hopcroft [73],) are extended into an ACO-based approach. Accordingly, the information of proposed and benchmark methods is summarized in Table 3.

As it is demonstrated in Table. 3, two categories of methods are adopted as the benchmark methods:

(1) Common neighbors (CN) [17], Jaccard index (J) [19], and resource allocation index (RA) [20, 21] as the local similarity-based measures.

(2) SHRA proposed by Soundarajan and Hopcroft [73], SHCN proposed by Soundarajan and Hopcroft [73], community-based resource allocation measure proposed by Bai et al. [74] (CMS-RA), and community-based common neighbors measure proposed by Bai et al. [74] (CMS-CN), as the extended similarity-based measures considering community structure. The difference between SHRA and SHCN with CMS-RA and CMS-CN is that the first two mentioned methods only calculate the similarity between the pair of nodes in the communities, while the last two mentioned methods calculate the similarity between the pair of nodes between different communities as well. The advantages of the first three proposed methods in this study, CRA, CNN, and CJ, over CMS-RA, and CMS-CN include: 1) adjusting the relative importance of inter-community neighbors compare to intra-community neighbors using two different weights which their sum is one, and 2) incorporating the parameter ϵ to consider the role of current weak links in the formation of future links.

In addition to the two aforementioned categories, an ACO-based version of the RA measure (RA-ACO) is also employed as a benchmark method. This approach differs from other ACO-based methods in this study in that its heuristic matrix is initialized using the RA values of node pairs, without considering their community co-membership. Similarly, the pheromone matrix is initialized by setting each entry to 1 if an edge exists between the two nodes, or to 0.25 otherwise. Consequently, community information is not incorporated into either matrix. RA measure was selected because its community-based version in ACO has performed better in most datasets.

4.4 Results and Discussion

To run the five versions of community-aware ACO-based methods, value of different parameters of the algorithm should be adjusted. Accordingly, $\rho = 0.1$, $\epsilon = 0.1$, $\acute{\alpha} = 0.6$, and $\acute{\beta} = 0.4$ are considered for all datasets. To adjust α and β, α is varied from 0.75 to 3 in increments of 0.25, while for each value, β is varied from 0.5 to 2 in increments of 0.25, and the best values are determined for every dataset (Table. 4).

The precision and AUC results obtained by running different methods on different datasets are presented in Tables 5 to 8. Considering 10 different runs of ACO and for every run, use of cross-validation, the mean value of the results is reported for each dataset. For the ACO-based results, the result intervals are also provided.

The bold values in Tables 5 to 8 represent the maximum values among all methods for each dataset. As shown, the mean precision and mean AUC obtained by the CRA-ACO method are higher than those of the other methods across all datasets except Dolphin

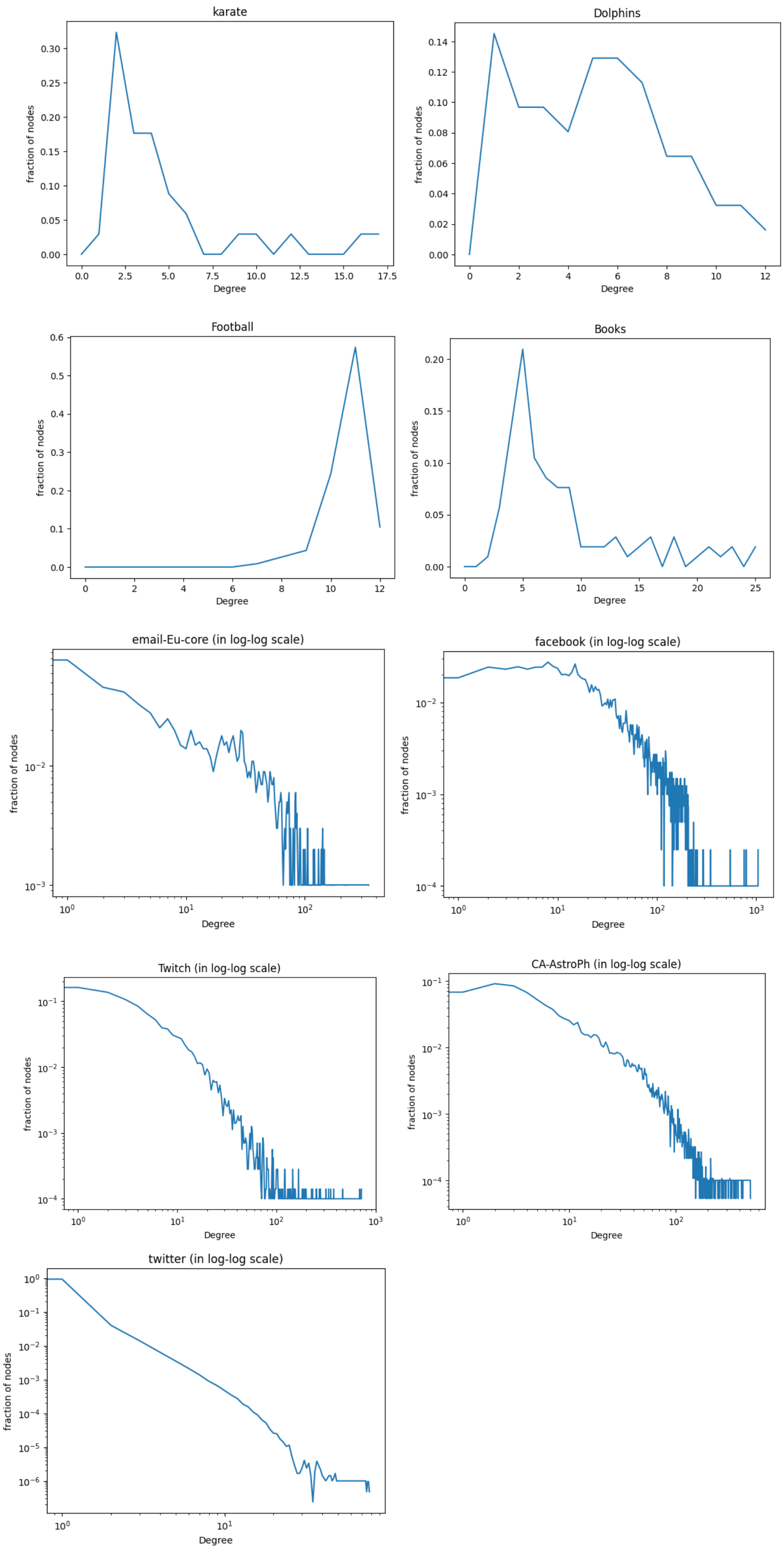


Figure 3. Degree Distributions of the Datasets.

Table 2. The Characteristics of the Applied Aatasets.

	dataset	Number of nodes	Number of links	Number of communities	Average degree	Diameter	Density	Clustering coefficient average
1	Karate	34	78	2	4.58	5	0.13	0.570
2	Dolphins	62	159	2	5.12	8	0.08	0.258
3	Football	115	613	12	10.66	4	0.09	0.403
4	Books	105	441	3	8.40	7	0.08	0.487
5	email-Eu-core	1005	25571	42	33.24	7	0.033	0.3994
6	Facebook	4039	88234	16	43.69	8	0.010	0.6055
7	Twitch of England	7126	35324	17	9.91	10	0.0013	0.1309
8	ca-AstroPh	18772	198110	38	21.1	14	0.0011	0.63
9	Twitter	81306	1342310	70	33.01	7	0.0004	0.56531

and Football. According to the degree distribution of the networks in Fig. 3, these two networks adhere more strongly to a power-law degree distribution, where not only the number of common neighbors but also their degrees play a significant role.

By comparing the values of local similarity-based measures and different versions of the proposed ACO-based method, it is evident that when a local similarity-based measure in these networks outperforms other methods, their corresponding ACO-based version outperforms other ACO-based methods.

CRA-ACO and CCN-ACO methods have better results than SHRA-ACO and SHCN-ACO methods. This issue demonstrates that considering inter-community relations is an essential factor in improving the performance of link prediction methods. These results are observed in community-based similarity measures, where CRA and CCN, and CMS-RA and CMS-CN outperform SHRA and SHCN due to considering the intra-community relationships. Moreover, CRA and CCN methods produce better results than CMS-RA and CMS-CN considering the applied improvements.

According to the precision and AUC results of different datasets, it can be concluded that the proposed approach outperforms others due to its combined nature (local and community structure) and also feedbacks from the ants at the end of every iteration.

5 Conclusions

Considering the importance of link prediction in the social network domain, i.e., predicting the links that will appear in the future in the network, extending the local similarity-based method based on 1) information of the community structure, and 2) the ability to get feedback from the ants in ACO, is of concern in this research to explore more accurate answers. Community structure is one of the main characteristics of social networks, based on which both intra-community and inter-community relationships have their importance. In the few studies, which have considered community structure in the link prediction, often only the relationships within the communities are of concern and the relationships between communities are not considered. By a combination of the mentioned approaches in the ACO algorithm, a new method is proposed for link prediction in social networks. Accordingly, three new similarity measures are presented to initialize the heuristics matrix of the ACO algorithm, which predicts the links in some iterations. To evaluate the performance of the proposed method, several benchmark methods and nine real data sets are applied. The results show the better effectiveness of this method for networks.

In addition to the importance of link structure and its related characteristics in link prediction, information flow dynamics in social networks are very important in this regard. Assessing the process of information propagation between non-adjacent nodes can play an important role in determining future links. Combining this information with community structure information in the proposed method is one of the future research directions. Extending the idea of this research to the link prediction in attributed social networks, specifically signed social networks, is also one of the future works. Applying the community-based ant-colony approach to link prediction in dynamic and temporal

Table 3. Summery of the Proposed and Benchmark Methods.

	Proposed / Benchmark	Method	Description
1	**Proposed methods, Not ACO-based**	Community-based Jaccard index (CJ)	community-based considering both inter-community and intra-community links
2		Community-based common neighbors (CCN)	community-based considering both inter-community and intra-community links
3		community-based resource allocation (CRA)	community-based considering both inter-community and intra-community links
4	**Proposed methods, ACO-based**	ACO-based version of the proposed CJ (CJ-ACO)	community-based considering both inter-community and intra-community links
5		ACO-based version of the proposed CCN (CCN-ACO)	community-based considering both inter-community and intra-community links
6		ACO-based version of the proposed CRA (CRA-ACO)	community-based considering both inter-community and intra-community links
7		ACO-based version of SHCN proposed by Soundarajan and Hopcroft [73] (SHCN-ACO)	community-based considering only intra-community links
8		ACO-based version of SHRA proposed by Soundarajan and Hopcroft [73] (SHRA-ACO)	community-based considering only intra-community links
9	**Benchmark methods**	Jaccard index[19] (J)	Non community-based
10		Common neighbors [17] (CN)	Non community-based
11		resource allocation index [20, 21] (RA)	Non community-based
12		community-based common neighbors proposed by Bai et al. [74] (CMS-CN)	community-based considering both inter-community and intra-community links
13		community-based resource allocation proposed by Bai et al. [74] (CMS-RA)	community-based considering both inter-community and intra-community links
14		community-based common neighbors proposed by Soundarajan and Hopcroft [73] (SHCN)	community-based considering only intra-community links
15		community-based resource allocation proposed by Soundarajan and Hopcroft [73] (SHRA)	community-based considering only intra-community links
16		ACO-based version of RA index (RA-ACO)	Non community-based

Table 4. Values of α and β for each Dataset.

dataset	Karate	Dolphins	Football	Books	email-Eu-core	Facebook	Twitch of England	ca-AstroPh	Twitter
α **value**	1	1.5	0.75	1.5	2	2	1.5	1	1.75
β **value**	0.75	1	0.75	1.25	2	1.5	1	0.75	1.5

Table 5. Results of Mean Precision for None-Community-Based and Community-Based Methods.

Datasets	J	CN	RA	CMS-CN	CMS-RA	SHCN	SHRA	CJ	CCN	CRA
Karate	0.021	0.130	0.213	0.133	0.216	0.131	0.213	0.047	0.141	**0.253**
Dolphin	0.120	0.177	0.160	0.185	0.180	0.180	0.168	0.146	**0.208**	0.192
Football	0.418	0.396	0.382	0.421	0.436	0.463	0.428	0.507	**0.509**	0.499
Book	0.158	0.242	0.272	0.259	0.288	0.242	0.276	0.189	0.267	**0.299**
Email-Eu-core	0.292	0.312	0.329	0.358	0.370	0.354	0.362	0.348	0.364	**0.371**
Facebook	0.276	0.297	0.315	0.338	0.350	0.326	0.332	0.328	0.371	**0.384**
Twitch	0.144	0.168	0.168	0.184	0.189	0.179	0.181	0.166	0.190	**0.194**
ca-AstroPh	0.230	0.262	0.276	0.316	0.321	0.303	0.314	0.289	0.319	**0.328**
Twitter	0.146	0.161	0.176	0.224	0.229	0.207	0.211	0.165	0.227	**0.236**

Table 6. Results of Precision for ACO-Based Methods.

Datasets		SHCN-ACO	SHRA-ACO	CJ-ACO	CCN-ACO	CRA-ACO	RA-ACO
Karate	Mean	0.224	0.257	0.034	0.244	**0.338**	0.230
	Interval	[0.190,0.243]	[0.226,0.272]	[0.028,0.038]	[0.215,0.264]	[0.294,**0.388**]	[0.194,0.255]
Dolphin	Mean	0.214	0.168	0.139	**0.238**	0.221	0.168
	Interval	[0.207,0.212]	[0.149,0.181]	[0.123,0.141]	[0.228,**0.249**]	[0.216,0.228]	[0.142,0.179]
Football	Mean	0.461	0.464	0.527	**0.539**	0.524	0.468
	Interval	[0.434,0.492]	[0.447,0.497]	[0.488,0.552]	[0.501,**0.556**]	[0.593,0.546]	[0.435,0.492]
Book	Mean	0.271	0.309	0.206	0.292	**0.345**	0.290
	Interval	[0.248,0.303]	[0.277,0.335]	[0.179,0.237]	[0.261,0.334]	[0.328,**0.381**]	[0.256,0,329]
Email-Eu-core	Mean	0.380	0.383	0.371	0.394	**0.404**	0.377
	Interval	[0.364,0.405]	[0.356,0,417]	[0.356,0.398]	[0.382,0.422]	[0.383,**0.426**]	[0.357,0.403]
Facebook	Mean	0.389	0.397	0.367	0.406	**0.410**	0.355
	Interval	[0.348,0.443]	[0.349,0.456]	[0.304,0.428]	[0.367,0.458]	[0.364,**0.463**]	[0.312,0.399]
Twitch	Mean	0.194	0.206	0.186	0.226	**0.237**	0.187
	Interval	[0.123,0.275]	[0.108,0.316]	[0.101,0.258]	[0.138,0.336]	[0.159,**0.340**]	[0.099,0.269]
ca-AstroPh	Mean	0.354	0.368	0.331	0.365	**0.387**	0.345
	Interval	[0.284,0.439]	[0.301,0.457]	[0.278,0.389]	[0.306,0.440]	[0.320,**0.465**]	[0.300,0.418]
Twitter	Mean	0.273	0.289	0.262	0.286	**0.305**	0.269
	Interval	[0.253,0.304]	[0.248,0.313]	[0.251,0.279]	[0.245,0.307]	[0.252,**0.328**]	[0.232,0.302]

Table 7. Results of Mean AUC for None-Community-Based and Community-Based Methods.

Datasets	J	CN	RA	CMS-CN	CMS-RA	SHCN	SHRA	CJ	CCN	CRA
Karate	0.600	0.635	0.680	0.663	0.682	0.624	0.660	0.608	0.702	**0.714**
Dolphin	0.704	0.746	0.725	0.758	0.748	0.734	0.732	0.712	**0.860**	0.732
Football	0.819	0.827	0.804	0.828	0.821	0.824	0.811	0.827	**0.838**	0.809
Book	0.826	0.816	0.832	0.840	0.855	0.818	0.840	0.841	0.852	**0.866**
Email-Eu-core	0.877	0.884	0.892	0.901	0.911	0.893	0.897	0.898	0.911	**0.915**
Facebook	0.770	0.783	0.794	0.839	0.842	0.813	0.825	0.840	0.856	**0.858**
Twitch	0.672	0.695	0.707	0.718	0.726	0.712	0.724	0.690	0.728	**0.731**
ca-AstroPh	0.785	0.808	0.811	0.824	0.830	0.814	0.827	0.823	0.830	**0.839**
Twitter	0.721	0.743	0.751	0.771	0.779	0.767	0.774	0.760	0.788	**0.798**

Table 8. Results of AUC for ACO-Based Methods.

Datasets		SHCN-ACO	SHRA-ACO	CJ-ACO	CCN-ACO	CRA-ACO	RA-ACO
Karate	Mean	0.668	0.712	0.635	0.725	**0.780**	0.659
	Interval	[0.600,0.724]	[0.656,0,730]	[0.565,0.702]	[0.678,0,758]	[0.673,**0.849**]	[0.561,0.749]
Dolphin	Mean	0.826	0.752	0.728	**0.870**	0.830	0.762
	Interval	[0.763,0.856]	[0.674,0.832]	[0.660,0.796]	[0.831,**0.910**]	[0.778,0.869]	[0.696,0.814]
Football	Mean	0.849	0.843	0.860	**0.879**	0.846	0.841
	Interval	[0.817,0.870]	[0.812,0.882]	[0.837,0.893]	[0.840,**0.904**]	[0.808,0.877]	[0.805,0.878]
Book	Mean	0.868	0.874	0.885	0.902	**0.923**	0.885
	Interval	[0.831,0.893]	[0.830,0,902]	[0.856,0.919]	[0.868,0.935]	[0.892,**0.953**]	[0.833,0.922]
Email-Eu-core	Mean	0.909	0.914	0.902	0.924	**0.939**	0.903
	Interval	[0.897,0.915]	[0.909,0.922]	[0.894,0.917]	[0.912,0.937]	[0.926,**0.945**]	[0.891,0.915]
Facebook	Mean	0.870	0.877	0.862	0.897	**0.911**	0.843
	Interval	[0.829,0.904]	[0.822,0.913]	[0.815,0.894]	[0.850,0.931]	[0.857,**0.955**]	[0.809,0.890]
Twitch	Mean	0.738	0.744	0.715	0.750	**0.761**	0.737
	Interval	[0.611,0.825]	[0.652,0.857]	[0.590,0.803]	[0.628,0.865]	[0.654,**0.896**]	[0.604,0.820]
ca-AstroPh	Mean	0.882	0.906	0.865	0.910	**0.914**	0.865
	Interval	[0.859,0.928]	[0.875,0.925]	[0.826,0.878]	[0.874,0.936]	[0.881,**0.937**]	[0.817,0.884]
Twitter	Mean	0.808	0.821	0.794	0.819	**0.832**	0.807
	Interval	[0.745,0.875]	[0.734,0.883]	[0.717,0.866]	[0.724,0.872]	[0.784,**0.896**]	[0.759,0.870]

social networks represents another valuable direction for future research. The application of deep learning approaches integrated with ACO for enhancing link prediction in social networks represents another promising direction for future research.

Availability of Data and Materials

The applied datasets in this study are publicly available at https://github.com/vlivashkin/community-graphs, https://snap.stanford.edu/data/email-Eu-core.html, https://snap.stanford.edu/data/egonets-Facebook.html, https://www.kaggle.com/datasets/andreagarritano/twitch-social-networks, https://snap.stanford.edu/data/ca-AstroPh.html, https://snap.stanford.edu/data/ego-Twitter.html.

References

[1] Virinchi Srinivas and Pabitra Mitra. *Link prediction in social networks: role of power law distribution*. Springer, 2016. doi:10.1007/978-3-319-28922-9.

[2] Caiyan Dai, Ling Chen, and Bin Li. Link prediction based on sampling in complex networks. *Applied Intelligence*, 47(1):1–12, 2017. doi:10.1007/s10489-016-0872-1.

[3] Santo Fortunato. Community detection in graphs. *Physics reports*, 486(3-5):75–174, 2010. doi:10.1016/j.physrep.2009.11.002.

[4] Kibae Kim and Jörn Altmann. Effect of homophily on network formation. *Communications in Nonlinear Science and Numerical Simulation*, 44:482–494, 2017. doi:10.1016/j.cnsns.2016.08.011.

[5] Yann Bramoullé, Sergio Currarini, Matthew O Jackson, Paolo Pin, and Brian W Rogers. Homophily and long-run integration in social networks. *Journal of Economic Theory*, 147(5):1754–1786, 2012. doi:10.1016/j.jet.2012.05.007.

[6] Víctor Martínez, Fernando Berzal, and Juan-Carlos Cubero. A survey of link prediction in complex networks. *ACM computing surveys (CSUR)*, 49(4):1–33, 2016. doi:10.1145/3012704.

[7] Herman Yuliansyah, Zulaiha Ali Othman, and Azuraliza Abu Bakar. Taxonomy of link prediction for social network analysis: A review. *IEEE Access*, 8:183470–183487, 2020. doi:10.1109/ACCESS.2020.3029122.

[8] Ajay Kumar, Shashank Sheshar Singh, Kuldeep Singh, and Bhaskar Biswas. Link prediction techniques, applications, and performance: A survey. *Physica A: Statistical Mechanics and its Applications*, 553:124289, 2020. doi:10.1016/j.physa.2020.124289.

[9] Abdul Samad, Mamoona Qadir, Ishrat Nawaz, Muhammad Arshad Islam, and Muhammad Aleem. A comprehensive survey of link prediction techniques for social network. *EAI Endorsed Trans. Ind. Networks Intell. Syst.*, 7(23):e3, 2020. doi:10.4108/eai.13-7-2018.163988.

[10] Djihad Arrar, Nadjet Kamel, and Abdelaziz Lakhfif. A comprehensive survey of link prediction methods: D. arrar et al. *The journal of supercomputing*, 80(3):3902–3942, 2024. doi:10.1007/s11227-023-05591-8.

[11] Ehsan Sherkat, Maseud Rahgozar, and Masoud Asadpour. Structural link prediction based on ant colony approach in social networks. *Physica a: statistical mechanics and its applications*, 419: 80–94, 2015. doi:10.1016/j.physa.2014.10.011.

[12] Anand Nayyar and Rajeshwar Singh. Ant colony optimization-computational swarm intelligence technique. In *2016 3rd International conference on computing for sustainable global development (INDIACom)*, pages 1493–1499. IEEE, 2016.

[13] Bolun Chen and Ling Chen. A link prediction algorithm based on ant colony optimization. *Applied Intelligence*, 41(3):694–708, 2014. doi:10.1007/s10489-014-0558-5.

[14] Zhiwei Cao, Yichao Zhang, Jihong Guan, and Shuigeng Zhou. Link prediction based on quantum-inspired ant colony optimization. *Scientific reports*, 8(1):13389, 2018. doi:10.1038/s41598-018-31254-3.

[15] Zhiwei Cao, Yichao Zhang, Jihong Guan, Shuigeng Zhou, and Guanghui Wen. A chaotic ant colony optimized link prediction algorithm. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(9):5274–5288, 2019. doi:10.1109/TSMC.2019.2947516.

[16] Shambhu Kumar, Arti Jain, and Dinesh Bisht. Ant colony optimization for link prediction in online social networks: A swarm intelligence approach. In *Proceedings of the 2024 Sixteenth International Conference on Contemporary Computing*, pages 632–641, 2024. doi:10.1145/3675888.3676123.

[17] David Liben-Nowell and Jon Kleinberg. The link prediction problem for social networks. In *Proceedings of the twelfth international conference on Information and knowledge management*, pages 556–559, 2003. doi:10.1145/956863.956972.

[18] Lada A Adamic and Eytan Adar. Friends and neighbors on the web. *Social networks*, 25(3):211–230, 2003. doi:10.1016/S0378-8733(03)00009-1.

[19] Paul Sathre, Atharva Gondhalekar, and Wu-chun Feng. Edge-connected jaccard similarity for graph link prediction on fpga. In *2022 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–10. IEEE, 2022.

doi:10.1109/HPEC55821.2022.9926326.

[20] Srinivas Virinchi and Pabitra Mitra. Similarity measures for link prediction using power law degree distribution. In *International Conference on Neural Information Processing*, pages 257–264. Springer, 2013. doi:10.1007/978-3-642-42042-9_33.

[21] Víctor Martínez, Fernando Berzal, and Juan-Carlos Cubero. Adaptive degree penalization for link prediction. *Journal of Computational Science*, 13:1–9, 2016. doi:10.1016/j.jocs.2015.12.003.

[22] David Liben-Nowell. *An algorithmic approach to social networks*. PhD thesis, Massachusetts Institute of Technology, 2005.

[23] Zhie Gao and Amin Rezaeipanah. A novel link prediction model in multilayer online social networks using the development of katz similarity metric. *Neural Processing Letters*, 55(4):4989–5011, 2023. doi:10.1007/s11063-022-11076-1.

[24] Mustafa Coşkun, Abdelkader Baggag, and Mehmet Koyutürk. Fast computation of katz index for efficient processing of link prediction queries. *Data Mining and Knowledge Discovery*, 35(4):1342–1368, 2021. doi:10.1007/s10618-021-00754-8.

[25] Weiping Liu and Linyuan Lü. Link prediction based on local random walk. *EPL (europhysics Letters)*, 89(5):58007, 2010. doi:10.1209/0295-5075/89/58007.

[26] Hanghang Tong, Christos Faloutsos, and Jia-Yu Pan. Fast random walk with restart and its applications. In *Sixth international conference on data mining (ICDM'06)*, pages 613–622. IEEE, 2006. doi:10.1109/ICDM.2006.70.

[27] Oron Vanunu and Roded Sharan. A propagation-based algorithm for inferring gene-disease associations. In *German conference on bioinformatics*, pages 54–63. Gesellschaft für Informatik e. V., 2008.

[28] Linyuan Lü, Ci-Hang Jin, and Tao Zhou. Similarity index based on local paths for link prediction of complex networks. *Phys. Rev. E*, 80:046122, 2009. doi:10.1103/PhysRevE.80.046122.

[29] Alexis Papadimitriou, Panagiotis Symeonidis, and Yannis Manolopoulos. Friendlink: link prediction in social networks via bounded local path traversal. In *2011 International Conference on Computational Aspects of Social Networks (CASoN)*, pages 66–71. IEEE, 2011. doi:10.1109/CASON.2011.6085920.

[30] Kamal Berahmand, Elahe Nasiri, Saman Forouzandeh, and Yuefeng Li. A preference random walk algorithm for link prediction through mutual influence nodes in complex networks. *Journal of king saud university-computer and information sciences*, 34(8):5375–5387, 2022. doi:10.1016/j.jksuci.2021.05.006.

[31] Shihu Liu, Xueli Feng, and Jin Yang. Asymmetric influence-based superposed random walk link prediction algorithm in complex networks. *International Journal of Modern Physics C*, 36(10): 2442002, 2025. doi:10.1142/S0129183124420026.

[32] Aaron Clauset, Cristopher Moore, and Mark EJ Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453 (7191):98–101, 2008. doi:10.1038/nature06830.

[33] Roger Guimerà and Marta Sales-Pardo. Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences*, 106(52):22073–22078, 2009. doi:10.1073/pnas.0908366106.

[34] Natalie Stanley, Thomas Bonacci, Roland Kwitt, Marc Niethammer, and Peter J Mucha. Stochastic block models with multiple continuous attributes. *Applied Network Science*, 4(1):54, 2019. doi:10.1007/s41109-019-0170-z.

[35] Srilatha Pulipati, Ramasubbareddy Somula, and Balakesava Reddy Parvathala. Nature inspired link prediction and community detection algorithms for social networks: a survey. *International Journal of System Assurance Engineering and Management*, pages 1–18, 2021. doi:10.1007/s13198-021-01125-8.

[36] Nitai B Silva, Ren Tsang, George DC Cavalcanti, and Jyh Tsang. A graph-based friend recommendation system using genetic algorithm. In *IEEE congress on evolutionary computation*, pages 1–7. IEEE, 2010. doi:10.1109/CEC.2010.5586144.

[37] Jeff Naruchitparames, Mehmet Hadi Güneş, and Sushil J Louis. Friend recommendations in social networks using genetic algorithms and network topology. In *2011 IEEE Congress of Evolutionary Computation (CEC)*, pages 2207–2214. IEEE, 2011. doi:10.1109/CEC.2011.5949888.

[38] Catherine A Bliss, Morgan R Frank, Christopher M Danforth, and Peter Sheridan Dodds. An evolutionary algorithm approach to link prediction in dynamic social networks. *Journal of Computational Science*, 5(5):750–764, 2014. doi:10.1016/j.jocs.2014.01.003.

[39] Abbas Ali Rezaee and Navid Abravan. A hybrid friend-based recommendation system using the combination of meta-heuristic invasive weed and genetic algorithms. In *2020 10th International Conference on Computer and Knowledge Engineering (ICCKE)*, pages 665–669. IEEE, 2020. doi:10.1109/ICCKE50421.2020.9303619.

[40] Akrivi Krouska, Christos Troussas, and Cleo Sgouropoulou. Applying genetic algorithms for recommending adequate competitors in mobile game-based learning environments. In *Interna-*

tional Conference on Intelligent Tutoring Systems, pages 196–204. Springer, 2020. doi:10.1007/978-3-030-49663-0_23.

[41] Guangfu Chen, Haibo Wang, Yili Fang, and Ling Jiang. Link prediction by deep non-negative matrix factorization. *Expert Systems with Applications*, 188:115991, 2022. doi:10.1016/j.eswa.2021.115991.

[42] Asan Agibetov. Neural graph embeddings as explicit low-rank matrix factorization for link prediction. *Pattern Recognition*, 133:108977, 2023. doi:10.1016/j.patcog.2022.108977.

[43] Zhili Zhao, Ahui Hu, Nana Zhang, Jiquan Xie, Zihao Du, Li Wan, and Ruiyi Yan. Mining node attributes for link prediction with a non-negative matrix factorization-based approach. *Knowledge-Based Systems*, 299:112045, 2024. doi:10.1016/j.knosys.2024.112045.

[44] Jing Yang, Xiaofen Wang, Laurence T Yang, Yuan Gao, Shundong Yang, and Xiaokang Wang. Learning schema embeddings for service link prediction: a coupled matrix-tensor factorization approach. *IEEE Transactions on Services Computing*, 2025. doi:10.1109/TSC.2025.3541552.

[45] Wei Yu, Jiale Fu, Yanxia Zhao, Hongjin Shi, Xue Chen, Shigen Shen, and Xiao-Zhi Gao. Link prediction in bipartite networks via deep autoencoder-like nonnegative matrix factorization. *Applied Soft Computing*, 169:112616, 2025. doi:10.1016/j.asoc.2024.112616.

[46] Na Shan, Longjie Li, Yakun Zhang, Shenshen Bai, and Xiaoyun Chen. Supervised link prediction in multiplex networks. *Knowledge-Based Systems*, 203:106168, 2020. doi:10.1016/j.knosys.2020.106168.

[47] Riccardo Giubilei and Pierpaolo Brutti. Supervised classification for link prediction in facebook ego networks with anonymized profile information. *Journal of Classification*, 39(2):302–325, 2022. doi:10.1007/s00357-021-09408-2.

[48] Anisha Kumari, Ranjan Kumar Behera, Kshira Sagar Sahoo, Anand Nayyar, Ashish Kumar Luhach, and Satya Prakash Sahoo. Supervised link prediction using structured-based feature extraction in social network. *Concurrency and Computation: practice and Experience*, 34(13):e5839, 2022. doi:10.1002/cpe.5839.

[49] Yinzuo Zhou, Weilun Chen, and Huangrong Zou. A link prediction algorithm based on support vector machine. *Physica A: Statistical Mechanics and its Applications*, 673:130674, 2025. doi:10.1016/j.physa.2025.130674.

[50] Hossien Ghorbanzadeh, Amir Sheikhahmadi, Mahdi Jalili, and Sadegh Sulaimany. A hybrid method of link prediction in directed graphs. *Expert Systems with Applications*, 165:113896, 2021. doi:10.1016/j.eswa.2020.113896.

[51] Costas Mavromatis and George Karypis. Graph infoclust: Maximizing coarse-grain mutual information in graphs. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 541–553. Springer, 2021. doi:10.1007/978-3-030-75762-5_43.

[52] Jinyin Chen, Jian Zhang, Xuanheng Xu, Chenbo Fu, Dan Zhang, Qingpeng Zhang, and Qi Xuan. E-lstm-d: A deep learning framework for dynamic network link prediction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(6):3699–3712, 2019. doi:10.1109/TSMC.2019.2932913.

[53] Mohammad Mehdi Keikha, Maseud Rahgozar, and Masoud Asadpour. Deeplink: A novel link prediction framework based on deep learning. *Journal of Information Science*, 47(5):642–657, 2021. doi:10.1177/0165551519891345.

[54] Unai Zulaika, Ruben Sanchez-Corcuera, Aitor Almeida, and Diego Lopez-de Ipina. Lwp-wl: Link weight prediction based on cnns and the weisfeiler–lehman algorithm. *Applied Soft Computing*, 120: 108657, 2022. doi:10.1016/j.asoc.2022.108657.

[55] Joshua Robinson, Rishabh Ranjan, Weihua Hu, Kexin Huang, Jiaqi Han, Alejandro Dobles, Matthias Fey, Jan Eric Lenssen, Yiwen Yuan, Zecheng Zhang, et al. Relbench: A benchmark for deep learning on relational databases. *Advances in Neural Information Processing Systems*, 37: 21330–21341, 2024. doi:10.52202/079017-0672.

[56] Sonia Choudhary and Gyanender Kumar. Enhancing link prediction in dynamic social networks through hybrid gcn-lstm models. *Knowledge and Information Systems*, 67(8):6717–6751, 2025. doi:10.1007/s10115-025-02430-5.

[57] Adnan Zeb, Summaya Saif, Junde Chen, Anwar Ul Haq, Zhiguo Gong, and Defu Zhang. Complex graph convolutional network for link prediction in knowledge graphs. *Expert systems with applications*, 200:116796, 2022. doi:10.1016/j.eswa.2022.116796.

[58] Hongxu Chen, Hongzhi Yin, Xiangguo Sun, Tong Chen, Bogdan Gabrys, and Katarzyna Musial. Multi-level graph convolutional networks for cross-platform anchor link prediction. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1503–1511, 2020. doi:10.1145/3394486.3403201.

[59] Peiliang Zhang, Jiatao Chen, Chao Che, Liang Zhang, Bo Jin, and Yongjun Zhu. Iea-gnn: Anchor-aware graph neural network fused with information entropy for node classification and link prediction. *Information Sciences*, 634:665–676, 2023. doi:10.1016/j.ins.2023.03.022.

[60] Chaobo He, Junwei Cheng, Xiang Fei, Yu Weng, Yulong Zheng, and Yong Tang. Community preserving adaptive graph convolutional networks for link prediction in attributed networks. *Knowledge-Based Systems*, 272:110589, 2023. doi:10.1016/j.knosys.2023.110589.

[61] Yuxin Wang, Xiannian Hu, Quan Gan, Xuanjing Huang, Xipeng Qiu, and David Wipf. Efficient link prediction via gnn layers induced by negative sampling. *IEEE Transactions on Knowledge and Data Engineering*, 37(1):253–264, 2024. doi:10.1109/TKDE.2024.3481015.

[62] Yuren Mao, Yu Hao, Xin Cao, Yunjun Gao, Chang Yao, and Xuemin Lin. Boosting gnn-based link prediction via pu-auc optimization. *IEEE Transactions on Knowledge and Data Engineering*, 2025. doi:10.1109/TKDE.2025.3525490.

[63] Marcus Lim, Azween Abdullah, and NZ Jhanjhi. Data fusion-link prediction for evolutionary network with deep reinforcement learning. *International Journal of Advanced Computer Science and Applications*, 11(6), 2020.

[64] Ling Chen, Jun Cui, Xing Tang, Yuntao Qian, Yansheng Li, and Yongjun Zhang. Rlpath: a knowledge graph link prediction method using reinforcement learning based attentive relation path searching and representation learning: Rlpath: A knowledge graph link prediction method using reinforcement learning based attentive relation path searching and representation learning. *Applied Intelligence*, 52(4):4715–4726, 2022. doi:10.1007/s10489-021-02672-0.

[65] Mingshuo Nie, Dongming Chen, Dongqi Wang, and Huilin Chen. Local optimization policy for link prediction via reinforcement learning. *IEEE Transactions on Network Science and Engineering*, 12(2):1224–1236, 2025. doi:10.1109/TNSE.2025.3526340.

[66] Jingwei Wang, Yunlong Ma, Min Liu, Han Yuan, Weiming Shen, and Ling Li. A vertex similarity index using community information to improve link prediction accuracy. In *2017 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 158–163. IEEE, 2017. doi:10.1109/SMC.2017.8122595.

[67] Mohamed Hassen Kerkache, Lamia Sadeg-Belkacem, Fatima Benbouzid-Si Tayeb, and Amri Ali. Supervised learning using community detection for link prediction. In *International Conference on Computing Systems and Applications*, pages 85–94. Springer, 2022. doi:10.1007/978-3-031-12097-8_8.

[68] Anupam Biswas and Bhaskar Biswas. Community-based link prediction. *Multimedia Tools and Applications*, 76(18):18619–18639, 2017. doi:10.1007/s11042-016-4270-9.

[69] Anisha Kumari, Ranjan Kumar Behera, Bibudatta Sahoo, and Satya Prakash Sahoo. Prediction of link evolution using community detection in social network. *Computing*, 104(5):1077–1098, 2022. doi:10.1007/s00607-021-01035-4.

[70] M Mohamed Iqbal and K Latha. An effective community-based link prediction model for improving accuracy in social networks. *Journal of Intelligent & Fuzzy Systems*, 42(3):2695–2711, 2022. doi:10.3233/JIFS-211821.

[71] Rong Kuang, Qun Liu, and Hong Yu. Community-based link prediction in social networks. In *International Conference in Swarm Intelligence*, pages 341–348. Springer, 2016. doi:10.1007/978-3-319-41009-8_37.

[72] Jingwei Wang, Yunlong Ma, Min Liu, and Weiming Shen. Link prediction based on community information and its parallelization. *IEEE Access*, 7:62633–62645, 2019. doi:10.1109/ACCESS.2019.2907202.

[73] Sucheta Soundarajan and John Hopcroft. Using community information to improve the precision of link prediction methods. In *Proceedings of the 21st international conference on world wide web*, pages 607–608, 2012. doi:10.1145/2187980.2188150.

[74] Shenshen Bai, Shiyu Fang, Longjie Li, Rui Liu, and Xiaoyun Chen. Enhancing link prediction by exploring community membership of nodes. *International Journal of Modern Physics B*, 33(31): 1950382, 2019. doi:10.1142/S021797921950382X.

[75] Longjie Li, Shiyu Fang, Shenshen Bai, Shijin Xu, Jianjun Cheng, and Xiaoyun Chen. Effective link prediction based on community relationship strength. *IEEE Access*, 7:43233–43248, 2019. doi:10.1109/ACCESS.2019.2908208.

[76] Zhao Yang, Rongjing Hu, and Ruisheng Zhang. An improved link prediction algorithm based on common neighbors index with community membership information. In *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pages 90–93. IEEE, 2016. doi:10.1109/ICSESS.2016.7883022.

[77] Akrati Saxena, George Fletcher, and Mykola Pechenizkiy. Hm-eiict: Fairness-aware link prediction in complex networks using community information. *Journal of Combinatorial Optimization*, 44(4):2853–2870, 2022. doi:10.1007/s10878-021-00788-0.

[78] Sara E Garza and Satu Elisa Schaeffer. Community detection with the label propagation algorithm: a survey. *Physica A: Statistical Mechanics and its Applications*, 534:122058, 2019. doi:10.1016/j.physa.2019.122058.

[79] Aric Hagberg, Pieter J Swart, and Daniel A Schult. Exploring network structure, dynamics,

and function using networkx. Technical report, Los Alamos National Laboratory (LANL), 2007.

[80] Wayne W Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, 33(4):452–473, 1977. doi:10.1086/jar.33.4.3629752.

[81] David Lusseau, Karsten Schneider, Oliver J Boisseau, Patti Haase, Elisabeth Slooten, and Steve M Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations: can geographic isolation explain this unique trait? *Behavioral ecology and sociobiology*, 54(4):396–405, 2003. doi:10.1007/s00265-003-0651-y.

[82] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826, 2002. doi:10.1073/pnas.122653799.

[83] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006. doi:10.1073/pnas.0601602103.

[84] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM transactions on Knowledge Discovery from Data (TKDD)*, 1(1): 2–es, 2007. doi:10.1145/1217299.1217301.

[85] Jure Leskovec and Julian Mcauley. Learning to discover social circles in ego networks. *Advances in neural information processing systems*, 25, 2012.

[86] Benedek Rozemberczki, Carl Allen, and Rik Sarkar. Multi-scale attributed node embedding. *Journal of Complex Networks*, 9(2):cnab014, 2021. doi:10.1093/comnet/cnab014.

[87] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008. doi:10.1088/1742-5468/2008/10/P10008.

Seyed Mohammad Hosseini received his B.Sc. and M.Sc. from the University of Isfahan, Isfahan, Iran, in 2016 and 2019 respectively. His main areas of research interest include graph analysis, and social networks.

Afsaneh Fatemi received her B.Sc. from the Isfahan University of Technology, Isfahan, Iran, in 1995, and her M.Sc. and Ph.D from the University of Isfahan, Isfahan, Iran in 2001 and 2012 respectively, all in Computer Engineering. Currently, she is an associate professor in the Faculty of Computer Engineering at the University of Isfahan, and a member of UI Big Data research group. Her main research interests lie in the fields of Complex Networks, Question Answering and Dialogue Systems, and applications of LLMs.

Maryam Hosseini-Pozveh received her B.Sc. degree in software engineering and her M.Sc. degree in software engineering from the University of Isfahan, Isfahan, Iran, in 2005 and 2009, respectively. She obtained her Ph.D. degree in the software engineering field from the University of Isfahan in 2016. Currently, she is an assistant professor in the Department of Computer Engineering, at the Shahreza Campus, University of Isfahan. Her main areas of research interest include Complex networks, Data science, and Advanced algorithms.