Complex & Intelligent Systems (2025) 11:307
https://doi.org/10.1007/s40747-025-01925-6

ORIGINAL ARTICLE

Strength prominence index: a link prediction method in fuzzy social network

Sakshi Dev Pandey[1] · Sovan Samanta[2,3,4] · A. S. Ranadive[1] · Leo Mrsic[4,5] · Antonios Kalampakas[6] · Tofigh Allahviranloo[2]

Received: 1 February 2025 / Accepted: 18 April 2025

Abstract
Link prediction is a field within social network studies that aims to forecast future connections based on the structure of a social network. This paper introduces a link prediction method based on the strength and prominence of seed node pairs, referred to as the strength prominence index. In this method, we get a consistent score for any pair of nodes, regardless of whether they share a common neighbour. Several key characteristics have been identified. In our experiments, we used three well-known estimators to evaluate the accuracy of link prediction algorithms: precision, area under the precision-recall curve, and area under the receiver operating characteristic curve. A comparative study with existing methods is also presented, supported by relevant graphs and tables. Validation using Facebook data sets demonstrates the effectiveness of the proposed method.

Keywords Link prediction · Similarity indices · Fuzzy social network · Strength prominence (SP) index

Introduction

The power of social networks in designing and analyzing a wide range of complex systems, including social, biotechnological, and communication networks, has been well-documented [30]. Research on social network systems has gained significant traction and has become a valuable tool across various scientific disciplines. One ongoing challenge in this field is link prediction in multiple networks, which has been extensively studied in recent years. The core idea of the link prediction model is to estimate the edges that will be added to a social media network between two time points, t and $t+1$. Increasing the number of edges within a social network is a critical challenge. Numerous implementations of link prediction exist in the literature. For example, online shopping networks like Amazon, Flipkart, Myntra, eBay, and Alibaba use link prediction to suggest related items [18]. Similarly, social media platforms use it to recommend new friends, and security applications use it to uncover networks of terrorists and criminals.

Motivation

Complex methods can provide insights into network evolution, link formation, and link predictability [36]. While these methods are praised for their elegance, they are often seen as

✉ Sovan Samanta
ssamantavu@gmail.com

Sakshi Dev Pandey
sdpandey.in@gmail.com

A. S. Ranadive
asranadive04@yahoo.co.in

Leo Mrsic
leo.mrsic@algebra.hr

Antonios Kalampakas
antonios.kalampakas@aum.edu.kw

Tofigh Allahviranloo
tofigh.allahviranloo@istinye.edu.tr

1 Department of Mathematics, Guru Ghasidas Vishwavidyalaya (A Central University), Bilaspur, CG 495009, India

2 Research Center of Performance and Productivity Analysis, Istinye University, Istanbul, Turkey

3 Department of Mathematics, Tamralipta Mahavidyalaya, Tamluk, WB 721636, India

4 Department of Technical Sciences, Algebra University, Gradiscanska 24, 10000 Zagreb, Croatia

5 Rudolfovo Science and Technology Centre, Podbreznik 15, 8000 Novo Mesto, Slovenia

6 College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait

Published online: 29 May 2025

proofs of concept rather than practical solutions. They typically estimate missing links based on the local graphical or topological structure of links and nodes, without considering the strength, nature, or characteristics of ties and nodes. In today's world of manufactured relationships, it is essential to record specific characteristics, such as the depth of relationships.

Therefore, research into efficient and reliable approaches must continue. Traditional methods, such as Common Neighbor [29], Jaccard Coefficient [12], Resource Allocation, Adamic Adar [1], and Preferential Attachment [5], focus on the quantity of common neighbours. In 2013, Cannistraci et al. [9] proposed an improved index called CAR, based on the local (neighborhood)-community-paradigm. They demonstrated that traditional neighborhood-based methods could be significantly enhanced by incorporating the local community link (LCL), which considers the number of links between common neighbours. However, in social networks, each actor and tie has unique properties and characteristics. These factors are not equally measurable, and sometimes we need to predict links between nodes with no common neighbours. Metrics that accurately predict node similarity are encouraging. Recently, fuzzy graph network models have shown essential characteristics visible in most real-life datasets, highlighting the crucial role of fuzzy graphs in link building. Fuzzy graphs can play a more generalized role and convey specific score functions.

Important contributions of the study

This paper emphasizes the strength of nodes and ties over the number of nodes or links, and demonstrates how fuzzy graphs can help predict missing links. Fuzzy graphs are also useful for dealing with vague information, such as when objects or their relationships are not precisely defined. This study introduces the Strength Prominence (SP) index for link prediction, a new similarity index that provides a more realistic and accurate score function compared to the CAR index. Two types of score functions are defined for scenarios where a common neighbor exists and where no common neighbor exists between the seed nodes.

Framework of the study

The review of literature on link predictions in social networks and improvements in their score functions is presented in "Literature review" section. "Preliminaries" section provides a list of notations and abbreviations (Table:1), preliminary definitions related to fuzzy graphs and basic similarity indices of link prediction. "Strength prominence (SP) index for link prediction in fuzzy graphs and evaluation metrics" section introduces the SP-index with two types of score functions for link predictions and describes some evaluation matrices. "Description and implementation of an algorithm for link prediction" section presents the algorithms and their implementation with numerical illustrations, it also includes a comparative analysis of existing score functions. "Area of applications" section presents applications of link prediction in different fields. Finally, "Conclusion" section concludes the manuscript.

Literature review

Liben-Nowell et al. [17] specifically suggested a link prediction approach to social networks. Each node on the network represents a person or other entity, and edges represent the connection or relationship between them. Based on the network's nodes' degree of proximity to one another, various link prediction methods are discussed. Hasan et al. [2] went on to elaborate on this work and show that having more topological network information available significantly improves prediction outcomes. They used supervised learning to perform binary classification tasks, which is comparable to link prediction in their framework while taking into account various similarity measures as features. A paper on link prediction on collaboration networks in physics and biology was presented by M. E. J. Newman [29]. In these networks, multiple authors are regarded as connected if they simultaneously coauthor at least one paper. The results provide experimental support for previously hypothesized clustering as well as power-law degree distribution throughout networks mechanisms.

In recent years, numerous link prediction techniques have been put forth in various contexts [3, 11, 16]. In this article, we just focus on similarity-based techniques using the degree of strength of the nodes and edges in social media networks. Therefore, the main challenge is to appropriately define node similarity. The network structure is the sole basis for one subset of similarity indices. The preferential attachment [5] index, which is merely the product of the degrees of two seed nodes, is the most basic. The Jaccard [12] index is a normalization of common neighbours [29], which counts the number of neighbours that seed nodes have in common. Adamic–Adar [1] and Resource Allocation [38] are characterized by utilizing the degree of information of common neighbours to obtain a better resolution. Based on the idea that the likelihood of a link existing between two nodes increases if their mutual neighbours are also inhabitants of the same local community, Cannistrai–Alanis–Ravi [9] suggested an improved index (CAR). Wu et al. [36] recently put forth a new kind of score function that is based on node and link clustering data. They tested their algorithms on three different categories of networks, including small-scale, middle-scale, and large-scale networks. In this follow-up, Kumar et al. [15] defines the concept of a level-2 common node as well as its associated clustering coefficient, which extracts clustering data from level-2 common neighbours of the seed node couple and uses this data to compute the similarity score. The aforementioned techniques primarily rely on common neighbours and only count the number of common nodes and edges between them. There are also a few proposed global and quasi-local techniques, including Kartz [14], Sim Rank [13], Random walk [19], Local Paths [20], and others. Although the aforementioned techniques can produce better outcomes than typical neighborhood-based techniques, it can be challenging to use them on intricate and expansive networks. Few related works on networks can be found in [10, 21, 28, 34].

Several fuzzy based link prediction techniques have been discussed in this paragraph. Bastani et al. [6] introduced similarity indices leveraging fuzzy logic and granular computing, incorporating Yager's clustering approach [37]. Bhawsar et al. [7] developed indices based on fuzzy soft sets and Markov models, enhancing prediction accuracy in uncertain environments. On the basis of semi-supervised fuzzy clustering, Tran et al. [35] proposed some new similarity measures for link prediction. The RSM index for link prediction was also developed by Mahapatra et al. [22], and it is based on the nature of the common neighbor, which is determined by the degree to which the common neighbor is a member of the fuzzy graph. However, it does not account for the falsity and indecency parameters often present in uncertain environments. To address these limitations within neutrosophic fuzzy graphs, again Mahapatra et al. [23] enhanced the RSM index by incorporating additional parameters such as the nature of the job, location, and others. An adaptive decision-making system called learning automation learns how to select the best course of action, which enhances performance. Behnaz et al. [26] suggests a method for link prediction in fuzzy social networks based on distributed learning automata. Fuzzy graphs undoubtedly play a crucial part in the development of complex networks. The aforementioned paragraph demonstrates some research on link prediction, but more work needs to be done. The research conducted in this study is also a step toward the same objective.

Preliminaries

Here we define some key terms associated with the fuzzy graphs and talk about the various similarity metrics that can be used for link predictions.

Basic definitions

Definition 1 [39] Let X be a collection of objects represented usually by x, then a fuzzy set A in X is a collection of ordered tuples, such that

$$A = \{(x, \mu_A(x)) | x \in X\}, \tag{3.1}$$

where $\mu_A : X \to [0, 1]$ is a membership function (generalized characteristic function). $\mu_A(x)$ represent the degree of membership of an element x in A.

Definition 2 [39] Let X be universe of discourse and A be a fuzzy set in X, then core of a fuzzy set is defined as

$$Core(A) = \{x | \mu_A(x) = 1\}. \tag{3.2}$$

Definition 3 [24] A fuzzy graph $G^* = (V, \mu, \nu)$ over the graph $G = (V, E)$ is characterized by membership functions $\mu : V \to [0, 1]$ and $\nu : V \times V \to [0, 1]$ such that

$$\nu(v_1, v_2) \le \mu(v_1) \wedge \mu(v_2)\ \forall v_1, v_2 \in V, \tag{3.3}$$

where μ is called the fuzzy vertex set of G and $\mu(v_1)$ represents the membership value of the vertex v_1. Similarly, ν is called the fuzzy edge set of G and $\nu(v_1, v_2)$ represents the membership value of the edge (v_1, v_2). Clearly ν is a fuzzy relation on μ. We suppose that V is finite and non empty, ν is reflexive and symmetric.

Definition 4 [25] A path between the pair of nodes x and y in a fuzzy graph G^* is a collection of distinct nodes $P : x = v_0, v_1, v_2, \ldots, v_n = y$ such that $\mu(v_i) > 0$ and $\min\{\mu(v_{i-1}), \mu(v_i)\} \ge \nu(v_{i-1}, v_i) > 0$ for all i.

Definition 5 [25] The minimal membership degree of an edge in the path is called the strength of the path that means $S(P_i) = \min_{1 \le i \le n}\{\nu(v_{i-1}, v_i)\}$ is called the strength of P. Strength of connectedness between two vertices x and y is defined as the maximum of strengths of all paths in G^* from x to y, represented by $CONN_{G^*}(x, y) = \max_{P_i \in \rho[x,y]} S(P_i)$, where $\rho[x, y]$ is the collection of all paths between x and y.

Definition 6 [27] Let $G^* = (V, \mu, \nu)$ be a fuzzy graph, then degree of a vertex v_i in a fuzzy graph G^* is defined as $d(v_i) = \sum_{v_i \ne v \in V} \nu(v_i, v)$.

In this article, we refer to the degree of v_i as prominence or prominence degree of node v_i.

Definition 7 [8] A fuzzy graph $G^* = (V, \mu, \nu)$ is said to be strong if all the edges in a fuzzy graph take the least membership value of their adjacent vertices. Such that,

$$\nu(x, y) = \min\{\mu(x), \mu(y)\} \quad \forall (x, y) \in E \subset V \times V.$$

Definition 8 [27] A fuzzy graph $G^* = (V, \mu, \nu)$ is said to be complete if $\nu(x, y) = \min\{\mu(x), \mu(y)\}\ \forall x, y \in V$.

Table 1 List of notations and abbreviations

Notation	Meaning
G	Graph
G*	Fuzzy graph
V	Set of vertices
E	Set of edges
$\mu(v_i)$	Membership value of the vertex v_i
$\nu(v_i, v_j)$	Membership value of the edge (v_i, v_j)
$S(v_i, v_j)$	Score value of the link prediction between v_i and v_j
$\Gamma(v_i)$	Collection of neighbours of the node v_i
k_z	Degree of the common neighbour z
$LCL(v_i, v_j)$	Local community link between v_i and v_j
MAP	Mean average precision
AUC	Area under receiver operating characteristic curve
AUP	Area under the precision-recall curve

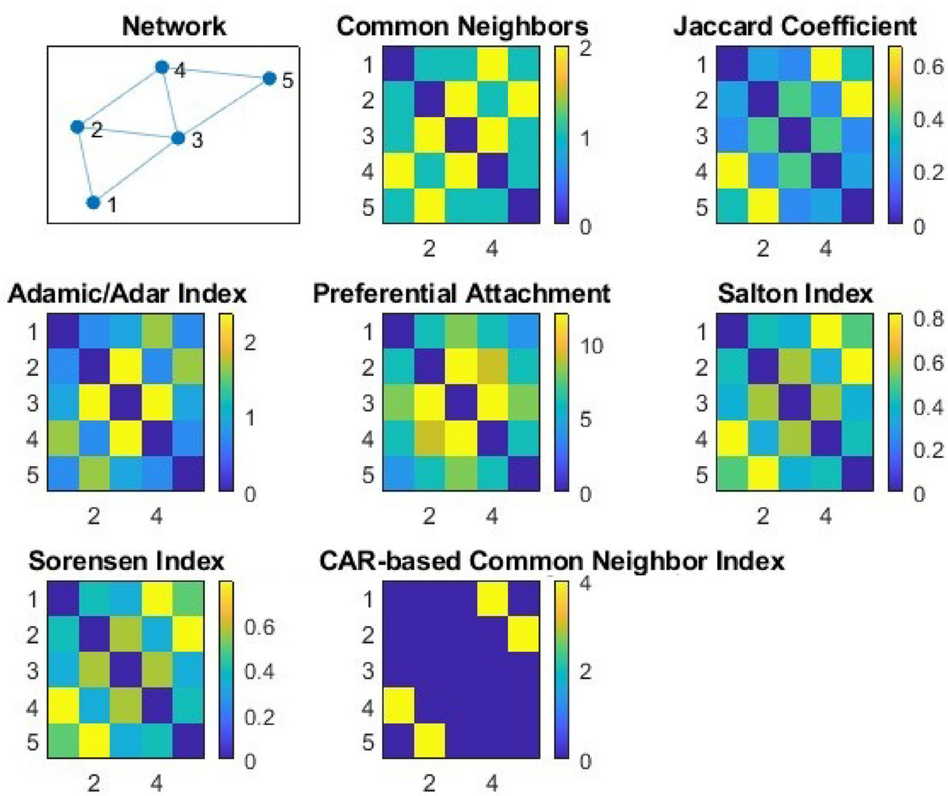

Fig. 1 A graph and its link prediction scores

Some basic similarity indices

- Common Neighbors(CN) [29]: The number of common neighbours for a given pair of seed nodes v_i and v_j in a particular network is determined as the size of the intersection of the two nodes' $(v_i \& v_j)$ neighborhoods.

$$S(v_i, v_j) = |\Gamma(v_i) \cap \Gamma(v_j)|, \tag{3.4}$$

where $\Gamma(v_i)$ and $\Gamma(v_j)$ are the collection of neighbours of the nodes v_i and v_j respectively.

- Jaccard Coefficient (JC) [12]: The Jaccard index is described to be the probability of selecting common neighbours of paired seed nodes from all of their neighbours. The amount of common neighbours between the two seed nodes increases the pairwise Jaccard score.

$$S(v_i, v_j) = \frac{|\Gamma(v_i) \cap \Gamma(v_j)|}{|\Gamma(v_i) \cup \Gamma(v_j)|}. \tag{3.5}$$

- Adamic/Adar Index [1]: Adamic and Adar proposed a metric for calculating the degree of similarity between two websites based on shared attributes, which can then be used in link prediction after a few modifications. It is founded on the idea that common neighbours with lower degrees are given more weight. This makes sense in the actual world as well; for instance, someone with more friends will spend less time with each friend individually than someone with fewer friends.

$$S(v_i, v_j) = \sum_{z \in \Gamma(v_i, v_j)} \frac{1}{logk_z}, \tag{3.6}$$

  in this case, k_z represents the degree of the common neighbor z.
- Preferential Attachment (PA) [5]: Using the concept of preferential attachment, an expanding scale-free network is created. Following is a formula for computing the preferential attachment value between two nodes.

$$S(v_i, v_j) = k_{v_i}.k_{v_j}. \tag{3.7}$$

  On most networks, according to research in the literature, this exhibits the worst efficiency. The main benefits of this metric are its simplicity (since it needs the least amount of data to calculate the score) and computational speed. The efficacy of the PA increases in assortative networks while it deteriorates dramatically in disassortative networks.
- Salton Index (SI) [31]: The Salton index, formerly known as Cosine similarity, can be used to calculate document commonalities in a vector space. By computing the cosine of the angle between two records, the similarity score between them is determined. The main measure here is orientation rather than magnitude. The Salton score is calculated as

$$S(v_i, v_j) = \frac{|\Gamma(v_i) \cap \Gamma(v_j)|}{\sqrt{k_{v_i}.k_{v_j}}}. \tag{3.8}$$

- Sorensen Index (SI) [33]: This similarity index, developed by Thorvald Sorensen in 1948, is particularly useful for ecological data specimens. This index calculates the score as a ratio of two times the number of neighbours that are common and the sum of the degrees of the nodes v_i and v_j. It is very close to the Jaccard index. It is more resistant to anomalies than Jaccard.

$$S(v_i, v_j) = \frac{2|\Gamma(v_i) \cap \Gamma(v_j)|}{k_{v_i} + k_{v_j}}. \tag{3.9}$$

- CAR-based Common Neighbour index (CAR) [9]: This index is founded on the idea that as the number of links among the common neighbours (local community links (LCLs)) of the seed node pairs rises, so does the likelihood that a link will exist between seed nodes.

$$S(v_i, v_j) = CN(v_i, v_j) \times LCL(z_i, z_j), \tag{3.10}$$

  where $CN(v_i, v_j)$ is the number of common neighbours between seed nodes and $LCL(z_i, z_j)$ is the number of links between common neighbours of seed nodes.
- RSM Index [22]: The fuzzy graph social network defines this similarity index, which concentrates on the nature/characteristics of common neighbours. The RSM index is determined by averaging the nature (N_i) of seed node common neighbours. The following is the definition of this index:

$$S(v_i, v_j) = \sum_{i=1}^{n} \frac{N_i}{n}, \tag{3.11}$$

  where $N_i = \min\{\nu(v_i, z_i), \nu(v_j, z_i)\}$ and $z_i \in \Gamma(v_i v_j)$, i=1,2,3,. . .,n.

Here is an example of link prediction using the basic similarity indices in Fig. 1.

Strength prominence (SP) index for link prediction in fuzzy graphs and evaluation metrics

SP index

The SP index, a new similarity index, is described in this section. This link prediction index is based on the strength of links, common neighbours, strength of connectedness and prominence degree of seed nodes, which is inspired by the CAR model. SP suggests that two seed nodes are more probable to link along if their common first neighbours have a good strength of connections to both seed node pairs and each other, as well as in the absence of common neighbor / neighbours SP uses the strength of path and prominence degree of seed nodes. As a result of the following formulation, SP provides higher discriminative resolutions between candidate connections with the same number of common-first neighbours than CAR and RSM index, and this improvement in resolution is unmistakably derived from the usage of the relationship strength perspective. The score value for two distinct circumstances is calculated using SP index and is defined as follows:

Case 1: Let $G^* = (V, \mu, \nu)$ be a fuzzy graph. If there is a common neighbor between the seed nodes x and y or euclidean distance (shortest path length) between seed nodes is equal to 2 (see Fig. 2) then the similarity based SP index is given by

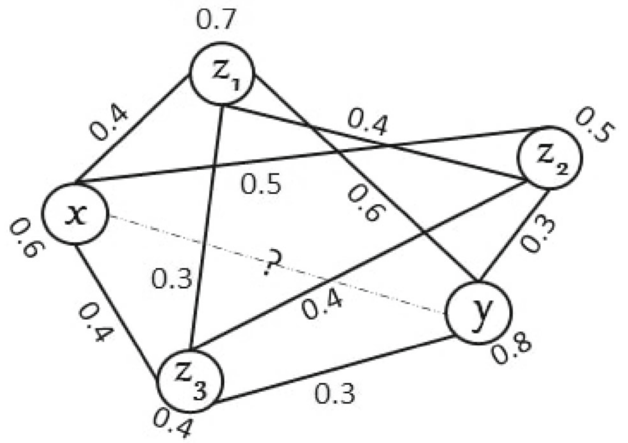

Fig. 2 The common neighbor (z_1, z_2, z_3) approach to predict links in a fuzzy graph

$$SP(x, y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu(x, z_i), \nu(y, z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu(x), \mu(z_i), \mu(y)\}} + \frac{\sum_{z_i, z_j \in \Gamma(xy)} \nu(z_i, z_j)}{\sum_{z_i, z_j \in \Gamma(xy)} \min\{\mu(z_i), \mu(z_j)\}}. \quad (4.1)$$

Specifically, $\Gamma(xy)$ is the set of vertices that are neighbours to both the x and y seed nodes. The first part of the Eq. (4.1), represents the degree of connection strength between seed nodes and common neighbours, while the second part represents the connection strength between the common neighbours themselves. When taken together, they illustrate the degree of "how likely it is that the seed nodes are connected". This score function is premised on the idea that strong ties between common neighbours increase the likelihood of a link existing between seed nodes.

Example 1 We consider a fuzzy graph (Fig. 2) with 5 nodes, in this network there exists common neighbours $\{z_1, z_2, z_3\}$ between seed nodes. We can compute the score of link prediction between seed nodes, $SP(x, y) = 0.67 + 0.85 = 1.52$.

Case 2: If there does not exists any common neighbor between the seed nodes x and y or the euclidean distance (shortest path length) between seed nodes is greater than 2 (see Fig. 3) then the similarity based SP index is determined by strength of connectedness between seed nodes and there prominence degrees which is defined as follows:

$$SP(x, y) = CONN_{G^*}(x, y) + \frac{d(x) + d(y) + 2d(x)d(y)}{2(1 + d(x) + d(y) + d(x)d(y))}, \quad (4.2)$$

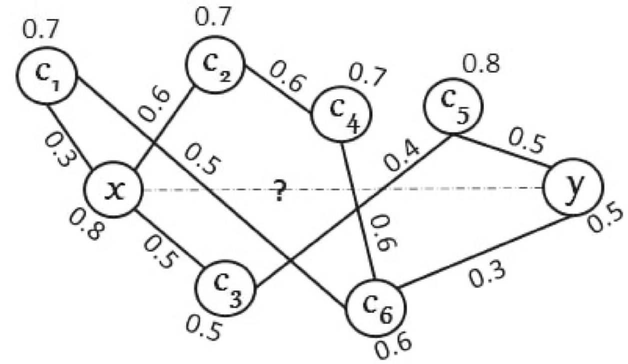

Fig. 3 Strength of connectedness approach to predict links in a fuzzy graph

where $CONN_{G^*}(x, y)$ is the strength of connectedness between seed nodes x and y, defined as the Max of the strengths of all paths between x and y, also $d(x), d(y)$ denotes the prominence of the seed nodes, which is determined similar to the degree of nodes. This definition is based on the assumption that there is a significant likelihood of a link existing between prominent actors in a social network. If the pair of seed nodes have different prominence degrees but their sum is equal, then the likelihood of existence of a link between seed nodes that have the same degree of prominence is high.

From the above formulations (equation 4.1 and equation 4.2) it is clear that SP index is bounded and lies between 0 and 2 ($0 \le SP(x, y) \le 2$).

Example 2 We consider a fuzzy graph (Fig. 3) with 8 nodes, in this network there does not exist any common neighbor between seed nodes. We can find the strength of all the paths between x and y and compute the score of link prediction between seed nodes, $P_1 = x, c_1, c_6, y$, $S(P_1) = 0.3$, $P_2 = x, c_3, c_5, y$, $S(P_2) = 0.4$, $P_3 = x, c_2, c_4, c_6, y$, $S(P_3) = 0.3$. Now, $CONN_{G^*}(x, y) = \max\{S(P_1), S(P_2), S(P_3)\} = 0.4$, also $d(x) = 0.8$, $d(y) = 0.8$. Thus the score of link prediction between seed nodes, $SP(x, y) = 0.844$.

Further, for another network we assume that if we have three seed node pairs (x, y), (g_1, h_1), and (g_2, h_2) with same strength of connectedness and varying degrees of prominence, $d(x) = 0.8, d(y) = 0.8$; $d(g_1) = 0.4, d(h_1) = 1.2$; $d(g_2) = 0.6, d(h_2) = 1.0$, but their sums are identical that is $0.8 + 0.8 = 0.4 + 1.2 = 0.6 + 1.0 = 1.6$. In this situation, the seed node pair with equal prominence has the highest likelihood of a link, we observe that from Eq. (4.2), $SP(x, y) = 0.444$, $SP(g_1, h_1) = 0.415$, and $SP(g_2, h_2) = 0.437$. This signifies that if both individuals are equally popular, i.e., have equal prominence degree, then the SP index provides more score value as compared to the other pairs of individuals who are not equally popular, however, all these pairs of individuals have identical sum of their prominence degrees.

The SP index, as stated in Eq. (4.1), is a generalization of the RSM index. RSM index simply takes into account the membership of links between seed nodes and common neighbours and ignores node potential and the strength of links between common neighbours, whereas SP index takes into account both of these qualities. The following theorem clarifies the concept of this generalization.

Theorem 1 *Let $G^* = (V, \mu, \nu)$ be the fuzzy graph. Suppose core of μ be the whole of V and if we omit the evaluation of links between common neighbours then SP index (having common neighbours) provides same prediction value as RSM index.*

Proof Given that core of the vertex fuzzy set μ be the whole of V,

$$\mu(x) = 1 \quad \forall x \in V. \tag{4.3}$$

Suppose $z_i's$, $i = 1, 2, 3 \ldots n$, are the common neighbor between seed nodes $a, b \in V$, then

$$SP(a, b) = \frac{\sum_{i=1}^{n} \min\{\nu(a, z_i), \nu(b, z_i)\}}{\sum_{i=1}^{n} \min\{\mu(a), \mu(z_i), \mu(b)\}} + \frac{\sum_{i \neq j=1}^{n} \nu(z_i, z_j)}{\sum_{i \neq j=1}\{\mu(z_i), \mu(z_j)\}}$$

Now, since $\mu(a) = \mu(b) = \mu(z_i) = 1$, for all $a, b, z_i \in V$.

$$SP(a, b) = \frac{\sum_{i=1}^{n} \min\{\nu(a, z_i), \nu(b, z_i)\}}{n} + \frac{\sum_{i \neq j=1}^{n} \nu(z_i, z_j)}{n}$$

Also, RSM index ignores the evaluation of links between common neighbours then $\nu(z_i, z_j) = 0, \quad i, j = 1, 2, \ldots n$.

$$SP(a, b) = \frac{\sum_{i=1}^{n} \min\{\nu(a, z_i), \nu(b, z_i)\}}{n}$$

From equation 3.11 (RSM index), it can be clearly seen that RHS is nothing but RSM(a,b). This completes the proof. □

Theorem 2 *Let $G^* = (V, \mu, \nu)$ be a strong fuzzy graph. Then the score of link prediction between seed node pairs (having common neighbours) possess the highest value.*

Proof Given that $G = (V, \mu, \nu)$ be a strong fuzzy graph. By the definition of strong fuzzy graph

$$\nu(x, y) = \min\{\mu(x), \mu(y)\}, \forall (x, y) \in E \subset V \times V. \tag{4.4}$$

Suppose a and b are the seed nodes in the fuzzy graph and $z_1, z_2, \ldots z_n$ are the common neighbours between a and b, then the score value of link prediction is calculated by

$$\begin{aligned} SP(a, b) &= \frac{\sum_{i=1}^{n} \min\{\nu(a, z_i), \nu(b, z_i)\}}{\sum_{i=1}^{n} \min\{\mu(a), \mu(z_i), \mu(b)\}} + \frac{\sum_{i \neq j=1}^{n} \nu(z_i, z_j)}{\sum_{i \neq j=1}^{n} \min\{\mu(z_i), \mu(z_j)\}} \\ &= \frac{\sum_{i=1}^{n} \min\{\min\{\mu(a), \mu(z_i)\}, \min\{\mu(b), \mu(z_i)\}\}}{\sum_{i=1}^{n} \min\{\mu(a), \mu(z_i), \mu(b)\}} + \frac{\sum_{i \neq j=1}^{n} \min\{\mu(z_i), mu(z_j)\}}{\sum_{i \neq j=1}^{n} \min\{\mu(z_i), \mu(z_j)\}} \\ &= \frac{\sum_{i=1}^{n} \min\{\mu(a), \mu(z_i), \mu(b)\}}{\sum_{i=1}^{n} \min\{\mu(a), \mu(z_i), \mu(b)\}} + \frac{\sum_{i \neq j=1}^{n} \min\{\mu(z_i), mu(z_j)\}}{\sum_{i \neq j=1}^{n} \min\{\mu(z_i), \mu(z_j)\}} \\ &= 2 \end{aligned}$$

Hence the score of link prediction in strong fuzzy graph takes its highest value. □

Remark 1 If $G^* = (V, \mu, \nu)$ is a strong fuzzy graph with each node has an equal membership then SP index approaches to the highest value whenever the prominence degree of seed nodes are equal and tends to infinity.

Remark 2 A complete fuzzy graph is strong fuzzy graph. Therefore, Theorem 2's conclusion also applies to a complete fuzzy graph, ensuring that every pair of nodes in a complete fuzzy graph has a maximum score.

The above remarks are applicable in both cases whether seed nodes have common neighbours or not.

Theorem 3 *Let $G_1 = (V_1, \mu_1, \nu_1)$ and $G_2 = (V_2, \mu_2, \nu_2)$ be two isomorphic fuzzy graphs then the score (link prediction value) for each corresponding node pair in G_1 and G_2 is equal.*

Proof Let $G_1 = (V_1, \mu_1, \nu_1)$ be isomorphic to $G_2 = (V_2, \mu_2, \nu_2)$. Let $f : V_1 \to V_2$ be the bijection such that

$$\mu_1(x) = \mu_2(f(x)) \text{ for all } x \in V_1,$$
$$\mu_1(x, y) = \mu_2(f(x), f(y)) \text{ for all } x, y \in V_1.$$

For case 1: When seed node pairs have common neighbours,

$$SP(x, y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu_1(x, z_i), \nu_1(y, z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu_1(x), \mu_1(z_i), \mu_1(y)\}} + \frac{\sum_{z_i, z_j \in \Gamma(xy)} \nu_1(z_i, z_j)}{\sum_{z_i, z_j \in \Gamma(xy)} \min\{\mu_1(z_i), \mu_1(z_j)\}},$$

$$= \frac{\sum_{z_i \in \Gamma(f(x)f(y))} \min\{\nu_2(f(x), f(z_i)), \nu_2(f(y), f(z_i))\}}{\sum_{z_i \in \Gamma(f(x)f(y))} \min\{\mu_2(f(x)), \mu_2(f(z_i)), \mu_2(f(y))\}} + \frac{\sum_{z_i, z_j \in \Gamma(f(x)f(y))} \nu_2(f(z_i), f(z_j))}{\sum_{z_i, z_j \in \Gamma(f(x)f(y))} \min\{\mu_2(f(z_i)), \mu_2(f(z_j))\}},$$
$$= SP(f(x), f(y)).$$

For case 2: When seed node pairs do not have common neighbours,

since G_1 is isomorphic to G_2, the strength of any path between x and y in G_1 is equal to the strength of any path between $f(x)$ and $f(y)$ in G_2. Therefore, $CONN_{G_1}(x, y) = CONN_{G_2}(f(x), f(y))$ for all $x, y \in V_1$, also by definition $d(x) = d(f(x))$ for all $x \in V_1$,

$$SP(x, y) = CONN_{G_1}(x, y) + \frac{d(x) + d(y) + 2d(x)d(y)}{2(1 + d(x) + d(y) + d(x)d(y))},$$
$$= CONN_{G_2}(f(x), f(y)) + \frac{d(f(x)) + d(f(y)) + 2d(f(x))d(f(y))}{2(1 + d(f(x)) + d(f(y)) + d(f(x))d(f(y)))},$$
$$= SP(f(x), f(y)).$$

Hence G_1 and G_2 have equal score values for link predictions. □

Theorem 4 (Boundedness of SP Index) *For any pair of nodes x and y in a fuzzy graph $G^* = (V, \mu, \nu)$, the SP index is bounded between 0 and 2, i.e.*

$$0 \leq SP(x, y) \leq 2.$$

Proof Case 1: When there are common neighbours between x and y:

$$SP(x, y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu(x, z_i), \nu(y, z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu(x), \mu(z_i), \mu(y)\}} + \frac{\sum_{z_i, z_j \in \Gamma(xy)} \nu(z_i, z_j)}{\sum_{z_i, z_j \in \Gamma(xy)} \min\{\mu(z_i), \mu(z_j)\}}.$$

Since ν and μ are membership functions with values in [0, 1], each term in the SP index is bounded by 1. Therefore, the sum of these terms is bounded by 2.

Case 2: When there are no common neighbours between x and y:

$$SP(x, y) = CONN_{G^*}(x, y) + \frac{d(x) + d(y) + 2d(x)d(y)}{2(1 + d(x) + d(y) + d(x)d(y))}.$$

The strength of connectedness $CONN_{G^*}(x, y)$ is bounded by 1, and the second term is a fraction that is also bounded by 1. Hence, the sum is bounded by 2.
Therefore, $0 \leq SP(x, y) \leq 2$. □

Theorem 5 (Symmetry of SP Index) *The SP index is symmetric for any pair of nodes x and y in a fuzzy graph $G^* = (V, \mu, \nu)$, i.e.*

$$SP(x, y) = SP(y, x).$$

Proof Case 1: When there are common neighbours between seed node pairs x and y:

$$SP(x, y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu(x, z_i), \nu(y, z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu(x), \mu(z_i), \mu(y)\}} + \frac{\sum_{z_i, z_j \in \Gamma(xy)} \nu(z_i, z_j)}{\sum_{z_i, z_j \in \Gamma(xy)} \min\{\mu(z_i), \mu(z_j)\}},$$

since $\min\{\nu(x, z_i), \nu(y, z_i)\} = \min\{\nu(y, z_i), \nu(x, z_i)\}$ and the sums are commutative, $SP(x, y) = SP(y, x)$.
Case 2: When there are no common neighbours between x and y:

$$SP(x, y) = CONN_{G^*}(x, y) + \frac{d(x) + d(y) + 2d(x)d(y)}{2(1 + d(x) + d(y) + d(x)d(y))},$$

the terms $CONN_{G^*}(x, y)$ and $\frac{d(x)+d(y)+2d(x)d(y)}{2(1+d(x)+d(y)+d(x)d(y))}$ are symmetric in x and y.
Therefore, $SP(x, y) = SP(y, x)$. □

Theorem 6 (Monotonicity with Respect to Strength of Connections) *If the strength of connections between nodes x and y increases, the SP index also increases.*

Proof Case 1: When there are common neighbours between x and y:

$$SP(x, y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu(x, z_i), \nu(y, z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu(x), \mu(z_i), \mu(y)\}} + \frac{\sum_{z_i, z_j \in \Gamma(xy)} \nu(z_i, z_j)}{\sum_{z_i, z_j \in \Gamma(xy)} \min\{\mu(z_i), \mu(z_j)\}}.$$

If $\nu(x, z_i)$ or $\nu(y, z_i)$ increases, the numerator in the first term increases, leading to an increase in $SP(x, y)$. Similarly, if $\nu(z_i, z_j)$ increases, the numerator in the second term increases, leading to an increase in $SP(x, y)$.
Case 2: When there are no common neighbours between x and y:

$$SP(x, y) = CONN_{G^*}(x, y) + \frac{d(x) + d(y) + 2d(x)d(y)}{2(1 + d(x) + d(y) + d(x)d(y))}.$$

If the strength of the path between x and y increases, $CONN_{G^*}(x, y)$ increases, leading to an increase in $SP(x, y)$. □

Theorem 7 (Monotonicity with Respect to Prominence Degree) *If the prominence degree of a node x increases, the SP index for any pair of nodes involving x also increases.*

Proof Case 1: When there are common neighbours between x and y:

$$SP(x,y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu(x,z_i), \nu(y,z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu(x), \mu(z_i), \mu(y)\}} + \frac{\sum_{z_i,z_j \in \Gamma(xy)} \nu(z_i,z_j)}{\sum_{z_i,z_j \in \Gamma(xy)} \min\{\mu(z_i), \mu(z_j)\}}.$$

The prominence degree $d(x)$ affects the membership values $\mu(x)$. If $\mu(x)$ increases, the denominators in the SP index decrease, leading to an increase in $SP(x, y)$.
Case 2: When there are no common neighbours between x and y:

$$SP(x,y) = CONN_{G^*}(x,y) + \frac{d(x)+d(y)+2d(x)d(y)}{2(1+d(x)+d(y)+d(x)d(y))}.$$

If $d(x)$ increases, the second term in the SP index increases, leading to an increase in $SP(x, y)$. □

Evaluation metrics

The SP index is a novel similarity measure that incorporates multiple factors such as the strength of links, common neighbours, strength of connectedness, and prominence degree of seed nodes. This comprehensive approach aims to provide a more accurate prediction of potential links in a network.

To validate the SP index, we compared its performance against established link prediction indices like Common Neighbors (CN) and Jaccard Coefficient (JC) using the Mean Average Precision (MAP) metric. MAP is a robust evaluation metric that considers the precision of predictions at different ranks, providing a comprehensive measure of the prediction quality.

The results (Fig. 4) showed that the SP index achieved a competitive MAP score, indicating its effectiveness in accurately predicting links. By plotting the MAP curves, we can visually compare the precision of the SP index with CN and JC indices across different ranks, further validating its performance.

Overall, the SP index demonstrates validity as a reliable link prediction method, offering improved resolution and accuracy by leveraging the strength of connections and prominence of nodes in the network.

We use **Precision, AUC** and **AUP** three separate estimators, to fully estimate the expected results. The fundamental steps in the computation of these estimators are identical. Let's use E to denote the entire link set, which is arbitrarily split into E_t and E_p. The probe set E_p is utilized for testing, and none of the data from the probe set may be used for prediction. The training set E_t is utilized as the known knowledge. Of course, $E = E_t \cup E_p$ and $E_t \cap E_p = null$. In this study, we use 15% of the links as test links. Our experiments used an average of 1500 independent runs to determine accuracy.

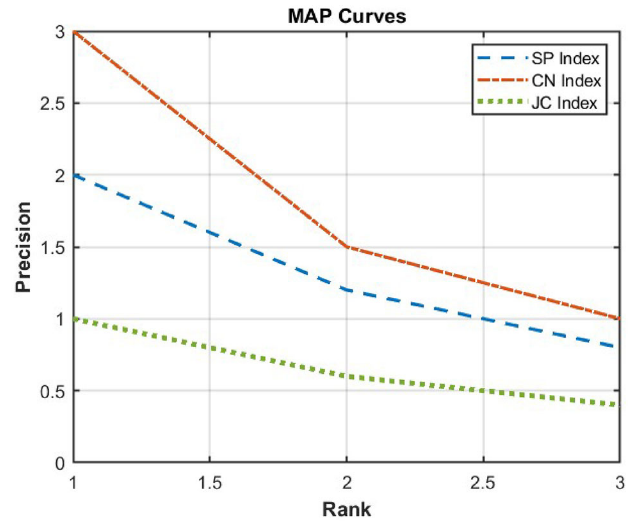

Fig. 4 Validation through MAP

Following the similarity matrix calculation, we apply a well-known technique to estimate how effective the suggested approach is. First, all non-observed links (those that are part of $U - E_t$, where U stands for the collection of all node pairs) are sorted in descending order. The precision is then determined by dividing the number of relevant items chosen by the total number of items chosen. This indicates that the precision can be defined as equation (4.5) if we pick the top-L connections as the predicted ones, among which the L_r links are correct (i.e., there are L_r links in the probe set E_p). Higher precision leads to greater accuracy in precision. For the sake of this paper, we set L to 21 (the size of E_p).

$$Precision = \frac{L_r}{L} \tag{4.5}$$

Cannistraci et al. [9] proposed the area under precision curve, which is designated as AUP. In this study, ten alternative values of L are used to produce the precision curve. As stated in the definition of precision above, the greatest value of L is that value, and all other values are an arithmetic sequence with a difference equal to one-fifth of the largest L.

An indicator of how accurately the prediction algorithms perform is the area under the receiver operating characteristic (AUC) curve. The likelihood that a randomly selected missing link (i.e., a link in E_p) will receive a higher score than a randomly selected non-existent link (i.e., a link in $U - E$) can be understood as its meaning. In reality equation (4.6)'s definition of the AUC computation is presented. where n represents the number of independent comparisons, n' represents instances in which the missing links have a higher score and n'' counts instances in which they have the same score. The AUC value should be close to 0.5 if each score was produced using a separate, identical distribution. As a result, the values exceeds 0.5 reflects how much better the algorithm performs than random chance. It is still a widely used index in link prediction estimates, the AUC findings are included in this article.

$$AUC = \frac{n' + 0.5n''}{n} \tag{4.6}$$

Description and implementation of an algorithm for link prediction

Algorithm description for determining the likelihood of link existence between seed nodes

Given a fuzzy graph $G^* = (V, \mu, \nu)$ with $|V| = n$. In [4], Banerjee proposed an optimal algorithm for determining the strength of connectedness between a pair of fuzzy graph vertices, which may be utilized to calculate the score for link prediction in the absence of common neighbor. Consider the steps below:

Step 1: Input the adjacency matrix of training graph of G^* and find the zero entries to identify the seed node pairs.

Step 2: For each seed node pair (x, y), determine whether they have common neighbours by checking non-zero entries in their adjacency matrix rows.

Step 3: If common neighbours exist, compute as following and stop; otherwise, proceed to Step 4.

$$SP(x, y) = \frac{\sum_{z_i \in \Gamma(xy)} \min\{\nu(x, z_i), \nu(y, z_i)\}}{\sum_{z_i \in \Gamma(xy)} \min\{\mu(x), \mu(z_i), \mu(y)\}} + \frac{\sum_{z_i, z_j \in \Gamma(xy)} \nu(z_i, z_j)}{\sum_{z_i, z_j \in \Gamma(xy)} \min\{\mu(z_i), \mu(z_j)\}}.$$

Step 4: (i) If there is no common neighbor between x and y, determine the strength of connectedness $CONN_{G^*}(x, y)$ by employing the algorithm in [4].
(ii) Find the prominence degree $d(x)$ and $d(y)$.
(iii) Compute the following,

$$SP(x, y) = CONN_{G^*}(x, y) + \frac{d(x) + d(y) + 2d(x)d(y)}{2(1 + d(x) + d(y) + d(x)d(y))}.$$

Following these procedures, one may determine the likelihood of the existence of links between seed nodes in a fuzzy graph.

Time complexity: The worst-case time complexity of the algorithm's is $O(n^3)$. This is primarily due to Step 4, which involves computing the strength of connectedness using an external algorithm (shortest path) [4].

- **Step 1** (Identifying seed node pairs)
  This step involves finding the seed node pairs in the adjacency matrix of the graph, which takes $O(n^2)$ time.
- **Step 2** (Finding common neighbours)
  For each pair of seed nodes, we check if they have any common neighbours by inspecting their adjacency matrix rows. This process takes $O(n^2\Delta)$, where Δ is the average node degree.
- **Step 3** (Computing link prediction score)
  If common neighbours are found, we calculate the link prediction score using the provided formula. This step has a time complexity of $O(n^2\Delta^2)$ in the worst case.
- **Step 4** (For pairs without common neighbours)
  If no common neighbours exist between a pair of nodes, we calculate the strength of connectedness and prominence degree. This step takes $O(n^3)$ time, if an all-pairs shortest path algorithm is used.

Since $O(n^3)$ dominates other terms in dense graphs, it represents the overall complexity. For sparse graphs, the complexity reduces to $O(n^2)$.

Implementation of the algorithm on a fuzzy social network

Consider the Facebook profiles of 20 people who are socially connected (see Table 2) and are included in the data from [32]. In the public profile, the number of friends is indicated. Each participant was approached offline and requested to supply information regarding the number of mutual friends she/he had with other members in order to gather this data. Table 3 reflects this.

The primary network of 20 nodes and 85 edges (see Table 2, Fig. 5) is used as an example of the suggested approach. The normalized value of the total number of friends is the membership value of the nodes that are being considered. Additionally, if two nodes possess mutual friends within the Facebook network, an edge connecting them is presumed to be present. If x and y are the names of the individuals in the Facebook data, m is the number of mutual friends, and k is the maximum number of mutual friends in the data, the edge membership value between x and y is given by $\nu(x, y) = \min\{\mu(x), \mu(y)\}\frac{m}{k}$ (see Table 4).

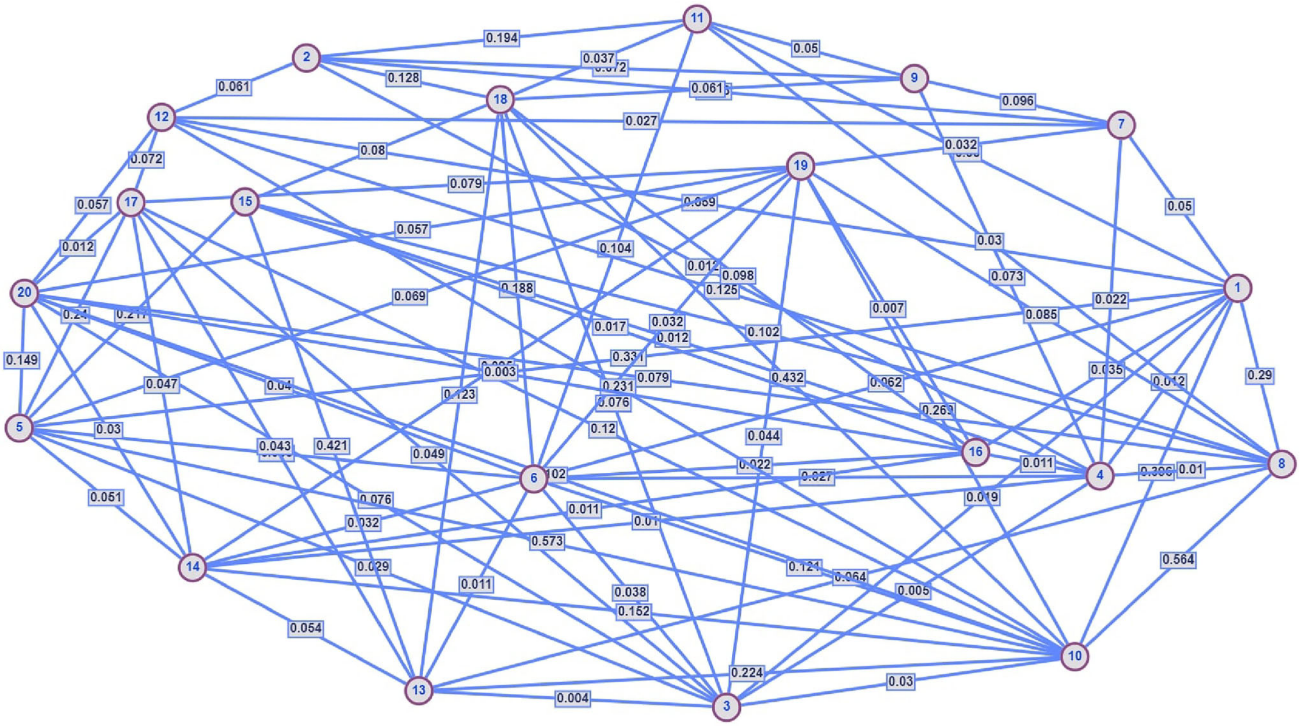

Fig. 5 Fuzzy social network (Facebook data)

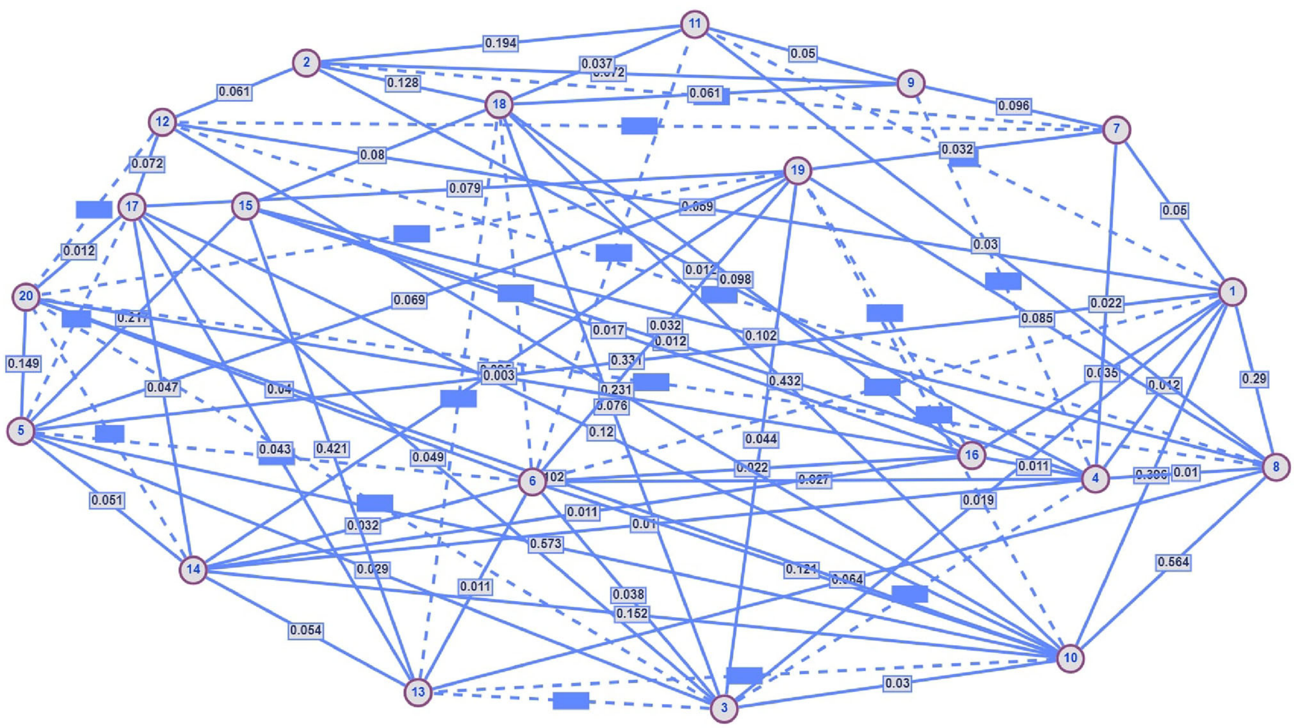

Fig. 6 Training graph (Facebook data)

Table 2 Facebook profile data

Sr. no	Names	Number of friends	Normalized value	Sr. no	Names	Number of friends	Normalized value
1	AS	2379	0.479	11	M2	1005	0.201
2	AB	5000	1	12	MU	4931	0.986
3	CA	770	0.154	13	PG	2159	0.431
4	CD	408	0.081	14	PN	1372	0.274
5	DJ	4463	0.892	15	RC	3659	0.731
6	DS	1085	0.217	16	S1	566	0.1132
7	DV	1076	0.215	17	SN	2432	0.486
8	K1	3168	0.633	18	S2	2158	0.431
9	K2	1536	0.307	19	TS	2142	0.428
10	M1	4988	0.997	20	UM	1004	0.201

Links in the network are randomly separated into two parts: probe set E_p and training set E_t. In this article, E_p has 21 edges, which are depicted in Fig. 6 (training graph) by dashed lines. Now, an algorithm can only use the information in the training graph to compute the score of the non-observed links (i.e. $U - E_t$).

With the help of the training graph, we can easily calculate the score of all node pairs that have common neighbours, but there are many node pairs in this data that do not have any common neighbour. In this case, we compute the maximum spanning tree (MST, Fig. 7) of the training graph to determine the strength of connectedness between individuals. We also compute the degree of individuals to determine their prominence.

Results and comparative analysis

We now compute the link prediction of all non-observed node pairs (See Table 5) using several similarity indices such as the SP index, RSM index, Common neighbour index, CAR index, Salton index, Jaccard index, Sorensen index, and Preferential attachment index. In these tables, node pairs that do not have common neighbours are highlighted. Table 5 shows the link prediction results using eight different similarity indices. The results that have been normalized are presented in boldface. When compared from the classical indices, it is clear that the SP index can only do the best discrimination of the link prediction score. The main reason for this is that most of these indices focus solely on the number of common neighbours and ignore the network's other topological properties, whereas the SP index can compute the score of link prediction of seed node pairs, regardless of whether common neighbours exist or not between them, and it also emphasizes the strength of the relations and the prominence of the individuals.

It has been observed that the SP index and RSM index have the highest Score for the same seed node pair (5,8), which contains four common neighbours and one LCL with strong relationships. The CAR index has the best score for a node pair (10,19) that contains six LCLs, but its strength is quite low. Common Neighbour, Jaccard, Salton and Sorensen indices have the highest score for the node pair (6,17) since it has six common neighbours, but its strength of relations with seed node pairs is not very strong. PA has the highest score for the node pair (4,10) since both nodes have the highest degree. However, when we sort these non-observed links according to their score obtain from various similarity indices, the SP index is the only one that can accurately discriminate all of these node pairs for link prediction.

Here, we present the accuracy findings for a number of similarity indices. Since it has been demonstrated that Salton, Jaccard, Sorensen, PA, and CN work better than other traditional similarity indices, we only select these indices, along with CAR and RSM, to compare with the SP index. Table 6 displays the link prediction results determined by AUC. According to this estimation, Salton and SP perform better on this data set than other similarity indices. We also provide a column SP/similarity indices, which displays the relative improvement over each other, for comparing SP with various similarity indices. If the value is positive, SP outperforms other similarity indexes, and vice versa. From the perspective of the improvement ratio, SP may boost each similarity index by more than ten per cent, with the exception of the Salton index. Here, we took into account a small network, but these percentages can be improved in a big network. Table 6 also shows the prediction results estimated based on precision; still, SP outperforms other similarity indexes. According to this estimation, SP can improve by several hundred percent the RSM similarity index. In both of these estimations, SP can outperform PA by more than forty per cent.

In Fig. 8, we also show the precision curve of these similarity indices. The precision curve is obtained in this case by utilizing ten alternative values of L, which are 5, 10, ..., 50 for all similarity indices. Nearly for all similarity indices

Table 3 Number of mutual friends between the concerned person

Names	AS (1)	AB (2)	CA (3)	CD (4)	DJ (5)	DS (6)	DV (7)	K1 (8)	K2 (9)	M1 (10)	M2 (11)	MU (12)	PG (13)	PN (14)	RC (15)	S1 (16)	SN (17)	S2 (18)	TS (19)	UM (20)
AS (1)	0	0	10	12	56	23	19	49	0	67	12	10	0	0	0	25	0	0	0	0
AB (2)	0	0	0	12	0	0	69	0	19	0	78	5	0	0	0	0	0	24	0	0
CA (3)	10	0	0	5	15	20	0	0	0	16	0	0	2	0	0	0	26	40	23	40
CD (4)	12	12	5	0	0	27	22	10	72	0	0	0	0	10	12	11	0	0	0	0
DJ (5)	56	0	15	0	0	13	0	0	0	52	0	0	0	15	24	0	40	0	13	60
DS (6)	23	0	20	27	13	0	0	0	0	45	42	0	4	12	0	16	0	70	12	16
DV (7)	19	69	0	22	0	0	0	0	36	0	0	10	0	0	0	0	0	0	12	0
K1 (8)	49	0	0	10	0	0	0	0	0	72	12	16	12	0	13	0	0	0	16	32
K2 (9)	0	19	0	72	0	0	36	0	0	0	20	0	0	0	0	0	0	16	0	0
M1 (10)	67	0	16	0	52	45	0	72	0	0	0	19	42	45	0	0	20	81	51	41
M2 (11)	12	78	0	0	0	42	0	12	20	0	0	0	0	0	0	0	0	15	0	0
MU (12)	10	5	0	0	0	0	10	16	0	19	0	0	0	0	0	0	12	0	0	23
PG (13)	0	0	2	0	0	4	0	12	0	42	0	0	0	16	79	0	8	23	0	0
PN (14)	0	0	0	10	15	12	0	0	0	45	0	0	16	0	0	8	14	0	28	12
RC (15)	0	0	0	12	24	0	0	13	0	0	0	0	79	0	0	12	0	15	0	0
S1 (16)	25	0	0	11	0	16	0	0	0	0	0	0	0	8	12	0	0	70	5	2
SN (17)	0	0	26	0	40	0	0	0	0	20	0	12	8	14	0	0	0	0	15	5
S2 (18)	0	24	40	0	0	70	0	0	16	81	15	0	23	0	15	70	0	0	0	0
TS (19)	0	0	23	0	13	12	12	16	0	51	0	0	0	28	0	5	15	0	0	23
UM (20)	0	0	40	0	60	16	0	32	0	41	0	23	0	12	0	2	5	0	23	0

Table 4 Edge membership value between individuals in Facebook Network

Names	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0.019	0.012	0.331	0.062	0.05	0.29	0	0.396	0.03	0.059	0	0	0	0.035	0	0	0	0
2	0	0	0	0.012	0	0	0.183	0	0.072	0	0.194	0.061	0	0	0	0	0	0.128	0	0
3	0.019	0	0	0.005	0.029	0.038	0	0	0	0.03	0	0	0.004	0	0	0	0.049	0.076	0.044	0.076
4	0.012	0.012	0.005	0	0	0.027	0.022	0.01	0.073	0	0	0	0	0.01	0.012	0.011	0	0	0	0
5	0.331	0	0.029	0	0	0.035	0	0	0	0.573	0	0	0	0.051	0.217	0	0.24	0	0.069	0.149
6	0.062	0	0.038	0.027	0.035	0	0	0	0	0.121	0.104	0	0.011	0.032	0	0.022	0	0.188	0.032	0.04
7	0.05	0.183	0	0.022	0	0	0	0	0.096	0	0	0.027	0	0	0	0	0	0	0.032	0
8	0.29	0	0	0.01	0	0	0	0	0	0.564	0.03	0.125	0.064	0	0.102	0	0	0	0.085	0.079
9	0	0.072	0	0.073	0	0	0.096	0	0	0	0.05	0	0	0	0	0	0	0.061	0	0
10	0.396	0	0.03	0	0.573	0.121	0	0.564	0	0	0	0.231	0.224	0.152	0	0	0.12	0.432	0.269	0.102
11	0.03	0.194	0	0	0	0.104	0	0.03	0.05	0	0	0	0	0	0	0	0	0.037	0	0
12	0.059	0.061	0	0	0	0	0.027	0.125	0	0.231	0	0	0	0	0	0	0.072	0	0	0.057
13	0	0	0.004	0	0	0.011	0	0.064	0	0.224	0	0	0	0.054	0.421	0	0.043	0.123	0	0
14	0	0	0	0.01	0.051	0.032	0	0	0	0.152	0	0	0.054	0	0	0.011	0.047	0	0.095	0.03
15	0	0	0	0.012	0.217	0	0	0.102	0	0	0	0	0.421	0	0	0.017	0	0.08	0	0
16	0.035	0	0	0.011	0	0.022	0	0	0	0	0	0	0	0.011	0.017	0	0	0.098	0.007	0.003
17	0	0	0.049	0	0.24	0	0	0	0	0.12	0	0.072	0.043	0.047	0	0	0	0	0.079	0.012
18	0	0.128	0.076	0	0	0.188	0	0	0.061	0.432	0.037	0	0.123	0	0.08	0.098	0	0	0	0
19	0	0	0.044	0	0.069	0.032	0.032	0.085	0	0.269	0	0	0	0.095	0	0.007	0.079	0	0	0.057
20	0	0	0.076	0	0.149	0.04	0	0.079	0	0.102	0	0.057	0	0.03	0	0.003	0.012	0	0.057	0

Table 5 Score of SP index along with other similarity indices

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods									
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX		
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Common neighbor index	Number of local community link	CAR index	Normalized CAR index
(1,2)	0.146	**0.112**	0.012	0.012	**0.047**	1	**0.167**	0	0.00	**0**
(1,6)	0.58	**0.447**	0.174	0.044	**0.172**	4	**0.667**	2	8.00	**0.222**
(1,9)	0.233	**0.18**	0.05	0.05	**0.195**	1	**0.167**	0	0.00	**0**
(1,11)	0.349	**0.269**	0.03	0.03	**0.117**	1	**0.167**	0	0.00	**0**
(1,13)	0.148	**0.114**	0.064	0.064	**0.25**	1	**0.167**	0	0.00	**0**
(1,14)	0.9	**0.693**	0.224	0.056	**0.219**	4	**0.667**	2	8.00	**0.222**
(1,15)	0.43	**0.331**	0.348	0.087	**0.34**	4	**0.667**	2	8.00	**0.222**
(1,17)	0.407	**0.314**	0.198	0.066	**0.258**	3	**0.5**	2	6.00	**0.167**
(1,18)	0.839	**0.646**	0.45	0.15	**0.586**	3	**0.5**	1	3.00	**0.083**
(1,19)	0.356	**0.274**	0.205	0.051	**0.199**	4	**0.667**	1	4.00	**0.111**
(1,20)	1.135	**0.874**	0.254	0.085	**0.332**	3	**0.5**	1	3.00	**0.083**
(2,3)	0.494	**0.381**	0.076	0.076	**0.297**	1	**0.167**	0	0.00	**0**
(2,5)	***0.566***	***0.436***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(2,6)	0.146	**0.112**	0.012	0.012	**0.047**	1	**0.167**	0	0.00	**0**
(2,7)	0.283	**0.218**	0.084	0.042	**0.164**	2	**0.333**	0	0.00	**0**
(2,8)	0.141	**0.109**	0.04	0.02	**0.078**	2	**0.333**	0	0.00	**0**
(2,10)	0.296	**0.228**	0.128	0.128	**0.5**	1	**0.167**	0	0.00	**0**
(2,12)	***0.405***	***0.312***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(2,13)	***0.459***	***0.354***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(2,14)	0.122	**0.094**	0.01	0.01	**0.039**	1	**0.167**	0	0.00	**0**
(2,15)	0.179	**0.138**	0.092	0.046	**0.18**	2	**0.333**	0	0.00	**0**
(2,16)	0.559	**0.431**	0.109	0.055	**0.215**	2	**0.333**	0	0.00	**0**
(2,17)	***0.413***	***0.318***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(2,19)	***0.391***	***0.301***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(2,20)	***0.39***	***0.3***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods									
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX		
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Common neighbor index	Number of local community link	CAR index	Normalized CAR index
(3,4)	0.341	**0.263**	0.039	0.02	**0.078**	2	**0.333**	0	0.00	**0**
(3,7)	0.166	**0.128**	0.051	0.026	**0.102**	2	**0.333**	0	0.00	**0**
(3,8)	0.602	**0.464**	0.093	0.031	**0.121**	3	**0.5**	1	3.00	**0.083**
(3,9)	0.396	**0.305**	0.061	0.061	**0.238**	1	**0.167**	0	0.00	**0**
(3,11)	0.24	**0.185**	0.037	0.037	**0.145**	1	**0.167**	0	0.00	**0**
(3,12)	0.565	**0.435**	0.098	0.033	**0.129**	3	**0.5**	2	6.00	**0.167**
(3,13)	0.246	**0.19**	0.054	0.027	**0.105**	2	**0.333**	0	0.00	**0**
(3,14)	0.608	**0.468**	0.182	0.036	**0.141**	5	**0.833**	6	30.00	**0.833**
(3,15)	0.341	**0.263**	0.105	0.053	**0.207**	2	**0.333**	0	0.00	**0**
(3,16)	0.345	**0.266**	0.117	0.039	**0.152**	3	**0.5**	0	0.00	**0**
(3,20)	0.697	**0.537**	0.109	0.027	**0.105**	4	**0.667**	3	12.00	**0.333**
(4,5)	0.138	**0.106**	0.034	0.011	**0.043**	3	**0.5**	0	0.00	**0**
(4,9)	0.388	**0.299**	0.034	0.017	**0.066**	2	**0.333**	0	0.00	**0**
(4,10)	0.643	**0.495**	0.059	0.015	**0.059**	4	**0.667**	0	0.00	**0**
(4,11)	0.134	**0.103**	0.022	0.011	**0.043**	2	**0.333**	0	0.00	**0**
(4,12)	0.146	**0.112**	0.012	0.012	**0.047**	1	**0.167**	0	0.00	**0**
(4,13)	0.278	**0.214**	0.043	0.011	**0.043**	4	**0.667**	2	8.00	**0.222**
(4,17)	0.122	**0.094**	0.01	0.01	**0.039**	1	**0.167**	0	0.00	**0**
(4,18)	0.292	**0.225**	0.035	0.012	**0.047**	3	**0.5**	1	3.00	**0.083**
(4,19)	0.357	**0.275**	0.069	0.017	**0.066**	4	**0.667**	3	12.00	**0.333**
(4,20)	0.378	**0.291**	0.03	0.015	**0.059**	2	**0.333**	1	2.00	**0.056**
(5,6)	0.653	**0.503**	0.254	0.051	**0.199**	5	**0.833**	5	25.00	**0.694**
(5,7)	0.191	**0.147**	0.082	0.041	**0.16**	2	**0.333**	0	0.00	**0**
(5,8)	1.298	**1**	1.025	0.256	**1**	4	**0.667**	1	4.00	**0.111**
(5,9)	***0.448***	***0.345***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(5,11)	***0.562***	***0.433***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(5,12)	0.494	**0.381**	0.29	0.145	**0.566**	2	**0.333**	1	2.00	**0.056**

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods									
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX		
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Com-mon neigh-bor index	Number of local community link	CAR index	Normalized CAR index
(5,13)	0.38	**0.293**	0.268	0.134	**0.523**	2	**0.333**	0	0.00	**0**
(5,16)	0.146	**0.112**	0.066	0.017	**0.066**	4	**0.667**	0	0.00	**0**
(5,17)	0.643	**0.495**	0.277	0.055	**0.215**	5	**0.833**	5	25.00	**0.694**
(5,18)	0.632	**0.487**	0.541	0.18	**0.703**	3	**0.5**	1	3.00	**0.083**
(6,7)	0.182	**0.14**	0.054	0.027	**0.105**	2	**0.333**	0	0.00	**0**
(6,8)	0.237	**0.183**	0.174	0.044	**0.172**	4	**0.667**	0	0.00	**0**
(6,9)	***0.303***	***0.233***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(6,11)	***0.462***	***0.356***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(6,12)	0.558	**0.43**	0.121	0.121	**0.473**	1	**0.167**	0	0.00	**0**
(6,15)	0.231	**0.178**	0.04	0.013	**0.051**	3	**0.5**	1	3.00	**0.083**
(6,17)	0.586	**0.451**	0.245	0.041	**0.16**	6	**1**	4	24.00	**0.667**
(6,18)	0.569	**0.438**	0.181	0.06	**0.234**	3	**0.5**	1	3.00	**0.083**
(7,8)	0.326	**0.251**	0.092	0.031	**0.121**	3	**0.5**	1	3.00	**0.083**
(7,10)	0.233	**0.18**	0.05	0.05	**0.195**	1	**0.167**	0	0.00	**0**
(7,11)	0.249	**0.192**	0.05	0.05	**0.195**	1	**0.167**	0	0.00	**0**
(7,12)	0.515	**0.397**	0.05	0.05	**0.195**	1	**0.167**	0	0.00	**0**
(7,13)	***0.341***	***0.263***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(7,14)	0.141	**0.109**	0.042	0.021	**0.082**	2	**0.333**	0	0.00	**0**
(7,15)	0.146	**0.112**	0.012	0.012	**0.047**	1	**0.167**	0	0.00	**0**
(7,16)	0.382	**0.294**	0.046	0.023	**0.09**	2	**0.333**	1	2.00	**0.056**
(7,17)	0.149	**0.115**	0.032	0.032	**0.125**	1	**0.167**	0	0.00	**0**
(7,18)	0.284	**0.219**	0.061	0.061	**0.238**	1	**0.167**	1	1.00	**0.028**
(7,20)	***0.272***	***0.21***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(8,12)	1.088	**0.838**	0.29	0.145	**0.566**	2	**0.333**	0	0.00	**0**

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods									
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX		
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Common neighbor index	Number of local community link	CAR index	Normalized CAR index
(8,9)	0.149	**0.115**	0.03	0.03	**0.117**	1	**0.167**	0	0.00	**0**
(8,14)	0.333	**0.257**	0.301	0.075	**0.293**	4	**0.667**	4	16.00	**0.444**
(8,16)	0.347	**0.267**	0.062	0.021	**0.082**	3	**0.5**	2	6.00	**0.167**
(8,17)	0.18	**0.139**	0.242	0.081	**0.316**	3	**0.5**	0	0.00	**0**
(8,18)	0.509	**0.392**	0.542	0.181	**0.707**	3	**0.5**	0	0.00	**0**
(8,20)	0.161	**0.124**	0.102	0.102	**0.398**	1	**0.167**	0	0.00	**0**
(9,10)	0.199	**0.153**	0.061	0.061	**0.238**	1	**0.167**	0	0.00	**0**
(9,12)	***0.314***	***0.242***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(9,13)	***0.362***	***0.279***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(9,14)	***0.337***	***0.26***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(9,15)	0.199	**0.153**	0.061	0.061	**0.238**	1	**0.167**	0	0.00	**0**
(9,16)	0.54	**0.416**	0.061	0.061	**0.238**	1	**0.167**	0	0.00	**0**
(9,17)	***0.329***	***0.253***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(9,19)	0.149	**0.115**	0.032	0.032	**0.125**	1	**0.167**	0	0.00	**0**
(9,20)	***0.298***	***0.23***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(10,11)	0.167	**0.129**	0.067	0.034	**0.133**	2	**0.333**	0	0.00	**0**
(10,13)	0.368	**0.284**	0.172	0.043	**0.168**	4	**0.667**	2	8.00	**0.222**
(10,15)	0.222	**0.171**	0.399	0.133	**0.52**	3	**0.5**	0	0.00	**0**
(10,16)	0.471	**0.363**	0.169	0.034	**0.133**	5	**0.833**	4	20.00	**0.556**
(10,19)	0.501	**0.386**	0.39	0.065	**0.254**	6	**1**	6	36.00	**1**
(11,12)	***0.385***	***0.297***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(11,13)	0.149	**0.115**	0.03	0.03	**0.117**	1	**0.167**	0	0.00	**0**
(11,14)	***0.402***	***0.31***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(11,15)	0.167	**0.129**	0.067	0.034	**0.133**	2	**0.333**	0	0.00	**0**
(11,16)	0.327	**0.252**	0.037	0.037	**0.145**	1	**0.167**	0	0.00	**0**
(11,17)	***0.387***	***0.298***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods									
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX		
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Common neighbor index	Number of local community link	CAR index	Normalized CAR index
(11,19)	0.149	**0.115**	0.03	0.03	**0.117**	1	**0.167**	0	0.00	**0**
(11,20)	***0.364***	***0.28***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(12,13)	0.1	**0.077**	0.043	0.043	**0.168**	1	**0.167**	0	0.00	**0**
(12,14)	0.61	**0.47**	0.199	0.1	**0.391**	2	**0.333**	1	2.00	**0.056**
(12,15)	***0.579***	***0.446***	***0***		***0***	***0***	***0***	***0***	***0.00***	***0***
(12,16)	0.31	**0.239**	0.035	0.035	**0.137**	1	**0.167**	0	0.00	**0**
(12,18)	0.535	**0.412**	0.231	0.231	**0.902**	1	**0.167**	0	0.00	**0**
(12,19)	0.168	**0.129**	0.072	0.072	**0.281**	1	**0.167**	0	0.00	**0**
(12,20)	0.531	**0.409**	0.114	0.057	**0.223**	2	**0.333**	1	2.00	**0.056**
(13,16)	0.262	**0.202**	0.039	0.013	**0.051**	3	**0.5**	1	3.00	**0.083**
(13,18)	0.185	**0.143**	0.08	0.08	**0.313**	1	**0.167**	0	0.00	**0**
(13,19)	0.289	**0.223**	0.172	0.043	**0.168**	4	**0.667**	2	8.00	**0.222**
(13,20)	0.057	**0.044**	0.023	0.012	**0.047**	2	**0.333**	0	0.00	**0**
(14,15)	0.304	**0.234**	0.126	0.032	**0.125**	4	**0.667**	1	4.00	**0.111**
(14,18)	0.421	**0.324**	0.163	0.082	**0.32**	2	**0.333**	0	0.00	**0**
(14,20)	0.644	**0.496**	0.2	0.04	**0.156**	5	**0.833**	4	20.00	**0.556**
(15,17)	0.1	**0.077**	0.043	0.043	**0.168**	1	**0.167**	0	0.00	**0**
(15,19)	0.18	**0.139**	0.154	0.077	**0.301**	2	**0.333**	0	0.00	**0**
(15,20)	0.484	**0.373**	0.152	0.076	**0.297**	2	**0.333**	0	0.00	**0**
(16,17)	0.062	**0.048**	0.014	0.007	**0.027**	2	**0.333**	0	0.00	**0**
(16,19)	0.344	**0.265**	0.033	0.017	**0.066**	2	**0.333**	1	2.00	**0.056**
(17,18)	0.671	**0.517**	0.169	0.085	**0.332**	2	**0.333**	1	2.00	**0.056**
(18,19)	0.286	**0.22**	0.044	0.044	**0.172**	1	**0.167**	0	0.00	**0**
(18,20)	0.334	**0.257**	0.105	0.053	**0.207**	2	**0.333**	0	0.00	**0**
(19,20)	0.187	**0.144**	0.113	0.038	**0.148**	3	**0.5**	0	0.00	**0**

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods										
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX			
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Com-mon neigh-bor index	Number of local community link	CAR index	Normalized CAR index	
(1,2)	8	4	0.177	**0.221**	11	0.091	**0.136**	0.167	**0.209**	32	**0.4**
(1,6)	8	8	0.5	**0.623**	12	0.333	**0.497**	0.5	**0.625**	64	**0.8**
(1,9)	8	4	0.177	**0.221**	11	0.091	**0.136**	0.167	**0.209**	32	**0.4**
(1,11)	8	4	0.177	**0.221**	11	0.091	**0.136**	0.167	**0.209**	32	**0.4**
(1,13)	8	5	0.158	**0.197**	12	0.083	**0.124**	0.154	**0.193**	40	**0.5**
(1,14)	8	8	0.5	**0.623**	12	0.333	**0.497**	0.5	**0.625**	64	**0.8**
(1,15)	8	6	0.577	**0.719**	10	0.4	**0.597**	0.571	**0.714**	48	**0.6**
(1,17)	8	7	0.401	**0.5**	12	0.25	**0.373**	0.4	**0.5**	56	**0.7**
(1,18)	8	7	0.401	**0.5**	12	0.25	**0.373**	0.4	**0.5**	56	**0.7**
(1,19)	8	7	0.535	**0.667**	11	0.364	**0.543**	0.533	**0.666**	56	**0.7**
(1,20)	8	5	0.474	**0.591**	10	0.3	**0.448**	0.462	**0.578**	40	**0.5**
(2,3)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**
(2,5)	***4***	***7***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	*0.35*
(2,6)	4	8	0.177	**0.221**	11	0.091	**0.136**	0.167	**0.209**	32	**0.4**
(2,7)	4	4	0.5	**0.623**	6	0.333	**0.497**	0.5	**0.625**	16	**0.2**
(2,8)	4	7	0.378	**0.471**	9	0.222	**0.331**	0.364	**0.455**	28	**0.35**

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods										
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX			
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Common neighbor index	Number of local community link	CAR index	Normalized CAR index	
(2,10)	4	10	0.158	**0.197**	13	0.077	**0.115**	0.143	**0.179**	40	**0.5**
(2,12)	***4***	***3***	***0***	***0***	***7***	***0***	***0***	***0***	***0***	***12***	***0.15***
(2,13)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***
(2,14)	4	8	0.177	**0.221**	11	0.091	**0.136**	0.167	**0.209**	32	**0.4**
(2,15)	4	6	0.408	**0.509**	8	0.25	**0.373**	0.4	**0.5**	24	**0.3**
(2,16)	4	7	0.378	**0.471**	9	0.222	**0.331**	0.364	**0.455**	28	**0.35**
(2,17)	***4***	***7***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	***0.35***
(2,19)	***4***	***7***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	***0.35***
(2,20)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***
(3,4)	7	4	0.378	**0.471**	9	0.222	**0.331**	0.364	**0.455**	28	**0.35**
(3,7)	***7***	4	0.378	**0.471**	9	0.222	**0.331**	0.364	**0.455**	28	**0.35**
(3,8)	***7***	7	0.429	**0.535**	11	0.273	**0.407**	0.429	**0.536**	49	**0.613**
(3,9)	***7***	4	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**
(3,11)	***7***	4	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**
(3,12)	***7***	3	0.655	**0.817**	7	0.429	**0.64**	0.6	**0.75**	21	**0.263**
(3,13)	***7***	5	0.338	**0.421**	10	0.2	**0.299**	0.333	**0.416**	35	**0.438**
(3,14)	***7***	8	0.668	**0.833**	10	0.5	**0.746**	0.667	**0.834**	56	**0.7**
(3,15)	***7***	6	0.309	**0.385**	11	0.182	**0.272**	0.308	**0.385**	42	**0.525**
(3,16)	***7***	7	0.429	**0.535**	11	0.273	**0.407**	0.429	**0.536**	49	**0.613**
(3,20)	***7***	5	0.676	**0.843**	8	0.5	**0.746**	0.667	**0.834**	35	**0.438**
(4,5)	8	7	0.401	**0.5**	12	0.25	**0.373**	0.4	**0.5**	56	**0.7**
(4,9)	8	4	0.354	**0.441**	10	0.2	**0.299**	0.333	**0.416**	32	**0.4**
(4,10)	8	10	0.447	**0.557**	14	0.286	**0.427**	0.444	**0.555**	80	**1**
(4,11)	8	4	0.354	**0.441**	10	0.2	**0.299**	0.333	**0.416**	32	**0.4**
(4,12)	8	3	0.204	**0.254**	10	0.1	**0.149**	0.182	**0.228**	24	**0.3**
(4,13)	8	5	0.632	**0.788**	9	0.444	**0.663**	0.615	**0.769**	40	**0.5**
(4,17)	8	7	0.134	**0.167**	14	0.071	**0.106**	0.133	**0.166**	56	**0.7**
(4,18)	8	7	0.401	**0.5**	12	0.25	**0.373**	0.4	**0.5**	56	**0.7**
(4,19)	8	7	0.535	**0.667**	11	0.364	**0.543**	0.533	**0.666**	56	**0.7**

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods										
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX			
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Common neighbor index	Number of local community link	CAR index	Normalized CAR index	
(4,20)	8	5	0.316	**0.394**	11	0.182	**0.272**	0.308	**0.385**	40	**0.5**
(5,6)	7	8	0.668	**0.833**	10	0.5	**0.746**	0.667	**0.834**	56	**0.7**
(5,7)	7	4	0.378	**0.471**	9	0.222	**0.331**	0.364	**0.455**	28	**0.35**
(5,8)	7	7	0.571	**0.712**	10	0.4	**0.597**	0.571	**0.714**	49	**0.613**
(5,9)	***7***	***4***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	***0.35***
(5,11)	***8***	***4***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	***0.35***
(5,12)	7	3	0.436	**0.544**	8	0.25	**0.373**	0.4	**0.5**	21	**0.263**
(5,13)	7	5	0.338	**0.421**	10	0.2	**0.299**	0.333	**0.416**	35	**0.438**
(5,16)	7	7	0.571	**0.712**	10	0.4	**0.597**	0.571	**0.714**	49	**0.613**
(5,17)	7	7	0.714	**0.89**	9	0.556	**0.83**	0.714	**0.893**	49	**0.613**
(5,18)	7	7	0.429	**0.535**	11	0.273	**0.407**	0.429	**0.536**	49	**0.613**
(6,7)	8	4	0.354	**0.441**	10	0.2	**0.299**	0.333	**0.416**	32	**0.4**
(6,8)	8	7	0.535	**0.667**	11	0.364	**0.543**	0.533	**0.666**	56	**0.7**
(6,9)	***8***	***4***	***0***	***0***	***12***	***0***	***0***	***0***	***0***	***32***	***0.4***
(6,11)	***8***	***4***	***0***	***0***	***12***	***0***	***0***	***0***	***0***	***32***	***0.4***
(6,12)	8	3	0.204	**0.254**	10	0.1	**0.149**	0.182	**0.228**	24	**0.3**
(6,15)	8	6	0.433	**0.54**	11	0.273	**0.407**	0.429	**0.536**	48	**0.6**
(6,17)	8	7	0.802	**1**	9	0.667	**0.996**	0.8	**1**	56	**0.7**
(6,18)	8	7	0.401	**0.5**	12	0.25	**0.373**	0.4	**0.5**	56	**0.7**
(7,8)	4	7	0.567	**0.707**	8	0.375	**0.56**	0.545	**0.681**	28	**0.35**
(7,10)	4	10	0.158	**0.197**	13	0.077	**0.115**	0.143	**0.179**	40	**0.5**
(7,11)	4	4	0.25	**0.312**	7	0.143	**0.213**	0.25	**0.313**	16	**0.2**
(7,12)	4	3	0.289	**0.36**	6	0.167	**0.249**	0.286	**0.358**	12	**0.15**
(7,13)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***
(7,14)	4	8	0.354	**0.441**	10	0.2	**0.299**	0.333	**0.416**	32	**0.4**
(7,15)	4	6	0.204	**0.254**	9	0.111	**0.166**	0.2	**0.25**	24	**0.3**
(7,16)	4	7	0.378	**0.471**	9	0.222	**0.331**	0.364	**0.455**	28	**0.35**
(7,17)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods											
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX				
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Com-mon neigh-bor index	Number of local community link	CAR index	Normalized CAR index		
(7,18)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**	
(7,20)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***	
(8,12)	7	3	0.436	**0.544**	8	0.25	**0.373**	0.4	**0.5**	21	**0.263**	
(8,9)	7	4	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**	
(8,14)	7	8	0.535	**0.667**	11	0.364	**0.543**	0.533	**0.666**	56	**0.7**	
(8,16)	7	7	0.429	**0.535**	11	0.273	**0.407**	0.429	**0.536**	49	**0.613**	
(8,17)	7	7	0.429	**0.535**	11	0.273	**0.407**	0.429	**0.536**	49	**0.613**	
(8,18)	7	7	0.429	**0.535**	11	0.273	**0.407**	0.429	**0.536**	49	**0.613**	
(8,20)	7	5	0.169	**0.211**	11	0.091	**0.136**	0.167	**0.209**	35	**0.438**	
(9,10)	4	10	0.158	**0.197**	13	0.077	**0.115**	0.143	**0.179**	40	**0.5**	
(9,12)	***4***	***3***	***0***	***0***	***7***	***0***	***0***	***0***	***0***	***12***	***0.15***	
(9,13)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***	
(9,14)	***4***	***8***	***0***	***0***	***12***	***0***	***0***	***0***	***0***	***32***	***0.4***	
(9,15)	4	6	0.204	**0.254**	9	0.111	**0.166**	0.2	**0.25**	24	**0.3**	
(9,16)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**	
(9,17)	***4***	***7***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	***0.35***	
(9,19)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**	
(9,20)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***	
(10,11)	10	4	0.316	**0.394**	12	0.167	**0.249**	0.286	**0.358**	40	**0.5**	
(10,13)	10	5	0.566	**0.706**	11	0.364	**0.543**	0.533	**0.666**	50	**0.625**	
(10,15)	10	6	0.387	**0.483**	13	0.231	**0.345**	0.375	**0.469**	60	**0.75**	
(10,16)	10	7	0.598	**0.746**	12	0.417	**0.622**	0.588	**0.735**	70	**0.875**	
(10,19)	10	7	0.717	**0.894**	11	0.545	**0.813**	0.706	**0.883**	70	**0.875**	
(11,12)	***4***	***3***	***0***	***0***	***7***	***0***	***0***	***0***	***0***	***12***	***0.15***	
(11,13)	4	5	0.224	**0.279**	8	0.125	**0.187**	0.222	**0.278**	20	**0.25**	
(11,14)	***4***	***8***	***0***	***0***	***12***	***0***	***0***	***0***	***0***	***32***	***0.4***	
(11,15)	4	6	0.408	**0.509**	8	0.25	**0.373**	0.4	**0.5**	24	**0.3**	
(11,16)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**	
(11,17)	***4***	***7***	***0***	***0***	***11***	***0***	***0***	***0***	***0***	***28***	***0.35***	

Table 5 continued

Non observed node pairs (x,y) for link prediction	Link prediction in verious methods											
	SP INDEX		RSM INDEX			COMMON NEIGHBOR INDEX		CAR INDEX				
	SP index	Normalized SP index	Nature of common neighbours	RSM index	Normalized RSM index	Common neighbours	Normalized Com-mon neigh-bor index	Number of local community link	CAR index	Normalized CAR index		
(11,19)	4	7	0.189	**0.236**	10	0.1	**0.149**	0.182	**0.228**	28	**0.35**
(11,20)	***4***	***5***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***20***	***0.25***
(12,13)	3	5	0.258	**0.322**	7	0.143	**0.213**	0.25	**0.313**	15	**0.188**
(12,14)	3	8	0.408	**0.509**	9	0.222	**0.331**	0.364	**0.455**	24	**0.3**
(12,15)	***3***	***6***	***0***	***0***	***9***	***0***	***0***	***0***	***0***	***18***	***0.225***
(12,16)	3	7	0.218	**0.272**	9	0.111	**0.166**	0.2	**0.25**	21	**0.263**
(12,18)	3	7	0.218	**0.272**	9	0.111	**0.166**	0.2	**0.25**	21	**0.263**
(12,19)	3	7	0.218	**0.272**	9	0.111	**0.166**	0.2	**0.25**	21	**0.263**
(12,20)	3	5	0.516	**0.643**	6	0.333	**0.497**	0.5	**0.625**	15	**0.188**
(13,16)	5	7	0.507	**0.632**	9	0.333	**0.497**	0.5	**0.625**	35	**0.438**
(13,18)	5	7	0.169	**0.211**	11	0.091	**0.136**	0.167	**0.209**	35	**0.438**
(13,19)	5	7	0.676	**0.843**	8	0.5	**0.746**	0.667	**0.834**	35	**0.438**
(13,20)	5	5	0.4	**0.499**	8	0.25	**0.373**	0.4	**0.5**	25	**0.313**
(14,15)	8	6	0.577	**0.719**	10	0.4	**0.597**	0.571	**0.714**	48	**0.6**
(14,18)	8	7	0.267	**0.333**	13	0.154	**0.23**	0.267	**0.334**	56	**0.7**
(14,20)	8	5	0.791	**0.986**	8	0.625	**0.933**	0.769	**0.961**	40	**0.5**
(15,17)	6	7	0.154	**0.192**	12	0.083	**0.124**	0.154	**0.193**	42	**0.525**
(15,19)	6	7	0.309	**0.385**	11	0.182	**0.272**	0.308	**0.385**	42	**0.525**
(15,20)	6	5	0.365	**0.455**	9	0.222	**0.331**	0.364	**0.455**	30	**0.375**
(16,17)	7	7	0.286	**0.357**	12	0.167	**0.249**	0.286	**0.358**	49	**0.613**
(16,19)	7	7	0.286	**0.357**	12	0.167	**0.249**	0.286	**0.358**	49	**0.613**
(17,18)	7	7	0.286	**0.357**	12	0.167	**0.249**	0.286	**0.358**	49	**0.613**
(18,19)	7	7	0.143	**0.178**	13	0.077	**0.115**	0.143	**0.179**	49	**0.613**
(18,20)	7	5	0.338	**0.421**	10	0.2	**0.299**	0.333	**0.416**	35	**0.438**
(19,20)	7	5	0.507	**0.632**	9	0.333	**0.497**	0.5	**0.625**	35	**0.438**

Pair of nodes without common neighbours are highlighted in bold and italic

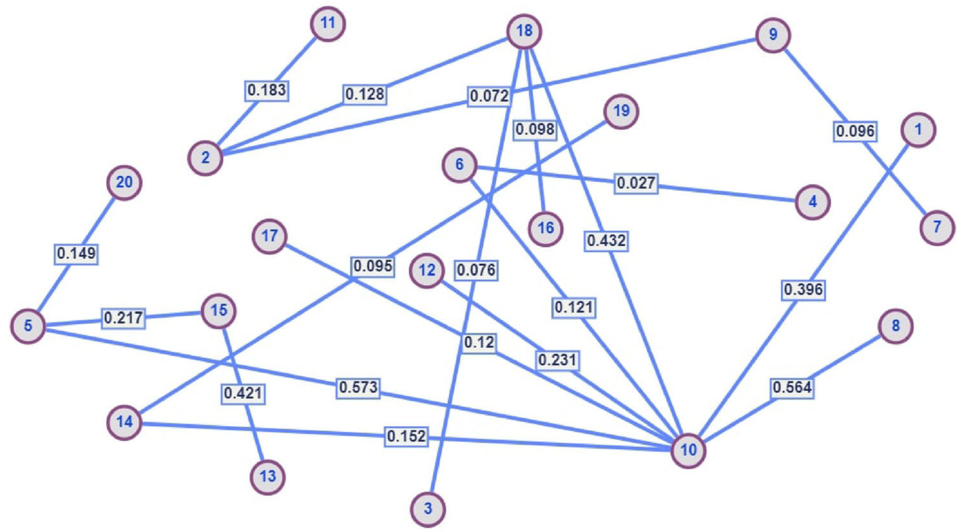

Fig. 7 Maximal spanning tree of the training graph

Table 6 Link prediction accuracy estimated by AUC and Precision

Estimations	AUC	SP/similariy indices	Precision	SP/similariy indices
Salton	0.689	0	0.286	+ 33.22%
Jaccard	0.625	+ 10.24%	0.286	+ 33.22%
Sorensen	0.625	+ 10.24%	0.286	+ 33.22%
PA	0.491	+ 40.33%	0.238	+ 60.08%
CN	0.601	+ 14.64%	0.333	+ 14.41%
RSM	0.597	+ 15.41%	0.143	+ 166.43%
CAR	0.624	+ 10.42%	0.333	+ 14.41%
SP	0.689		0.381	

have decreasing precision as L increases, with the exception of the PA index, which has fairly steady prediction outcomes despite L change. Overall, it supports the idea that raising L makes the prediction problem more complex. Figure 8 shows that the SP index performs consistently and better than other similarity indices. It suggests that the SP index is a solid supplement to existing similarity indices of link prediction when dealing with a typical network with strength of relationships and no common neighbours.

Area of applications

Link prediction is an essential method for forecasting future connections in complex networked systems. The literature presents various similarity indices for link prediction, each with distinct methodologies and applications. The Strength Prominence Similarity Index (SP index) is a novel approach in this domain, incorporating both the strengths and limitations of relationships. Unlike conventional indices that rely on common neighbours, the SP index ensures a consistent scoring mechanism even when seed node pairs lack common neighbours by considering the strength of path connectivity and their prominence degrees. Through the analysis of existing structures and relationships, link prediction (SP index) facilitates informed decision-making across multiple disciplines.

- **Social Network Analysis:** In social media platforms like Facebook, Twitter, and LinkedIn, link prediction is widely used to recommend new friends or professional connections. The algorithm suggests connections based on mutual friends, shared interests, or interaction history. Additionally, it helps in detecting influential users and analysing community structures, which is beneficial for marketing and user engagement strategies.
- **Biological and Medical Networks:** In bioinformatics and healthcare, link prediction is used to forecast interactions between proteins, which helps in understanding biological functions and disease mechanisms. Similarly, in drug discovery, it assists in identifying potential drug-target interactions, which speeds up the development of new treatments and minimizes experimental costs.
- **Cybersecurity and Fraud Detection:** In cybersecurity, link prediction helps in identifying suspicious activities, such as unusual patterns of communication that may indi-

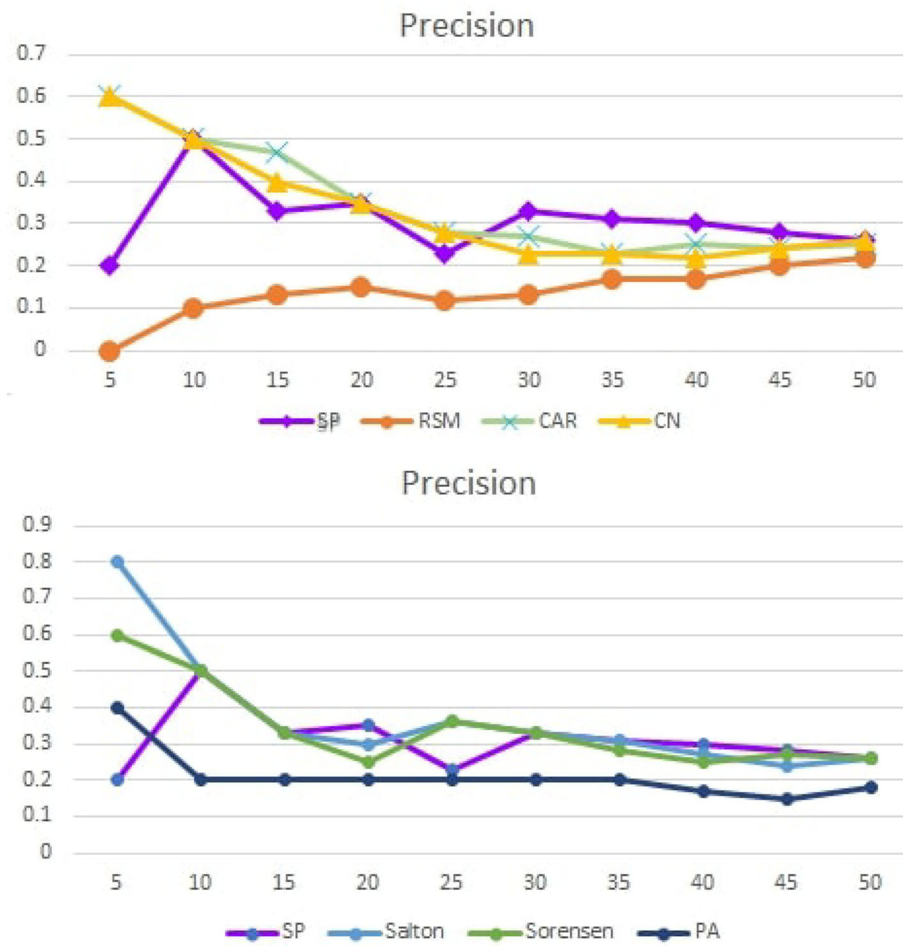

Fig. 8 Precision curve for similarity indices on different values of L

cate a cyber threat. It is also used in financial networks to detect fraudulent transactions by predicting abnormal links in banking and credit card networks, thereby preventing financial crimes.

- **Financial and Business Networks:** Link prediction is useful in analysing business partnerships by identifying potential collaborations based on past trends and market behaviour. In the financial sector, it helps assess credit risk by predicting links between borrowers and financial institutions, ensuring better decision-making in loan approvals and investment planning.
- **Scientific Collaboration and Citation Networks:** In academia, link prediction helps in predicting future collaborations between researchers based on their co-authorship history and shared research interests. It also assists in recommending relevant research papers by analysing citation patterns, helping researchers stay updated with important advancements in their field.
- **Transportation and Communication Networks:** In transportation systems, link prediction is used to improve route optimization and traffic flow by analysing movement patterns. In communication networks, such as the internet and telecommunication systems, it helps predict future links between devices or nodes, ensuring better connectivity and efficiency in data transfer.

Conclusion

Due to the present vast amount of network statistics in electronic form, various fields such as computer science, sociology, information science, and anthropology have recently given close attention to the topic of link prediction. All prior link prediction approaches were developed based on the number of neighbours and the number of links between them. The strength of the linkages and the significance of the seed node pairs in link prediction were recorded in this study. We proposed a new similarity index termed SP that may be used in both circumstances, whether there are common neighbours between seed node pairs or not. We used Facebook data to test the algorithms on three estimations: Precision, AUP, and AUC. Our experimental results show that the link prediction algorithm performs consistently and delivers a better discriminative score value for seed node pairs.

The primary contribution of this work is that we discover a score function based on the strength and prominence of seed nodes, which is not normally included in other methods, is particularly useful in forecasting missing links. Another advantage of the SP index is the fact that it is parameters-free, which makes it easier to apply to different types of networks, as picking an acceptable parameter for a certain network always requires additional information, which may be difficult to obtain.

Nonetheless, our work has some limitations, such as not taking into account the negative strength (weakness) of the links and the negative prominence of the seed nodes. We assume a modest-size network here, but manually calculating the score for complex or big networks is quite tough. For this computation, we must devise an appropriate pseudo code. These restrictions will be addressed in future research. Still, we hope that our findings will help future researchers deal with the uncertainty, strength, and weaknesses of ties in a fuzzy social network.

Author contributions Sakshi Dev Pandey: Conceptualization, Methodology, Writing—Original Draft. Sovan Samanta: Formal Analysis, Writing—Review and Editing. A. S. Ranadive: Supervision. Leo Mrsic: Data Curation, Software, Visualization, Funding Acquisition. Antonios Kalampakas: Investigation, Resources. Tofigh Allahviranloo: Project Administration, Validation.

Data Availability The data that support the findings of this study are available within the article. Additional data or materials are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no Conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Adamic LA, Adar E (2003) Friends and neighbors on the web. Soc Netw 25(3):211–230
2. Al Hasan M, Chaoji V, Salem S, Zaki M (2006) Link prediction using supervised learning. In: SDM06: workshop on link analysis, counter-terrorism and security, vol 30, pp 798–805
3. Arrar D, Kamel N, Lakhfif A (2024) A comprehensive survey of link prediction methods. J Supercomput 80(3):3902–3942
4. Banerjee S (1991) An optimal algorithm to find the degrees of connectedness in an undirected edge-weighted graph. Pattern Recogn Lett 12(7):421–424
5. Barabási A, Jeong H, Néda Z, Ravasz E, Schubert A, Vicsek T (2002) Evolution of the social network of scientific collaborations. Physica A 311(3):590–614
6. Bastani S, Jafarabad AK, Zarandi MHF (2013) Fuzzy models for link prediction in social networks. Int J Intell Syst 28(8):768–786
7. Bhawsar Y, Thakur G (2016) Performance evaluation of link prediction techniques based on fuzzy soft set and Markov model. Fuzzy Inf Eng 8(1):113–126
8. Bhutani KR, Rosenfeld A (2003) Strong arcs in fuzzy graphs. Inf Sci 152:319–322
9. Cannistraci CV, Alanis-Lobato G, Ravasi T (2013) From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. Sci Rep 3(1):1–13
10. Ezadi S, Allahviranloo T (2017) Numerical solution of linear regression based on z-numbers by improved neural network. Intell Autom Soft Comput 24(1):1–11
11. Guo F, Zhou W, Wang Z, Ju C, Ji S, Lu Q (2023) A link prediction method based on topological nearest-neighbors similarity in directed networks. J Comput Sci 69:102002
12. Jaccard P (1901) Distribution de la flore alpine dans le bassin des dranses et dans quelques régions voisines. Bull Soc Vaudoise Sci Nat 37:241–272
13. Jeh G, Widom J (2002) Simrank: a measure of structural-context similarity. In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 538–543
14. Katz L (1953) A new status index derived from sociometric analysis. Psychometrika 18(1):39–43
15. Kumar A, Singh SS, Singh K, Biswas B (2019) Level-2 node clustering coefficient-based link prediction. Appl Intell 49:2762–2779
16. Kumar A, Singh SS, Singh K, Biswas B (2020) Link prediction techniques, applications, and performance: a survey. Physica A 553:1–46
17. Liben-Nowell D, Kleinberg J (2003) The link prediction problem for social networks. In: Proceedings of the twelfth international conference on information and knowledge management, pp 556–559
18. Liu G (2022) An ecommerce recommendation algorithm based on link prediction. Alex Eng J 61(1):905–910
19. Liu W, Lü L (2010) Link prediction based on local random walk. Europhys Lett 89(5):1–6
20. Lü L, Jin C-H, Zhou T (2009) Similarity index based on local paths for link prediction of complex networks. Phys Rev E 80(4):1–9
21. Mahapatra R, Samanta S, Allahviranloo T, Pal M (2019) Radio fuzzy graphs and assignment of frequency in radio stations. Comput Appl Math 38:1–20
22. Mahapatra R, Samanta S, Pal M, Xin Q (2019) Rsm index: a new way of link prediction in social networks. J Intell Fuzzy Syst 37(2):2137–2151
23. Mahapatra R, Samanta S, Pal M, Xin Q (2020) Link prediction in social networks by neutrosophic graph. Int J Comput Intell Syst 13(1):1699–1713

24. Mathew S, Mordeson JN, Malik DS (2018) Fuzzy graph theory, vol 363. Springer, Berlin
25. Mathew S, Sunitha M (2009) Types of arcs in a fuzzy graph. Inf Sci 179(11):1760–1768
26. Moradabadi B, Meybodi MR (2017) Link prediction in fuzzy social networks using distributed learning automata. Appl Intell 47:837–849
27. Mordeson JN, Nair PS (2012) Fuzzy graphs and fuzzy hypergraphs, vol 46. Physica, Heidelberg
28. Mosleh M, Allahviranloo T, Otadi M (2012) Evaluation of fully fuzzy regression models by fuzzy neural network. Neural Comput Appl 21:105–112
29. Newman ME (2001) Clustering and preferential attachment in growing networks. Phys Rev E 64(2):1–4
30. Pandey SD, Ranadive A, Samanta S, Sarkar B (2022) Bipolar-valued fuzzy social network and centrality measures. Discrete Dyn Nat Soc 1–13:2022
31. Salton G, McGill M (1986) Introduction to modern information retrieval. Mcgraw-Hill, New York
32. Samanta S, Dubey VK, Sarkar B (2021) Measure of influences in social networks. Appl Soft Comput 99:1–42
33. Sorensen TA (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. Biol Skar 5:1–34
34. Sun M, Tang M (2025) A review of link prediction algorithms in dynamic networks. Mathematics 13(5):807
35. Tuan TM, Chuan PM, Ali M, Ngan TT, Mittal M, Son LH (2019) Fuzzy and neutrosophic modeling for link prediction in social networks. Evol Syst 10:629–634
36. Wu Z, Lin Y, Wang J, Gregory S (2016) Link prediction with node clustering coefficient. Physica A 452:1–8
37. Yager RR (2008) Intelligent social network analysis using granular computing. Int J Intell Syst 23(11):1197–1219
38. Zhou T, Lü L, Zhang YC (2009) Predicting missing links via local information. Eur Phys J B 71:623–630
39. Zimmermann HJ (2010) Fuzzy set theory. Wiley Interdiscip Rev Comput Stat 2(3):317–332