Journal of Computational Science 24 (2018) 313–328

Contents lists available at ScienceDirect

Journal of Computational Science

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

Link prediction in stochastic social networks: Learning automata approach

Behnaz Moradabadi*, Mohammad Reza Meybodi

Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 12 January 2017
Received in revised form 7 June 2017
Accepted 8 August 2017
Available online 15 August 2017

Keywords:
Stochastic social network
Link prediction
Learning automata

ABSTRACT

Link prediction is a main social network challenge that uses the network structure to predict future links. The common link prediction approaches to predict hidden links use a static graph representation where a snapshot of the network is analyzed to find hidden or future links. For example, similarity metric based link predictions are a common traditional approach that calculates the similarity metric for each non-connected link and sort the links based on their similarity metrics and label the links with higher similarity scores as the future links. Because people activities in social networks are dynamic and uncertainty, and the structure of the networks change over time, using deterministic graphs for modeling and analysis of the social network may not be appropriate. This paper proposes a new link prediction method based on learning automata for stochastic social networks. In a stochastic social network, the weights associated with the links are random variables. To do this, we first redefine some of the similarity metrics for link prediction in stochastic graphs and then propose a method based on learning automata to calculate the distribution of the proposed similarity metrics assuming that the probability distributions of the link weights are unknown. Also, the proposed method has capability to use in online stochastic social networks where the social network changes online and the future links must be predicted. To evaluate the proposed method we use different synthetic stochastic social networks and present that the stochastic link prediction achieves better results in comparison to the classical link prediction algorithms in the stochastic social networks.

© 2017 Published by Elsevier B.V.

1. Introduction

The predicting linkage between data objects is an interesting task in data mining research area. For examples predict web hyperlink creation, genetic prediction, protein–protein interactions, and the record linkage problem. In link prediction problem the data are represented with a network/graph representation. These data can be visualized as graphs, where a vertex corresponds to a person and a link represents some form of association between the corresponding persons [1,2]. The concept of a link in a social network usually is a common interest of the corresponding social network. All link prediction methods address the following question: "Given a pair of nodes u and v in the current social network, how likely it is that u will interact with v in the future?" [1]. Link prediction can be applied in very areas such as the Internet and web based applications [3,4], recommender systems [5,6] and in the bibliography application for deduplication [7], record linkage [8].

Most of the previous link prediction methods have been proposed based on a static network representation, where a snapshot of the network structure is available and the goal is to predict the future links. In such a static network, each link occurrence is represented by a one-time event and all interest is only the existence of the link. For example, one may be interested to know whether a customer will purchase a product in the future or whether an author will ever collaborate with another author in the future. But in many applications, the social networks are really online, non-deterministic and unpredictable, and the structure of the network and its parameters change over time; so using deterministic social network models with fixed values for links are restrictive in solving real social network problems. In other words, link prediction methods based on the static graph representation fail when the social network has online and non-deterministic behavior. One of the solutions to overcome this problem is link prediction using online and stochastic social networks in which each link is a random variable. In [9] Rezvanian proposed the stochastic social network

* Corresponding author.
E-mail addresses: moradabadi@aut.ac.ir (B. Moradabadi), mmeybodi@aut.ac.ir (M.R. Meybodi).

https://doi.org/10.1016/j.jocs.2017.08.007
1877-7503/© 2017 Published by Elsevier B.V.

and showed that the stochastic social network with a random variable for each link is a good approach to overcome this problem. By choosing a stochastic social network, they redefined some measures of social networks such as degree distribution, betweenness, cluster coefficient and strongness measures. Also, it is be noticed that some problems in stochastic networks were also proposed by other studies: path in the stochastic network [10], cover in the stochastic network [11], clique in the stochastic network [12], and spanning tree in the stochastic network [13]. So in this paper, we propose a learning automata based link prediction in stochastic social networks, called SLP, to predict the link occurrence using the uncertainty of the data. Also the proposed method can be extended as an online link prediction to use in the online stochastic social networks where the links can be added or removed over time and the future links must be predicted again.

Learning Automaton (LA) is an adaptive decision-making unit that tries to learn the optimal action from a set of allowable actions by interacting with a random environment [14]. In each step, it selects an action from its action set. The action selection in the LA is based on a probability distribution over the action set. The selected action is applied to the environment and then a reinforcement signal is produced by the environment. LA updates the probability distribution of its actions according to both reinforcement signal and a learning algorithm and again chooses an action. These steps are repeated until LA converges to some action.

This paper tries to propose a link prediction method based on similarity metrics in stochastic networks, in which each link has a random variable with an unknown probability distribution function. Traditionally in similarity based link prediction, the starting point is to extract the values/scores of a chosen similarity metric that represent the proximity of the pairs of nodes. Then, the pairs of non-connected nodes are at first ranked according to a chosen metric (for instance the number of common neighbors). After that, the top L ranked pairs are assigned as predicted links. To put it another way, it is always assumed that the links have the highest scores are most likely to occur. To obtain the goal of this paper, after present a brief overview of recent studies for link prediction in a social network, we first redefine some similarity metrics for stochastic graphs and then design a learning automata based algorithm for calculating the similarity metric under the situation that the probability distribution function of the weight of each link is unknown. In other words, the proposed algorithm tries to estimate the probability distribution function of the similarity metric using learning automata and sampling. The process of sampling from the links of the graph is guided by the aid of learning automata in such a way that the number of samples needed to be taken from the links of the stochastic graph and the computational complexity for estimating the similarity metrics probability distributions be reduced as much as possible. Similar to traditional similarity methods that the output of link prediction is the set of links with higher similarity, the output of the proposed link prediction is also a set of stochastic links with their similarity probability distribution functions. So, to determine a link's existence, we should sample from its similarity probability distribution function and based on this sample and a predefined threshold we make the decision about the link existence. This threshold is chosen based on empirical results and experts. It is also should be noticed that the proposed method can be applied in online stochastic social networks. In many recent studies, researchers use synthetic social networks similar to real networks in order to study and evaluate their methods using a computer generated graph models in order to exploit the structures and dynamics of networks. Synthetic data has no personal information and cannot be traced back to any individual; therefore, the use of synthetic data reduces confidentiality and privacy issues. So in order to evaluate the performance of the proposed algorithm in stochastic social networks, we conduct several experiments using different synthetic stochastic social networks and show that the stochastic models achieve better link prediction performance in comparison common stochastic link prediction method.

The rest of the paper is organized as follows. Section 2 dedicated to material and methods including a brief introduction of link prediction problem, an overview of recent works about dynamics of user behaviors in social network and a brief description of learning automata theory. In Section 3, the proposed similarity metrics for stochastic graphs are described. Section 4 introduces the proposed link prediction method for stochastic social networks based on learning automata. Section 5 presents the simulation results and finally section 6 summarizes the main conclusion of the proposed method.

2. Material and methods

In this section, to provide the necessary background, we present a brief description of link prediction. We also briefly review the studies performed by the researchers about the distributions of user behaviors in social networks. At the end of this section, learning automata and variable action set learning automata are introduced.

2.1. Link prediction

A classic definition of the link prediction problem is expressed by: "Given a snapshot of a social network at time t, we seek to accurately predict the links that will be added to the network during the interval from time t to a given future time $t+1$" [15]. The most widespread approach to the problem is to calculate the topological/structural patterns of the social network of interest [16]. Different metrics to describe node pairs have already been adopted in previous studies [17] and they explore the structural patterns of the network and commonly provide a degree of proximity/similarity between the nodes. There are many similarity metrics include [15]: 1) Local similarity metrics that only use the local information of a link to calculate similarity metric: Common Neighbors, Salton Index, Jaccard Index, Hub Depressed Index, Hub Promoted Index, Leicht-Holme Newman Index (LHN1), Preferential Attachment Index, Adamic-Adar Index and Resource Allocation Index, 2) Global similarity metrics that can use all information in the network to calculate the similarity metric between two nodes: Katz Index, Leicht-Holme-Newman Index (LHN2), Matrix Forest Index (MFI) and 3) Quasi-local metrics that do not require global topological information but use more information than local indices: Local Path Index, Local Random Walk, Superposed Random Walk, Average Commute Time, Cos+, random walk with restart, SimRank, Resource Allocation index and Local Path index. As previously mentioned, the starting point in these approaches is to extract the values/scores of different metrics that represent the proximity of the pairs of nodes. Then, the pairs of non-connected nodes are at first ranked according to a chosen metric (for instance the number of common neighbors) [15]. After that, the top L ranked pairs are assigned as predicted links. To put it another way, it is always assumed that the links have the highest scores are most likely to occur. In the following we briefly present the eight commonly similarity measures that we use them to redefine stochastic similarity metrics. In the rest of this paper, notation $\Gamma(x)$ denotes the neighbors set of node x.

1) Common Neighborhood: In this measure, two nodes v_i and v_j are more likely to have a link if they have many common neighbors. This score is defined as:

$$CN\left(v_i, v_j\right) = |\Gamma\left(v_i\right) \cap \Gamma\left(v_j\right)|$$

2) Salton: Because the common neighborhood metric is not normalized, some other metrics are proposed to normalize the

common neighborhood metric. Salton score is one of these metrics that normalizes the CN metrics using $\sqrt{|\Gamma(v_i)| \times |\Gamma(v_j)|}$ and it is defined as:

$$Salton(v_i, v_j) = \frac{|\Gamma(v_i) \cap \Gamma(v_j)|}{\sqrt{|\Gamma(v_i)| \times |\Gamma(v_j)|}}$$

3) Jaccard Index [18]: This index was proposed by Jaccard to normalize the CN metric and it is defined as:

$$Jaccard(v_i, v_j) = \frac{|\Gamma(v_i) \cap \Gamma(v_j)|}{|\Gamma(v_i) \cup \Gamma(v_j)|}$$

This metric is defined in such a way that for high number of common neighbors, the score would be higher.

4) Preferential Attachment (PA) [19]: The preferential attachment (PA) algorithm is based on the preferential attachment phenomena rule [20] that is discovered in a variety of social networks. In this method the link score is set to be the product of the degrees of the involved nodes and it is defined as follows:

$$PA(v_i, v_j) = |\Gamma(v_i)| \times |\Gamma(v_j)|$$

5) Adamic-Adar Index (AA) [21]: This index is an extension of common neighborhood method such that the less-connected neighbors have more weight and it is defined as:

$$AA(v_i, v_j) = \sum_{z \in \Gamma(v_i) \cap \Gamma(v_j)} \frac{1}{\log(\Gamma(z))}$$

This metric first proposed as a metric of similarity between two web pages.

6) Resource Allocation Index (RA): This index and Adamic-Adar Coefficient have similar formulas but they come from different motivations. RA is based on physical processes of resource allocation and can be applied to networks formed by airports (flow of aircraft and passengers) or networks formed by electric power stations and is defined as:

$$RA(v_i, v_j) = \sum_{z \in \Gamma(v_i) \cap \Gamma(v_j)} \frac{1}{|\Gamma(z)|}$$

7) Katz Index: In this metric, a similarity is defined as the sum of a number of paths with different lengths such that shorter paths have more weights. It is defined by the following equation:

$$Katz(v_i, v_j) = \sum_{l=1}^{\infty} \beta^l . |Path(v_i, v_j)^{<l>}|$$

Where $|Path(x, y)^{<l>}|$ is the number of paths between v_i and v_j with length l. The parameter $\beta(\leq 1)$ can be used to regularize this feature. The challenge of KATZ metric is its computationally complexity. It is also shown that the Katz metric can be calculated based on the following equation:

$$Katz = (I - \beta A)^{-1} - I$$

where $a = 0.05$, $b = 0.003$ is the adjacency matrix and I is an identity matrix of proper size. This method also has high complexity in large social networks.

8) LP Index: This index is a restricted version of Katz metric such that only paths in length 2 and 3 is considered. This metric has a lower computational complexity in comparison to Katz and it is defined as the following:

$$LPIndex(v_i, v_j) = A^2 + \varepsilon A^3$$

where A is the adjacency matrix and ε is a free parameter. If $\varepsilon = 0$ the LP is the same as CN metric.

For more information about other similarity metrics please refer to Liben- Nowell and Kleinberg [15]. In the rest of this section, we go on to review recently proposed link prediction methods:

In [22] a new data mining process, called Interaction Prediction, is proposed for the link prediction problem in dynamic networks. Interaction Prediction uses dynamic social networks time series forecast in addition to similarity metrics and network community structure to predict the links that may appear in the future. The proposed method focuses on the links between the communities and their experiments on real-world social networks show that the Interaction Prediction outputs interesting results in link prediction.

The authors of [23] have provided a link prediction method that uses the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) to determine the importance of each similarity metric in a linear format. They used a large dynamic social network and their method presents fast convergence and high accuracy for link prediction problem. The authors of [24] proposed a new link prediction using information theory and mutual information of network structure (MI-LP). They compared the model with six typical prediction methods on ten networks. Their results show two good features: improving the link prediction accuracy and the reasonable computational complexity. In [25] a new method called Multivariate Time Series Link Prediction is proposed for link prediction in evolving networks. This method combines temporal evolution of the network in addition to node similarities and node connectivity information. This method uses links and a calculated similarity metric for each time. The authors finally compare different similarity metrics in their experiments.

In [26] a weighted approach for modeling the occurrence of time is used to generate a different similarity metric and in our previous research we have proposed a new link prediction method based on temporal similarity metrics and Continuous Action set Learning Automata (CALA) [27]. The proposed method takes advantage of using different similarity metrics as well as different time periods. In the proposed algorithm, we have modeled the link prediction problem as a noisy optimization problem and use a team of CALAs to solve the noisy optimization problem. The obtained link prediction results show satisfactory of the proposed method for some social network data sets. Finally in [28], to predict the future links we have proposed a new link prediction which uses both fuzzy social network and DLA (FLP-DLA). In the proposed method we first create a fuzzy social network where each link has a fuzzy strength and then use this network to predict future links using distributed learning automata based on graph navigation and the fuzzy strength of links. The experiments showed that the proposed approach performs better than other strategies. Based on our research there is no recent research about the link prediction problem in stochastic social networks and most of the researches including our recent works focus only on the static and deterministic graphs. So in this paper we try to study the behavior of stochastic social networks in link prediction problem using LA optimization tool; because in real-world applications the social networks are really online, nondeterministic and unpredictable and the study of social networks problem in stochastic networks is valuable.

2.2. User behavior in social networks

Since generating and sharing information through social networks are done by users, studying how users behave and interact with their friends in online social networks play a significant role in social network analysis. Based on valuable research in social network dynamics, now we can produce synthetic graphs using computer programs that are similar to real networks in order to study the structures and dynamics of networks and their user activities in a systematic manner. In recent years, several research works regarding user activity modeling, characterizing the distribution of

user activities and statistical observations from users and in the rest of this section, we brief review some of them. Reference [29] analyzed the interactions among Facebook users. They found that the distribution of the number of messages sent by a user follows a power-law distribution. Reference [30] studied thousands of blogs and millions of posts, they showed that the popularity of posts drops with a Power-law distribution. Leskovec et al. in [31], also focused directly on the microscopic node behavior of four large social network datasets (Flickr, Delicious, Yahoo Answers, and LinkedIn) and observed that user's lifetime follows an exponential distribution. Nazir et al. in [32] showed that the popularity of the applications follows a power-law distribution with an exponential decay. Reference [33] analyzed the patterns of user content generation in online social networks and they found that the overall user lifetime follows the stretched exponential distribution. Kwak et al. in [34] collect user profiles, social relations, trending topics and tweets on Twitter and they found the distribution of the users in a Retweet tree follows a power-law distribution. Gyarmati et al. in [35] analyzed the degree distribution of the nodes for online sessions of users in Bebo, MySpace, Netlog and they revealed that the time spent online by users follows a Weibull distribution. Galuba et al. in [36] analyzed the information propagation in Twitter and they found that posting frequencies followed a power-low distribution across users and URLs. Yan et al. in [37] studied the interval time distribution of people posting in microblogs systems and found that it follows a power-law distribution. Liu et al. in [38] found that the number of daily active users using some applications in Facebook follows a power-law distribution. Vongsingthong et al. in [39] tracked the activities of some users in a Facebook through a questionnaire and they found questioning on products follows a Power-law distribution. Bild et al. in [40], studied empirically aggregate user behavior on Twitter and they found that the lifetime tweet distribution is a type-II discrete Weibull stemming from a Power-law hazard function. For more information about related studies please refer to [9].

2.3. Learning automata

A learning automaton [41] is an agent that can make a decision and learn how to choose the true decision by interacting with a random environment iteratively. Each LA has a probability distribution over its finite action-set and at each iteration it selects an action based on the corresponding probability and sends the action to the random environment. The environment sends a reinforcement signal to the LA based on the evaluation of the input action. The action probability vector of the LA is updated based on the given reinforcement signal from the random environment. The goal of a learning automaton is to find the optimal action such that the average penalty received from the environment is minimized as much as possible. The environment can be described by a triple $E \equiv \{\alpha, \beta, c\}$, where $\alpha \equiv \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ represents the finite set of the inputs, $\beta \equiv \{\beta_1, \beta_2, \ldots, \beta_m\}$ denotes the set of the values that can be taken by the reinforcement signal, and $c \equiv \{c_1, c_2, \ldots, c_m\}$ denotes the set of the penalty probabilities, where the element c_i is associated with the given action α_i. If the penalty probabilities are constant, the random environment is said to be a stationary random environment, and if they vary with the time, the environment is called a non-stationary environment. The random environments based on their reinforcement signal β can be classified into P-model, Q-model and S-model. The reinforcement signal can only take two binary values 0 and 1 in P-model environments, take one of a finite number of values in the interval [0,1] in Q-model environment, and take a value in the interval $[a, b]$ in S-model. The learning automata we used it in the proposed algorithm is S-model LA. The relationship between the learning automaton and its random environment has been shown in Fig. 1 [41]. Learning automata can be classified into two main families [41]: fixed structure learning automata and variable structure learning automata. Variable structure learning automata are represented by a triple $\langle \beta, \alpha, T \rangle$, where β is the set of inputs, α is the set of actions, and T is the learning algorithm. The learning algorithm is a recurrence relation which is used to modify the action probability vector. Various learning algorithms have been proposed. The linear reward-ε penalty algorithm, $L_{R-\varepsilon P}$ [41], is one of the learning schemas that we used it in the proposed algorithm. Let $\alpha(n) = \alpha_i$, be the action chosen at step n from the probability distribution p. Then the learning algorithm for updating the action probabilities in the $L_{R-\varepsilon P}$ is as follows:

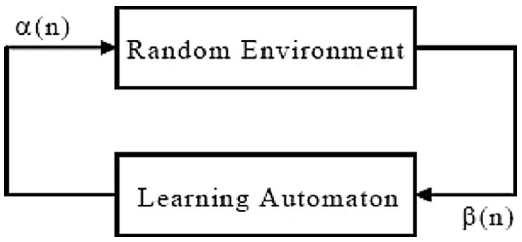


Fig. 1. The relationship between a learning automata and its random environment.

$$\begin{aligned} p_i(n+1) &= p_i(n) + \text{a.}\left(\beta(n)\right).(1 - p_i(n)) - \text{b.}\left(1 - \beta(n)\right).p_i(n) \\ p_j(n+1) &= p_j(n) + \text{a.}\left(\beta(n)\right).p_j(n) + \frac{b.\beta(n)}{r-1} - \text{b.}\left(1 - \beta(n)\right).p_j(n) \text{j} \neq \text{i} \end{aligned} \quad (1)$$

where a and b, 0 < b << a < 1, denote the reward and penalty parameters and determine the amount of increase and decrease of the action probabilities, respectively. Based on the above equation if $\beta(n) = 0$ (unfavorable response) we penalize the action α_i and reward other actions and if $\beta(n) = 1$ (favorable response) then we reward action α_i and penalize other actions. In other words based on the fact that $\beta(n)$ is in [01] based on S-model and the above equation, the$\beta(n)$ determines the amount of goodness of the received reinforcement signal and it is used to update the action probabilities. Based on Reference [42] the $L_{R-\varepsilon P}$ learning algorithm is $\varepsilon - optimal$ which means for every $\varepsilon > 0$, we have:

$$\lim_{n\to} E\left[\beta_n\right] > \max_i \left\{c_i\right\} - \varepsilon$$

where $\max_i \left\{c_i\right\}$ is the maximum expected reinforcement signal.

Learning automata have been found to be useful in systems where incomplete information about the environment, wherein the system operates, exists and recently, several learning automata based approaches have been presented for improving the performance of many applications [43] [44] [45].

3. Proposed similarity metrics in stochastic graphs

Generally, a stochastic graph G can be described by a triple $<V, G, W>$, where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of nodes, $E = \left\{e_{ij}\right\} \subseteq V \times V$ is the set of links, and W is a matrix in which w_{ij} is a random variable associated to link e_{ij} if such a link exists. In this section we redefine some similarity metrics for link prediction problem in stochastic network as the following:

1) Stochastic Common Neighborhood: In this measure, the common neighborhood random variable is a random variable that represents the sum of shared stochastic weights that the two nodes v_i and v_j have in their common neighbors and it is defined by:

$$CN(v_i, v_j) = \sum_{\{k | e_{ik} \in E, e_{jk} \in E\}} \min(w_{ik}, w_{jk})$$

2) Stochastic Salton: This score in stochastic social networks is defined as the following random variable:

$$Salton(v_i, v_j) = \frac{\sum_{\{k|e_{ik}\in E, e_{jk}\in E\}} \min(w_{ik}, w_{jk})}{\sqrt{\sum_{\{k|e_{ik}\in E\}} w_{ik} \times \sum_{\{l|e_{jl}\in E\}} w_{jl}}}$$

3) Stochastic Jaccard Index: This index in stochastic social networks is defined as the following random variable:

$$Jaccard(v_i, v_j) = \frac{\sum_{\{k|e_{ik}\in E, e_{jk}\in E\}} \min(w_{ik}, w)}{\sum_{\{k|e_{ik}\in E, e_{jk}\in E\}} \max(w_{ik}, w_{jk})}$$

4) Stochastic Preferential Attachment: The preferential attachment (PA) random variable in the stochastic network is defined as follows:

$$PA(v_i, v_j) = \sum_{\{k|e_{ik}\in E\}} w_{ik} \times \sum_{\{l|e_{jl}\in E\}} w_{jl}$$

5) Stochastic Adamic-Adar Index [21]: This index is an extension of common neighborhood method and in stochastic network is defined as the following random variable:

$$AA(v_i, v_j) = \sum_{\{k|e_{ik}\in E, e_{jk}\in E\}} \frac{1}{\log(\sum_{\{z|e_{zk}\in E\}} w_{zk})}$$

6) Resource Allocation Index: this index in the stochastic social network is defined as the following random variable:

$$RA(v_i, v_j) = \sum_{\{k|e_{ik}\in E, e_{jk}\in E\}} \frac{1}{\sum_{\{z|e_{zk}\in E\}} w_{zk}}$$

7) Stochastic Katz Index: In this metric, the similarity is defined as the sum of the stochastic paths weight between v_i and v_j with different lengths such that shorter paths have more weights and it is defined as the following random variable:

$$Katz(v_i, v_j) = \sum_{l=1}^{\infty} \beta^l \times \sum_{\{e_{zk}\in path(v_i, v_j)^{\langle l\rangle}\}} w_{zk}$$

Where $path(v_i, v_j)^{<l>}$ is any path between v_i and v_j with length l.

8) Stochastic LP Index: This index is a restricted version of Katz metric such that only paths in length 2 and 3 is considered. This metric in the stochastic social network is a random variable and it is defined as the following:

$$LP - Index(v_i, v_j) = \sum_{l=2}^{3} \beta^l \times \sum_{\{e_{zk}\in path(v_i, v_j)^{\langle l\rangle}\}} w_{zk}$$

4. Proposed link prediction method in stochastic graphs

In the previous section, we redefined some similarity metrics for stochastic graphs. In this section, a link prediction method based on learning automata (SLP) is proposed for stochastic social networks under the situation that the probability distribution functions of the weights associated with the links of the graph are unknown. The proposed algorithms take advantage of using learning automata to estimate the distribution of some chosen similarity metric in order to predict future links. In the algorithm after initialization phase, it iterates the sampling phase, updating phase and feedback phase until one of the stopping conditions are met. Similar to traditional similarity methods that the output of link prediction is the set of links with higher similarity, the output of the proposed link prediction is also a set of stochastic links with their similarity probability distribution functions. So, to determine a link's existence, we should sample from its similarity probability distribution function and based on this sample and a predefined threshold we make the decision about the link existence. This threshold is chosen based on empirical results and experts. The proposed method also can have an additional phase called changes phase that can be used in online stochastic social networks to predict the future links after a change in the network has occurred. The details of initialization of the proposed method, sampling phase, updating phase, feedback phase, stopping phase and changes phase are given below.

4.1. Initialization

Let $G(V, E, W)$ be the input stochastic graph, where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of nodes,

$E = \{e_{ij}\} \subseteq V \times V$ is the set of links, and $W = \{w_{ij}\}$ is the set of random variables each of which is associated with a link weight of the input stochastic graph. It is assumed that the weight of each edge is a positive random variable with an unknown probability distribution function. The proposed algorithm uses two sets of learning automata LA_{Links} and LA_{Tests}:

1) LA_{Links} is a set of LAs for each of existed links in the network and tries to estimate the importance of sampling from the corresponding links in calculation of similarity metrics. Each LA_L in LA_{Links} have two actions: $\{\alpha_L^1$ = "take sample", α_L^2 = "do not take sample"$\}$. Let $p_{L_i} = \{p_{L_i}^1, p_{L_i}^2\}$ be the action probability vector of learning automaton LA_{L_i}, and initialized equally $p_{L_i}^1 = p_{L_i}^2 = 1/2$ for all $v_i \in V$. The process of taking samples from a link is determined by the learning automaton assigned to that link. So this set of LAs is learned in a such way that the number of samples of the network that is required for calculating the similarity metric be reduced as much as possible. At the beginning of the algorithm, the weight of each link is initialized with some random samples in order to provide a coarse estimate of the weight of that link.

2) LA_{Tests} is a set of LAs for each of test links that must be predicted and tries to estimate if the distribution of the similarity metric corresponding to the test link must be updated or not. Each LA_T in LA_{Tests} also has two actions: $\{\alpha_T^1$ = "update the similarity metric probability distribution", α_T^2 = "do not update the similarity metric probability distribution"$\}$. Let $p_{T_i} = \{p_{T_i}^1, p_{T_i}^2\}$ be the action probability vector of learning automaton LA_{T_i}, and initialized equally $p_{T_i}^1 = p_{T_i}^2 = 1/2$ for all $t_i \in T$(Test set). The action of each LA_T in LA_{Tests} determines that if we should update the similarity metric probability distribution of the corresponding link or not. So the LA_{Tests} is guided such that the computational complexity of calculating similarity metrics is decreased as much as possible.

4.2. Sampling phase: sampling the stochastic social network

In each iteration of the proposed method, in the first phase, each LA_L in the LA_{Links} chooses an action in a parallel manner. The action of each LA_L determines that we should sample a new weight or use the previous weight as the weight of corresponding link in the current iteration. If the chosen action be α_L^1 = "take sample", then we sample a new weight for the corresponding link and if it be α_L^2 = "do not take sample" then we use the weight of previous iteration as the weight of corresponding link in the current iteration. So this set of LAs is learned in a such way that the number of samples of the network that is required for calculating final similarity metrics probability distributions be reduced as much as possible. In other words in SLP the sampling process implemented by learning automata

aims to take more samples from the promising region of the graph, the regions that reflect higher rate of activities, instead of walking around and taking unnecessary samples from non-promising regions of the graph. Now we have a weighted social network and the first phase of the SLP is finished here.

4.3. Updating phase: update the similarity metrics probability distributions

In this phase, we try to update the probability distribution of the similarity metrics using the weighted graph that is generated in the previous phase. To do this, each LA_T in the LA_{Tests} selects a new action. If the action of the LA_T be α_T^1 = "update the similarity metric probability distribution", the calculation of the new estimates for the probability distribution of the similarity metric for the corresponding link are then performed. But if the chosen action be α_T^2 ="do not update the similarity metric probability distribution"} then we do not update the similarity metric probability distribution of the corresponding link and we use its previous probability distribution as its current probability distribution.

4.4. Feedback phase: update the action probabilities of LAs

The goal of this phase is to reward or penalize each LA in the network to update their action probability distributions. To do this first we calculate the distance between the current similarity metric probability distribution and the one obtained in the previous iteration for every test link i and called it d_i. To calculate the distance we use Skew divergence distance that is described in the experiment setting. This metric is chosen with empirical experiments (Experiment 6) and based on the trade-off of computational complexity and accuracy. Now we use this value to generate reinforcement signal for every LA in the network based on the following equations:

$$\beta_{T_i} = d_i \quad \text{for every i} \in LA_{Test},$$

$$\beta_{L_j} = \frac{\sum_{i \in m} d_i}{|m|} \quad \text{for every i} \in LA_{Links},$$

where m is the set of test links that use link j to calculate their similarity metrics. Finally we update each LA according to the $L_{R-\varepsilon P}$ learning algorithm and the generated reinforcement signal. In other words for each LA_{T_i} in LA_{Tests} we use d_i as its reinforcement signal and penalize or reward the action α_T^1 (update the similarity metric probability distribution) of the corresponding LA based on the closeness of the d_i to 0 and 1, respectively. In other words if the new distribution is close to the previous distribution, it means that the distribution of the link is not changed in respect to the previous iteration and so we reward the action α_T^2 (do not update the similarity metric probability distribution) and if the new distribution is different from the previous one then we reward the action α_T^1 (update the similarity metric probability distribution). Also for each LA_{L_i} in LA_{Links} we use the average reinforcement signal of the test links that use the corresponding link to calculate their similarity metrics. So the proposed algorithm tries to investigate the sample of links that are more important to generate new information about similarity metrics by using LA_{Links} and reduce the cost of calculating similarity metrics in each iteration by using LA_{Tests}.

4.5. Stopping phase

In the proposed algorithm sampling, updating and feedback phases is repeated until the average of entropy of probability vector of learning automata reaches a predefined value T_{min} or the maximum number of iteration, K, is reached to a threshold k. The information entropy of a learning automaton with r actions can be defined as follows [46]:

$$H = -\sum_i^r p_i.\log(p_i) \tag{2}$$

where p_i, is the probability of choosing ith action of a learning automaton. The entropy for a learning automaton has maximum value of one when all the actions have equal probabilities of choosing and has minimum value of zero when the action probability vector is a unit vector. After stopping the SLP, we have a similarity probability distribution function for each test link as the output of the proposed method. Now because of the similarity metrics are random variable, to determine a link's existence, we should sample from its similarity probability distribution function and based on this sample and a predefined threshold we make the decision about the link existence. This threshold is chosen based on empirical results and experts.

4.6. Changes phase

This phase of SLP is proposed to adapt the SLP in online stochastic social networks. In the online stochastic social networks the links can be occurred or removed over the time and the future links must be predicted again. In the proposed algorithm after the stopping phase is finished, we have a link prediction result for the structure of the network until time T. Now for adding the online link prediction capability to the SLP, assume that in time T + 1 we have two available actions: a link is added to the network or a link is removed from the network. In the following we proposed two scenarios to handle these actions for reconsidering future links:

1) A new link is added to the network: in this case for the new link l we consider a new LA_{Link} learning automata,LA_l, to determine the importance of the link l' sampling in the prediction task. Also we reset every LA_t of the LA_{Test} set in the neighbors of link l with a special depth d to the initial configuration. Then we repeat the learning procedure for the set of LA_l and the LA_t in order to reconsider the link prediction result.
2) A new link is removed from the network: in this case for the removed link l we only reset every LA_t of the LA_{Test} set in the neighbors of link l with a special depth d to initial configuration. Then we repeat the learning procedure for the set of LA_t in order to reconsider the link prediction result.

This online learning procedure reduces the computational complexity of calculating link prediction by using local learning instead of considering total network structure. The pseudo-code of the proposed SLP is shown in Algorithm 1.

Algorithm 1: Pseudo code of the proposed algorithm (SLP)

Let t be the iteration counter and initially set to **0.**

Let K and T_{min} be the maximum number of iterations and the entropy threshold, respectively.

Let a, b be the reward and punishment rate to update learning automaton in the learning algorithm.

Set LA_{Link} be the set of learning automata, one learning automaton for each link in the network.

Set LA_{Test} be the set of learning automata, one learning automaton for each link that must be predicted.

Set the initial action probability distribution for each learning automaton in LA_{Link} and LA_{Test} using $p_{L_i}^1 = p_{L_i}^2 = \frac{1}{2}, p_{T_i}^1 = p_{T_i}^2 = 1/2$.

while $t<K$ **or** $P<T_{min}$ **do**

// Sampling Phase

Select action for each learning automaton LA_{L_i} in the LA_{Link} set based on their action probability distributions.

For each LA_{L_i} that chooses the action α_L^1= "take sample" we sample a new weight for the corresponding link as the weight of current iteration.

For each LA_{L_i} that chooses the action α_L^2= "do not take sample" we use the previous weight of the corresponding link as the weight of current iteration.

// Updating Phase

Select action for each learning automaton LA_T in the LA_{Test} set based on their action probability distributions.

For each LA_{T_i} that chooses the action α_T^1= "update the similarity metric probability distribution" we update the similarity metric probability distribution using the stochastic similarity metrics proposed in section 3 and the generated weights in the previous phase.

For each LA_{T_i} that chooses the action α_T^2= "do not update the similarity metric probability distribution" we use the previous similarity metric probability distribution as the current probability distribution.

// Feedback Phase

For each LA_{T_i} we calculate the skew divergence distance d_i.

For each LA_{T_i} we set $\beta_{T_i} = d_i$ and update the action probability distribution function of LA_{T_i} based on β_{T_i} and equation (1).

For each LA_{L_i} we set $\beta_{L_j} = \frac{\sum_{i \in m} d_i}{|m|}$ and update the action probability distribution function of LA_{L_i} based on β_{L_i} and equation (1).

//**Calculate LAs Information Entropy**

Set P = Calculate average entropy of all LAs in the network using equation (2).

Set t = t + 1

end while

For each test link j

Output the similarity metric probability distribution function as its output.

5. Experimental results

In this section, in order to evaluate the performance of the SLP, several computer experiments are conducted on different synthetic stochastic graphs and the proposed method has been compared in term of performance and accuracy. In the rest of this section, we first give the experiments materials that we used in our experiments and then give a set of four experiments.

5.1. Experiment materials

This sub-section describe the method and materials that we used in our experiments. To do this first we review the methods we used to produce synthetic graphs and then present the evaluation parameters and metrics that we used them in our experiments.

5.1.1. Random methods to generate synthetic graphs

This sub-section presents the three methods we used them to generate synthetic graphs:

- Barabasi-Albert model (BA model) as a scale-free network with heavy-tailed degree distribution and parameters $N \in \{2000, 5000, 10000\}$ and $m_0 = m = 5$.
- Watts-Strogatz model (WS model) as a small world network is which reflects a common property of many real networks such as short average path length with parameters $N \in \{2000, 5000, 10000\}$ $k = 4$ and $p = 0.2$.
- Erdds-Renyi model (ER model) with parameters $N \in \{2000, 5000, 10000\}$ and $p = 0.2$.

In used social network each link's weight is a random variable with an exponential distribution whose mean is selected randomly from set $p \in \{1, 1.5, 2, 2.5\}$. This setting is adopted from the empirical experimtns [47].

5.1.2. Evaluation metrics

Here we briefly review four common distance measures that try to estimate the distance of two distributions:

1) Kolmogorov-Smirnov distance statistic [48]: for estimate the distance between two cumulative distribution functions (CDFs) and it is in the interval [01] and values closer to 0 means more similarity and lower difference. This metric is defined as the follow:

$$\mathrm{KS}(\mathrm{P}, \mathrm{Q}) = \max_x |P(x) - Q(x)|$$

where P and Q are two CDFs of original and estimated data, respectively, and x represents the range of the random variable.

2) Skew divergence [48]: for estimate the distance between two probability distribution functions (PDFs) and defined as follows:

$$\mathrm{SD}(\mathrm{P}, \mathrm{Q}, \alpha) = D[\alpha \mathrm{P} + (1-\alpha)\mathrm{Q} || \alpha \mathrm{Q} + (1-\alpha)\mathrm{P}]$$

where D is the Kiillback-Leibler (KL) divergence, which measures the similarity between two PDFs P and O that do not have continuous support over the full range of values and $a = 0.99$. The KL divergence is defined as follows:

$$\mathrm{KL}(\mathrm{P}(\mathrm{x}) || \mathrm{Q}(\mathrm{x})) = \sum\nolimits_x p(x) \log \frac{P(x)}{Q(x)}$$

3) Pearson's correlation coefficient [49]: to measure the similarity between the estimated parameter values and the original parameters values and it is in the interval [01] and values closer to 1 mean more similarity. This distance defined as follow:

$$\mathrm{PCC}(\mathrm{P}, \mathrm{Q}) = \frac{n\sum_i p_i q_i - \sum_i p_i \sum_i q_i}{\sqrt{n\sum_i p_i^2 - (\sum_i p_i)^2} - \sqrt{n\sum_i q_i^2 - (\sum_i q_i)^2}}$$

where p_i, and q_i, are the values of the original parameter P and estimated parameter O and n is the number of parameters.

4) Normalized L1 distance [50]: to measure the distance between two positive m-dimensional real vectors P (original distribution) and Q (estimated distribution) and defined as below:

$$L_1(\mathrm{P}, \mathrm{Q}) = \frac{1}{n} \sum_i \frac{|p_i - q_i|}{p_i}$$

5.1.3. Experiment settings

In the proposed method for stopping criteria we used the following parameters: $T_{min} = 0.05$, $k = n \times 50$. Also for the learning automata parameters we used $a = 0.05$, $b = 0.003$. It also should be noted that these parameters are obtained based on empirical experiments. Also, all the results reported in following experiments are based on the averages taken over 30 runs.

5.2. Experiments

In this section a set of experiments is conducted as the follow: in the first experiment, we evaluate the performance of the proposed method in term of stochastic evaluation metrics. In the second experiment we try to compare the final result of the proposed link prediction method with some recent link predictions. In the third experiment, we report the final probability distribution of some random nodes in an instance run to get an idea about the output of the SLP. In the forth experiment we analyze the number of samples that is taken by the proposed LA to reach a certain accuracy and its execution time. In the fifth experiment we test the proposed SLP using three different learning parameters to obtain the best learning parameter. Finally in the sixth experiment we compare the proposed SLP with different distance metrics to show the reason we chose skew divergence distance metric.

5.2.1. Experiment 1

In this experiment, we try to compare the proposed method to SSM (standard sampling method) in term of stochastic evaluation metric. The SSM in stochastic social networks operates simply as follow: in each iteration each random variable in the network is sampled and the distribution of the interested metric is calculated based on sampled values. So the SSM based link prediction based on some chosen similarity metric operates as follow: in each iteration, the algorithm samples each stochastic link and then it calculates the similarity metric for each test link and updates the distributions of the similarity metrics of the test links based on sampled links. Finally, it outputs the distribution of similarity metrics as the output of link prediction. For this experimentation, different synthetic stochastic graphs (BA, WS, and ER) with size from 1000 to 10000 are used and the results of this experimentation for each type of network are averages taken over different used sizes. To do this we consider the different distribution distance metrics: Kolmogorov-Smimov (KS) distance, skew divergence (SD), normalized L distance (ND) and Pearson's correlation coefficient (PCC) for estimated similarity metrics including: Stochastic Common Neighborhood (SCN), Stochastic Salton (SS), Stochastic Jaccard Index (SJ), Stochastic Preferential Attachment (SPA), Stochastic Adamic-Adar Index (SAA), Resource Allocation Index (SRA), Stochastic Katz Index (SKatz), and Stochastic LP Index (SLP) to compare the proposed algorithm with the SSM. The average distance of different similarity metrics is given in Figs. 2–4 for different synthetic stochastic graphs. From Figs. 2–4, our proposed method (SLP) outperforms SSM for all type of datasets, and for all of the proposed similarity metrics. Based on the learning ability of our algorithm to sample important regions, it is not surprising that the proposed algorithm does better in comparison to the blind SSM algorithm.

5.2.2. Experiment 2

This experiment tries to compare the final non-stochastic link prediction result of the proposed method with some classical and recent link prediction methods. In order to improve the comparison, a set of algorithms in two categories is chosen:

- Common Similarity-Based Link Predictions: in this group of methods, the following common similarity-based link prediction methods are chosen for comparison with the proposed method: CN, Salton, Jaccard, PA, AA as local similarity metrics, Katz as the global similarity metric and LP as the quasi-local similarity metric [51].
- Supervised Link Predictions: in this group of algorithms, three recent link prediction algorithms are chosen, Interaction Prediction (IP) [52], CMA-ES [53] and MI-LP [24], that try to predict future links.

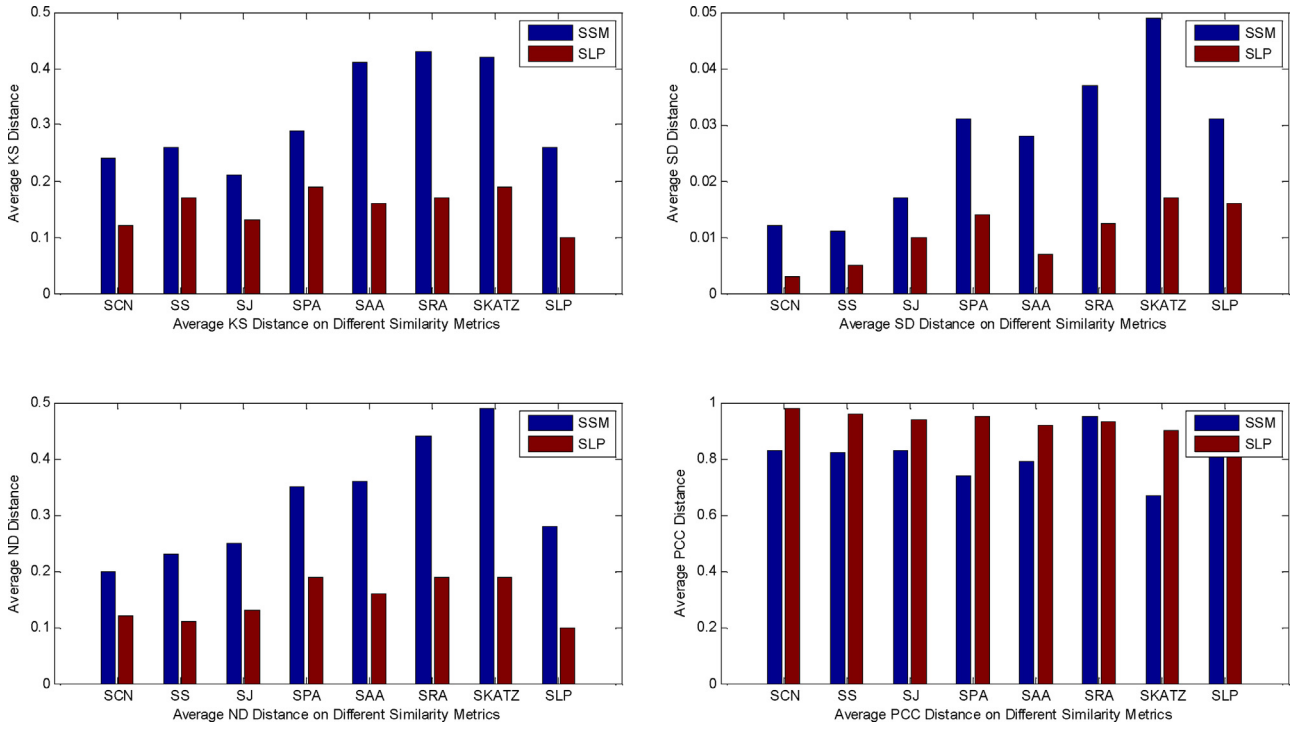


Fig. 2. Comparing average distance metrics of different similarity metrics for synthetic BA graph.

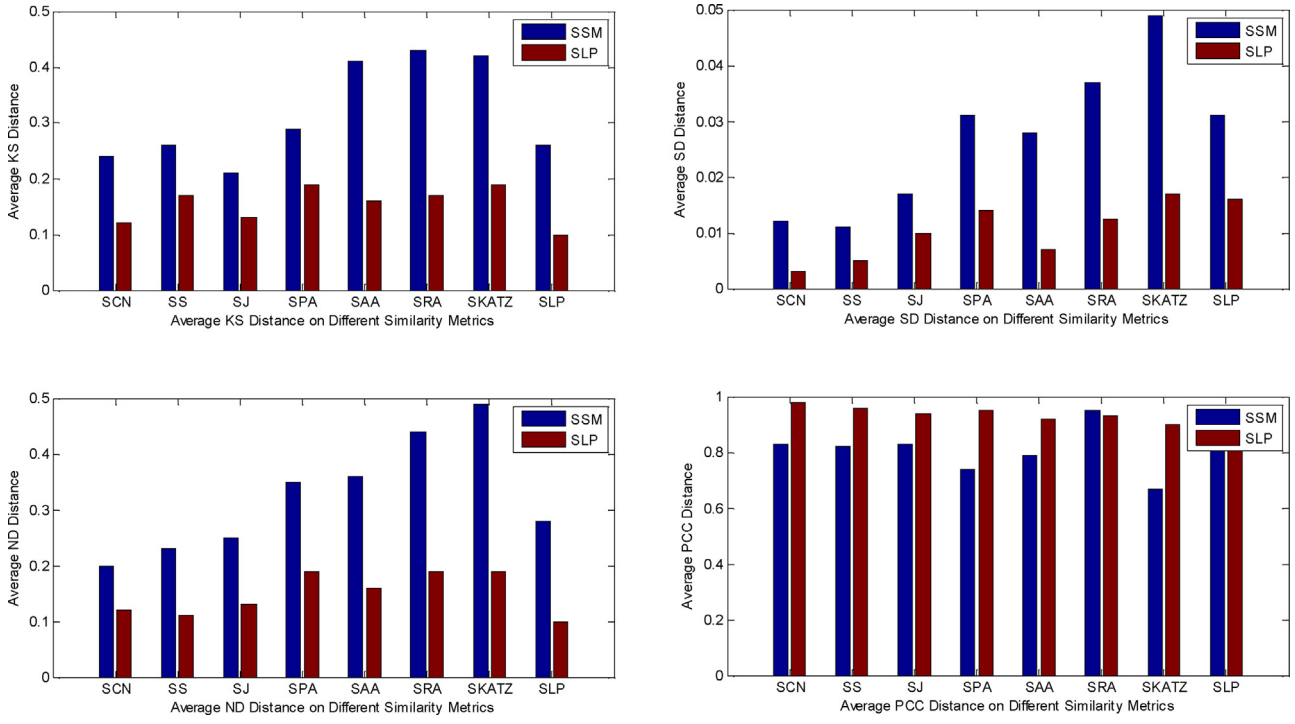


Fig. 3. Comparing average distance metrics of different similarity metrics for synthetic WS graphs.

To evaluate the proposed method using other link prediction methods, we use the two common evaluation metrics as follows:

AUC Metric [51]: If we rank all of the non-existent links based on their scores, the AUC metric can be interpreted as the probability that a random missing link has a higher score than a random non-existent link. In the algorithmic implementation, at each point in time it randomly picks a missing link and a nonexistent link and compares their scores. If from n independent comparisons, there are n′ times that missing links have a higher score and n̈times that they have the same score, the AUC value is:

$$\mathrm{AUC} = \frac{n' + 0.5n''}{n}$$

If the AUC value has a value of more than 0.5, it is better than the random link prediction algorithm; the farther the value from 0.5, the more accurate the algorithm.

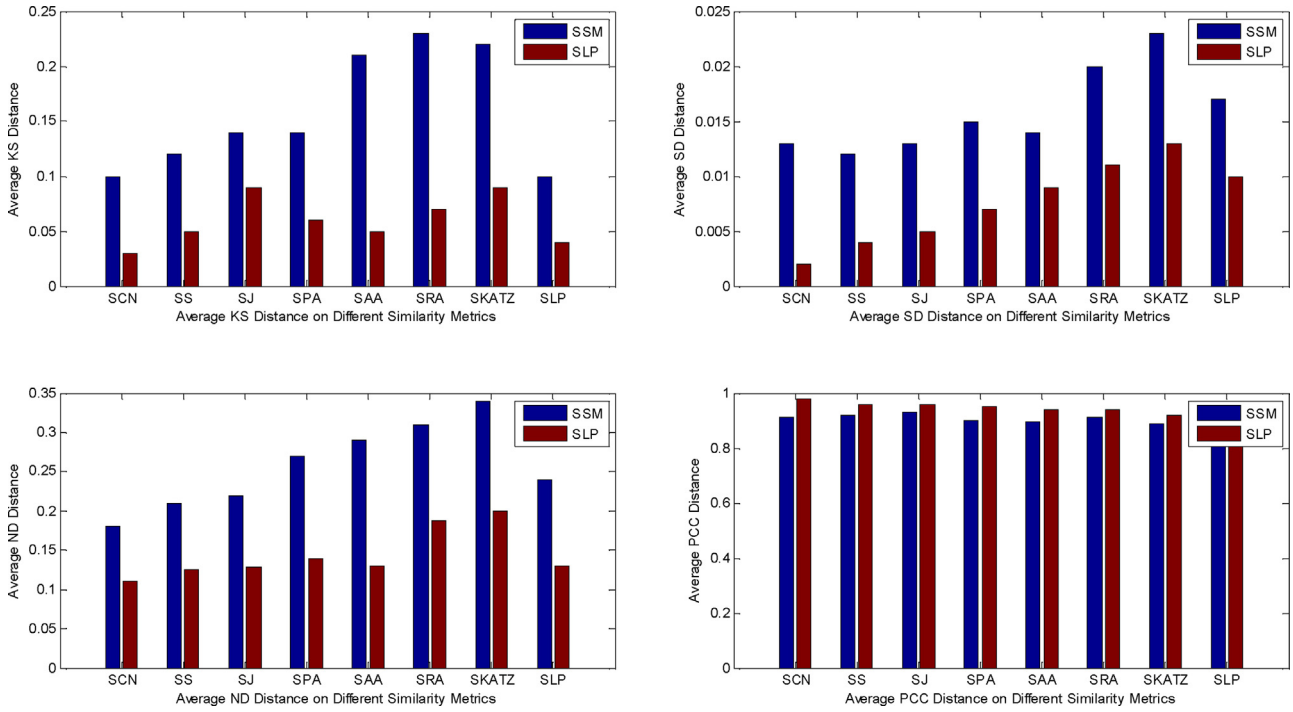


Fig. 4. Comparing average distance metrics of different similarity metrics for synthetic ER graphs.

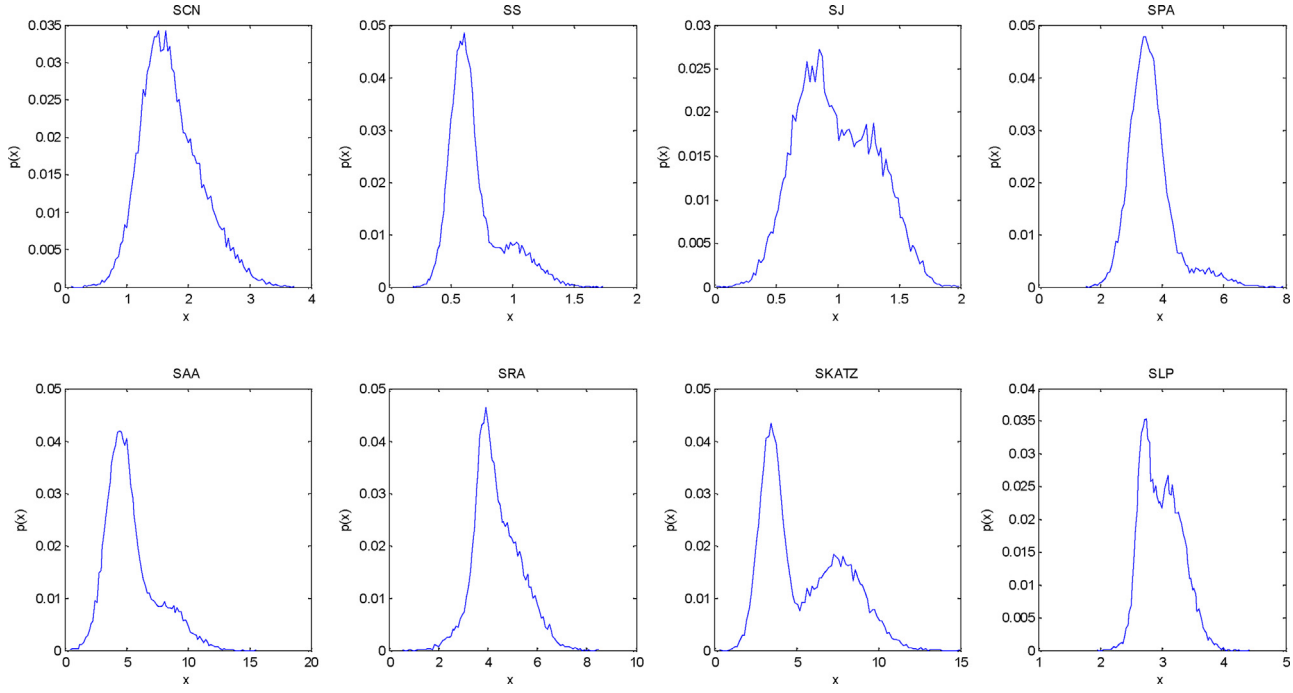


Fig. 5. Different similarity metrics probability distributions of the random test node 1 in an instant run.

Precision [51]: If we predict L links to be connected and L_r links from L links are right, the precision is defined as:

$$\text{Precision} = \frac{L_r}{L}$$

Clearly, higher precision means higher prediction accuracy.

It should be noted that the parameters of the used algorithms are borrowed from their references. Tables 1 and 2 present the average AUC and precision scores based on 30 random runs of the SLP, respectively. It should also be mentioned that for the following tables given in this section, the best results are highlighted. Tables 1 and 2 demonstrate that the proposed SLP is able to achieve AUC (0.9263) and precision (0.8238) measures which are significantly better than local similarity-based algorithms: CN (AUC = 0.7372, precision = 0.5513), Salton (AUC = 0.6946, precision = 0.5873), Jaccard (AUC = 0.7178, precision = 0.6219), PA (AUC = 0.7645, precision = 0.6303), AA (AUC = 0.7617, precision = 0.6672). Given that the proposed stochastic link prediction uses the stochastic information of the social network to predict future links, it is not surpris-

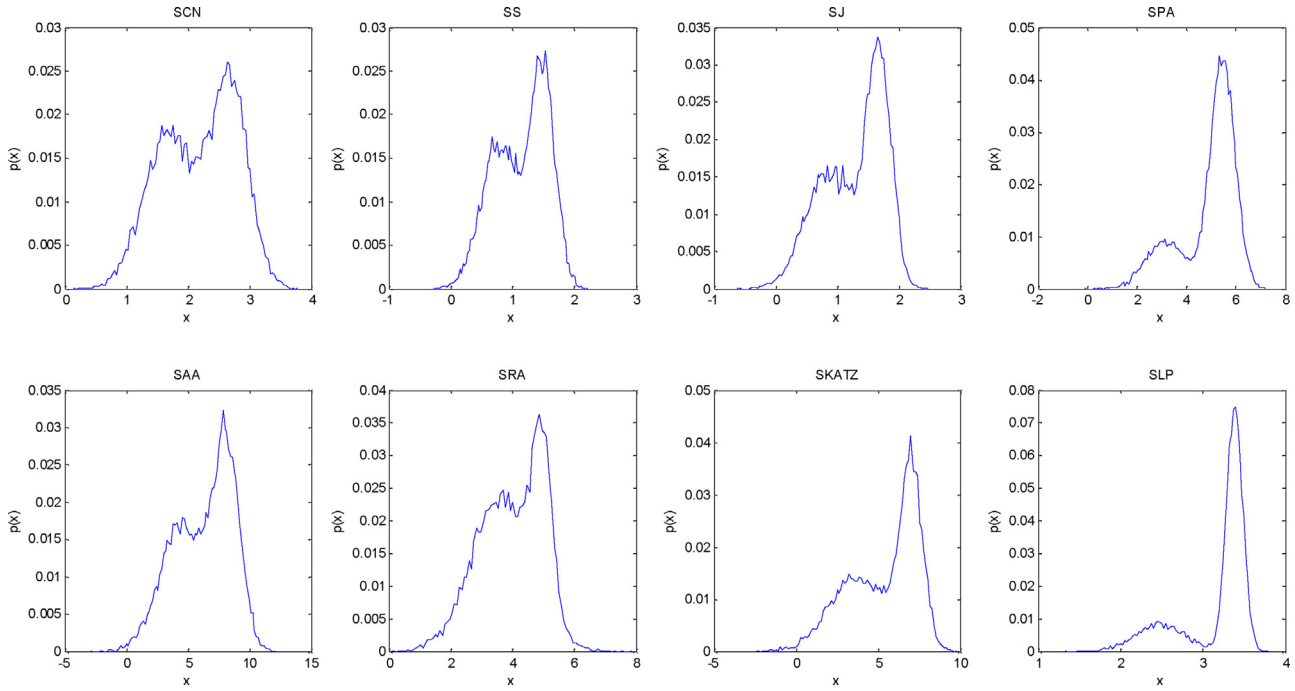


Fig. 6. Different similarity metrics probability distributions of the random test node 2 in an instant run.

Table 1
AUC measure of proposed SLP and other link prediction methods.

Method/Graph	BA-Graph	WS-Graph	ER-Graph	Average
CN	0.7154	0.7541	0.7421	0.7372
Salton	0.6942	0.7055	0.6843	0.6946
Jaccard	0.7328	0.7191	0.7015	0.7178
PA	0.7645	0.7895	0.7684	0.7741
AA	0.7599	0.7611	0.7641	0.7617
Katz	0.8344	0.8457	0.8462	0.8421
LP	0.8001	0.8199	0.8278	0.8159
IP	0.8575	0.8697	0.8799	0.8690
CMA-ES	0.8547	0.8676	0.8723	0.8648
MI-LP	0.8975	0.8978	0.8976	0.8976
SLP	**0.9351**	**0.9145**	**0.9295**	**0.9263**

Table 2
Precision measure of the proposed SLP and other link prediction methods.

Method/Graph	BA-Graph	WS-Graph	ER-Graph	Average
CN	0.5472	0.5571	0.5498	0. 5513
Salton	0.5972	0.5847	0.5801	0. 5873
Jaccard	0.6174	0.6282	0.6201	0. 6219
PA	0.6281	0.6354	0.6274	0. 6303
AA	0.6784	0.6654	0.6579	0. 6672
Katz	0.7185	0.7249	0.7346	0. 7260
LP	0.6978	0.6900	0.6907	0. 6928
IP	0.7549	0.7582	0.7541	0. 7557
CMA-ES	0.7786	0.7719	0.7694	0. 7733
MI-LP	0.7958	0.7964	0.7913	0. 7945
SLP	0.8271	0.8245	0.8200	**0. 8238**

ing that the SLP performs better than local similarity metrics. The proposed method is also superior to the Katz (AUC = 0.8421, precision = 0.7260) and LP (AUC = 0.8159, precision = 0.6928) link predictions. In addition, from Tables 1 and 2 it can be seen that the SLP is able to achieve an AUC and precision which is much better than IP (AUC = 0.8690, precision = 0.7557), CMA-ES (AUC = 0.8976, precision = 0.7733) and MI-LP (AUC = 0.8238, precision = 0.5710).

5.2.3. Experiment 3

In this experiment, we choose three random test nodes in BA graphs and report the final probability distribution of each similarity metric of the corresponding node in an instant run. These reports are presented in Figs. 5–7. Because the links and so the similarity metrics in the stochastic social networks are random variables and the proposed link prediction tries to estimate the true distribution of the similarity metric for each test link so we report the final similarity metric distributions for three instant nodes. The goal of these figure is to visualize the output of the proposed method as a stochastic method. As it mentioned before at the end of SLP, we have a similarity probability distribution function for each test link and because of the similarity metrics are random variable, to determine a link's existence, we should sample from its similarity probability distribution function and based on this sample and a predefined threshold we make the decision about the link existence. Also we try to present the fitted distribution for each link using non-parametric kernel method. So Figs. 8–10 show the histogram of each random link with fitted distribution using non-parametric kernel method. This experiment is conducted here to get an idea of the final output of the SLP.

5.2.4. Experiment 4

In this experiment we try to compare the number of samples that is taken by the proposed LA with the number of samples taken by SSM to reach a certain accuracy. To do this we consider three accuracy values: {0.65, 0.80, 0.95} and run the SLP and SSM on different synthetic networks BA, WS, and ER with different sizes. We report the average of required samples of the SLP and SSM for each synthetic graphs that is obtained using 30 instant runs (10 instant runs for each size of {2000, 5000, 10000}) in Table 3. From the results it can be found the SLP decreases the rate of required samples by 50% which shows that the proposed SLP is completely better than SSM method. Also to evaluate the performance of the LA_{Tests} set, we compare execution time of the proposed SLP with the SSM and the SLP-Without-LA_{Tests} (SLP that don't have LA_{Tests} set and do not learn anything about if we should calculate and update the similarity metrics probability distributions in each iteration

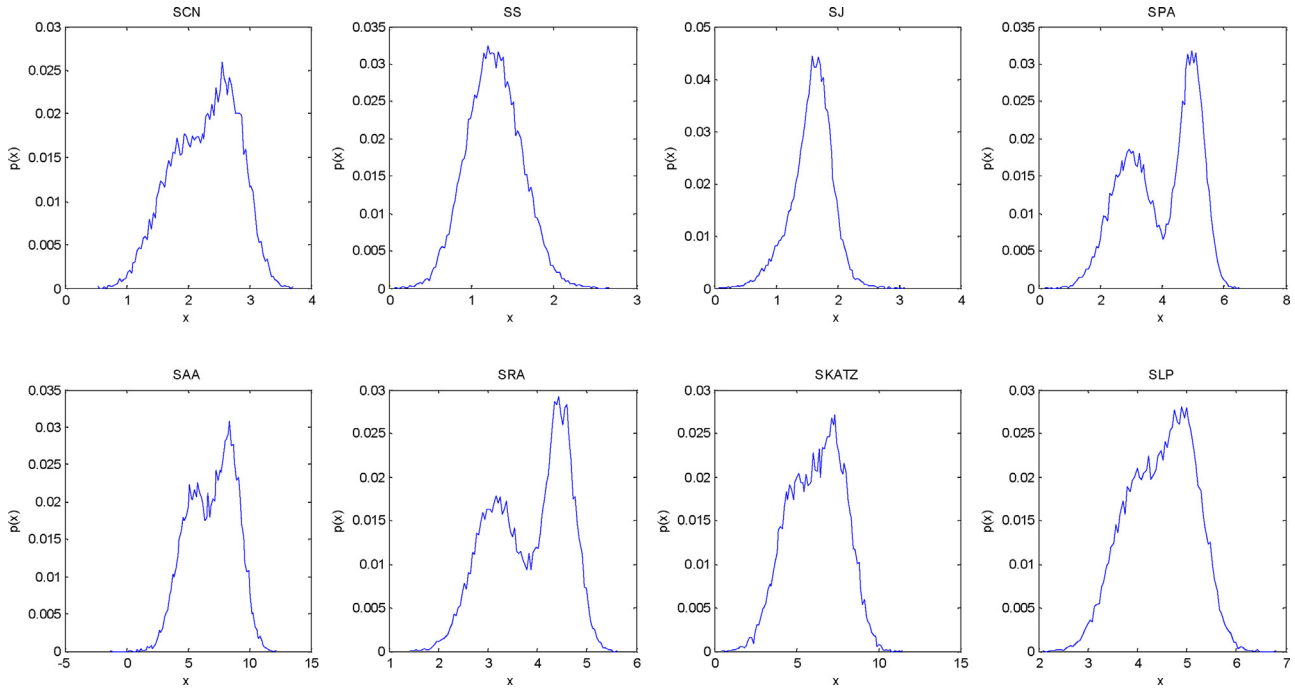


Fig. 7. Different similarity metrics probability distributions of the random test node 3 in an instant run.

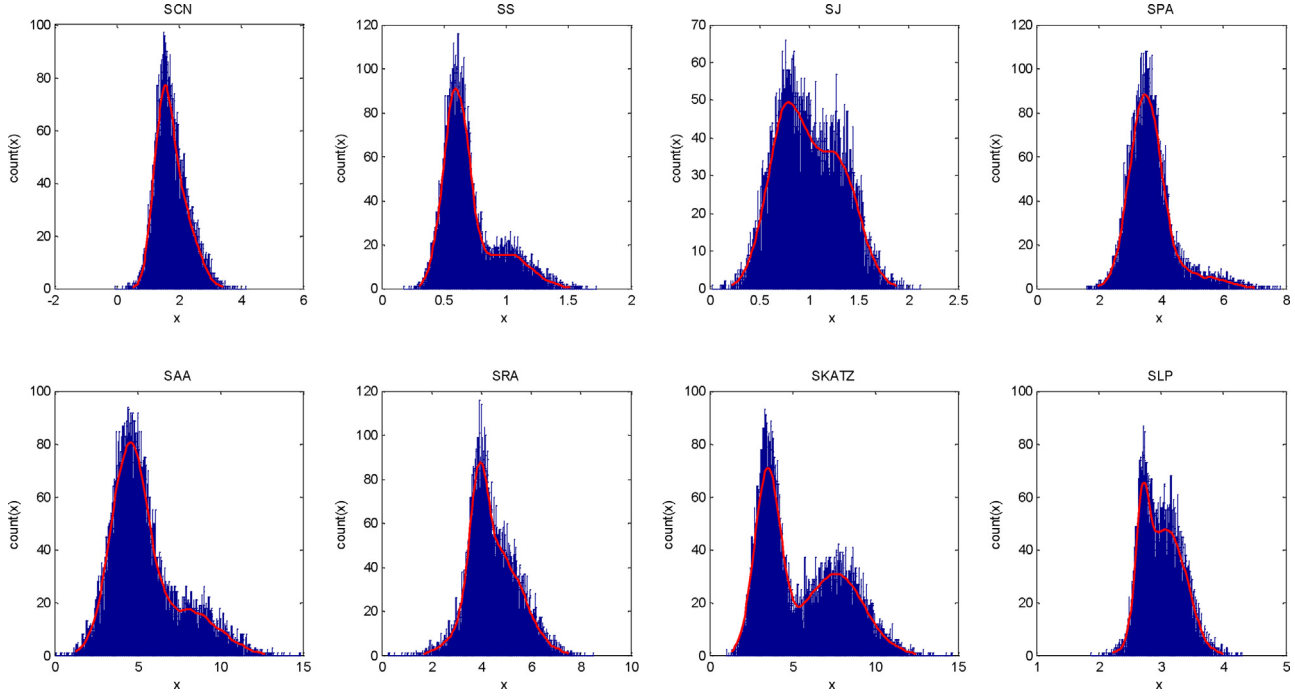


Fig. 8. Different similarity metrics histogram with a fitted distribution using kernel method for node 1.

Table 3
Comparison of the number of samples required by SLP and SSM based on different synthetic graphs and different accuracy levels.

Method/Graph	SLP Accuracy = 0.65	SSM Accuracy = 0.65	SLP Accuracy = 0.80	SSM Accuracy = 0.80	SLP Accuracy = 0.95	SSM Accuracy = 0.95
BA-Graph	$7.84 \times 10^4 \pm 9.7 \times 10^2$	$19.25 \times 10^4 \pm 31 \times 10^2$	$12.87 \times 10^4 \pm 11 \times 10^2$	$24.12 \times 10^4 \pm 45 \times 10^2$	$16.93 \times 10^4 \pm 15 \times 10^2$	$35 \times 10^4 \pm 61 \times 10^2$
WS-Graph	$7.12 \times 10^4 \pm 8.8 \times 10^2$	$21.41 \times 10^4 \pm 45 \times 10^2$	$12.14 \times 10^4 \pm 10 \times 10^2$	$26.11 \times 10^4 \pm 49 \times 10^2$	$15.55 \times 10^4 \pm 14 \times 10^2$	$39 \times 10^4 \pm 58 \times 10^2$
ER-Graph	$7.99 \times 10^4 \pm 6.9 \times 10^2$	$20.89 \times 10^4 \pm 24 \times 10^2$	$13.02 \times 10^4 \pm 9.1 \times 10^2$	$24.82 \times 10^4 \pm 52 \times 10^2$	$16.12 \times 10^4 \pm 16 \times 10^2$	$38 \times 10^4 \pm 63 \times 10^2$

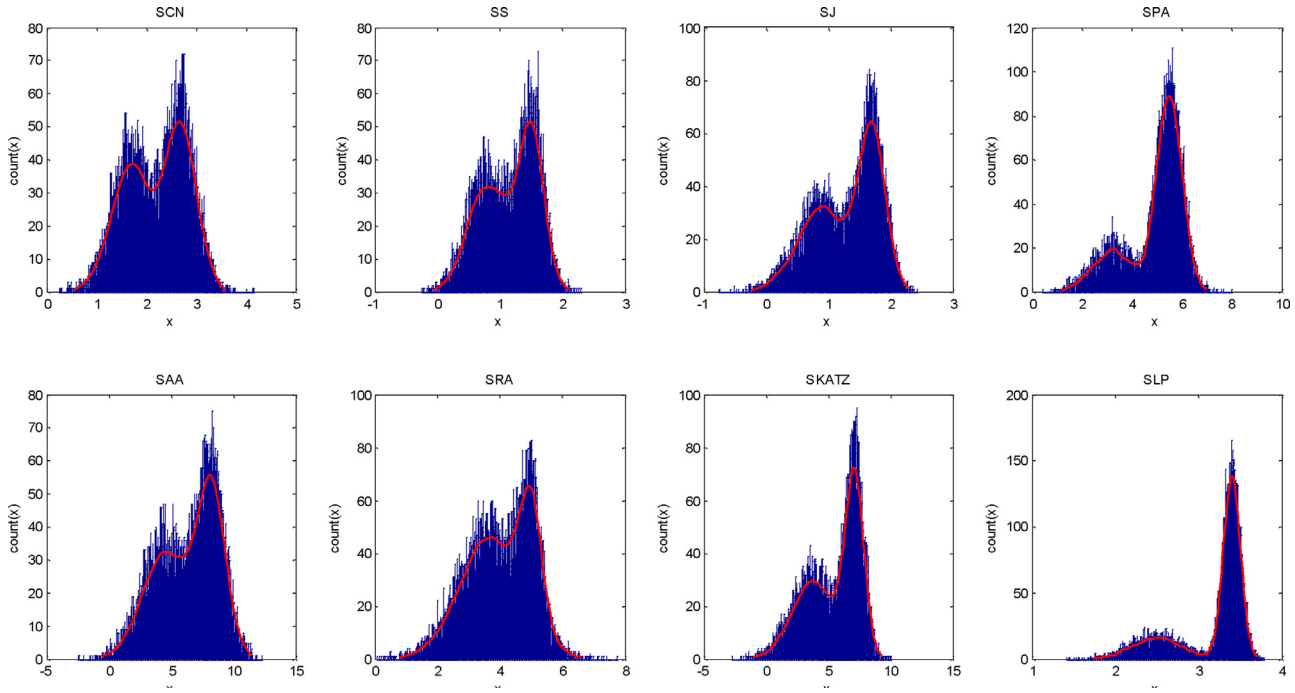


Fig. 9. Different similarity metrics histogram with a fitted distribution using kernel method for node 2.

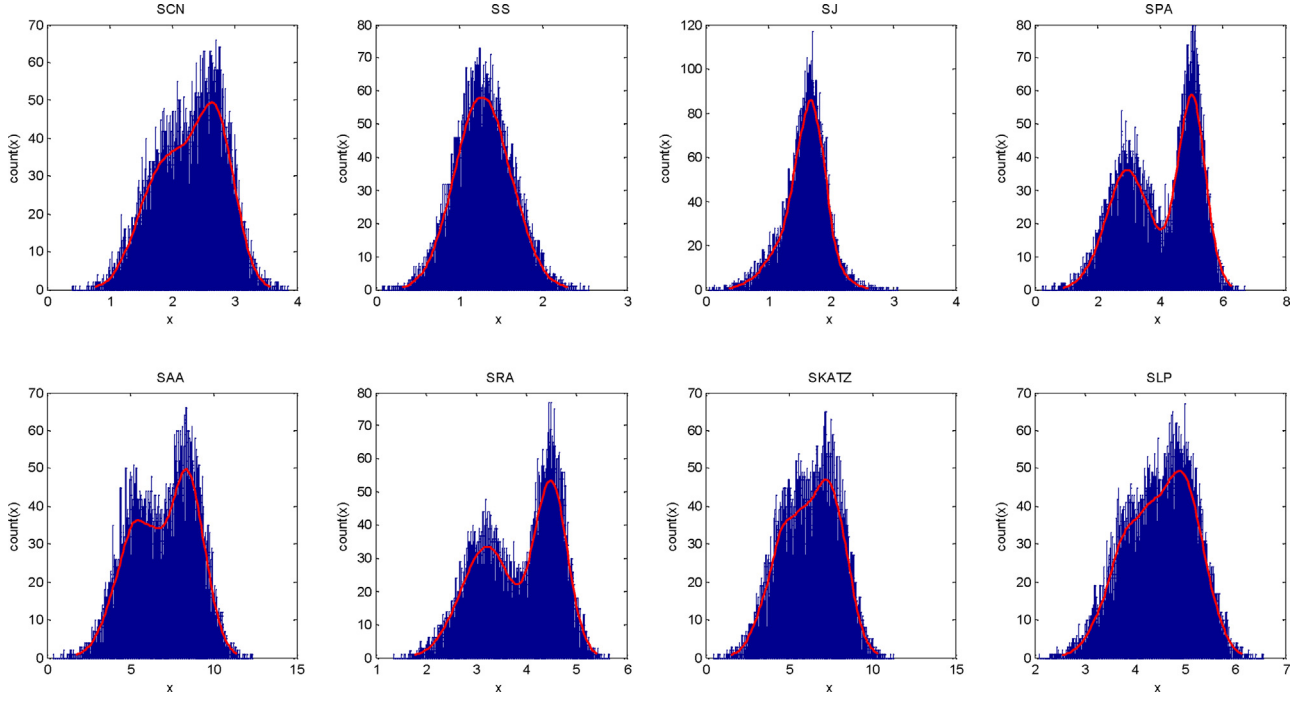


Fig. 10. Different similarity metrics histogram with a fitted distribution using kernel method for node 3.

or not) and report the average of execution time over 30 instant runs in Table 4 for different sizes of synthetic graphs and accuracy level 0.95. From the results reported here we see that using LA_{Tests} completely decreases the computational complexity of calculating similarity metric probability distributions.

5.2.5. Experiment 5

In this experiment, we evaluate the convergence behavior of the SLP using information entropy and three different LA configurations. To do this we use the three following configurations and calculate the average information entropy of the probability vector of the LAs: $a = 0.05$, $b = 0.005$, $a = 0.05$, $b = 0.003$, $a = 0.01$, $b = 0.001$. So, Fig. 11 plots the average information entropy taken over LAs versus iteration number for three kinds of synthetic stochastic graphs (BA-2000, BA-5000, BA-10000; WS-2000, WS-5000, WS-10000; ER-2000, ER-5000, ER-10000). These figures show that the average entropy of SLP decreases generally where the algorithm proceeds. Also based on these figures we can conclude that the con-

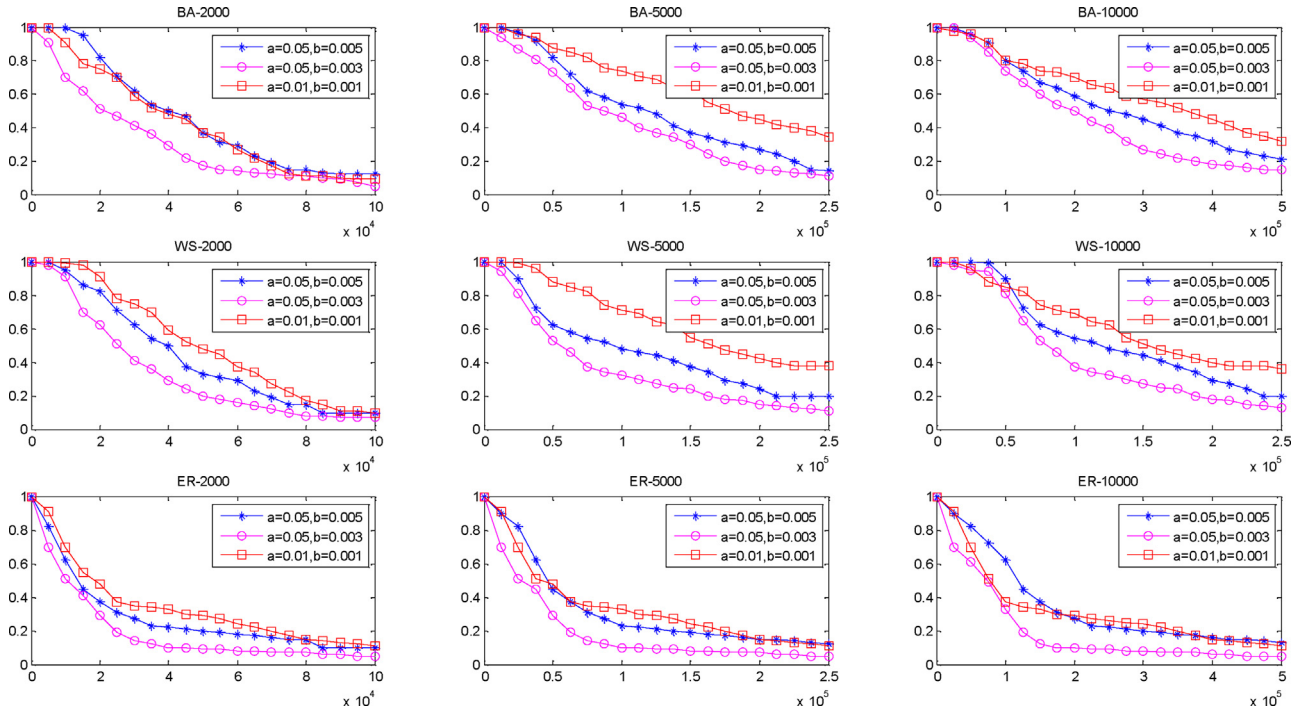


Fig. 11. The convergence diagram of the proposed SLP using different LA configurations.

Table 4
Average execution time (in minutes) of the SLP, SLP-Without LA_{Tests}, and the SSM based on different synthetic graphs.

Method/Graph	SLP	SLP-Without LA_{Tests}	SSM
BA-2000	10.52 ± 2.10	20.45 ± 3.40	45.10 ± 5.51
BA-5000	14.13 ± 2.33	26.22 ± 3.23	60.47 ± 6.47
BA-10000	18.39 ± 3.45	33.29 ± 4.20	91.42 ± 9.30
WS-2000	11.41 ± 1.58	21.21 ± 3.10	42.17 ± 4.26
WS-5000	15.46 ± 2.58	27.01 ± 3.40	62.13 ± 6.59
WS-10000	20.00 ± 3.51	35.18 ± 4.37	93.39 ± 9.51
ER-2000	10.35 ± 2.23	20.48 ± 3.22	44.24 ± 5.10
ER-5000	14.40 ± 3.01	27.32 ± 3.58	34.11 ± 7.15
ER-10000	21.12 ± 3.40	34.50 ± 4.51	95.02 ± 10.03

figuration $a = 0.05$, $b = 0.003$ has the best convergence results and for this reason we use this configuration for the other experiments as mentioned before.

5.2.6. Experiment 6

In this experiment, we evaluate the performance of the proposed SLP using different distance metrics. To do this we use the following distance metrics and calculate the running time and the AUC metric of the proposed SLP: Kolmogorov-Smirnov distance, Skew divergence, Pearson's correlation coefficient, and Normalized L1 distance. So, Table 5 represents the AUC metric and the running time of the proposed SLP for three synthetic stochastic graphs BA-5000, WS-5000, and ER-5000. Table 5 shows that the SLP with skew divergence has the best performance in term of accuracy and running time and for this reason we use this metric for the other experiments as mentioned before.

6. Conclusion

Because of people activities in social networks are dynamic and uncertainty, and the structure of the networks change over time, deterministic graphs may not be appropriate for modeling and analysis of the social network. The one of solutions is using stochastic social networks as a model to simulate the dynamics and uncertainty of the social networks. This paper tries to present a new similarity based link prediction method for stochastic social networks to overcome the problem of traditional link prediction that only uses a static snapshot of the graph to predict future links. The proposed method takes advantage of using learning automata to estimate the distribution of some chosen similarity metric in such a way that the number of required samples and also the computational complexity be decreased as much as possible. Also, the proposed method has capability to use in online stochastic social networks where the social network changes online and the future links must be predicted again. We consider different synthetic stochastic social networks as the test bench mark and conduct some experiments. The experiments showed that the proposed stochastic link prediction improves the link prediction accuracy and the computational complexity in the stochastic networks.

Table 5
AUC and Running time (in minutes) of the SLP using different distance metrics.

Metric/Graph	BA-5000		WS-5000		ER-5000	
	Running Time	AUC	Running Time	AUC	Running Time	AUC
Kolmogorov-Smirnov	14.23	0.9015	15.51	0.8774	15.16	0.9001
Skew divergence	13.59	0.9481	14.47	0.9249	13.51	0.9399
Pearson's correlation coefficient	17.51	0.9359	18.12	0.9314	17.52	0.9263
Normalized L1	10.51	0.8723	11.27	0.8641	11.02	0.8794

References

[1] Lü Linyuan, Tao Zhou, Link prediction in complex networks: a survey, Physica A 390 (6) (2011) 1150–1170.
[2] Mohammad Al Hasan, J. Zaki Mohammed, A Survey of Link Prediction in Social Networks. Social Network Data Analytics, Springer, US, 2011, pp. 243–275.
[3] Sisay Fissaha Adafre, Maarten de Rijke, Discovering missing links in Wikipedia, in: Proceedings of the 3rd International Workshop on Link Discovery, ACM, 2005.
[4] Jianhan Zhu, Jun Hong, John G. Hughes, Using Markov models for web site link prediction, in: Proceedings of the Thirteenth ACM Conference on Hypertext and Hypermedia, ACM, 2002.
[5] Li Xin, Hsinchun Chen, Recommendation as link prediction: a graph kernel-based machine learning approach, in: Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM, 2009.
[6] Zan Huang, Xin Li, Hsinchun Chen, Link prediction approach to collaborative filtering, in: Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries, ACM, 2005.
[7] Bradley Malin, Edoardo Airoldi, M. Kathleen, Carley A network analysis model for disambiguation of names in lists, Comput. Math. Organiz. Theory 11 (2) (2005) 119–139.
[8] Ahmed K. Elmagarmid, Panagiotis G. Ipeirotis, Vassilios S. Verykios, Duplicate record detection: a survey, IEEE Trans. Knowl. Data Eng. 19 (1) (2007) 1–16.
[9] A. Rezvanian, M.R. Meybodi, Stochastic graph as a model for social networks, Comput. Human Behav. 64 (2016) 621–640.
[10] H. Beigy, M.R. Meybodi, Utilizing distributed learning automata to solve stochastic shortest path problems International Journal of Uncertainty, Fuzziness Knowl.-Based Syst. 14 (05) (2006) 591–615.
[11] A. Rezvanian, M.R. Meybodi, Finding minimum vertex covering in stochastic graphs: a learning automata approach, Cybern. Syst. 46 (8) (2015) 698–727.
[12] A. Rezvanian, M.R. Meybodi, Finding maximum clique in stochastic graphs using distributed learning automata, Int. J. Uncertainty Fuzziness Knowledge Based Syst. 23 (01) (2015) 1–31.
[13] J.A. Torkestani, M.R. Meybodi, A learning automata-based heuristic algorithm for solving the minimum spanning tree problem in stochastic graphs, J. Supercomput. 59 (2) (2012) 1035–1054.
[14] Thathachar Mandayam A.L, P. Shanti Sastry, Varieties of learning automata: an overview, IEEE Trans. Syst. Man Cybern., Part B (Cybern.) 32 (6) (2002) 711–722.
[15] M. Al Hasan, M.J. Zaki, A Survey of Link Prediction in Social Networks In Social Network Data Analytics, Springer, 2011, pp. 243–275.
[16] T. Murata, S.1 Moriyasu, Link prediction based on structural properties of online social networks, New Gen. Comput. 26 (3) (2008) 245–257.
[17] M. Al Hasan, V. Chaoji, S. Salem, M. Zaki, Link prediction using supervised learning, SDM'06: Workshop on Link Analysis, Counter-terrorism and Security (2006).
[18] Gerard Salton, Michael J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, Inc., New York, NY USA, 1986.
[19] Albert-László Barabási, Réka Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[20] Mark EJ. Newman, Clustering and preferential attachment in growing networks, Phys. Rev. E 64 (2) (2001) 025102.
[21] Lada A. Adamic, Eytan Adar, Friends and neighbors on the web, Soc. Networks 25 (3) (2003) 211–230.
[22] G. Rossetti, R. Guidotti, D. Pennacchioli, D. Pedreschi, F. Giannotti, Interaction prediction in dynamic networks exploiting community discovery, Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2015 (2015) 553–558 (s.n.).
[23] C.A. Bliss, M.R. Frank, C.M. Danforth, P.S. Dodds, An evolutionary algorithm approach to link prediction in dynamic social networks, J. Comput. Sci. 5 (5) (2014) 750–764.
[24] Fei Tan, Yongxiang Xia, Boyao Zhu, Link prediction in complex networks: a mutual information perspective, PLoS One 9 (9) (2014) e107056.
[25] A. Ozcan, S.G. Oguducu, Multivariate temporal Link Prediction in evolving social networks, Computer and Information Science (ICIS), 2015 IEEE/ACIS 14th International Conference (2015) 185–190.
[26] S. Huang, Y. Tang, F. Tang, J. Li, Link prediction based on time-varied weight in co-authorship network in Computer Supported Cooperative Work in Design (CSCWD), Proceedings of TheIEEE 18th International Conference On, 2014 (2014) 706–709.
[27] B. Moradabadi, M.R. Meybodi, Link prediction based on temporal similarity metrics using continuous action set learning automata, Physica A 460 (2016) 361–373.
[28] B. Moradabadi, M.R. Meybodi, Link prediction in fuzzy social networks using distributed learning automata, Appl. Intell. (2017) 1–13.
[29] S.A. Golder, D.M. Wilkinson, B.A. Huberman, Rhythms of social interaction: messaging within a massive online network, in: Communities and Technologies, Springer, 2007, pp. 41–66.
[30] J. Leskovec, M. McGlohon, C. Faloutsos, N.S. Glance, M. Hurst, Patterns of Cascading behavior in large blog graphs, SDM 7 (2007) 551–556.
[31] J. Leskovec, L. Backstrom, R. Kumar, A. Tomkins, Microscopic evolution of social networks, Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2008) 462–470.
[32] A. Nazir, S. Raza, C.-N. Chuah, Unveiling facebook: a measurement study of social network based applications, Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement (2008) 43–56.
[33] L. Guo, E. Tan, S. Chen, X. Zhang, Y.E. Zhao, Analyzing patterns of user content generation in online social networks, Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2009) 369–378.
[34] H. Kwak, C. Lee, H. Park, S. Moon, What is Twitter, a social network or a news media? In Proceedings of the 19th International Conference on World Wide Web (2010) 591–600.
[35] L. Gyarmati, T.A.1 Trinh, Measuring user behavior in online social networks, IEEE Network 24 (5) (2010) 26–31.
[36] W. Galuba, K. Aberer, D. Chakraborty, Z. Despotovic, W. Kellerer, Outtweeting the twitterers-predicting information cascades in microblogs, WOSN 10 (2010) 3–11.
[37] Q. Yan, L. Wu, L. Zheng, Social network based microblog user behavior analysis, Physica A 392 (7) (2013) 1712–1723.
[38] H. Liu, A. Nazir, J. Joung, C.-N. Chuah, Modeling/predicting the evolution trend of osn-based applications, Proceedings of the 22nd International Conference on World Wide Web (2013) 771–780.
[39] S. Vongsingthong, S. Boonkrong, M. Kubek, H. Unger, On the distributions of user behaviors in complex online social networks, in: Recent Advances in Information and Communication Technology 2015, Springer, 2015, pp. 237–246.
[40] D.R. Bild, Y. Liu, R.P. Dick, Z.M. Mao, D.S.1 Wallach, Aggregate characterization of user behavior in twitter and analysis of the retweet graph, ACM Trans. Internet Technol. (TOIT) 15 (1) (2015) 4.
[41] Mandayam A.L. Thathachar, S. Pidaparty, Sastry Networks of Learning Automata: Techniques for Online Stochastic Optimization, Springer, 2003.
[42] Kumpati S. Narendra, Mandayam A.L. Thathachar, Learning Automata: an Introduction, Courier Corporation, 2012.
[43] A. Rezvanian, M.R. Meybodi, An adaptive mutation operator for artificial immune network using learning automata in dynamic environments, Nature and Biologically Inspired Computing (NaBIC), 2010 Second World Congress (2010) 479–483.
[44] A. Rezvanian, M.R. Meybodi, LACAIS: learning automata based cooperative artificial immune system for function optimization, in: Contemporary Computing, Springer, 2010, pp. 64–75.
[45] A. Rezvanian, M. Rahmati, M.R. Meybodi, Sampling from complex networks using distributed learning automata, Physica A 396 (2014) 224–234.
[46] A. Mousavian, A. Rezvanian, M.R. Meybodi, Cellular learning automata based algorithm for solving minimum vertex cover problem, 2014 22nd Iranian Conference on Electrical Engineering (ICEE) (2014) 996–1000.
[47] D.R. Bild, Y. Liu, R.P. Dick, Z.M. Mao, D.S.1 Wallach, Aggregate characterization of user behavior in twitter and analysis of the retweet graph, ACM Trans. Internet Technol. (TOIT) 15 (1) (2015) 4.
[48] Z.S. Jalali, A. Rezvanian, M.R. Meybodi, Social network sampling using spanning trees, Int. J. Mod. Phys. C 27 (05) (2016) 1650052.
[49] P. Luo, Y. Li, C. Wu, G. Zhang, Toward cost-efficient sampling methods, Int. J. Mod. Phys. C 26 (05) (2015) 1550050.
[50] Z.S. Jalali, A. Rezvanian, M.R. Meybodi, A two-phase sampling algorithm for social networks, 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI) (2015) 1165–1169.
[51] Lü Linyuan, Tao Zhou, Link prediction in complex networks: a survey, Physica A 390 (6) (2011) 1150–1170.
[52] G. Rossetti, R. Guidotti, D. Pennacchioli, D. Pedreschi, F. Giannotti, 'Interaction prediction in dynamic networks exploiting community discovery', Proceedings Ofthe 2015 IEEE/ACM International Conference on Advances in Social Networks Analysi S and Mining, 2015 (2015) 553–558 (s.n).
[53] C.A. Bliss, M.R. Frank, C.M. Danforth, P.S. Dodds, An evolutionary algorithm approach to link prediction in dynamic social networks, J. Comput. Sci. 5 (no. 5) (2014) 750–764.



Behnaz Moradabadi received the B.S. from Tabriz University (2009), Tehran, Iran, M.S. from Sharif University (2011), Tehran, Iran, and she is now Ph.D. student in computer engineering in the Amirkabir University of Technology, Tehran,I ran (from 2013). Her current research interests include social networks, learning systems, soft computing, and information retrieval.



Mohammad Reza Meybodi received the B.S. and M.S. degrees in economics from Shahid Beheshti University, Tehran, Iran, in 1973 and 1977, respectively, and the M.S. and Ph.D. degrees in computer science from Oklahoma University, Norman, OK, USA, in 1980 and 1983, respectively. He is currently a Full Professor with Computer Engineering Department, Amirkabir University of Technology, Tehran. Prior to his current position, he was an Assistant Professor with Western Michigan University, Kalamazoo, MI, USA, from 1983 to 1985, and an Associate Professor with Ohio University, Athens, OH, USA, from 1985 to 1991. His current research interests include channel management in cellular networks, learning systems, parallel algorithms, soft computing, and software development.