

Research Article

Link Prediction Based on the Derivation of Mapping Entropy

Hefei Hu ¹, Yanan Wang,¹ Zheng Li,² Yang Tian ³ and Yuemei Ren⁴

¹School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

²Network Department of China Mobile Wenzhou Branch, Wenzhou 325000, China

³State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

⁴School of Electronic Information Engineering, Henan Polytechnic Institute, Nanyang 473000, China

Correspondence should be addressed to Hefei Hu; hufefei@bupt.edu.cn

Received 10 June 2021; Accepted 27 July 2021; Published 2 August 2021

Academic Editor: Fei Xiong

Copyright © 2021 Hefei Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The algorithms based on topological similarity play an important role in link prediction. However, most of traditional algorithms based on the influences of nodes only consider the degrees of the endpoints which ignore the differences in contribution of neighbors. Through generous explorations, we propose the DME (derivation of mapping entropy) model concerning the mapping relationship between the node and its neighbors to access the influence of the node appropriately. Abundant experiments on nine real networks suggest that the model can improve precision in link prediction and perform better than traditional algorithms obviously with no increase in time complexity.

1. Introduction

A large number of complicated systems in nature can be described by complex networks [1]. The nodes in the network represent the individuals in the real system, and the edges connecting two nodes represent the relationships of the individuals in the real system. The existing networks in the world can be divided into social networks, biological networks [2, 3], and so on. Link prediction evaluates the possibility of the link between two nodes in the network by the known network structure or node attributes. Through link prediction, we can find the links existing but unknown in the network which misses some data. Besides, we can also predict the possible links in the coming evolution of the network [4]. Link prediction plays an important role in practical application. For instance, through link prediction, the unknown interaction between proteins is predicted, which avoids the high experimental cost [5]. Furthermore, it also plays a role in user recommendation [6].

In the early period, most researchers engaged in link prediction focused on the similarity of the attributes of nodes such as age, occupation, interest, and so on [2, 7] to

judge the possibility of links. This method can achieve high-precision prediction. However, it is difficult to extract the attributes of nodes in complex networks, and the reliability of information is hard to assure [8]. So, researchers turned their attention to the study of network structure [9, 10] which has relatively low computational complexity.

Algorithms based on the similarity of the network structure can be divided into three categories: local similarity algorithm, global similarity algorithm, and quasilocal similarity algorithm according to the path length [11]. The core idea of local similarity is common neighbors. On this basis, considering the influence of endpoints from different angles, a variety of local similarity indices is derived. For instance, common neighbor (CN) index [12] considers that if two nodes have more common neighbors, then they are more likely to have connected edges. Adamic-Adar (AA) index [13] holds the idea that the common neighbor with small degree has a greater contribution. Accordingly, each node is given a weight. Because these local similarity indices only consider the local structure of the network which lead to low precision, the third-order and higher-order path similarity indices were proposed such as Katz index [14]. Katz index

considers all paths of two unconnected nodes in the case of short path priority. There is no doubt that this greatly increases the computational complexity. Compared with the above two kinds of algorithms, the quasilocal similarity algorithm with moderate complexity and precision is more and more widely applied. Superposed random walk (SRW) index [15] is one of the indices based on the Markov model.

In traditional algorithms, only the degree of endpoint is considered when we evaluate the influence of it. This considers the influences of neighbor nodes to the same extent which loses the impacts of indirect neighbors [16, 17]. In fact, due to the different degrees of neighbors, their influence on the endpoint should be different. The larger the degree, the greater the influence. However, taking global nodes into account will increase the complexity of the algorithm, and the result is not necessarily good. Because the influence of endpoint is limited, it only has great influence on nearby neighbors. Therefore, this paper proposes to use the derivation of mapping entropy (DME) of node to represent the influence, which represents the mapping relation between a node and its neighbors. It considers not only the weight of the node but also the weight of its neighbors. Figure 1 shows a clear illustration.

On the basis of above discussion, we improve the SRW model, taking the influence of indirect neighbors into account. Through extensive experiments on nine complex networks, the results show that DME can achieve higher precision than traditional algorithms in most cases.

The rest of paper is organized as follows. In Section 2, we propose a new model based on the DME index. In Section 3, we introduce 9 complex networks and experimental approaches. In Section 4, five classic models are introduced as reference. In Section 5, results and analysis are presented. In Section 6, we arrive at a conclusion of our study.

2. Model Based on the Derivation of Mapping Entropy

2.1. Network Model. $G(V, E)$ is defined as a network, where V is the node set and E is the edge set. The total number of nodes is N and the total number of edges is E . The universal set U can have $(N \times (N - 1))/2$ links. The method of link prediction is to give a score s_{xy} to each pair of unconnected nodes which indicates the likelihood of connecting the two nodes. Then, all unconnected nodes are arrayed in descending order of score. The node pair in the top represents that the two nodes are the most likely to generate a connection. In order to test the precision of the algorithm, the known edge set E is divided into training set E^T and testing set E^P . Only the testing set can be used to calculate scores. Obviously, $E = E^T \cup E^P$ and $E^T \cap E^P = \phi$. We define an edge belonging to U but not to E as an inexistent edge. In this paper, we use precision [18] to measure the accuracy of link prediction algorithm, which describes the proportion of real links in the top- L links with highest scores. If there are m real links in top- L links, the precision of the algorithm can be expressed as

$$\text{precision}(L) = \frac{l}{L}. \quad (1)$$

In order to simplify the model, we use undirected and unweighted networks.

2.2. Superposed Random Walk (SRW) Model. The SRW model inspired from the LRW model considers random walk between endpoint x and y , making the nodes nearby more likely to connect to the target node [19]. It is defined as

$$s_{xy}^{\text{SRW}}(t) = \sum_{l=2}^t \left[\frac{k_x}{2|E|} \times \pi_{xy}(l) + \frac{k_y}{2|E|} \times \pi_{yx}(l) \right], \quad (2)$$

where the initial density vector $\vec{\pi}_{xy}(0) = \vec{e}_x$ and it evolves as $\vec{\pi}_{xy}(t+1) = P^T \times \vec{\pi}_{xy}(t)$. P represents the probability transition matrix with $p_{xy} = (a_{xy}/k_x)$, and $a_{xy} = 1$ when the link exists; if not, $a_{xy} = 0$. Besides, t denotes the time steps.

2.3. Derivation of Mapping Entropy (DME) Model. Inspired by Shannon entropy, the information entropy [20] of the network can be expressed as

$$E = - \sum_{i=1}^N \text{DC}_i \log \text{DC}_i, \quad (3)$$

where DC_i is the degree centrality of node i . A node and its neighbors construct a subnetwork. The local entropy (LE) [21] of the subnetwork originated from endpoint v_i is shown in the following formula:

$$\text{LE}_i = - \sum_{j=1}^M \text{DC}_j \log \text{DC}_j, \quad (4)$$

where DC_j is the degree centrality of node v_j , which belongs to the neighbor set M of node v_i . Taking the mapping relation between a node and its neighbors into account, we can obtain the mapping entropy (ME):

$$\text{ME}_i = -\text{DC}_i \sum_{j=1}^M \log \text{DC}_j, \quad (5)$$

where DC_i is the degree centrality of node v_i and DC_j is the degree centrality of one of the neighbors of node v_i .

Inspired by ME index, we introduce the derivation of mapping entropy: DME, which is defined by interleaving the degrees of node v_i and v_j .

$$\text{DME}_i = k_i \sum_{j=1}^M \lg k_j. \quad (6)$$

The definition considers both the degrees of the node and the degrees of its neighbors which takes the influence of indirect neighbors into account. This may be useful for distinguishing the importance of neighbors. Based on the SRW model, we consider using the DME index to replace the influence of the endpoint, which can perform better than the ME model introduced later through experiments based on the superposed random walk. The model is defined as

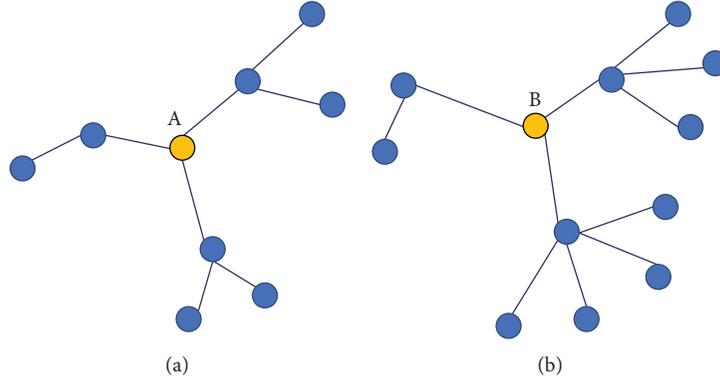


FIGURE 1: Sketch maps of influence based on the derivation of mapping entropy. As is shown, the degree of node A in subgraph (a) is equal to the degree of node B in subgraph (b). In traditional models, they are considered to have the same influence. Nevertheless, in the DME model, the differences in the contribution of their neighbors are taken into account. We think that node B has more influence. By calculation, we quantify the influence of A and B as 3.77, 4, 81.

$$s_{xy}^{\text{DME}}(t) = \sum_{l=2}^t \left[\frac{k_x \sum_{j=1}^M \lg k_j}{2|E|} \times \pi_{xy}(l) + \frac{k_y \sum_{i=1}^N \lg k_i}{2|E|} \times \pi_{yx}(l) \right]. \quad (7)$$

As mentioned above, for better comparison, we also apply the ME index into the SRW model and the ME model as shown below.

$$s_{xy}^{\text{ME}}(t) = \sum_{l=2}^t \left[\frac{-\text{DC}_x \sum_{j=1}^M \log \text{DC}_j}{2|E|} \times \pi_{xy}(l) + \frac{-\text{DC}_y \sum_{i=1}^N \log \text{DC}_i}{2|E|} \times \pi_{yx}(l) \right]. \quad (8)$$

3. Experimental Data

In order to confirm the validity of the DME model, we conduct abundant experiments on nine real networks. They are listed as follows: (1) US Air (USAir), describing the network of the US air transportation system [22]; (2) Yeast PPI (Yeast), expressing the protein-protein relationship network of yeast [23]; (3) Food Web (Food), describing the association of carbon exchanges in the cypress wetlands of Florida ecosystem [24]; (4) Power Grid (Power), expressing the electrical power transportation network in the west of the US [25]; (5) UC Irvine, representing social network (Ucsocial), describing an online social network composed of students of the University of California, Irvine [26]; (6) Jazz, indicating the collaborative relationships among jazz musicians [27]; (7) EuroSiS Web (EuroSiS), showing the interactions between Science in Society actors from 12 European countries [28]; (8) Router, referring to the transmission of data packets between the routers in Internet [4]; and (9) King James, coming from the datasets Lü et al. collected. The fundamental topological features of the mentioned networks are listed in Table 1.

Our model is applied to the undirected and unweighted connected networks. Accordingly, we make arcs turn into undirected links. Besides, we delete the loops and multiple connections. Subsequently, the maximal edge-connected graph is extracted from each raw dataset to guarantee the connectivity of the whole.

Before the experiment, the edge set E of the nine networks is divided into two parts E^T and E^P randomly. The training set E^T contains 90% of the whole edge set. The testing set E^P contains 10%. The connectivity of the E^T is guaranteed by the means of adding edges randomly to the minimum spanning tree until the training set contains 90% links. Next, 30 groups of separate experimental data for each network are divided in the same size. Then, they are applied for the averaged precision by statistical methods to avoid the randomness of results.

4. Reference Standard

In order to highlight the superiority of our algorithm, five classic methods are listed as follows.

- (1) In common neighbor (CN) [12], the similarity is judged by the number of neighbors shared by node x and node y , which is defined as

$$s_{xy}^{\text{CN}} = |\Gamma(x) \cap \Gamma(y)|, \quad (9)$$

where $\Gamma(x)$ represents the set of neighbors of node x . Besides, $|\Gamma(x) \cap \Gamma(y)|$ refers to the amount of common neighbors of nodes x and y .

- (2) Preferential attachment (PA) [29] considers that the probability of a new link linked to the node x is proportional to k_x , so the probability between node x

TABLE 1: The basic statistical features of the 9 benchmark networks.

Nets	$ V $	$ E $	$\langle k \rangle$	$\langle d \rangle$	C	r	H
USAir	332	2126	12.81	2.74	0.749	-0.208	3.36
Yeast	2364	10898	9.2	5.16	0.378	0.470	3.35
Food	128	2075	32.42	1.78	0.335	-0.112	1.24
Power	4941	6594	2.669	15.87	0.107	0.003	1.45
Ucsocial	1893	13835	14.62	3.06	0.138	-0.188	3.81
Jazz	198	2742	27.7	6.45	0.618	0.02	1.4
EuroSiS	1272	6454	10.15	3.86	0.382	-0.012	2.46
Router	5021	6257	2.49	6.45	0.033	-0.138	5.50
King James	1707	9059	10.61	3.38	0.710	-0.052	3.92

$|V|$ denotes the number of nodes in the network, $|E|$ represents the total number of links, $\langle k \rangle$ is the average degree of all nodes, $\langle d \rangle$ indicates the average distance among nodes, C represents the clustering coefficient, r is the associativity coefficient, and H denotes the degree heterogeneity defined as $H = (\langle k^2 \rangle / \langle k \rangle^2)$.

and y is proportional to $k_x \times k_y$. The index is defined as

$$s_{xy}^{\text{PA}} = k_x \times k_y. \quad (10)$$

This index does not require the information of the neighborhood of each endpoint. Therefore, it has low computational complexity.

- (3) In Adamic-Adar (AA) [13], the idea is that the contribution of the node with small degree is greater. So, each node is given a weight value equaling to $1/(\log k_z)$ where k_z is the degree of a node from common neighbor set. The similarity is defined as

$$s_{xy}^{\text{AA}} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{\log(k_z)}, \quad (11)$$

where k_z represents the degree of common neighbor z .

- (4) Resource allocation (RA) [30], derived from AA, considers the resource allocation of network. Each node is given a weight value which is equal to $1/(k_z)$, and the similarity is defined as

$$s_{xy}^{\text{RA}} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{k_z}. \quad (12)$$

- (5) Superposed random walk (SRW) has been discussed in Section 2 in detail.

5. Results and Analysis

In order to prove the effectiveness of the DME model, abundant experiments have been carried out in nine real networks. The results are shown as follows.

In Figure 2, we plot the variation of the average precision with random walk steps obtained by SRW, ME, and DME in nine networks in the case of $L=100$. We can see that DME performs better obviously in 8 of the 9 networks than SRW. Furthermore, compared to both SRW and ME models, the DME model achieves the maximum precision in 6 of the 9 networks. Because the ME index reflects the

robustness of local network, it is more suitable for applying in network attacks to represent the importance of nodes. Therefore, we arrive at a conclusion that DME can achieve the highest accuracy in most cases when the random walk step t is optimal. Besides, it can reach the maximum precision in the minimum number of steps so that it can reduce the computation with the same precision.

Table 2 contains the detailed description of Figure 2. Furthermore, it also compares our model with other five classical models. The maximum precision is emphasized in bold and the corresponding step is in the parentheses. As is shown, the DME model reaches the highest precision in 6 of the 9 networks under the condition of $L=100$ compared with other five traditional models.

For ensuring the integrity of the experiment, we also conduct experiments in the case of $L=50$. The results are shown in Table 3. We italicize the values when the DME model is more exact than SRW. There are still 6 networks. Nevertheless, the advantage is not obvious when compared with other five models comprehensively. This means the DME model performs better in the top 100 links than 50 links. Actually, L is often defined as a large number to avoid random error.

The reason why the DME model can have an excellent performance is that it takes the mapping relationship between a node and its neighbors into comprehensive consideration. In this way, the differences in contribution of neighbors (i.e., the influences of indirect neighbors) are included, so that the model can assess the importance of endpoint better.

Though the DME model can achieve preferable performance in most datasets experimented by us, it also has no superiority in a few networks such as Jazz. By analyzing the topological characteristics of these networks, we find that they usually have the same features. The model we propose may be not suitable for the networks with good associativity coefficient and high clustering coefficient. We infer the reason is that the differences in contribution of neighbor nodes in such networks cannot be well reflected.

Furthermore, time complexity is also a significant factor to evaluate an index. For instance, CN index has $O(N^3)$

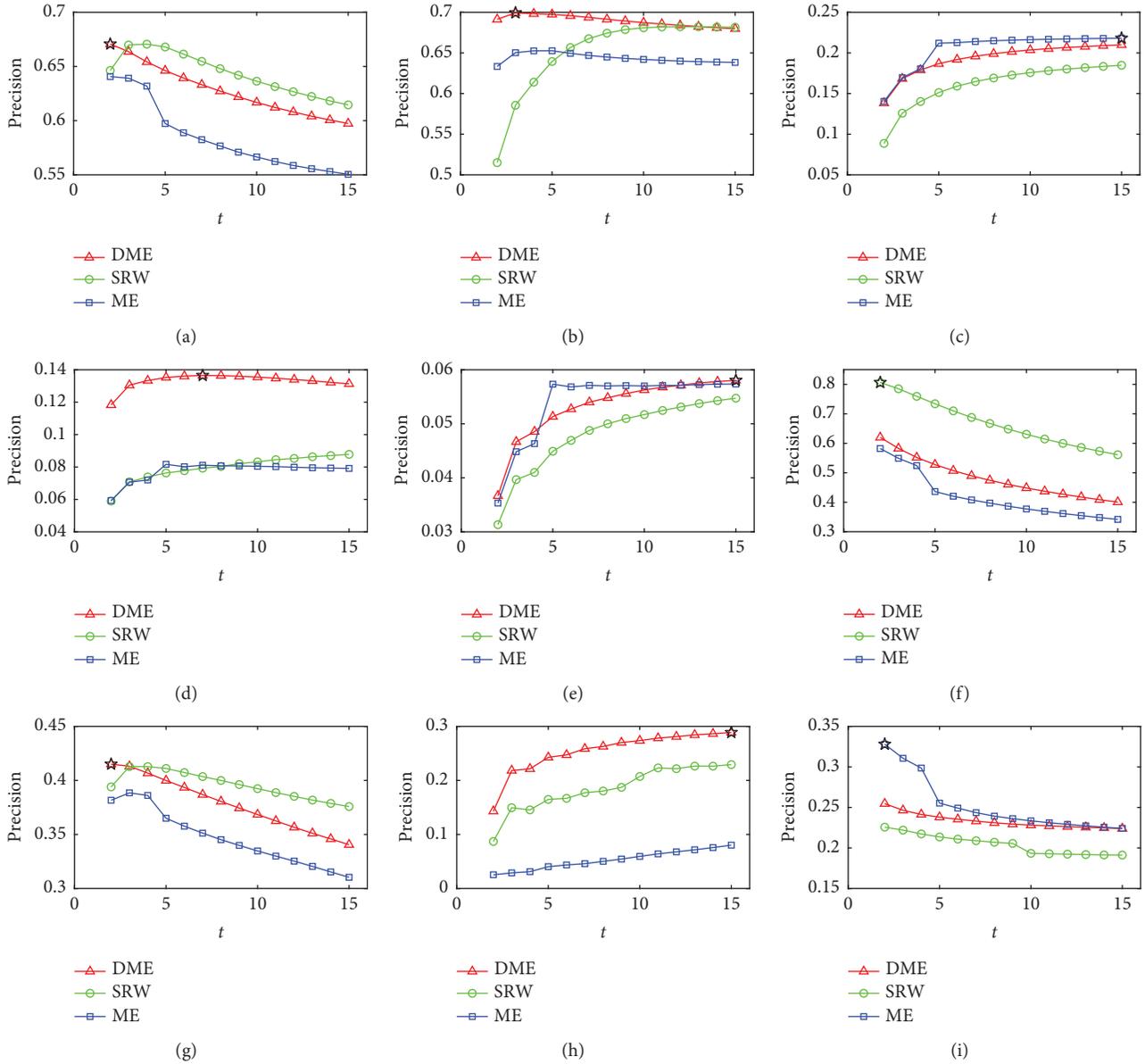


FIGURE 2: Precision of SRW (green circles), ME (blue rectangles), and DME (red triangles) versus number of random walk steps t with $L = 100$ in 9 real datasets. The highest precision in each network is marked by a black five-angled star. As is shown, DME performs better obviously in 8 of the 9 networks than SRW. Furthermore, compared to SRW and ME models, the DME model achieves the maximum precision in 6 of the 9 networks. (a) USAir. (b) Yeast. (c) Food. (d) Power. (e) Ucsocial. (f) Jazz. (g) EuroSiS. (h) Router. (i) King James.

TABLE 2: Precision on the nine benchmark networks in the case of $L = 100$.

Precision	CN	PA	AA	RA	SRW	DME
USAir	0.635	0.479667	0.654333	0.64	0.670583 (4)	0.670667 (2)
Yeast	0.728	0.555	0.722667	0.515	0.682242 (12)	0.699333 (3)
Food	0.086	0.211333	0.089667	0.088667	0.184810 (15)	0.209952 (15)
Power	0.124333	0.017333	0.081333	0.057333	0.087810 (15)	0.136500 (7)
Ucsocial	0.039333	0.057333	0.037333	0.031333	0.054738 (15)	0.058048 (15)
Jazz	0.816	0.168667	0.835667	0.806333	0.806333 (2)	0.62 (2)
EuroSiS	0.456333	0.049333	0.46667	0.394	0.412833 (3)	0.415 (2)
Router	0.132667	0.016	0.102333	0.075	0.229278 (15)	0.288690 (15)
King James	0.404333	0.190667	0.554333	0.82867	0.225667 (2)	0.254667 (2)

Each data point is an average over 30 independently divided datasets. The number in bold means the highest precision. The value enclosed in the parentheses represents the step corresponding to optimal precision.

TABLE 3: Precision on the nine benchmark networks with $L = 50$.

Precision	CN	PA	AA	RA	SRW	DME
USAir	0.821333	0.646667	0.82533	0.753333	0.8016 (6)	0.806889 (4)
Yeast	0.779333	0.594667	0.763333	0.470667	0.813429 (15)	0.817333 (15)
Food	0.118667	0.274	0.120667	0.116667	0.268 (15)	0.275762 (15)
Power	0.212	0.008667	0.059333	0.045333	0.119333 (15)	0.187778 (7)
Ucsocial	0.050667	0.077667	0.043333	0.041333	0.070667 (15)	0.064667 (15)
Jazz	0.895333	0.222667	0.906667	0.876667	0.0.876667 (2)	0.716667 (2)
EuroSiS	0.593333	0.074	0.560667	0.437333	0.466667 (5)	0.507333 (2)
Router	0.17	0.014667	0.121333	0.065333	0.134 (15)	0.196810 (15)
King James	0.552667	0.269333	0.69	0.898667	0.898667 (2)	0.502 (2)

Each data point is an average over 30 independent datasets, each of which is randomly divided into training set and test set with 90% and 10% probability. The value enclosed in the parentheses represents the step corresponding to optimal precision. Besides, the italicized values illustrate that the DME model performs better than the SRW model.

computational complexity while the complexity of RA is $2 \times O(N^3)$. The complexity of SRW which considers local path is $M \times O(N^3)$, and M is far less than N^3 . The model we introduce has the same time complexity of $M \times O(N^3)$ as SRW but can realize higher precision.

6. Conclusions

Existing link prediction algorithms on the basis of structural similarity mostly focus on the paths or the influences of nodes with only their degrees considered. Because the differences in contribution of neighbors are not considered, the precision of algorithm is limited. Through analysis, we propose the derivation of mapping entropy (DME) model, which interleaves the degrees of node and its neighbors. We investigate our model in comparison of CN, PA, AA, RA, SRW, and ME models on nine real datasets. The results indicate that the DME model prominently performs better than other six models and can achieve maximum precision in the minimum number of steps which reduces the computation with the same precision. Furthermore, the DME model does not increase time complexity.

The DME model proposed in our study reveals the effectiveness of distinguishing the differences in neighbors' contribution. This finding can provide a reference for future research. However, we only take the influence of indirect neighbors into account and ignore other factors such as coreness and H-index which can describe the maximal connected subgraph. Besides, we do not know the performance of the DME model in weighted and directed networks.

The results of our research are meaningful, and they are of great significance to the practical application of academic research. We can apply it in recommendation system, social cooperation network, information and communication technology, potential interactions in biological networks, and so on. Significantly, this work can inspire further work to add other factors such as H-index on the basis of our model and optimize the DME model in weighted and directed networks.

Data Availability

The datasets used in this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (nos. 61821007 and 62090015) and Science and Technology Key Project of Henan Province (Research on Key Technologies of Link Prediction in Complex Networks) (no. 202102311007).

References

- [1] W. Wang, M. Tang, H. F. Zhang, and Y. C. Lai, "Dynamics of social contagions with memory of nonredundant information," *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 92, Article ID 012820, 2015.
- [2] L. Lu and T. Zhou, "Link prediction in complex networks: a survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2010.
- [3] W. Wang, Y. Ma, T. Wu, Y. Dai, X. Chen, and L. A. Braunstein, "Containing misinformation spreading in temporal social networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 12, Article ID 123131, 2019.
- [4] R. Pech, D. Hao, L. Pan, H. Cheng, and T. Zhou, "Link prediction via matrix completion," *EPL*, vol. 117, Article ID 38002, 2016.
- [5] X. Zhu, W. Li, H. Tian, and S. Cai, "Hybrid influence of degree and h-index in the link prediction of complex networks," *EPL (Europhysics Letters)*, vol. 122, 2018.
- [6] Y. Li, P. Luo, Z. P. Fan, K. Chen, and J. Liu, "A utility-based link prediction method in social networks," *European Journal of Operational Research*, vol. 260, 2017.
- [7] Q. Sun, R. Hu, Y. Zhao, Y. Yao, and Y. Fan, "An improved link prediction algorithm based on degrees and similarities of nodes," in *Proceedings of IEEE/ACIS International Conference on Computer Information Science*, Wuhan, China, May 2017.
- [8] A. De, N. Ganguly, and S. Chakrabarti, "Discriminative link prediction using local links, node features and community structure," 2013, arXiv.
- [9] J.-X. Yang and X.-D. Zhang, "Revealing how network structure affects accuracy of link prediction," *The European Physical Journal B*, vol. 90, no. 8, p. 157, 2017.
- [10] X. Chen, K. Gong, R. Wang, S. Cai, and W. Wang, "Effects of heterogeneous self-protection awareness on resource-

- epidemic coevolution dynamics,” *Applied Mathematics and Computation*, vol. 385, Article ID 125428, 2020.
- [11] H. Liu, Z. Hu, H. Haddadi, and H. Tian, “Hidden link prediction based on node centrality and weak ties,” *EPL (Europhysics Letters)*, vol. 101, no. 1, Article ID 18004, 2013.
- [12] F. Lorrain and H. C. White, “Structural equivalence of individuals in social networks,” *Social Networks*, vol. 1, pp. 67–98, 1977.
- [13] Y. Yao, R. Zhang, F. Yang et al., “Link prediction via layer relevance of multiplex networks,” *International Journal of Modern Physics C*, vol. 28, no. 8, Article ID 1750101, 2017.
- [14] Q. Zhang, M. Li, and Y. Deng, “Measure the structure similarity of nodes in complex networks based on relative entropy,” *Physica A Statistical Mechanics Its Applications*, vol. 491, 2017.
- [15] W. Liu and L. Lü, “Link prediction based on local random walk,” *EPL (Europhysics Letters)*, vol. 89, Article ID 58007, 2010.
- [16] X. Chen, Q. Liu, R. Wang, Q. Li, and W. Wang, “Self-awareness-based resource allocation strategy for containment of epidemic spreading,” *Complexity*, vol. 2020, Article ID 3256415, 12 pages, 2020.
- [17] X. Chen, R. Wang, D. Yang, J. Xian, and Q. Li, “Effects of the awareness-driven individual resource allocation on the epidemic dynamics,” *Complexity*, vol. 2020, Article ID 8861493, 12 pages, 2020.
- [18] L. Lin-Yuan, “Link prediction on complex networks,” *Journal of University of Electronic Science and Technology of China*, vol. 39, pp. 651–661, 2010.
- [19] Y. Liu, T. Li, and X. Xu, “Link prediction by multiple motifs in directed networks,” *IEEE Access*, vol. 8, pp. 174–183, 2020.
- [20] F. Liu, B. Liu, C. Sun, M. Liu, and X. Wang, “Deep belief network-based approaches for link prediction in signed social networks,” *Entropy*, vol. 17, no. 4, pp. 2140–2169, 2015.
- [21] T. Nie, Z. Guo, K. Zhao, and M. Zhe, “Using mapping entropy to identify node centrality in complex networks,” *Physica, A. Statistical Mechanics and Its Applications*, vol. 453, 2016.
- [22] Z. Shan, “Link prediction based on local information considering preferential attachment,” *Physica A: Statistical Mechanics and Its Applications*, vol. 443, pp. 537–542, 2016.
- [23] R. Pech, D. Hao, Y. L. Lee, Y. Yuan, and T. Zhou, “Link prediction via linear optimization,” *Physica A: Statistical Mechanics and Its Applications*, vol. 528, 2018.
- [24] Z. Wu, Y. Lin, J. Wang, and S. Gregory, “Link prediction with node clustering coefficient,” *Physica A: Statistical Mechanics and its Applications*, vol. 452, no. C, pp. 1–8, 2015.
- [25] L. Jiao, F. Liu, J. Wu, and L. Ding, “Prediction of missing links based on community relevance and ruler inference,” *Knowledge Based Systems*, vol. 98, pp. 200–215, 2016.
- [26] W. Chao, V. Satuluri, and S. Parthasarathy, “Local probabilistic models for link prediction,” in *Proceedings of IEEE International Conference on Data Mining*, Omaha, NE, USA, October 2007.
- [27] L. Lü, C. H. Jin, and T. Zhou, “Similarity index based on local paths for link prediction of complex networks,” *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 80, Article ID 046122, 2009.
- [28] B. Ermiş, E. Acar, and A. T. Cemgil, “Link prediction in heterogeneous data via generalized coupled tensor factorization,” *Data Mining and Knowledge Discovery*, vol. 29, no. 1, pp. 203–236, 2015.
- [29] V. Martinez, F. Berzal, and J. C. Cubero, “A survey of link prediction in complex networks,” *Acm Computing Surveys*, vol. 49, pp. 69.1–69.33, 2017.
- [30] L. Ji and G. Deng, “Link prediction in a user–object network based on time-weighted resource allocation,” *Physica A Statistical Mechanics Its Applications*, vol. 388, pp. 3643–3650, 2009.