Article

A Link Prediction Method Based on Neural Networks

Keping Li, Shuang Gu * and Dongyang Yan

State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China; kpli@bjtu.edu.cn (K.L.); dongyangyan@bjtu.edu.cn (D.Y.)

* Correspondence: 17114248@bjtu.edu.cn; Tel.: +86-188-4587-4546

Abstract: Link prediction to optimize network performance is of great significance in network evolution. Because of the complexity of network systems and the uncertainty of network evolution, it faces many challenges. This paper proposes a new link prediction method based on neural networks trained on scale-free networks as input data, and optimized networks trained by link prediction models as output data. In order to solve the influence of the generalization of the neural network on the experiments, a greedy link pruning strategy is applied. We consider network efficiency and the proposed global network structure reliability as objectives to comprehensively evaluate link prediction performance and the advantages of the neural network method. The experimental results demonstrate that the neural network method generates the optimized networks with better network efficiency and global network structure reliability than the traditional link prediction models.

Keywords: link prediction; global network structure reliability; neural network; network evolution; network structure optimization

Citation: Li, K.; Gu, S.; Yan, D. A Link Prediction Method Based on Neural Networks. *Appl. Sci.* **2021**, *11*, 5186. <https://doi.org/10.3390/app11115186>

Academic Editor: Davide Careglio

Received: 19 May 2021

Accepted: 1 June 2021

Published: 3 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A network is a special case to express the relationship between systems. In recent years, network systems have attracted more and more attention from various disciplines in modeling, network topology and structure optimization [1]. Generally, network system is often seen in reliability engineering practice and real life, such as communication networks, computer networks, circuit networks, traffic networks, urban water supply networks, power supply networks, etc. However, we know that these real networks are constantly evolving in actual operation to ensure the sustainable development of various fields. Therefore, a series of network evolution models and algorithms aiming at optimizing the structure are emerged as the time requires [2].

As an essential mechanism for network evolution, link prediction has received widespread attention as soon as it was proposed, and it has been practically applied in many fields [3]. At present, the research on link prediction mainly focuses on evaluating the accuracy and precision of missing links. For many real networks, network evolution aims to extend the network topology to ensure the long-term development and regular operation of networks. Therefore, it is necessary to study link prediction with the purpose of improving reliability in network evolution.

However, the structure complexity and diversity of real networks bring great difficulties to link prediction aiming at improving reliability. In recent years, various disciplines have devoted themselves to intelligent research in related fields. Artificial intelligence has been at the forefront of the development of science and technology around the world with its advantages of improving productivity and reducing the difficulty of work. It has exerted a profound influence on all fields and promoted the evolution of society. We have to say that artificial intelligence of networks is the general trend. Among them, the neural network has become the representative of artificial intelligence methods because of the powerful self-learning function and the ability to find optimal solutions. It provides strong support for link prediction based on improving reliability.

Therefore, in order to ensure the safe operation and sustainable development of networks, an artificial intelligence method based on neural networks is used to realize link prediction in order to improve reliability in network evolution. The main contributions of this paper are as follows.

1. We try to combine artificial intelligence method with complex network theory, and propose a link prediction method based on neural networks.
2. A link pruning strategy based on the greedy algorithm is applied to solve the problem of neural network generalization on experiments.
3. According to the traditional reliability, the quantitative formula of global network structure reliability is given to measure the performance of extended networks.
4. By conducting two kinds of experiments on several networks with $N = 30, 50, 80$ and 100 , we prove that the neural network method is the best in improving network efficiency and global network structure reliability compared with different link prediction models.

The structure of the remainder of the paper is as follows. The work related to link prediction and reliability is reviewed in Section 2. Section 3 introduces the research base, including traditional link prediction models and the evaluation indexes. In Section 4, the neural network algorithm is described. In Section 5, we present the experimental results and conclusions are given in Section 6.

2. Literature Review

In network evolution, the primary methods to improve reliability include adding links, reconnecting links and adding nodes with changing network topology. At present, the most widely used method is link prediction based on adding edges [4]. Barabási and Albert [5] regarded the node degree as the basis of link possibility. Newman [6] introduced the concept of the common neighbor to link prediction for the first time. As soon as this method came out, it caused a boom in the field of link prediction. Ma et al. [7] proposed a local friend recommendation index to facilitate link prediction by a typical structural property: nodes were preferentially linked to the nodes with the weak clique structure in many real networks. By considering both the nodes' types of effects and structural similarities, Fan et al. [8] proposed a combined link prediction index and studied the robustness. Gao et al. [9] proposed a link prediction algorithm suitable for bipartite networks. Aghabozorgi and Khayyambashi [10] tested the similarity measure based on structural units through a supervised learning experiment framework, and found that the relations built through the paths between the endpoints determine the influence of the new links. Yang et al. [11] proposed a novel link prediction index significant influence, which was modeled by distinguishing the strong influence of the weak. Pech et al. [12] assumed that the likelihood of a link between two nodes can be realized by a linear summation of neighbor nodes' contributions. As a result, they obtained the analytical solution of the optimal likelihood matrix and better performance. Balls-Barker and Webb [13] proposed a new method to predict the formation of unobserved links in the real-world networks, relying on the theory of isospectral matrix reduction to calculate the probability and finally transitioning from one vertex to another in a random walk on the network. Li et al. [14] combined the four similarity indexes, including Common Neighbors, Leicht–Holme–Newman, Cosine based on the Laplacian matrix, and Matrix Forest based on the Logistic regression algorithm and the Xgboost algorithm, and introduced the idea of stacking into the link prediction of complex networks. Bai et al. [15] constructed a new model that regarded link prediction in multiplex networks as a multi-attribute decision-making problem, in which the potential links in the target layer are alternatives, the layers are viewed as attributes, and the similarity score of a potential link in each layer are an attribute value.

Artificial intelligence method, a hot research topic in recent years, is also used in link prediction. Liu et al. [16] believed that different common neighbors played various roles so that they have different contributions, and therefore proposed a local naive

Bayes model. Wu [17] proposed a generalized tree augmented naive Bayesian probability model by exploiting mutual information to quantify neighbors' neighborhood influence. Moreover, it was easily adapted to other common neighbors-based methods including Common Neighbors, Adamic/Adar and Resource Allocation. Considering the differences between latent features, Wang et al. [18] proposed the variational Bayesian probabilistic matrix factorization with Student-t prior model. Xiao et al. [19], who studied the internal and external factors that affect the link formation, integrated the user behavior and user relationships into link prediction and proposed a three-level hidden Bayes model. Yuan et al. [20] calculated the graph kernel similarities between subgraphs to train the Support Vector Machine (SVM) classifier to achieve link prediction. Shan et al. [21] regarded link prediction as a binary classification problem and proposed a supervised method to achieve link prediction in multiplexed networks. Experiments conducted on six networks have confirmed the effectiveness of this method. In order to fully explore and utilize the public opinion characteristics of social network node users, Wang et al. [22] established a multidimensional network model oriented to the topology of "We the Media" networks and designed a link prediction algorithm suitable for multidimensional networks. Neural network technology is a mature intelligent computing technology, and its application has achieved great success at present [23]. Li et al. [24] proposed a multiple-types link prediction model for heterogeneous networks based on back propagation neural networks. Ozcan and Oguducu [25] proposed a multivariate link prediction method in heterogeneous network evolution by using a nonlinear autoregressive neural network with external inputs. Cai et al. [26] proposed a link prediction method based on recurrent neural networks for the time-varying characteristics and the historical information of node pairs in opportunistic networks that had a vital influence on the future link state. Lee and Sohn [27] combined an artificial neural network algorithm and hill climbing algorithm to propose a new intelligent link prediction method in scale-free networks. Wang et al. [28] inferred the possible regulatory relationship between two genes through calculations, which can be expressed as a link prediction problem between two nodes in the graph. An end-to-end gene regulatory graph neural network method is proposed under the framework of supervised and semi-supervised. Xiao et al. [29] proposed an attention-based convolutional neural network link prediction method that can effectively improve the performance by fusing and mining structural features and text features.

Reliability, which plays an essential role in ensuring network security [30], is the goal of network development. In other words, reliability is an important index to evaluate network structure or link prediction in network evolution. Shi et al. [31] thought that adding redundant edges to key nodes can effectively improve the network robustness. Shargel et al. [32] proposed an optimized network with better robustness and connectivity by parametrizing two aspects of network construction: growth and preferential attachment. Hayashi and Matsukubo [33] propose a method to improve scale-free networks' invulnerability by randomly adding the shortest path to some nodes. Xiao et al. [34] studied the controllability based on edge directions. Yan et al. [35] proposed an algorithm based on local information to optimize the network controllability. Sohn [36] regenerated scale-free networks into more optimized robust networks based on neural networks.

At present, the main content of link prediction is to evaluate the prediction accuracy of missing links. However, link prediction is an important step in network evolution, and its crucial task is to ensure network performance and improve reliability for most real networks. Reliability has always been used as an index for the comprehensive evaluation of structural safety, but there is no clear definition. Therefore, to solve these problems, we propose a new link prediction method implemented by the neural network algorithm to improve reliability in network evolution. The input training data are based on the scale-free networks and the output desired data are based on the optimized networks obtained by link prediction models. The network efficiency is used as an index to evaluate the quality of the methods. To verify the network performance comprehensively, the global network structure reliability is proposed in the paper.

3. Problem Description and Formula

Link prediction refers to predicting the possibility of a link between two nodes that have not yet been connected in the network through known nodes and structural information [37]. The undirected scale-free network $G(V, L)$ is studied in this paper, where V and L denote sets of nodes and links, respectively. In this section, we introduce link prediction models and evaluation indexes for the undirected scale-free networks. Link prediction methods need to provide a score (S) to predict the probability of a link existing between a pair of unconnected nodes. For example, S_{xy}^{CN} represents the score of possible links between two nodes x and y in a network calculated by the Common Neighbors (CN) model. Similarly, S_{xy}^{RA} is the score of the Resource Allocation (RA) model, and S^{MF} is a matrix of the scores of all links obtained using the Matrix Forest (MF) model. All links not generated in the original network are sorted in decreasing order according to their scores, and the links at the top are most likely to exist.

3.1. Link Prediction Models

There exist many methods applied in the link prediction field. Some of the well-known techniques [38], including local link prediction methods and global link prediction methods, are selected as the necessary link prediction models to compare with the neural network methods.

Common Neighbors (CN) [39] takes the number of common neighbors as a score to determine whether or not a link can be generated between two nodes. The more common neighbors two nodes share, the greater the probably they are related with a link. For nodes x, y , the sets of their neighbors are $\Gamma(x), \Gamma(y)$; their common neighbors' set is $\Gamma(x) \cap \Gamma(y)$. CN is defined as:

$$S_{xy}^{CN} = |\Gamma(x) \cap \Gamma(y)| \quad (1)$$

Resource Allocation (RA) [40] measures the possibility of a link occurs by common neighbors' degree. Here, z represents the common neighbor node of node x and y , k_z is the degree of node z . Then, the expression of RA is:

$$S_{xy}^{RA} = \sum_{z \in \Gamma(x) \cap \Gamma(y)} \frac{1}{k_z} \quad (2)$$

The Jaccard Coefficient (JC) [41] is proposed by Jaccard, which considers the neighbor nodes and common neighbors of two nodes. It is defined as:

$$S_{xy}^{JC} = \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|} \quad (3)$$

Hub Promoted (HP) [42] is proposed for quantifying the topological overlap of pairs of substrates in metabolic networks, and defined as:

$$S_{xy}^{HP} = \frac{|\Gamma(x) \cap \Gamma(y)|}{\min\{k_x, k_y\}} \quad (4)$$

Leicht–Holme–Newman (LHN) [43] considers the number of common neighbors and degree of two nodes. The expression of LNH is:

$$S_{xy}^{LHN} = \frac{|\Gamma(x) \cap \Gamma(y)|}{k_x \times k_y} \quad (5)$$

Average Commute Time (ACT) [44] thinks that the shorter the time, the more similar the two nodes, then the possibility of link between the nodes x and y can be defined as:

$$S_{xy}^{ACT} = \frac{1}{l_{xx}^+ + l_{yy}^+ - 2l_{xy}^+} \quad (6)$$

where l_{xy}^+ denotes the corresponding entry in terms of the Laplacian matrix's pseudoinverse, L^+ ($L = D - A$) [45].

Cosine Time (CT) [45] is an inner-product-based measure, which is based on L^+ by calculating the possibility of two vectors. The cosine time is defined as the cosine of the node vectors:

$$S_{xy}^{CT} = \cos(x, y)^+ = \frac{l_{xy}^+}{\sqrt{l_{xx}^+ \cdot l_{yy}^+}} \quad (7)$$

Matrix Forest (MF) [46] thinks the similarity between x and y can be understood as the ratio of the number of spanning rooted forests, such that nodes x and y belong to the same tree rooted at x to all spanning rooted forests of the network. MF is defined as:

$$S^{MF} = (I + L)^{-1} \quad (8)$$

3.2. Reliability Indexes

3.2.1. Network Efficiency (E)

At present, the direct definition or measurement of network reliability is not clear. Network efficiency is usually used to indirectly evaluate it. For nodes x, y in a network, the efficiency is the reciprocal of the shortest distance between them [47]. For an undirected network, it is expressed as:

$$E = \frac{2 \sum_{x < y} \frac{1}{d_{xy}}}{N(N-1)} \quad (9)$$

where d_{xy} is the shortest distance between the nodes x and y ; N is the number of nodes in the network. It reflects the overall connection in the whole range of the network itself. The greater the value is, the better the network structure establishes.

3.2.2. Global Network Structure Reliability (P)

Gupta and Kundu [48] first proposed the generalized exponential distribution of reliability index and applied it to the machine, components and other equipment [49]. Its form is:

$$P = e^{-\lambda t} \quad (10)$$

here, λ is the failure rate. The failure rate of the systems is difficult to obtain, so the formula evolves into:

$$P = 1 - e^{-\lambda C} \quad (11)$$

where λ is the reliability factor and C is the equivalent system cost of destructions after achieving the object.

In researches, many complex systems are regarded as networks. Therefore, we consider changing some expression forms to realize the calculation of global network structure reliability to comprehensively characterize the optimization degree in the process of network evolution. Generally, the parameter λ is related to the knowledge level, the computing resources, human resources, time and other measurable factors [50]. Similarly, the critical network structural characteristic indexes include average shortest path length, betweenness centrality structural entropy and average clustering coefficient [51,52]. Therefore, we quantify global network structure reliability within a weighted summation method to describe the parameter λ . One of the most straightforward indexes to measure a network is the degree. We consider that the average degree ($K = \frac{\sum_{x=1}^N k_x}{N}$) has the same effect as

C in calculating the global network structure reliability, so the global network structure reliability can be expressed in the following form:

$$P = 1 - e^{-(\varepsilon_1 L + \varepsilon_2 E_B + \varepsilon_3 C) \times K} \quad (12)$$

where L is the average shortest path length. E_B is the betweenness centrality structural entropy. C is the average clustering coefficient. $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are the weights of parameters, and $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 1$. Here, we assume that $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1/3$ in a general scale-free network. The following is the specific description and formula explanation for these indexes.

A. Average Shortest Path Length (L).

The average shortest path length is an important feature measure in network analysis, which is the average distance of all node pairs. For an undirected network, it is defined as follows:

$$L = \frac{2 \sum_{x < y} d_{xy}}{N(N-1)} \quad (13)$$

The smaller the L , the greater the connection intensity between nodes. In order to solve the difference between L and other indexes, let $L = 1/L$.

B. Betweenness Centrality Structural Entropy (E_B).

Node betweenness centrality characterizes the importance of the nodes. Network structure entropy is introduced to characterize the dispersion of the nodes' importance [51]. Let d_{yk} denote the total number of the shortest path starting from the nodes y to k ; $d_{yk}(x)$ is the number of the shortest path through the node x from node y to k . If it goes through the node x , then $d_{yk}(x) = 1$, otherwise $d_{yk}(x) = 0$. The betweenness centrality of node x is B_x :

$$B_x = \frac{\sum_{y \neq x \neq k} \frac{d_{yk}(x)}{d_{yk}}}{(N-1)(N-2)} \quad (14)$$

The betweenness centrality structural entropy is expressed as:

$$E_B = - \sum_{x=1}^N B_x \ln B_x \quad (15)$$

C. Average Clustering Coefficient (C).

The clustering coefficient is a characteristic index to describe the tightness of the network [52]. A node with a larger clustering coefficient is closer to its neighbor nodes in the network. The clustering coefficient of node x (C_x) is:

$$C_x = \frac{2E_x}{k_x(k_x-1)} \quad (16)$$

Therefore, the network's average clustering coefficient is:

$$C = \frac{1}{N} \sum_{x=1}^N C_x \quad (17)$$

here, E_x is the number of edges connected to neighbor nodes of node x .

4. Method

4.1. Neural Networks (NN)

Neural networks have many successful applications in solving various problems related to prediction, classification and regression [53]. There are three basic layers in neural networks, including input layer, hidden layer and output layer, as shown in Figure 1. Usually, the data form of neural networks is arrays. As we know, a network is always expressed as an adjacency matrix. It is impossible to train the neural network as input data because using the adjacency matrix as the input of neural networks would require

the dimension size that can represent 2^{N^2} adjacency matrices, where N is the number of nodes in the network. Here, a new network representation called link list (LL) is applied that contains binary elements representing existence of node pairs among all possible combinations $\binom{N}{2}$. Generally, the possible links are $L = \{(1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 4), (3, 5), (4, 5)\}$ for an undirected network with $N = 5$. If four of the links exist by $L = \{(1, 4), (2, 3), (2, 4), (3, 5)\}$, then $LL = [0\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 1\ 0]$. In this way, the network's adjacency matrix is transformed into an array, and link prediction of a network is achieved by the neural network method. The neural network is trained based on original scale-free networks as the input data and optimized networks by link prediction models as the desired output data. Specifically, we first generate scale-free networks with specified requirements as original data, and the link prediction models obtain the optimized networks. Then, the adjacency matrix of the networks is transformed into the link list. In the training of the neural network, the input is the link list of the original scale-free networks and the output is the link list of the optimized networks generated by the link prediction models. It solves the problem that neural networks cannot train network data.

Figure 1. Neural network model.

The original networks are scale-free networks with the initial number of nodes $N_0 = 3$ and link density $M_0 = 2$ for $N = 30$, $N = 50$, $N = 80$ and $N = 100$. The link prediction method based on neural networks is evaluated by the network efficiency and global network structure reliability. In order to verify the feasibility and advantages of link prediction based on neural networks, we use the traditional link prediction models as the performance reference models. In experiments, we assume that the number of links increased by link prediction models is equal to the number of network's nodes. We randomly generate a database of 10 data sets which have scale-free networks with the same number of nodes as the input data of the neural network; 80% of them are training data and 20% are test data. In the neural network method, the number of hidden layers is set to 4 [54]. The maximum training epochs is 1000, the training goal is 0.00001 and the neural network's learning rate is set to 0.001.

We performed two kinds of experiments on scale-free networks with $N = 30, 50, 80$ and 100 , and analyze the results by MATLAB. First, we randomly generated several scale-free networks with a specified number of nodes as data sets. In Experiment 1, the input data are different networks with the same number of nodes, and the output data are optimized networks obtained by the same link prediction model. We found that the traditional link prediction models have different advantages for different networks. Therefore, in order to give full play to the advantages of different link prediction models to obtain the more optimized networks, we carried out Experiment 2, in which the input data are the

same network and the output data are the optimized networks generated by different link prediction models.

4.2. Link Pruning

The number of links of the optimized networks produced by neural networks cannot be controlled to be equal to the number of links of the optimized network produced by link prediction models, due to neural networks' generalization properties. To solve this problem, there is an additional stage named link pruning. Link pruning is the process of removing redundant links generated by neural networks according to the goal, so that the optimized network generated by the neural network has the same network topology complexity as the optimized network generated by link prediction. Here, in order to maximize the reliability, we apply the idea of the greedy algorithm to achieve link pruning. The greedy algorithm means that when solving the problem, the algorithm always makes what seems to be the best choice at the moment. Our specific approach is to add all the links generated by the neural network into the original network to form a new network. Then, one of all the links generated by the neural network is randomly deleted from the new network at a time to calculate the network efficiency and global network structure reliability. After all the links generated by the neural network have been deleted once, the data are sorted from large to small according to the target. Finally, the link pruning process is completed when the number of links in the network is equal to that generated by the link prediction models.

4.3. General Scheme

Next, we describe the general scheme for the above the neural network method.

Step 1—The scale-free networks that meet the specified requirements are randomly generated as the original networks.

Step 2—The optimized networks are obtained by link prediction models.

Step 3—Convert the original networks and optimized networks to the link list.

Step 4—Train the neural network where the inputs are scale-free networks and the outputs are optimized networks.

Step 5—The predicted networks are input into the trained neural network to obtain the new optimized networks.

Step 6—According to the goals, the final optimized networks are obtained by link pruning.

Step 7—According to Equations (11) and (14), the network efficiency and global network structure reliability are calculated.

5. Experiment Results

We found that the hubs of the optimized networks produced by the link prediction models are obviously observed and the degree difference between nodes is large through comparison and analysis. For example, Figure 2 shows the network topology of the original network, the optimized network generated by the RA model and the optimized network generated by the neural network. This structure is unstable when the network is targeted attacked. As for the optimized networks produced by neural networks, compared to the optimized network produced by the link prediction models, the number of medium-sized hubs increased. The structure can withstand greater attacks. Therefore, the performance of the network structure optimized by the neural network is better.

Figure 2. Network topology with $N = 100$. (a) The original network, $E = 0.36789562$, $P = 0.45966444$; (b) the optimized network produced by RA model, $E = 0.43532997$, $P = 0.87918564$; (c) the optimized network produced by neural networks based on RA model, $E = 0.47471380$, $P = 0.90748775$.

In order to further illustrate and verify our conclusion, we calculate the network efficiency and global network structure reliability and plot the experimental results of the networks with $N = 30, 50, 80$ and 100 into Figures 3–6. As shown in Figure 3, Figure 3a is the network efficiency of the networks with $N = 30$ for different methods. Figure 3b shows the global network structure reliability of the networks with $N = 30$ for different methods. Similarly, we can explain Figure 4a to Figure 6b. It is found that the networks optimized by neural networks are significantly higher than those by the traditional link prediction models, regardless of network efficiency or global network structure reliability. Besides, by comparing Figures 3–6, we find that the neural network has a more obvious effect on the traditional link prediction models with poor optimization results. For example, the network efficiency of the optimized network obtained by LHN and MF models in Figure 3a is the minimum, but the network efficiency of the optimized network obtained by the neural network is almost the maximum. In Figure 4b, CN and RA models have achieved high global network structure reliability. Although the trained neural network is higher than the traditional link prediction models, the improvement is not apparent. The conclusion is also reflected in Figures 5 and 6 for networks with $N = 80$ and 100 . Comparing (a) and (b) of the four figures we can see that the difference in network efficiency is more prominent. In contrast, the results of global network structure reliability proposed are more similar for different methods. This is because the global structure network reliability is more comprehensive than network efficiency in evaluating network structure. At the same time, the availability of the global network structure reliability can be explained. We can also see that the original scale-free networks' global network structure reliability is generally low, which indicates that the structure of the randomly generated scale-free networks is unstable.

Figure 3. The link prediction results of all methods for networks with $N = 30$. (a) Network efficiency; (b) global network structure reliability.

Figure 4. The link prediction results of all methods for networks with $N = 50$. (a) Network efficiency; (b) global network reliability.

Figure 5. The link prediction results of all methods for networks with $N = 80$. (a) Network efficiency; (b) global network reliability.

Figure 6. The link prediction results of all methods for networks with $N = 100$. (a) Network efficiency; (b) global network reliability. **Note:** 1. ON is the original network; 2. CN is the optimized network produced by CN model; 3. NN-CN is the optimized network produced by neural networks based on CN model; 4. NN is the optimized network produced by neural networks.

From the above results, we find that the traditional link prediction models' performance ranking is inconsistent for the networks with different node numbers and structures. For example, the optimized network generated by ACT model has the maximum reliability for $N = 30$ in Figure 3b. The optimized network generated by RA model has the maximum network efficiency for $N = 80$ in Figure 5a. In Figure 6, for the prediction network 1 with $N = 100$, the network efficiency and global network structure reliability obtained by the CN model are greater than that obtained by the RA model. However, the network efficiency and global network structure reliability obtained by CN model are lower than that obtained by RA model for the prediction network 2 with $N = 100$. Thus, we believe that different link prediction models have different advantages in various networks. Therefore, in order to achieve the best results, we carried out new experiments considering the benefits of different models. In the experiments, the input data are the same original network each time, and the output data are the optimized network generated by different link prediction models. After training, we input the original network again, and output a higher performance network. The experimental results shown in Table illustrate that the optimized networks generated by neural networks have the best performance. Moreover, the method based on neural networks has obvious advantages over the traditional link prediction models in improving network efficiency and global network structure reliability. The network efficiency and global network structure reliability in different methods are illustrated in Table 1 and Figures 7 and 8. Table 1 shows the specific data about network efficiency and global network structure reliability of all methods for all networks. In Figures 7 and 8, the X-axis is the number of nodes, and the Y-axis is the network efficiency or global network structure reliability. We want to observe how network efficiency and global network structure reliability vary with the number of nodes. We can see intuitively that the optimized network generated by neural networks always has the most excellent network efficiency and global network structure reliability. In addition, we find that the more nodes the network has, the less network efficiency and global network structure reliability is. This can be attributed to that a large number of nodes indicates a much complex network, the performance of which becomes difficult to improve.

Table 1. Network efficiency and global network structure reliability of all methods.

Index	N	Predicted Network	ON	CN	RA	JC	HP	LHN	ACT	CT	MF	NN
E	30	1	0.4869732	0.5559387	0.5647510	0.5570881	0.5490422	0.5354406	0.5524904	0.5352490	0.5385058	0.5839081
		2	0.4809962	0.5634100	0.5593870	0.5643678	0.5590038	0.5528736	0.5582376	0.5450192	0.5442529	0.5827586
	50	1	0.4360408	0.5011565	0.5127211	0.5011565	0.4937415	0.4678776	0.4991837	0.4768027	0.4756463	0.5202721
		2	0.4320952	0.5051701	0.5163265	0.5051701	0.5123401	0.4806122	0.4945578	0.4691157	0.4679592	0.5418367
	80	1	0.3796730	0.4395359	0.4525053	0.4395886	0.4390243	0.4144937	0.4388133	0.4128006	0.4117880	0.4672996
		2	0.3963555	0.4646097	0.4595464	0.4603376	0.4575791	0.4274895	0.4539821	0.4290348	0.4251319	0.4782437
	100	1	0.3678956	0.4413333	0.4353300	0.4413333	0.4646667	0.3918721	0.4238990	0.4037037	0.3999360	0.4889899
		2	0.3652727	0.4240774	0.4320842	0.4240774	0.4365152	0.3928552	0.4192862	0.4024478	0.3980572	0.4520875
	P	30	1	0.6823815	0.9154475	0.9168245	0.9104210	0.9042422	0.8596296	0.9315846	0.9055676	0.8974670
			2	0.6763270	0.9278472	0.9103899	0.9289363	0.9165503	0.8806171	0.9315667	0.9089169	0.9048870
		50	1	0.6612709	0.9276401	0.9239963	0.9276401	0.9053636	0.8396845	0.9272298	0.8934019	0.8909736
			2	0.5693640	0.9087218	0.9129838	0.9087218	0.9083342	0.8359719	0.9047767	0.8777054	0.8767607
		80	1	0.4570695	0.8651847	0.8807530	0.8633415	0.8425382	0.7697039	0.8694548	0.8234131	0.8140307
			2	0.5195350	0.9115835	0.9004921	0.9037903	0.8911566	0.8022490	0.9034039	0.8425871	0.8441484
		100	1	0.4596644	0.8791856	0.8640235	0.8791856	0.8941482	0.7420475	0.8421491	0.8128243	0.8105283
			2	0.3797415	0.8563602	0.8617217	0.8563602	0.8548664	0.7152076	0.8246843	0.8174176	0.8009378

Figure 7. Network efficiency in different methods. (a) Predicted network 1; (b) predicted network 2.

Figure 8. Global network structure reliability in different methods. (a) Predicted network 1; (b) predicted network 2.

6. Conclusions

The study of optimization network performance by artificial intelligence methods is important and challenging research. In this paper, we propose a new method that produces

optimized networks with high network efficiency and reliability in network evolution. The neural network method based on link prediction models greatly reduces the training complexity but improves the performance. In order to solve the generalization problem of neural networks, we propose link pruning, which applies the thought of greedy algorithm to complete missing links and remove redundant links generated by neural networks according to the goal. Furthermore, we also give the quantitative index of global network structure reliability to evaluate the network optimization performance.

We designed two kinds of experiments via MATLAB to explore the link prediction method based on neural networks to improve the reliability. The experimental results show that the optimized networks generated by neural networks have better performance in network efficiency and global network structure reliability than traditional link prediction models. Therefore, the neural network method has higher accuracy and performance than popular link prediction models, and the artificial intelligence method offers a more positive contribution to link prediction.

Author Contributions: Conceptualization, K.L. and S.G.; methodology, S.G.; software, S.G. and D.Y.; validation, S.G.; formal analysis, K.L. and S.G.; investigation, S.G. and D.Y.; resources, K.L.; data curation, S.G. and D.Y.; writing—original draft preparation, S.G.; writing—review and editing, K.L.; S.G. and D.Y.; visualization, S.G.; supervision, K.L.; project administration, K.L.; funding acquisition, K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Beijing Natural Science Foundation (Grant No. 8202039), the National Natural Science Foundation of China (Grant Nos. 71942006, 71621001) and Research Foundation of State Key Laboratory of Railway Traffic Control and Safety, Beijing Jiaotong University (Grant No. RCS2021ZT001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sohn, I. Small-world and scale-free network models for IoT systems. *Mob. Inf. Syst.* **2017**, *61*, 1–9. [\[CrossRef\]](#)
2. Li, G.; Deng, L.; Xiao, G.; Tang, P.; Wen, C.Y.; Hu, W.H.; Pei, J.; Shi, L.P.; Stanley, H.E. Enabling controlling complex networks with local topological information. *Sci. Rep.* **2018**, *8*, 1–10. [\[CrossRef\]](#)
3. Guimerà, R.; Sales-Pardo, M. Missing and spurious interactions and the reconstruction of complex networks. *Proc. Natl. Acad. Sci. USA* **2009**, *106*. [\[CrossRef\]](#)
4. Sherkat, E.; Rahgozar, M.; Asadpour, M. Structural link prediction based on ant colony approach in social networks. *Physica A Stat. Mech. Appl.* **2015**, *41*, 80–94. [\[CrossRef\]](#)
5. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Newman, M.E.J. Clustering and preferential attachment in growing networks. *Phys. Rev. E* **2001**, *64*, 025102. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Ma, C.H.; Zhou, T.; Zhang, H.F. Playing the role of weak clique property in link prediction: A friend recommendation model. *Sci. Rep.* **2016**, *6*. [\[CrossRef\]](#)
8. Fan, C.; Liu, Z.; Lu, X.; Xiu, B.; Chen, Q. An efficient link prediction index for complex military organization. *Physica A Stat. Mech. Appl.* **2017**, *469*, 572–587. [\[CrossRef\]](#)
9. Gao, M.; Chen, L.; Li, B.; Li, Y.; Liu, W.; Xu, Y. Projection-based link prediction in a bipartite network. *Inform. Sci.* **2017**, *376*, 158–171. [\[CrossRef\]](#)
10. Aghabozorgi, F.; Khayyambashi, M.R. A new similarity measure for link prediction based on local structures in social networks. *Physica A Stat. Mech. Appl.* **2018**, *501*, 12–23. [\[CrossRef\]](#)
11. Yang, Y.; Zhang, J.; Zhu, X.; Tian, L. Link prediction via significant influence. *Physica A Stat. Mech. Appl.* **2018**, *492*, 1523–1530. [\[CrossRef\]](#)
12. Pech, R.; Hao, D.; Lee, Y.L.; Yuan, Y.; Zhou, T. Link prediction via linear optimization. *Physica A Stat. Mech. Appl.* **2019**, *528*. [\[CrossRef\]](#)
13. Balls-Barker, B.; Webb, B. Link prediction in networks using effective transitions. *Linear Algebra Appl.* **2020**, *599*, 79–104. [\[CrossRef\]](#)
14. Li, K.Y.; Tu, L.L.; Chai, L. Ensemble-model-based link prediction of complex networks. *Comput. Netw.* **2020**, *166*. [\[CrossRef\]](#)

15. Bai, S.S.; Zhang, Y.K.; Li, L.J.; Shan, N.; Chen, X.Y. Effective link prediction in multiplex networks: A TOPSIS method. *Expert Syst. Appl.* **2021**, *177*. [\[CrossRef\]](#)
16. Liu, Z.; Zhang, Q.M.; Lü, L.Y.; Zhou, T. Link prediction in complex networks: A local naive Bayes model. *Europhys. Lett.* **2011**, *96*. [\[CrossRef\]](#)
17. Wu, J.H. A generalized tree augmented naive Bayes link prediction model. *J. Comput. Sci.* **2018**, *27*, 206–217. [\[CrossRef\]](#)
18. Wang, Y.S.; Liu, F.B.; Xia, S.T.; Wu, J. Link sign prediction by variational Bayesian probabilistic matrix factorization with student-t prior. *Inform. Sci.* **2017**, *405*, 175–189. [\[CrossRef\]](#)
19. Xiao, Y.P.; Li, X.X.; Wang, H.H.; Xu, M. 3-HBP: A three-level hidden Bayesian link prediction model in social networks. *IEEE Trans. Comput. Soc. Syst.* **2018**, *5*, 430–443. [\[CrossRef\]](#)
20. Yuan, W.W.; He, K.Y.; Guan, D.H.; Zhou, L.; Li, C.L. Graph kernel based link prediction for signed social networks. *Inform. Fusion* **2019**, *46*, 1–10. [\[CrossRef\]](#)
21. Shan, N.; Li, L.J.; Zhang, Y.K.; Bai, S.S.; Chen, X.Y. Supervised link prediction in multiplex networks. *Knowl. Based Syst.* **2020**, *203*. [\[CrossRef\]](#)
22. Wang, G.H.; Wang, Y.F.; Li, J.M.; Liu, K.D. A multidimensional network link prediction algorithm and its application for predicting social relationships. *J. Comput. Sci.* **2021**, *53*. [\[CrossRef\]](#)
23. Sanchez, E.N.; Rodriguez-Castellanos, D.I.; Chen, G.R.; Ruiz-Cruz, R. Pinning control of complex network synchronization: A recurrent neural network approach. *Int. J. Control Autom.* **2017**, *15*, 1405–1414. [\[CrossRef\]](#)
24. Li, J.C.; Zhao, D.L.; Ge, B.F.; Yang, K.W. A link prediction method for heterogeneous networks based on BP neural network. *Physica A Stat. Mech. Appl.* **2018**, *495*, 1–17. [\[CrossRef\]](#)
25. Ozcan, A.; Oguducu, S.G. Link prediction in evolving heterogeneous networks using the NARX neural networks. *Knowl. Inf. Syst.* **2018**, *55*, 333–360. [\[CrossRef\]](#)
26. Cai, X.L.; Shu, J.; Al-Kall, M. Link prediction approach for opportunistic networks based on recurrent neural network. *IEEE Access* **2019**, *7*, 2017–2025. [\[CrossRef\]](#)
27. Lee, Y.H.; Sohn, I. Reconstructing damaged complex networks based on neural networks. *Symmetry* **2017**, *9*, 310. [\[CrossRef\]](#)
28. Al-Kuwaiti, M.; Kyriakopoulos, N.; Hussein, S. A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability. *IEEE Commun. Surv. Tut.* **2009**, *11*, 106–124. [\[CrossRef\]](#)
29. Shi, C.; Peng, Y.; Zhuo, Y.; Tang, J.Y.; Long, K.P. A new way to improve the robustness of complex communication networks by allocating redundancy links. *Phys. Scripta* **2012**, *85*. [\[CrossRef\]](#)
30. Shargel, B.; Sayama, H.; Epstein, I.R.; Bar-Yam, Y. Optimization of robustness and connectivity in complex networks. *Phys. Rev. Lett.* **2003**, *90*. [\[CrossRef\]](#)
31. Hayashi, Y.; Matsukubo, J. Improvement the robustness on geographical networks by adding shortcuts. *Physica A Stat. Mech. Appl.* **2007**, *380*, 552–562. [\[CrossRef\]](#)
32. Xiao, Y.D.; Lao, S.Y.; Hou, L.L.; Small, M.; Bai, L. Effects of edge directions on the structural controllability of complex networks. *PLoS ONE* **2015**, *10*, e0135282. [\[CrossRef\]](#)
33. Yan, H.Y.; Hou, L.L.; Ling, Y.X.; Wu, G.H. Optimizing complex networks controllability by local structure information. *Int. J. Mod. Phys. C* **2016**, *27*. [\[CrossRef\]](#)
34. Sohn, I. A robust complex network generation method based on neural networks. *Physica A Stat. Mech. Appl.* **2019**, *523*, 593–601. [\[CrossRef\]](#)
35. Kwon, H.; Yoon, H.; Choi, D. Restricted evasion attack: Generation of restricted-area adversarial example. *IEEE Access* **2019**, *7*, 60908–60919. [\[CrossRef\]](#)
36. Kwon, H.; Yoon, H.; Park, K.W. Acoustic-decoy: Detection of adversarial examples through audio modification on speech recognition system. *Neurocomputing* **2020**, *417*, 357–370. [\[CrossRef\]](#)
37. Lü, L.Y.; Jin, C.H.; Zhou, T. Similarity index based on local paths for link prediction of complex networks. *Phys. Rev. E* **2009**, *80*. [\[CrossRef\]](#)
38. Dong, L.Y.; Li, Y.L.; Yin, H.; Le, H.; Rui, M. The algorithm of link prediction on social network. *Math. Probl. Eng.* **2013**, *2013*. [\[CrossRef\]](#)
39. Yao, L.; Wang, L.N.; Pan, L.; Yao, K. Link prediction based on common-neighbors for dynamic social network. *Procedia Comput. Sci.* **2016**, *83*, 82–89. [\[CrossRef\]](#)
40. Adamic, L.A.; Adar, E. Friends and neighbors on the web. *Soc. Netw.* **2003**, *25*, 211–230. [\[CrossRef\]](#)
41. Zhou, T.; Lü, L.; Zhang, Y.C. Predicting missing links via local information. *Eur. Phys. J. B* **2009**, *71*, 623–630. [\[CrossRef\]](#)
42. Wang, P.; Xu, B.W.; Wu, Y.R.; Zhou, X.Y. Link prediction in social networks: The state-of-the-art. *Sci. China Inform. Sci.* **2015**, *58*. [\[CrossRef\]](#)
43. Ravasz, E.; Somera, A.L.; Mongru, D.A.; Oltvai, Z.N.; Barabási, A.L. Hierarchical organization of modularity in metabolic networks. *Science* **2002**, *297*, 1551–1555. [\[CrossRef\]](#)
44. Fouss, F.; Piroette, A.; Renders, J.M.; Saerens, M. Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE T. Knowl. Data En.* **2007**, *19*, 355–369. [\[CrossRef\]](#)
45. Lü, L.Y.; Zhou, T. Link prediction in complex networks: A survey. *Physica A Stat. Mech. Appl.* **2011**, *390*, 1150–1170. [\[CrossRef\]](#)
46. Chebotarev, P.; Shamis, E.V. The matrix-forest theorem and measuring relations in small social groups. *Automat. Rem. Contr.* **1997**, *58*, 1505–1514.

-
47. Zanin, M.; Sun, X.Q.; Wandelt, S. Studying the topology of transportation systems through complex networks: Handle with care. *J. Adv. Transport.* **2018**, 3156137. [[CrossRef](#)]
 48. Gupta, R.D.; Kundu, D. Generalized exponential distribution: Existing results and some recent developments. *J. Stat. Plan. Infer.* **2007**, 137, 3525–3536. [[CrossRef](#)]
 49. Xu, T.Q.; Chen, Y.P. Two-sided M-Bayesian credible limits of reliability parameters in the case of zero-failure data for exponential distribution. *Appl. Math. Model.* **2014**, 38, 2586–2600. [[CrossRef](#)]
 50. Liu, N.; Zhang, J.; Zhang, H.; Liu, W.X. Security assessment for communication networks of power control systems using attack graph and MCDM. *IEEE Trans. Power Deliver.* **2010**, 25, 1492–1500. [[CrossRef](#)]
 51. Xiao, Y.H.; Wu, W.T.; Wang, H.; Xiong, M.M.; Wang, W. Symmetry-based structure entropy of complex networks. *Physica A Stat. Mech. Appl.* **2008**, 387, 2611–2619. [[CrossRef](#)]
 52. Liu, W.Y.; Li, X.; Liu, T.; Liu, B. Approximating betweenness centrality to identify key nodes in a weighted urban complex transportation network. *J. Adv. Transport.* **2019**, 9024745. [[CrossRef](#)]
 53. Krishnaveni, S.; Pethalakshmi, A. Toward automatic quality detection of Jasmenum flower. *ICT Express* **2017**, 3, 148–153. [[CrossRef](#)]
 54. Wang, R.B.; Xu, H.Y.; Li, B.; Feng, Y. Research on method of determining hidden layer nodes in BP neural network. *Comput. Technol. Dev.* **2018**, 28, 31–35.