Cognitive Computation
https://doi.org/10.1007/s12559-021-09873-2

A Novel Functional Network Based on Three-way Decision for Link Prediction in Signed Social Networks

Qun Liu[1] · Ying Chen[1] · Gangqiang Zhang[1] · Guoyin Wang[1]

Received: 20 November 2020 / Accepted: 29 April 2021

Abstract
Aiming to reveal the potential relationships between users, link prediction has been considered as a fundamental research issue in signed social networks. The key of the link prediction is to measure the similarity between users. Many existing researches use connections between users and their common neighbors to measure the similarities, and these methods rely too much on the structure of social networks. Most of them use the deep neural network to enhance the prediction accuracy. However, the complete structure of the huge social network cannot be captured easily, and the models learnt by the deep neural network are unexplainable and uncontrolled. As an explainable model, functional network is a recent replacement for standard neural network. Therefore, we revise the traditional strategy of functional network and propose a novel functional network framework. Firstly, the attributes are preprocessed through the cloud model to define their importance before inputting them into the functional network. Then the association algorithm is used to do aggregate computation in computing neurons for defining the connections between neurons well. Finally, we use three-way decisions to process the samples in the boundary to optimize the performance of model. Experiments executed on six real datasets show that our method has significantly higher link prediction precision than the state-of-the-art works. From our discussions, the improved functional network can be a valid replacement for neural networks in some fields.

Keywords Functional network · Three-way decisions · Cloud model · Link prediction

Introduction

The task of the link prediction is to predict the positive or negative sign for connections between users in the signed social networks [1]. By revealing the potential relationships between users, link prediction can be used in many applications. For example, most consumers are willing to share their shopping experiences with their friends in the recommendation systems [2]. If two users do not trust each other, they are unlikely to share with each other.

How to predict the unknown relationships in a signed social network? Most of existing researches use the connections between users to measure their similarities, which divided into traditional methods and deep learning methods [3].

✉ Qun Liu
liuqun@cqupt.edu.cn

1 Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China

1) Traditional Methods. Common Neighbors (CN) [4] used the number of users' common neighbors to evaluate their similarities. Two users are more similar if they have more common neighbors. In order to help authors decide where they should submit their manuscripts, the Content-based Journals & Conferences Recommender System is put forward in [5]. Logistic Regression (LR) method in [6] used the triangles including the connections between users and the number of their common neighbors to measure the similarity. Overlapping community detection is more complex due to the ambiguous of nodes which may be partitioned to different communities simultaneously. The rough set theory based on uncertain similarity between nodes is defined in dual-nucleus subspace by fully considering the topological structure [7]. In [8], the authors put forward the Label Propagation (LP), it begins from the initially collected labels, and propagates trust scores along the network. Label Spreading (LS) method defines a smooth function which is inspired by the spreading activation networks and diffusion kernels to calculate the similarity at

Published online: 01 June 2021

each iteration [9]. A graphical model is designed named E-Trust [10]. It incorporates type-based dyadic and triadic correlations. These methods are genuinely dependent on the network structures.

2) Deep Learning Methods. Based on transfer learning, BP tri-train used the idea of collaborative training to predict the social relationships of users in cross-social networks [11]. Recently, Graph Neural Network (GNN) has been received more and more research attentions and achieved good results in many machine learning tasks, e.g., semi-supervised node classification [12], network embedding [13], and link prediction [14]. It introduces neural network (NN) into graph data by defining convolution [12], attention [15], and other mechanisms. A semi-supervised graph convolutional/attention model is proposed in [14] which describes a node's embedding by performing convolutional operations on its neighbors' embedding. Though the deep learning methods can receive more excellent performances than the traditional ways do, they are based on a NN with a large number of layers and nonlinear activation functions and suffer from unexplainable and uncontrollable problems.

Three-way decisions (3WD) model [16, 17] was proposed by Yao in the study of the decision-theoretic rough sets model [18], which provided a reasonable semantic interpretation for the positive, negative, and boundary regions of the rough set model. In recent years, the theoretical research and practical application of the three-way decisions have been greatly developed, such as the three-way granular computing [19], the three-way classification [20], the three-way cluster analysis [21], the three-way medical decision making [22], and three-way recommendation [23]. The three-way decisions provide an effective strategy for solving complex problems [24].

As we all know, one of the key issue is how to process the uncertain information in artificial intelligence field. At present, there are many theories to cope with it from different points, such as fuzzy set, rough set, probability statistics, evidence theory, and so on. The cloud model theory as a new method was proposed by Prof. Li in 1995, which provides a new way to realize the bidirectional cognitive transformation between qualitative concept and quantitative data [25]. The authors in [26] put forward a new way called cloud transform to fit the real distribution of data. It can map the distribution of data and discover the association rules that can be understood easily. In [27], the concept tree of shapes is constructed, and the method of shape description of time-series, named time-series representation with cloud model, is presented. By mapping qualitative linguistic word into quantitative values using cloud models, the literature [28] proposes a new evaluating way showing lifetime difference of products under different working conditions.

To solve the problems of existing works, we use functional network (FN), a recent replacement for standard NN, to predict the sign of links through learnt neuron functions. In this paper, due to the different value range of each attribute, we first prepare the attributes with a cloud model to transform their values into the interval [0, 1]. The preprocessing not only discretizes the attributes but also mining its influence on link prediction results. Then they will be input into the FN model whose network architecture is gained by the Apriori algorithm. Furthermore, the unknown parameters which are used to build the model for link prediction are learnt by solving the following constrained least squares (CLS) problem. To further optimize the effect of FN model, we use three-way decisions to process the samples of the boundary. Experiments held on six real datasets show that our proposed method has significantly higher precision than the existing works.

The main contributions of our work are as follows:

1) To depict the characteristics of statistics and fuzziness on attributes, we discretize the attributes with a cloud model to define their importance during the prediction procedure before inputting.
2) The interpretable framework FN is constructed to predict links based on the Apriori algorithm. Our method not only improves accuracy but also overcomes the highly dependence of the network structure.
3) Incorporating the three-way decisions into a FN model, we propose 3WD-FN, which significantly improves the inferring precision(+7.73-31.60%) against the state-of-the-art methods.

The rest of this paper is organized as follows: Section 2 introduces the preliminaries; Section 3 presents the details of FN and 3WD-FN; in section 4, the experimental results of our proposed method are shown, which are compared with some existing advance works; finally, conclusions and the future works are given in section 5.

Preliminaries

Functional Network

Since 1998, functional networks(FNs) have been proposed which have no connection weights between neurons, and activation functions are not fixed, but learnable [29]. Note that FN can bring domain knowledge together to determine its structure, and the training data are used to estimate the unknown parameters of the neurons in the input layer. During the FN learning procedure, a method based on minimizing a least squares error function is used to obtain the

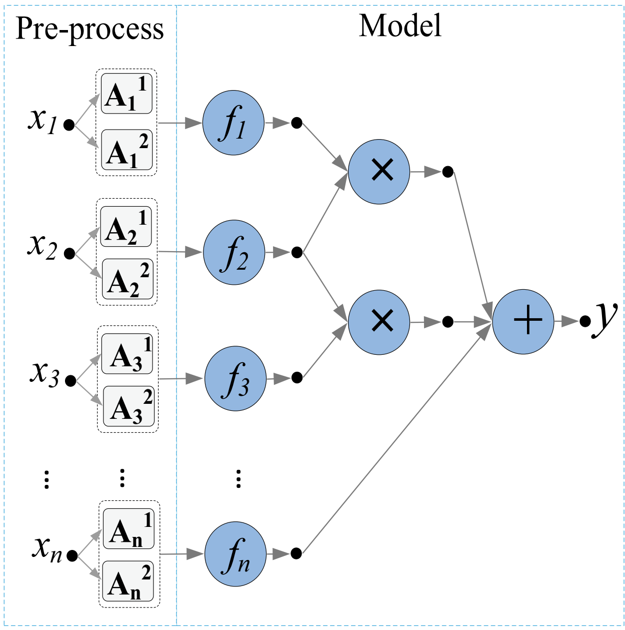


Fig. 1 The structure of functional network

hyper-parameters. One of the advantages of FN is that the learning procedure is based on solving linear equations.

The FN model structure shown in Fig.1 consists of the following elements [30]:

1) Layers of Storing Units. These units are plotted by small filled circles.
2) Layers of Neurons (Computing or Functional Units). Each neuron is plotted by a blue circle with its name in it. These neurons are the computing units that calculate a set of the input values and pass a set of output values to the next layer of storing units. Note that the activation functions of these neurons are not given but learnt.
3) Directed Links. The neurons are connected to the storing units by a set of directed links, which indicate the direction of information flow.

The standard NN and FN have many similarities, but they also differ in many ways. The main differences between NN and FN are shown in Table 1 [30].

Cloud Model

In the natural world, it is necessary to study the randomness and fuzziness of uncertain information to make the computers have comprehension and judgement capabilities similar to humans, especially the formal representation of knowledge [31].

Probability theory and fuzzy set theory are the most two effective tools in the research of uncertainty knowledge representation. In probability theory, the normal distribution is the most important probability distribution, which can be used as an approximation for a large number of random phenomena. As the foundation of fuzzy sets, membership functions are used to measure fuzzy degree, and normal (Gaussian) membership function is regarded as one of the most suitable membership functions for many fuzzy concepts [32].

The cloud model makes it possible to obtain the distributing range of a qualitative concept, which is defined as Eq. (1).

$$\mu : U \to [0,1], \quad \forall x \in U \quad x \to \mu(x). \tag{1}$$

Let C be a qualitative concept in a quantitative domain U that is described by precise numeric and x be a randomly realization of C. If there is a quantitative value $x \in U$, the certainty degree of $\mu(x) \in [0,1]$ has a stabilization tendency of random value. Then the distribution of x on U is defined as a cloud, and every x is defined as a cloud drop.

According to different probability distributions, there are many ways to implement cloud models. Due to the important position of Gaussian distribution in probability theory and random theory, Gaussian cloud model is a commonly used cognitive model to describe qualitative concepts.

Table 1 Differences between neural and functional

Neural network	Functional network
Black box.	Can reproduce physical properties.
Have the ability to learn from data.	Have the ability to learn from data and domain knowledge.
Weights are learnt.	Neural functions are learnt.
All the active functions in each layer are identical.	Different neurons can have different neural functions.
Each active function has only a single argument.	Each neural functions can be multi-argument.
The network structure is selected by trial and error.	The network structure is problem-driven.
Loss function is minimized using back propagation methods.	Loss function is minimized by solving linear systems of equations.

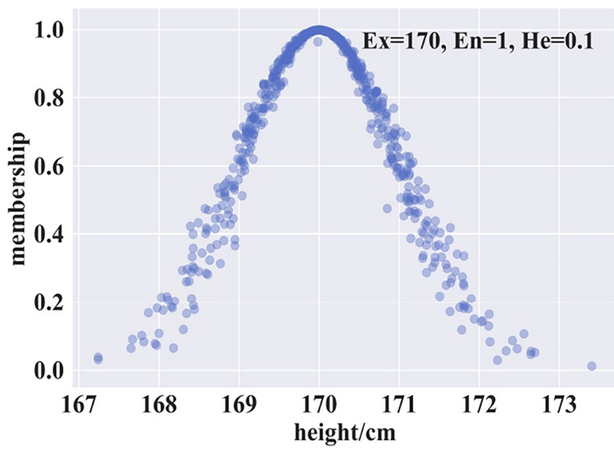


Fig. 2 A cloud map of the concept "around one hundred and seventy centimeters" composed of cloud drops

As shown in Fig. 2, the Gaussian cloud model uses three numerical features, namely *Ex* (expectation), *En* (entropy), and *He* (hyper entropy), to describe the qualitative concept, which just accords with human thought [31]. Among them, *Ex* is the measure of a basic certainty of the concept, *En* describes the uncertainty measure of the concept, and *He* represents the uncertainty measure of *En*. In the Gaussian cloud model, for concept C, the uncertainty of the k-th cloud drop in one-dimensional space is defined as Eq. (2).

$$\delta(x)_k = e^{-\frac{(x-Ex)^2}{2y^2}}, \quad y \sim N(En, He^2). \tag{2}$$

Three-way Decisions

Definition 1 Let $\Omega = \{s_1, s_2, \ldots, s_m\}$ be a finite set of m states and $A = \{a_1, a_2, \cdots, a_n\}$ represents a finite set of n possible actions. Let $\lambda(a_j \mid s_i)$ be the loss, for taking action a_j when the state is s_i. Let $P(s_i \mid z)$ represent the conditional probability that the object z owns the state s_i, while supposing action a_j is taken. The expected loss from taking action on a_j can be given as Eq. (3).

$$R(a_j \mid z) = \sum_{i=1}^{m} \lambda(a_j \mid s_i) P(s_i \mid z). \tag{3}$$

In rough set theory [18], a set Z is approximated by three regions, namely, the positive region $POS(Z)$ includes the objects that are sure to belong to Z, the boundary region $BND(Z)$ includes the objects that are possible belong to Z, and the negative region $NEG(Z)$ includes the objects that are not belong to Z.

In the link prediction, we have a set of two states $\Omega = \{Z, Z^C\}$ indicating that is friend relationship between users or foe relationship, respectively. With respect to these three regions, the set of actions is given by $A = \{a_P, a_B, a_N\}$, where a_P , a_B and a_N represent the three actions in classifying a link z, namely, deciding $z \in POS(Z)$, deciding $z \in BND(Z)$, and deciding $z \in NEG(Z)$. Six losses functions are imported, λ_{PP}, λ_{BP} and λ_{NP} denote the losses incurred for taking actions a_P , a_B, a_N , respectively, when a link belongs to Z. And λ_{PN} , λ_{BN} and λ_{NN} denote the losses incurred for taking these actions when the link does not belong to Z.

The expected losses associated with taking different actions for link z can be expressed as:

$$\begin{aligned} R(a_P \mid z) &= \lambda_{PP} P(Z \mid z) + \lambda_{PN} P(Z^C \mid z), \\ R(a_B \mid z) &= \lambda_{BP} P(Z \mid z) + \lambda_{BN} P(Z^C \mid z), \\ R(a_N \mid z) &= \lambda_{NP} P(Z \mid z) + \lambda_{NN} P(Z^C \mid z). \end{aligned} \tag{4}$$

From the Bayesian decision criterion, the set of actions with the least expected loss needs to be chosen:

(P) If $R(a_P \mid z) \le R(a_B \mid z)$ and $R(a_P \mid z) \le R(a_N \mid z)$ both hold, then $z \in POS(Z)$;

(B) If $R(a_B \mid z) \le R(a_P \mid z)$ and $R(a_B \mid z) \le R(a_N \mid z)$ both hold, then $z \in BND(Z)$;

(N) If $R(a_N \mid z) \le R(a_P \mid z)$ and $R(a_N \mid z) \le R(a_B \mid z)$ both hold, then $z \in NEG(Z)$.

Then a reasonable assumption can be made:

$$\begin{aligned} 0 \le \lambda_{PP} \le \lambda_{BP} < \lambda_{NP}, \\ 0 \le \lambda_{NN} \le \lambda_{BN} < \lambda_{PN}. \end{aligned} \tag{5}$$

That is, the loss of classifying a link z in Z into the positive region $POS(Z)$ is less than or equal to the loss of classifying z into the boundary region $BND(Z)$, and both of the above losses are strictly less than the loss of classifying z into the negative region $NEG(Z)$. The reverse order of losses is used for classifying a link not in Z. Since $P(Z \mid z) + P(Z^C \mid z) = 1$, under above condition, the three new decision rules ($P1$)-($N1$) are as follows:

($P1$) If $P(Z \mid z) \ge \alpha$ and $P(Z \mid z) \ge \gamma$ both hold, then $z \in POS(Z)$;

($B1$) If $\beta \le P(Z \mid z) \le \alpha$ both hold, then $z \in BND(Z)$;

($N1$) If $P(Z \mid z) \le \beta$ and $P(Z \mid z) \le \gamma$ both hold, then $z \in NEG(Z)$.

Among them,

$$\begin{aligned} \alpha &= \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}, \\ \beta &= \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}, \\ \gamma &= \frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}. \end{aligned} \tag{6}$$

The conditions of rule ($B1$) suggest that $\alpha > \beta$ may be a reasonable constraint. If we obtain the following condition on the loss functions [16]:

$$\frac{(\lambda_{NP}-\lambda_{BP})}{(\lambda_{BN}-\lambda_{NN})} > \frac{(\lambda_{BP}-\lambda_{PP})}{(\lambda_{PN}-\lambda_{BN})}. \tag{7}$$

then $0 \le \beta < \gamma < \alpha \le 1$. In this case, the following simplified rules are obtained:

($P2$) If $P(Z \mid z) \ge \alpha$, then $z \in POS(Z)$;

($B2$) If $\beta < P(Z \mid z) < \alpha$, then $z \in BND(Z)$;

($N2$) If $P(Z \mid z) \le \beta$, then $z \in NEG(Z)$.

Methods

Functional Network

As an interpretable model, FN can be the replacement of the NN. It is constructed by the functional equations with domain prior knowledge initially. In this paper, we use the Apriori algorithm to get the initial network structure. The model of constructing the network is shown in Fig. 1.

A general scheme of FN proposed in literature [33] used the fuzzy membership function to discretize the attributes in the input layer for the network structure constructing. However, the membership function used in literature [33] leads to distortion of the input easily. The cloud model considers not only fuzziness but also statistics of inputting attributes. Therefore, the preprocessing for the attributes with a cloud model can describe the distribution of each attribute accurately. After discretizing and normalizing, the attributes can be fitted to be used in Apriori algorithm well.

The object of the link prediction task is to predict users' relationships, so it belongs to a binary classification. After normalizing the attributes, all attributes satisfy the normal distribution, and the forward Gaussian cloud transform method can be used to granulate each attribute value into the interval [0, 1] as shown in Eq. (8) [34]. We use A_i^j, the certainty degree of attribute, to represent original data.

$$\begin{aligned} &A_i^j(x_i) = e^{-\frac{(x-Ex)^2}{2En_j'^2}}, \\ &En_j' \sim N\left(En_j, He_j^2\right), \quad i = 1,2,\dots n, j = 0,1. \end{aligned} \tag{8}$$

A_i^j describes the certainty degree that the ith attribute belongs to the concept j.

For simplicity, we substitute x to the certainty degree value A in Eq. (8). From Fig. 1, we can see there are two layers in our FN model. In the first layer, each neuron is a basis function whose parameter needs to be learnt by training data. The basis function is defined as Eq. (9).

$$f(x_i) = w_i\phi(x_i), i = 1,2,\cdots n. \tag{9}$$

where ϕ are the common families of linearly independent functions, n indicates the number of the input attributes and the coefficients w_i are the parameters of the FN. Note that if the basis functions are different, the prediction effect will be greatly affected. In our work, the basis function of polynomial family is chosen $\phi = \{x\}$.

We use the Apriori algorithm to find out frequent item sets. Then a model is built according to these frequent item sets, and the connections between neurons in the network can be constructed. Firstly, some attributes can construct the frequent 2-item set, and the others cannot. For those belonging to the frequent 2-item set, their values will be as the input for a neuron in the second layer, and then calculated by Eq. (10). The obtained results will be as the input for the neuron in the output layer. Others not belonging to the frequent 2-item sets will be as the input for a neuron in the second layer, its value will also be the input of the neuron in the output layer directly.

Taking the DBLP dataset as an example. Firstly, 15 attributes of the node have been extracted from the input data as $a_1, a_2, a_3, \cdots, a_{15}$, respectively, and they will be calculated to obtain the output of first layer as $f_1, f_2, f_3, \cdots, f_{15}$, respectively. Then, the Apriori algorithm is used to get the corresponding frequent 2-item, such as $\{1,6\}, \{3,4\}, \{7,11\}, \{7,12\}, \{7,13\}$. After computed by Eq. (10), the inputs of third layer as g_1, g_2, g_3, g_4, g_5 are generated. As shown in Fig. 3, the frequent 2-item sets and the non-frequent item-sets of neuron functions are finally accumulated to get the final value of y.

$$g(x_i, x_j) = f(x_i)f(x_j), (i,j) \in M_2. \tag{10}$$

where M_2 means the frequent 2-item set.

For the output layer, the accumulation operation will be completed in the neuron according to Eq. (11).

$$y = \sum_{i=1}^{I} f_i + \sum_{j=1}^{m_2} g_j. \tag{11}$$

where y is the predicted value of the output. I represents neuron functions of non-frequent item-sets. m_2 represents the number of frequent 2-item set obtained.

Hence, the loss function of the FN model is defined as Eq. (12):

$$e_k = y^{(k)} - \left(\sum_{i=1}^{I} f_i + \sum_{j=1}^{m_2} g_j\right). \tag{12}$$

where k represents the kth sample. $y^{(k)}$ is the ground truth. To find the optimal functional parameters w_i in Eq. (9), we need to minimize the sum of squared errors(SSE) by Eq. (13).

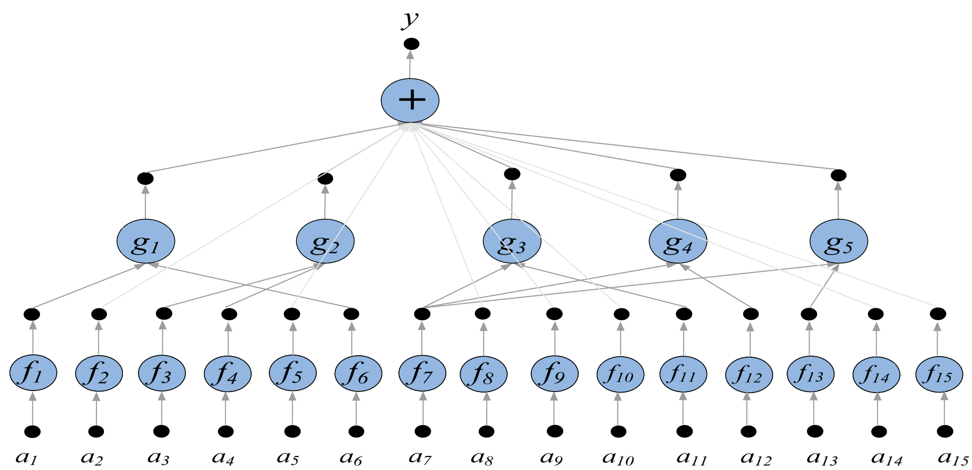


Fig. 3 The structure of the FN model constructed on the DBLP dataset

$$
\begin{aligned}
Q &= \sum_{k=1}^{P} e_k^2 \\
&= \sum_{k=1}^{P} \left[y^{(k)} - \left(\sum_{i=1}^{I} f_i + \sum_{j=1}^{m_2} g_j \right) \right]^2 \\
&= \sum_{k=1}^{P} \left[y^{(k)} - \left(\sum_{i=1}^{I} w_i \phi(x_k) + \sum_{j=1}^{m_2} g_j \right) \right]^2 .
\end{aligned} \tag{13}
$$

Given the topology of an FN, we need to know the conditions for it's structure uniqueness. It is whether several sets of functions (neurons) lead to the same output for the same input [35]. To guarantee the uniqueness of representation, a restriction condition as $\lambda\left(w_i^2 - 1\right)$ is added to Eq. (13). Applying the Lagrange multiplier method, Eq. (13) can be rewritten as Eq. (14).

$$
\begin{aligned}
Q_\lambda &= Q + \sum_{i=1}^{n} \lambda\left(w_i^2 - 1\right) \\
&= \sum_{k=1}^{P} \left[y^{(k)} - \left(\sum_{i=1}^{I} w_i \phi(x_k) + \sum_{j=1}^{m_2} g_j \right) \right]^2 + \\
&\quad \sum_{i=1}^{n} \lambda\left(w_i^2 - 1\right).
\end{aligned} \tag{14}
$$

Eq. (14) is calculated partial derivatives of w_i and λ as shown in Eq. (15).

$$
\begin{aligned}
\frac{\partial Q_\lambda}{\partial w_i} &= -2 \sum_{k=1}^{P} e_k \left(\phi(x_k) + \sum_{j=1}^{m_2} \frac{\partial g_j}{\partial w_i} \right) + 2\lambda w_i, \\
&\qquad P = 1, 2, 3 \ldots n, \\
\frac{\partial Q_\lambda}{\partial \lambda} &= \sum_{i}^{n} \left(w_i^2 - 1\right).
\end{aligned} \tag{15}
$$

Let the partial derivative be 0,

$$
\begin{aligned}
\frac{\partial Q_\lambda}{\partial w_i} &= -2 \sum_{k=1}^{P} e_k \left(\phi(x_k) + \sum_{j=1}^{m_2} \frac{\partial g_j}{\partial w_i} \right) + 2\lambda w_i = 0, \\
&\qquad P = 1, 2, 3 \ldots n, \\
\frac{\partial Q_\lambda}{\partial \lambda} &= \sum_{i}^{n} \left(w_i^2 - 1\right) = 0.
\end{aligned} \tag{16}
$$

After solving Eq. (16), the optimal functional parameters can be received. This method is simple and with fast convergence speed.

Three-way Decision Functional Network

Usually, the expression of relationships in social network is ambiguous. If the two-way decisions method is used to analyze the social relationships directly, we will not be able to obtain the desired results. Hence as a compromised decision-making method, the three-way decisions can effectively solve the uncertainty of social relationships classification.

In this paper, the set of links can be divided into three regions, $POS(Y)$ includes friend relationship between users, $BND(Y)$ includes relationship that need further-exam, and $NEG(Y)$ includes relationship that is foe.

Given the loss functions as Eq. (5), we can make the proper decisions for links based on the parameters (α, β), which are computed by loss functions.

For a relationship y, if the probability of being a friend relationship is $p(Y \mid y)$, where Y is approximated by three regions. Then the three-way decisions solution is:

If $P(Y \mid y) \geq \alpha$, then y is friend relationship;

If $P(Y \mid y) \leq \beta$, then y is foe relationship;

If $\beta < P(Y \mid y) < \alpha$, then y needs further-exam.

In this paper, to further improve the accuracy of prediction, the three-way decisions are used to filter out the boundary data sample. As shown in Fig 4, we add the third layer that accepts the input of frequent tri-item sets. For

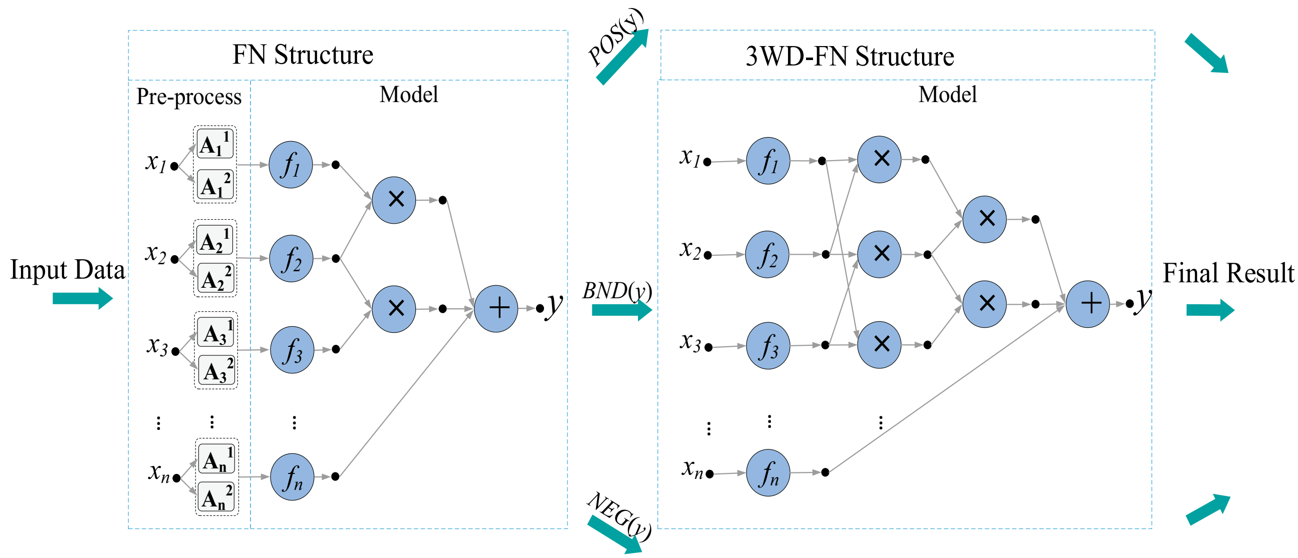


Fig. 4 The three-way decisions model of regions and decisions

those belonging to the frequent tri-item sets, their values will be as the input for a neuron in the third layer, and then calculated by Eq. (17).

$$h(x_i, x_j, x_k) = g(x_i)g(x_j) \times g(x_j)g(x_k) \times g(x_i)g(x_k), \quad (i,j,k) \in M_3. \tag{17}$$

where M_3 means the frequent tri-item set. Furthermore, the neuron in output layer will be computed by Eq. (18).

$$y = \sum_{i=1}^{I} f_i + \sum_{j=1}^{m_2} g_j + \sum_{k=1}^{m_3} h_k. \tag{18}$$

where m_3 represents the number of frequent tri-item sets. The loss function of the model can be rewritten as Eq. (19).

$$Q_\lambda = \sum_{k=1}^{P} \left[y^{(k)} - \left(\sum_{i=1}^{I} w_i \phi(x_k) + \sum_{j=1}^{m_2} g_j + \sum_{k=1}^{m_3} h_k \right) \right]^2 + \sum_{i}^{n} \lambda (w_i^2 - 1). \tag{19}$$

Note that after training with FN and 3WD-FN, we can obtain the parameters by Eq. (14) and Eq. (19), respectively. It means that we predict the parameters of FN with the whole dataset in the first prediction phase, then predict the parameters of the 3WD-FN with the data in the boundary domain during the second prediction phase. Finally, the prediction results are obtained by a unified integration procedure. The algorithm of constructing network structure model is the following Algorithm 1.

Algorithm 1: The algorithm of constructing network structure model

Input: Training Data D, Training label l, Confidence Coefficient $cc1$, $cc2$, Threshold t
Output: Complete Model
Calculate the classification result D_cl and the membership value A by GCT based on the D;
Calculate the frequent itemset $fi1$ based on $cc1$ and D_cl;
Build FN-model based on the $fi1$;
Set $loss\_1 = 0, d = 0$;
while $d < len(A)$ **do**
  Calculate prediction $y[d]$ base on $A[d]$;
  $loss\_1 = loss\_1 + (y[d] - l[d])^2$;
  $d = d + 1$;
end
Calculate parameter W_1 by polynomial solving based on $loss\_1$;
Set $d = 0$;
Create empty $listA'$;
while $d < len(A)$ **do**
  Calculate prediction $y[d]$ base on $A'[d]$ and FN-model ;
  if $y[d] < t$ *and* $y[d] > -t$ **then**
    append $A[d]$ into A';
end
Calculate the frequent itemset $fi2$ based on $cc2$ and D_cl ;
Build 3WD-FN-model based on the $fi2$;
Set $loss\_2 = 0, d = 0$;
while $d < len(A)$ **do**
  calculate prediction $y[d]$ base on $A'[d]$;
  $loss\_2 = loss\_2 + (y[d] - l[d])2$;
  $d = d + 1$;
end
Calculate parameter W_2 by polynomial solving based on $loss\_2$;
Model is built;

Table 2 Statistics of six datasets

Relationship	Dataset	Nodes	Relationships	Support of Apriori	Thresholds (α)
Trust	Alpha	1,718	9,983	0.280	0.25
Trust	OTC	2,614	15,087	0.315	0.30
Trust	Epinions	934	60,408	0.295	0.15
Friends	Slashdot	5,670	39,585	0.240	0.25
Support	WikiVote	3,557	97,971	0.315	0.15
Collaboration	DBLP	3,714	19,709	0.370	0.25

Experiments

Datasets We evaluate our method on the following six real datasets, and the statistical information of the datasets is listed in Table 2.

1) Epinions [36] is a product review site that users can rate the products by 1 to 5 scores. Each rating contains the user name, product name, product category, and the rating score. We only collect binary labels for their relationships.
2) Slashdot [37] is a website of technology-related network, where users share technologies with each other. After introducing, the Slashdot Zoo allows users to tag each other as friends or foes.
3) Alpha and OTC [38] are social platforms that users can use the encrypted currency to conduct anonymous transactions on the Internet. Both of them allow users to rate others on a scale of -10 to +10. In our model, we consider the score which is less than 0 is regarded as distrust, the score greater than 0 represents trust.
4) DBLP [39] is a collaboration network. The dataset contains the top 400 high-quality community information and illustrates cooperative relations or not. Our purpose is to predict the collaborative relationships between two voters.
5) The Wikipedia [40] community elects administrators through open discussions or voting. It has two types of relationships: support or opposes.

In order to protect the privacy of the users, the datasets used only include the structural information of the node and does not contain any private information, such as name, occupation, and shopping information, etc. The input attributes of these six datasets are uniformly 15 attributes as shown in Tables 3 which includes in/out-degree of node, number of neighbors, and clustering coefficients, etc.

Baseline Methods In this section, we will check the efficiency of the three-way decisions solution in FN. We compare LR [6], LP [8], LS [9], e-Trust [10], and SiGAT [14] methods. The labeled datasets are divided into the training and test set to learn and evaluate the comparison methods, respectively.

LR: it trains a multi-label classifier to predict the label of a relationship based on the features extracted from heterogeneous users' behaviors.

LP: it begins from the initially collected labels and propagates trust score along the network.

LS: it is similar to LP, but uses initial values to smooth the inferred values at each iteration.

e-Trust: by incorporating the discovered correlation patterns into a factor graph model, this method formalizes trust into multiple types and proposes a graphical model to incorporate type-based dyadic and triadic correlations.

SiGAT (Signed Graph Attention Network): it incorporates graph motifs into GAT to capture two well-known theories in signed network research, i.e., balance theory and

Table 3 Features defined in social network with relationship $e(u, v)$

Feature	Description
$d_{in}(u)$, $d_{in}(v)$	Represent the in-degree of node u and v, respectively.
$d_{out}(u)$, $d_{out}(v)$	Represent the out-degree of node u and v, respectively.
$d_{network}(u)$, $d_{network}(v)$	Represent the total-degree of node u and v in network, respectively.
$totalNeighbor(u, v)$	Represents the total number of neighbors of u and v in an undirected network.
$commonNeighbor(u, v)$	Represents the number of common neighbors of u and v in an undirected network.
$structuralBalance(a)$	Represents triad relations between u, v, and their common neighbor, and three nodes are mutual friends.
$structuralBalance(b)$	Represents two nodes are friends, and they are mutual enemies of the third node.
$structuralBalance(c)$	Represents two nodes are enemies, and they are mutual friends of the third node.
$structuralBalance(d)$	Represents three nodes are mutual enemies.
$structValue(u)$, $structValue(v)$	Represent the structure value of node u and v, respectively.
$edgeClusteringCoefficient(u, v)$	Represents the sum of the node u and v clustering coefficient value.

Table 4 Prediction performance of our proposed method comparing with existing works (%)

Dataset	Method	Precision	Recall	F1-Score	Accuracy
DBLP	LR	60.91	100.00	66.37	60.91
	LP	91.44	76.81	82.27	81.18
	LS	91.09	75.24	82.18	80.13
	e-Trust	80.62	86.59	82.29	78.62
	FN	76.47	80.79	77.19	72.82
	3WD-FN	74.33	93.19	82.49	73.55
Slashdot	LR	68.87	100.00	81.57	68.87
	LP	80.14	92.81	86.00	79.19
	LS	78.12	93.80	85.29	77.72
	e-Trust	80.25	90.17	84.85	77.81
	FN	91.99	69.15	78.67	66.17
	3WD-FN	92.61	96.43	94.43	90.42
Alpha	LR	90.67	100.00	95.11	90.67
	LP	91.25	99.54	95.21	90.92
	LS	90.87	99.69	95.07	90.63
	e-Trust	92.11	98.78	95.33	91.21
	FN	98.40	91.25	94.59	89.89
	3WD-FN	97.62	99.30	98.40	96.96
Wikivote	LR	79.13	100.00	88.35	79.13
	LP	81.54	99.26	89.53	81.62
	LS	80.24	99.73	88.93	80.36
	e-Trust	83.37	96.4	89.39	81.89
	FN	93.48	78.61	85.19	74.80
	3WD-FN	99.98	80.25	89.13	80.23
Epinions	LR	83.33	100.00	90.91	83.33
	LP	86.72	99.35	92.64	86.84
	LS	64.46	79.81	92.00	85.55
	e-Trust	89.77	96.67	93.11	88.07
	FN	81.06	85.27	82.53	70.83
	3WD-FN	96.06	93.30	94.30	89.26
OTC	LR	84.33	100.00	91.50	84.33
	LP	88.94	99.05	93.63	88.83
	LS	87.23	99.56	92.93	87.31
	e-Trust	90.28	93.32	93.61	88.74
	FN	99.96	84.39	91.51	84.37
	3WD-FN	99.95	84.43	91.54	84.40

status theory. And motifs offer the flexible structural pattern to aggregate and propagate messages on the signed network to generate node embeddings.

Our methods include FN and 3WD-FN. They are the models with or without using three-way decisions. For the type of link, it is determined by the positive and negative value of the final output y.

Hyper-parameters The support thresholds of Apriori algorithm are shown in Table 2 which indicates the parameter value of the FN model.

In order to optimize the effect of the 3WD-FN model, we first need to obtain a suitable threshold. So we tried different parameters and obtained their accuracy, precision, recall, and F1 measure results shown in Fig. 5. From the results, we selected five optimal values of 0.1, 0.15, 0.2, 0.25, and 0.3 for the six datasets to perform the tenfold crossover experiments, respectively. The thresholds for the boundary data of six datasets are set to $\beta = -\alpha$ as in Table 2. For example, when the threshold $\alpha = 0.1, \beta = -0.1$. We take the predicted value between $-0.1 < y < 0.1$ while the tuples are within the boundary, and then the 3WD-FN model is used to predict.

Evaluation Metrics In our experiments, we use four metrics to evaluate the performance. They are given as Eq. (20).

$$\begin{aligned}
&Precision = \frac{TP}{TP+FP},\\
&Recall = \frac{TP}{TP+FN},\\
&Accuracy = \frac{TP+TN}{TP+TN+FP+FN},\\
&F1-Score = \frac{2\times precision\times recall}{precision+recall}.
\end{aligned} \tag{20}$$

where TP represents the positive tuples that are correctly labeled by the model. Similarly, TN represents the negative tuples that are correctly labeled by the model. FP represents the negative tuples that are incorrectly labeled as positive. FN represents the positive tuples that are mislabeled as negative.

Performance Analysis Our experiment platform uses window10 and Intel Core i5-5200 CPU 2.20 GHz. The programming language is python, and the package networkx is used in the program to preprocess data. We randomly set the proportion of training and test data as 9:1 and conduct a ten-fold crossover experiment.

From the experimental results, the performance analysis is summarized as follows:

1) Prediction Performance: Table 4 shows the prediction performance for all methods, which indicates that our model has received better prediction results on six datasets. Especially, we can see the 3WD-FN model can raise in the range from 7.73% to 31.60% compared with the baseline methods. Figures 6 and 7 show the ten-fold crossover results of Alpha and Slashdot dataset. We can see that our method 3WD-FN excels to the other four methods in Alpha and Slashdot datasets. Comparing the 3WD-FN with FN, the former one is better than the latter. Moreover, with the decrease of the amount of data,

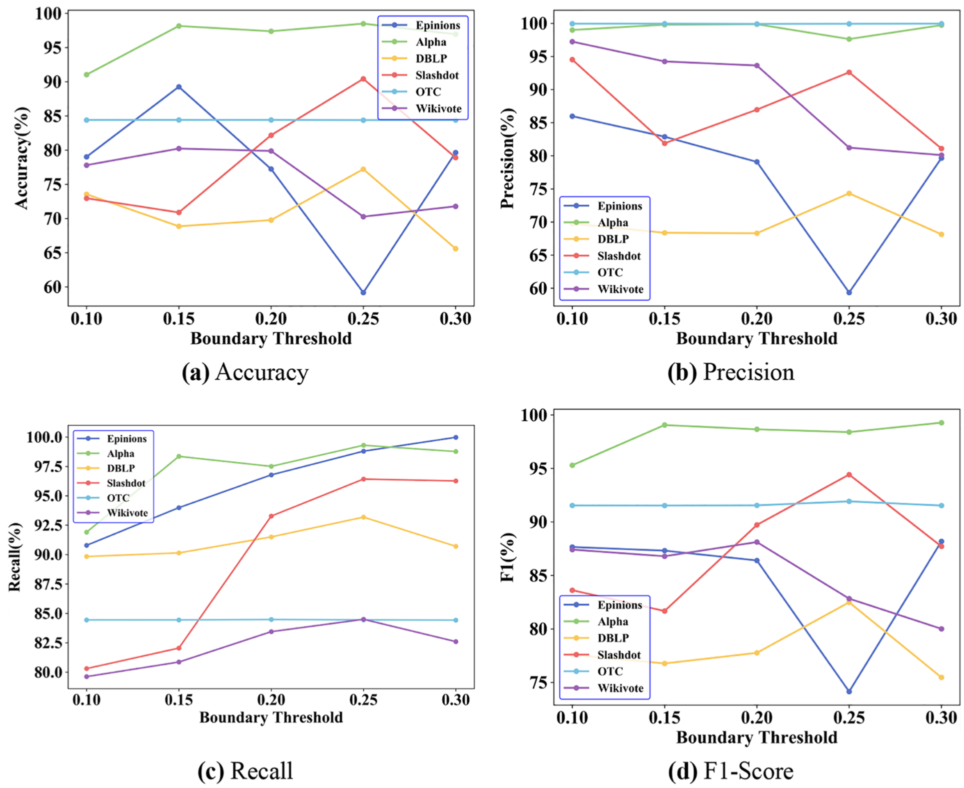


Fig. 5 The performing results with five thresholds for the six datasets

parameters of the 3WD-FN model can be solved easily and the performance is further improved. However, the reason why the effect of 3WD-FN is lower than the FN in OTC dataset is that there are fewer boundary samples in this dataset and leads to the purple line moves smoothly in Fig. 5. The merits of 3WD-FN cannot be exhibited.

2) Reasons: According to the results and strategies of baseline models, we analyzed some reasons. LR only considers the attribute features of the labeled relationships but ignores the unlabeled relationships. So, it cannot receive

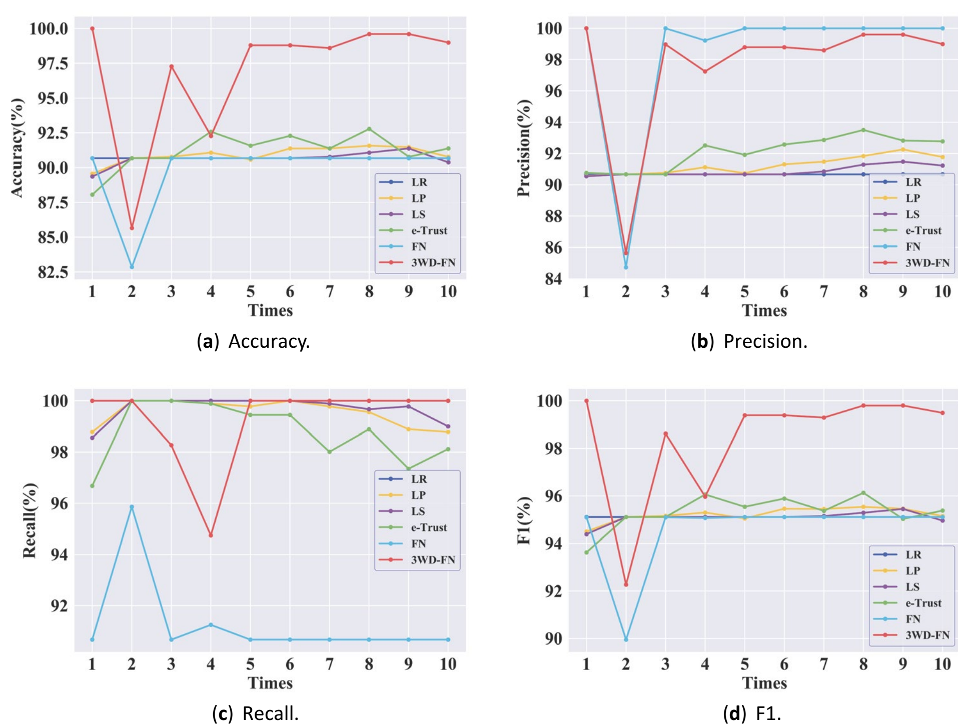
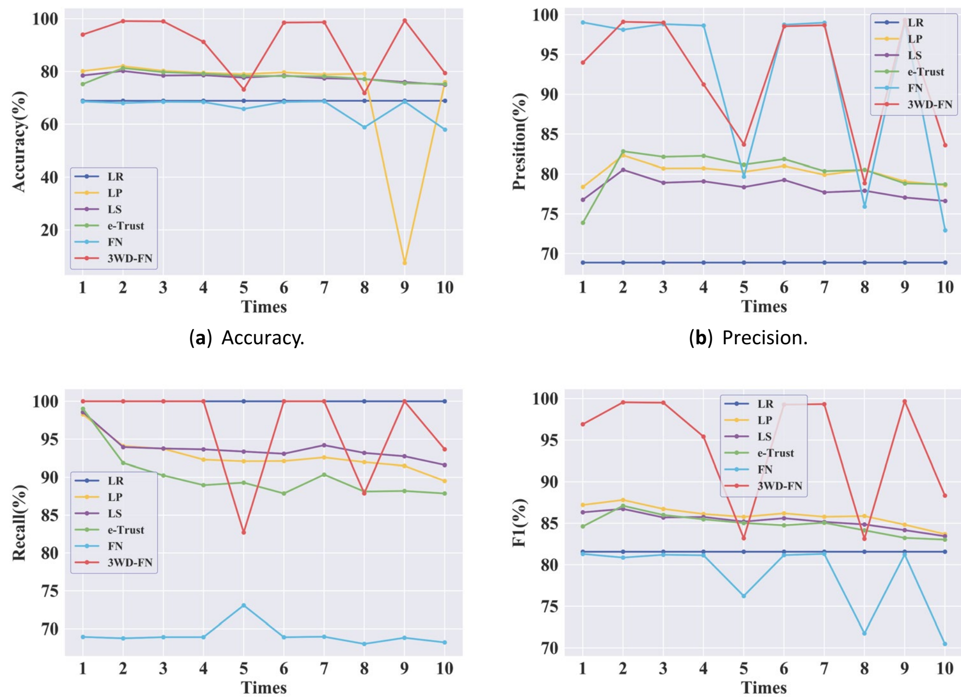


Fig. 6 The ten-fold crossover results of Alpha dataset

Fig. 7 The ten-fold crossover results of Slashdot dataset



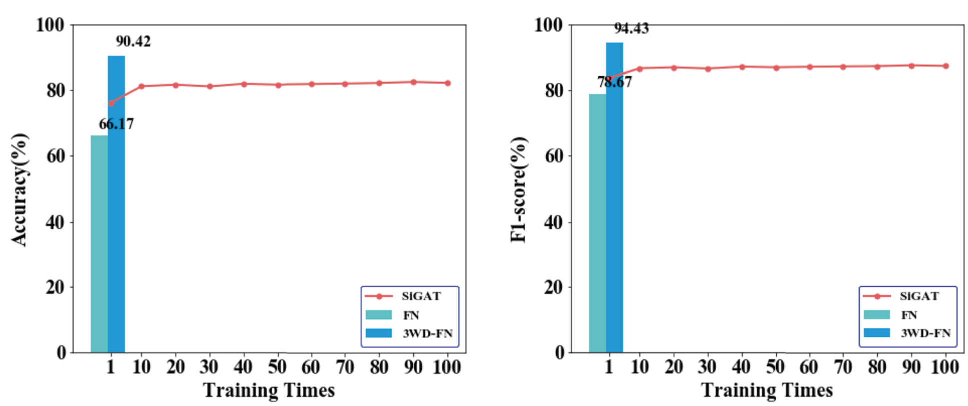
(a) Accuracy. (b) Precision. (c) Recall. (d) F1.

the complete structure of the social network. When the network is too sparse, the performance of e-Trust is very poor. Semi-supervised methods LP and LS leverage the correlations between labeled and unlabeled relationships, their performances are similar to the supervised method LR on Alpha, but much better than it on DBLP. Because the relationships of the neighbors are more likely to be the same type in DBLP than in the other datasets, the merit of LP and LS cannot be released.

3) Shortage of our model: From the results in Table 4, we can also see that our precision is relatively high, but the recall is lower. We analyze the datasets and find that the number of positive edges in the dataset is much larger than the negative edges. Our model relies on data very much. So, the prediction effect on the positive edges is better in our prediction results.

4) Comparison with the GNN: The results of Slashdot dataset are shown in Fig. 8 where the abscissa represents the training times of the model, and the ordinate represents the experimental results. It shows that our model is better than the advanced GNN model SiGAT. After training 90 times, the best accuracy result of SiGAT reaches 82.41%. But 3WD-FN achieves the best effect for only once, and the effect is obviously improved. In addition, compared with the "black box" structure of the NN model, our model structure is not only simpler than the neuron model structure, but also interpretable and controllable.

Fig. 8 The experimental results of Slashdot datasets of FN, 3WD-FN, and SiGAT



(a) Accuracy. (b) F1-Score.

5) Complexity: The time complexity of the NN is O(n). Our model is composed of a limited number of neurons, and the linear calculation of each neuron is $wx + b$. Therefore, the time complexity of our method is O(n), and the space complexity of our method is lower than traditional NNs. At the same time, the number of neurons and the connection mode are determined by the association rules between attributes. In the worst case, when the attributes are related to each other, the space complexity of our model is O(n^2).

Conclusions and Future Works

This paper studies how to infer relationships in signed social networks. We propose a method predicting links based on the functional networks and three-way decisions. Besides our model structure is more interpretable than the neural network, the time complexity of our model is also lower since it is a general linear model. Experimental results on six real-world datasets show that our method outperforms some advanced methods. Compared with NN, the structure of FN model is simple, the number of neurons and layers are relatively small. But it cannot cope with samples which are difficult to be distinguished. The introduce of 3WD can provide a good solution to help them be analyzed continually. It not only reduces the processing time but also enhances the performance of our model.

However, during the ten-fold crossover experiments, sometimes our model parameters cannot be solved in some folds. At this time, we use the average value of the parameters solved in other folds to replace in these folds. It means the shortcomings of our model is relying on the data sensitively.

In the future, we will continue to optimize the functional network model to make it more generalized. We will try to introduce the expert knowledge into our framework, and avoid relying on the data too much. Furthermore, we will consider a more effective way to construct and optimize our model and design better equations for calculating parameters in the neurons.

Acknowledgements This work is supported by the State Key Program of National Nature Science Foundation of China (61936001), the National Key Research and Development Program of China (2016QY01W0200), partly funded by National Nature Science Foundation of China (61772096), the Key Research and Development Program of Chongqing (cstc2017zdcy-zdyfx0091), and the Key Research and Development Program on AI of Chongqing (cstc2017rgzn-zdyfx0022).

Declarations

Conflicts of interest No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication. I would like to declare on behalf of my co-authors that the work described is original research. This work has not been published previously, and not under consideration for publication elsewhere, in whole or in part. All authors listed have approved the manuscript that is enclosed.

References

1. Wang S, Tang J, Aggarwal C, Chang Y, Liu H. Signed network embedding in social media. In Society for Indus Appl Math. 2017; pp. 327-335.
2. Liu S, Jiang C, Lin Z, Ding Y, Duan R, Xu Z. Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach. Info Sci. 2015;306:34–52.
3. Liu H, Hu Z, Mian A, Tian H, Zhu X. A new user similarity model to improve the accuracy of collaborative filtering. Knowl-Based Syst. 2014;56:156–66.
4. Libe-Nowell D, Kleinberg J. The link-prediction problem for social networks. J Am Soc Info Sci Tech. 2007;58(7):1019–31.
5. Wang D, Liang Y, Xu D. A content-based recommender system for computer science publications. Knowl-Based Syst. 2018;157:1–9.
6. Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ. LIBLINEAR: A library for large linear classification. J Mach Learn Res. 2008;1871-1874.
7. Feng Y, Chen H. An Improved Density Peaks Clustering based on Rough Set Theory for Overlapping Community Detection. Int Conf Intell Syst Knowl Eng. 2019;21-28.
8. Zhu X, Ghahramani Z. Learning from labels and unlabeled data with label propagation. Tech Report. 2004;2002(3175):237–44.
9. Zhou D, Bousquet O, Lal TN, Weston J, Schkopf B. Learning with local and global consistency. In Advances in neural information processing systems. 2004; pp. 321-328.
10. Cen Y, Zhang J, Wang G, Qian Y, Meng C, Dai Z, Tang J. Trust Relationship Prediction in Alibaba E-Commerce Platform. IEEE Trans Knowl Data Eng. 2019;32(5):1024–35.
11. Qun Liu, Shuxin Liu, Guoyin Wang. Social Relationship Prediction across Networks Using Tri-training BP Neral Networks. Neurocomputing. 2020;401:377–91.
12. Kipf TN, Welling M. Semi-supervised classification with graph convolutional network. https://arxiv.org/abs/1609.02907.
13. Kipf TN, Welling M. Variational graph auto-encoders. https://arxiv.org/abs/1611.07308.
14. Huang J, Shen H, Hou L, Cheng X. Signed graph attention network. In International Conference on Artificial Neural Network. 2019; pp. 566-577.
15. Veli P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. Graph attention network. https://arxiv.org/abs/1710.10903.
16. Yao Y. An outline of a theory of three-way decisions. In International Conference on Rough Sets and Current Trends in Computing. 2012; pp. 1-17.
17. Yao Y. Three-way decision: an interpretation of rules in rough set theory. In International Conference on Rough Sets and Knowledge Technology. 2009; pp. 642-649.
18. Yao Y. Decision-theoretic rough set models, in: International Conference on Rough Sets and Knowledge Technology, Springer. 2007; pp. 1-12.
19. Yao Y. Three-way granular computing, rough sets, and formal concept analysis. Int J Approx Reas. 2020;2020(116):106–25.
20. Yao Y. Three-way decision and granular computing. Int J Approx Reas. 2018;2018(103):107–23.
21. Yu H, Chen Y, Lingras P, Wang GY. A three-way cluster ensemble approach for large-scale data. Int J Approx Reas. 2019;2019(115):32–49.
22. Afridi MK, Azam N, Yao JT, Alanazi E. A three-way clustering approach for handling missing data using GTRS. Int J Approx Reas. 2018;2018(98):11–24.

23. Yao JT, Azam N. Web-based medical decision support systems for three-way medical decision making with game-theoretic rough sets. IEEE Trans Fuzzy Syst. 2015;23(1):3-15.
24. Zhang HR, Min F, Shi B. Regression-based three-way recommendation. Info Sci. 2017;378(2017):444-461.
25. Li D. Membership clouds and membership cloud generators. Comp Res Dev. 1995;32(6):15–20.
26. Du Y, Li DY. Concept partition based on cloud and its application to mining association rules. J Soft. 2001;12(2):196–203.
27. Jiang R, Li DY. Similarity search based on shape representation in time-series data sets. J Comp Res Dev. 2000;37(5):601–8.
28. Song Y, Li D. Reliability evaluation of electronic products based on cloud models. Acta Electronica Sinica. 2000;28(12):74–6.
29. Castillo E, Gutierrez JM. Nonlinear time series modeling and prediction using functional network. Extracting information masked by chaos. Phys Lett A. 1998; 244;(1-3):71-84.
30. Castillo E, Hadi AS. Functional network. Wiley StatsRef: Statistics Reference Online; 2014.
31. Li D, Han J, Shi X, Chan MC. Knowledge representation and discovery based on linguistic atoms. Knowl-Based Syst. 1998;10(7):431–40.
32. Wang G, Xu C, Li D. Generic normal cloud model. Info Sci. 2014;280:1–15.
33. Loia V, Parente D, Pedrycz W, Tomasiello S. A granular functional network with delay: some dynamical properties and application to the sign prediction in social network. Neurocomputing. 2018;321:61–71.
34. Yang J, Wang G, Liu Q, Guo Y, Liu Y, Gan W, Liu Y. Retrospect and Prospect of Research of Normal Cloud Model. Chinese J Comp. 2018;41(03):724–44.
35. Zhou G, Zhou Y, Huang H, Tang Z. Functional networks and applications: A survey. Neurocomputing. 2019;335:384–99.
36. Tang J, Gao H, Liu H, Das Sarma A. eTrust: Understanding trust evolution in an online world. In Proceedings of the ACM SIGKDD international conference on Knowledge discovery and data mining. 2012; pp. 253-261.
37. Leskovec J, Huttenlocher D, Kleinberg J. Signed network in social media. In Proceedings of the SIGCHI conference on human factors in computing systems. 2010; pp. 1361-1370.
38. Kumar S, Spezzano F, Subrahmanian VS, Faloutsos C. Edge weight prediction in weighted signed network. In IEEE International Conference on Data Mining (ICDM). 2016; pp. 221-230.
39. Yang J, Leskovec J. Defining and evaluating network communities based on ground-truth. Knowl Info Syst. 2015;42(1):181–213.
40. West R, Paskov HS, Leskovec J, Potts C. Exploiting social network structure for person-to-person sentiment analysis. Trans Assoc Comput Linguist. 2014;2:297–310.