Attribute-Aware Graph Recurrent Networks for Scholarly Friend Recommendation Based on Internet of Scholars in Scholarly Big Data

Chunyou Zhang, Xiaoqiang Wu, Wei Yan, Lukun Wang, and Lei Zhang

***Abstract*—The academic society is stepping into the age of scholarly big data, where finding suitable scholars for collaboration has become ever difficult. Scholarly recommendation approaches are designed to overcome the information overload problems. However, previous methods mainly consider network topology without considering scholars' academic information and the manually designed similarity measurements may not have a good performance when applying to large-scale sparse networks. To this end, this article proposes to design a scholarly friend recommendation system by taking advantages of network embedding and scholar attributes. It is worth mentioning that different from traditional scientific collaborator recommendations, our goal is to recommend potential friends for scholars using academic social networks. We first construct an attributed social network by extracting scholars' academic attributes from digital libraries. Then, we perform an attributed random walk which can jointly model network structure and scholar attributes. Finally, a novel graph recurrent neural framework is adopted to embed attributed scholar interactions within the model for recommendations. Experimental results on two real-world scholarly datasets demonstrate the effectiveness of our proposed method.**

***Index Terms*—Internet of things, Network representation learning, trajectory data mining, vehicular ad hoc network.**

Manuscript received July 23, 2019; revised September 8, 2019; accepted September 22, 2019. Date of publication October 14, 2019; date of current version January 17, 2020.This work was supported in part by the National Natural Science Foundation of Shandong Province under Grant ZR2018BF005, in part by the Science and Technology Innovation Guide Project of the Inner Mongolia Autonomous Region under Grant KCBJ2018028, in part by the Program for Young Talents of Science and Technology in Universities of Inner Mongolia Autonomous Region under Grant NJYT-19-B15. Paper no. TII-19-3279. *(Corresponding author: Lukun Wang.)*

C. Zhang, X. Wu, and W. Yan are with the College of Mechanical Engineering, Inner Mongolia University for Nationalities, Tongliao 028000, China (e-mail: zcy19801204@126.com; wangzai8402@163.com; ywimun@163.com).

L. Wang is with the Department of Information Engineering, Shandong University of Science and Technology, Taian 271019, China (e-mail: wanglukun@sdust.edu.cn).

L. Zhang is with the School of Mechanical Engineering, Tianjin University of Commerce, Tianjin 300134, China (e-mail: zhgraceli@163.com).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2019.2947066

I. Introduction

IN RECENT years, scholars are producing ever seen large-scale publications. The academic society is stepping in the age of scholarly big data [1], [2]. Due to the easy access to tremendous academic data, the information overload problem has drawn extensive attention. It is difficult for scholars to find needed publications to read and meet new scholars for collaboration. Against this background, scholarly recommendation systems, i.e., publication recommendation [3] and collaborator recommendation [4], [5] have been designed to help scholars find needed information timely and conveniently.

Scientific collaboration recommendation is one of the main tasks in scholarly recommendations. The goal is to find potential collaborators by measuring the similarity between scholars. It has been proven that similar scholars may collaborate with each other [6]. Thus, the key is to measure the similarity among scholars. Various approaches have been proposed. Since scientific collaborations are indexed by coauthorships so that a scientific collaboration network can be constructed. Many network-based similarity measurement approaches have been designed, including common neighbors (CN), random walk-based approaches (RW) [7], [8].

However, it is insufficient to merely consider network topology without considering scholars' academic information. It is no doubt that scholars occupy various scholarly attributes, which may benefit the recommendation systems. Therefore, many approaches have been designed by incorporating scholar attributes into networks. For examples, Xia *et al.* [5] design the MVCwalker model by considering various academic factors. Kong *et al.* [4] design the CCRec model by exploiting publication contents and collaboration networks for scientific collaborator recommendation.

Although previous methods have explored how to combine network topology and scholar attributes for scholarly recommendation systems, the adopted metrics are manually designed. These may lead to bias recommendation results. Meanwhile, these methods may not have a good performance when applying to large-scale sparse networks. Recently, network embedding technique [9], [10] has been extensively investigated to learn the low-dimensional representations of nodes in large-scale networks. Its effectiveness has been proven in many network-driven tasks, such as link prediction, node classification, and community detection [11], [12].

To this end, we aim to design a scholarly friend recommendation system by taking advantages of network embedding and scholar attributes. It is worth mentioning that different from traditional scientific collaborator recommendations, our goal is to recommend potential friend for scholars using academic social networks. First, an attributed social network is constructed by extracting scholars' academic attributes from digital libraries. Then, we get the representations of scholars by performing an attributed random walk, which can jointly model network structure and scholar attributes. We adopt a graph recurrent neural framework to embed attributed scholar interactions within the model for recommendations. Finally, we perform experimental results on two real-world scholarly datasets, which demonstrate that our proposed method outperforms several baseline methods.

The rest of this article is organized as follows. Section II defines the investigated problem. The details of the proposed model are introduced in Section III. Section IV presents the experimental setups and the experimental results are reported in Section V. Related works are introduced in Sections VI and Section VII concludes this article.

II. Related Works

In this section, we introduce the related works from two perspectives: scholarly data mining and scientific collaboration recommendation.

A. Scholarly Data Mining

The academic society is stepping into the age of scholarly big data [1], [2], where scholars are able to access tremendous academic information. Since scholarly information are well-structured in networks, academic social network analysis [13], [14] has been one of the main approaches for scholarly data mining. Specifically, many scholarly data-driven tasks have been extensively investigated, including academic relationship mining [15], scientific recommendation [3], [16], scholar profiling [17], and scientific impact evaluation [18]. Due to the large scale and sparsity characteristics of scholarly data, traditional social network analysis approaches can not perform well. Recently, network embedding technique [9], [10] has been extensively studied to learn the low-dimensional representations of nodes in large-scale networks. Its effectiveness has been proven in many network-driven tasks, such as link prediction, node classification, and community detection. The basic framework of our proposed model is the network embedding approach.

B. Scientific Collaboration Recommendation

Scientific collaboration recommendation is one of the main tasks in scholarly data mining [19]. Scientific collaboration aims to find potential collaborators by measuring the similarity between scholars based on the assumption that similar scholars may collaborate with each other in the future. The key is how to measure the similarity among scholars precisely. Since scientific collaborations can be indexed by coauthorships so that a scientific collaboration network can be constructed. Many link prediction-based approaches have been designed, including CN, RW. Some works have extended these approaches by considering academic factors. For examples, Xia *et al.* [5] design the MVCwalker model by considering various academic factors. Kong *et al.* [4] design the CCRec model by exploiting publication contents and collaboration networks for scientific collaborator recommendation. However, previous method can not well-applied to large-scale scholarly dataset because the selected factors for measuring scholar similarity are manually designed, which is time consuming and may lead to biased results. Our proposed method takes the advantages of network embedding and scholar attributes which is different from traditional methods.

TABLE I
Main Symbols and Their Definitions

Symbols	Definition
n=$\mid \mathbf{V} \mid$	number of scholars
m=$\mid \mathbf{A} \mid$	number of scholar attributes
d	dimension of represented scholar vector
$\mathbf{A}\in \mathbb{R}^{n\times m}$	schlar attribute matrix
$\mathbf{H}\in \mathbb{R}^{n\times m}$	represented scholar vector matrix
$\mathbf{G}\in \mathbb{R}^{n\times d}$	netowrk adjacency matrix
θ	probability of walking
$\mathcal{T}$	attributed random walk-based scholar sequence

III. Problem Definition

We give the definitions of the main notions in Table I. Let $\mathcal{G} = \{\mathbf{V}, \mathbf{G}\}$ denote the online academic social networks, where $\mathbf{V}$ denotes the set of n scholars and $\mathbf{G}$ denotes the set of links among V. Specifically, $\mathbf{E}$ is the adjacency matrix denoted as $\mathbf{E} \in \mathbf{R}^{n\times n}$. $\mathbf{A}$ denotes the academic attributes of scholars extracted from digital libraries such as DBLP. In details, $\mathbf{A} \in \mathbb{R}^{n\times m}$ is a matrix that represents all scholar's academic attribute information. $\mathbf{a}_i$ denotes the academic attribute information of scholar i. To be clear, we assume that both $\mathbf{G}$ and $\mathbf{A}$ are none-negative.

Before formulating our research problems, we first introduce the definitions of some important proximities.

Definition 1 (Internet of Scholars): The internet of scholars refers to the network constructed from the online academic social network, i.e., LinkedIn. Different from the traditional academic social network which is constructed based on coauthorships, the links in internet of scholars denote the real connections. If two scholars are friends of each other, there will be a link among them.

Definition 2 (Scholarly Friend Recommendation): Scholarly friend recommendation aims to recommend friend for scholars in online academic social networks. Traditional collaborator recommendation is to recommend collaborators based on publication history. Our goal is to recommend friends for scholars using online academic social networks

Definition 3 (Network Embedding): Network embedding aims to learn a low-dimensional representations of nodes in complex networks by a mapping function $f : \{\mathbf{V}, \mathbf{G}\} \rightarrow \mathbf{H}$, where $\mathbf{H} \in \mathbb{R}^{n\times d}$. Here, d is much smaller than n, and $\mathbf{h}_i$ denotes the representation of scholar $i \in \mathbf{V}$.

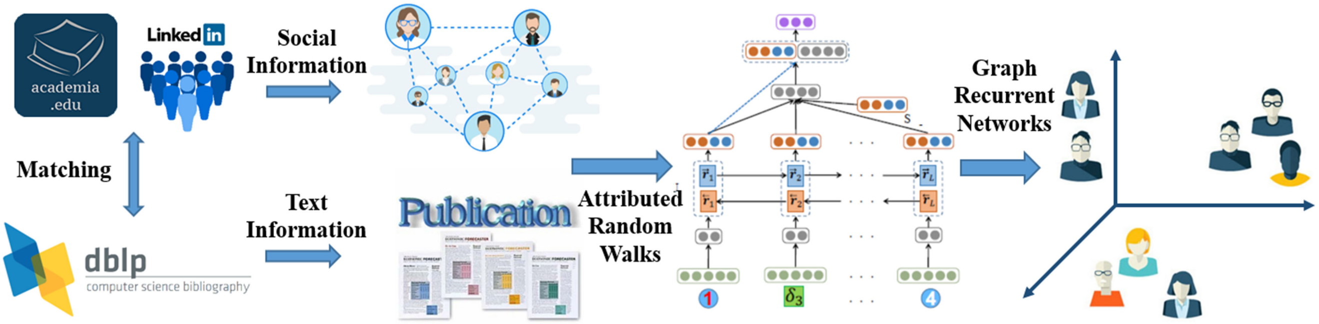


Fig. 1. Illustration of the proposed model.

Based on these definitions above, we can formulate our problems as follows: given an attributed internet of scholars $\mathcal{G} = \{\mathbf{V}, \mathbf{G}, \mathbf{A}\}$, our goal is to learn a low-dimensional representations of scholars by a mapping function $f : \{\mathbf{G}, \mathbf{A}\} \rightarrow \mathbf{H}$, and recommend scholarly friend based on the similarity of scholar's representation vectors.

IV. Proposed Model

In order to recommend friends for scholars using academic social networks, the key is to measure the similarity between scholars. We first construct an attributed social network by extracting scholars' academic attributes from digital libraries. Then, we perform an attributed random walk, which can jointly model network structure and scholar attributes. Finally, a novel graph recurrent neural framework is adopted to embed attributed scholar interactions within the model. The details of the proposed model are illustrated in Fig. 1.

A. Attributed Network Construction

In this section, we introduce how to construct an attributed social network based on internet of scholars. A plain social network can be gained by extracting the real friendship from online academic social network. That is, if two scholars are friends of each other, there will be a link among them. It is worth mentioning that such connection is un-weighted and un-directed. However, merely embedding such social network is not enough to measure the similarity between scholars, where scholars' attributes are overlooked. It has been proven that considering various academic factors can improve the performance of academic collaborator recommendation systems [4].

However, the academic information in online academic social networks is usually insufficient and inaccurate because most of the academic information should be uploaded by scholars themselves. For most cases, scholars are not willing to upload their academic information manually. Meanwhile, even scholars upload their information, these information are often outdated because it is impossible for scholars to update their information every day.

At the same time, the academic information, such as paper title, author list, and paper keywords can be retrieved from academic digital libraries or academic search engines. For example, Google Scholar and DBLP record scholars' academic information timely. Therefore, we first match scholars from academic social networks with his/her information on academic digital libraries. Such matching is designed based on the following two rules to avoid ambiguity:

1) The investigated scholar should have the same name or name abbreviation.
2) The scholar should have at least one same publication.

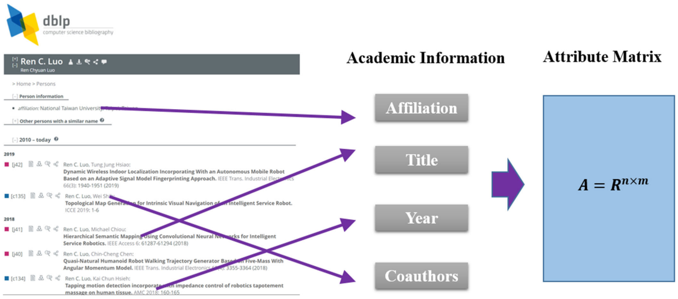

dblp
computer science bibliography
Ren C. Luo
Ren Chyuan Luo
> Home > Persons
[–] Person information
- *affiliation:* National Taiwan University, Taipei, Taiwan
[+] Other persons with a similar name
[–] 2010 – today
2019
[j42] Ren C. Luo, Tung Jung Hsiao: Dynamic Wireless Indoor Localization Incorporating With an Autonomous Mobile Robot Based on an Adaptive Signal Model Fingerprinting Approach. IEEE Trans. Industrial Electronics 66(3): 1940-1951 (2019)
[c135] Ren C. Luo, Wei Shih: Topological Map Generation for Intrinsic Visual Navigation of an Intelligent Service Robot. ICCE 2019: 1-6
2018
[j41] Ren C. Luo, Michael Chiou: Hierarchical Semantic Mapping Using Convolutional Neural Networks for Intelligent Service Robotics. IEEE Access 6: 61287-61294 (2018)
[j40] Ren C. Luo, Chin-Cheng Chen: Quasi-Natural Humanoid Robot Walking Trajectory Generator Based on Five-Mass With Angular Momentum Model. IEEE Trans. Industrial Electronics 65(4): 3355-3364 (2018)
[c134] Ren C. Luo, Kai Chun Hsieh: Tapping motion detection incorporate with impedance control of robotics tapotement massage on human tissue. AMC 2018: 160-165

Fig. 2. Illustration of scholarly attributes extraction.

After matching the scholar in academic social networks with the academic digital libraries, we can gain the academic information of the scholar. Take the DBLP as an example. As illustrated in Fig. 2, we can gain a scholar's academic information, including paper title, author list, year of publication, and author affiliation. Based on the basic academic information, we can calculate various academic attributes. Specifically, the academic attributes we investigate include research interest, number of collaborators, and number of publications within the recent five years. Specifically, the research interest attribute is calculated by performing topic models on the collections of scholars' paper title information. The number of collaborators denotes the extent to which scholars are willing to collaborate with others and the number of publications denotes the reputation of the scholar. Based on these academic attributes, we can gain the attribute matrix $\mathbf{A}$. Therefore, we can construct the attributed network $\mathcal{G} = \{\mathbf{V}, \mathbf{G}, \mathbf{A}\}$ for scholars. Each scholar in the $\mathbf{A}$ has m attributes and the set of scholar attribute is denoted as $\mathcal{A}$.

B. Embedding Via Attributed Random Walks

In order to learn the low-dimensional representations of scholars, we need to both embed the network structure of $\mathbf{G}$ and the rich scholar attributes of matrix $\mathbf{A}$. It is obvious that considering both the network $\mathbf{G}$ and auxiliary attribute information $\mathbf{A}$ can lead to better recommendation performance, which has been proven effective in many recommendation tasks.

To jointly embed scholar attributes and network topology, we employ the attributed random walk model. One of the plain approaches is to build a new network $\mathbf{L}$ based on the attribute matrix $\mathbf{A}$. That is, using the attributes information in $\mathbf{A}$ to weight the edges between scholar i and scholar j. The attribute similarity between scholars i and j can be calculated by the cosine similarity between $\mathbf{a}_i$ and $\mathbf{a}_j$. In this way, the original attributed network $\mathcal{G} = \{\mathbf{V}, \mathbf{G}, \mathbf{A}\}$ can be transformed to be a plain network $\mathbf{L}$. Although this approach is simple and easy to apply, the high time complexity of computing $\mathbf{L}$ makes it difficult to run it on large-scale networks.

To effectively embed the attributed network, the attributed random walks are utilized. The key of attributed random walks is to construct a new bipartite network $\mathcal{A}$ based on $\mathbf{A}$ so that random walks can be more diverse and reduce the time complexity. We first normalize the attribute matrix $\mathbf{A}$ and the network adjacency matrix $\mathbf{G}$. Assuming that each attribute category $\alpha_k \in \mathcal{A}$ as a vertex, we can gain a bipartite network $\mathfrak{G} \triangleq \{\mathbf{V}, \mathcal{A}, \mathcal{E}\}$, where $\mathcal{E}$ is the corresponding link set. Specifically, there will be a link between scholar $i \in \mathbf{V}$ and attribute $\alpha_k \in \mathcal{A}$ if scholar i has the attribute α_k.

Therefore, the attributed random walk will run among these $n + m$ nodes. Assuming that the random walker walks to scholar $i \in \mathbf{V}$, the next transition will be determined by a biased coin flipping, where the coin head appears with a probability θ and the coin tail probability is $1 - \theta$. Specifically, the random walk is biased by the following roles:

1) If the head appears, the walker will walk on the $\mathbf{G}$. It will walk to scholar $j \in \mathbf{V}$, with the following probability:

$$\mathrm{Pro}(i \to j) = \frac{\mathbf{g}_{ij}}{\sum_{P=1}^{n} \mathbf{g}_{ip}}. \tag{1}$$

2) If the tail appears, the walker will walk two steps on the $\mathfrak{G}$: It will first walks to vertex $\alpha_k \in \mathcal{A}$ with the following probability:

$$\mathrm{Pro}(i \to \alpha_k) = \frac{\mathbf{a}_{ik}}{\sum_{p=1}^{m} \mathbf{a}_{ip}}. \tag{2}$$

Then, the walker will walk back to a scholar $j \in \mathbf{V}$ with the following probability:

$$\mathrm{Pro}(\alpha_k \to j) = \frac{\mathbf{a}_{jk}}{\sum_{q=1}^{m} \mathbf{a}_{qk}}. \tag{3}$$

These walking strategies ensure that scholars in $\mathbf{V}$ can interact with each other not only accounting on the network topology, but also the scholar attributes. Specifically, if θ equals 1, the walker will run on the plain network. With the increasing of θ, the random walk will be biased by the scholar attributes. The transition probability matrix can be calculated as follows:

$$\mathbf{Pro} = \begin{bmatrix} \theta \mathbf{G} & (1-\theta)\mathbf{A} \\ (1-\theta)\mathbf{A}^T & 0 \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}. \tag{4}$$

We can see that the attributed random walker can combine the network $\mathbf{G}$ with the scholar attributes $\mathbf{A}$ comprehensively. The walks on $\mathbf{G}$ refer to the traditional random walks where the scholar attributes are not considered. The walks on $\mathbf{A}$ refers a new network, where the nodes are the scholar attributes. Such random walk is reasonable. When the coin head appears, the probability of transiting from scholar i to scholar j through any scholar attribute categories is determined by a similarity matrix $\mathbf{C}$, which is defined as follows:

$$\mathbf{C} = \mathbf{A}\mathbf{D}\mathbf{A}^T \tag{5}$$

where the matrix $\mathbf{D}$ is a diagonal matrix with $d_{kk} = \frac{1}{\sum_{q=1}^{n} \mathbf{a}_{qk}}$.

Assume that the attributed random walkers walk from scholar i to any attribute category α_k, the probability of walking from α_k to scholar j is not related with the one from i to α_k. Thus, we can calculate the $\mathbf{Pro}(i \to j)$ as follows:

$$\mathbf{Pro}(i \to j) = \sum_{k=1}^{m} \mathbf{Pro}\,(i \to \alpha_k)\,\mathbf{Pro}\,(\alpha_k \to j) \tag{6}$$

$$= \sum_{k=1}^{m} \frac{\mathbf{a}_{ik}}{\sum_{p=1}^{m} \mathbf{a}_{ip}} \frac{\mathbf{a}_{jk}}{\sum_{q=1}^{m} \mathbf{a}_{qk}} \tag{7}$$

$$= \left[\mathbf{a}_{i1}\mathbf{d}_{11}, \mathbf{a}_{i2}\mathbf{d}_{22}, \ldots, \mathbf{a}_{im}\mathbf{d}_{mm}\right] \mathbf{a}_j^T \tag{8}$$

$$= \mathbf{s}_{ij}. \tag{9}$$

Finally, based on above presented attributed random walks, we can sample both the network topology and the scholar attributes. Therefore, we can get the scholar index sequences as a set τ_i with a fixed length walks. The sequence set $\mathcal{T}$ can be gained by combing all the scholar sequence.

C. Embedding Via Graph Recurrent Networks

After gaining the scholar sequence set $\mathcal{T}$, our goal is to learn the low-dimensional representations of scholars based on it. A simple and widely used approach is to utilize the word embedding approach by treating the sequence as sentences and scholars as words. However, previous embedding approach can not well embed the scholar attributes [11], [12]. Inspired by previous work on graph convolutional networks (GCN) [20], we propose the incorporate the attributed random walks into a deep learning model.

Considering that the sequence set $\mathcal{T}$ is composed of both network topology and scholar attributes, the hidden state in recurrent neural networks (RNN) [21] can capture the interactions among scholars. Therefore, we proposed to employ the RNN model to embed the the sequence set $\mathcal{T}$. It mainly contains the following four steps.

1) Assume that the length of the sequence set $\mathcal{T}$ is L, our goal is to match any index in L to a low-dimensional vector. If the index is a scholar $j \in \mathbf{V}$, it will be mapped to his/her attribute $\mathbf{a}_j$. If the index is $\alpha_k \in \mathbf{A}$, we will map it to a new vector $\mathbf{e}_j$ which is a one-hot vector where the jth element is 1.

2) Then, a fully connected layer is utilized to reduce the dimension of scholar attributes. The final representation vectors $\{\mathbf{x}_j\}$ is calculated based on the following:

$$\mathbf{x}_j = \beta\left(\mathbf{a}_j \mathbf{W}_a + \mathbf{b}_a\right), \text{ or } \mathbf{x}_j = \beta\left(\mathbf{e}_j \mathbf{W}_a + \mathbf{b}_a\right). \tag{10}$$

Algorithm 1: Algorithms for GRN.

```
Input: G, A, d, B, L, $\alpha$ and lables
Output: Attributed representations $\mathbf{H}$ of scholars
Normalize G and A;
Compute P based ;
Construct a probability table;
for node i in V do
    for j=1 do
        Set i as the initial node;
        Perform a (L-1) length attributed random walks;
        Append the learn sequence to $t_i$;
    Append $t_i$ to $\tau_i$;
for ephch=1:$epoch_{m}ax$ do
    Shuffle the order of $t_i$ in $\tau_i$ ;
    for Set $t_i$ in $\tau_i$ do
        Take B sequences in $t_i$ as input and nodel i' label
        as output to train GRN;
        Update weighted matrix and bias terms to
        miminize the objective function;
for $t_i$= $t_1$ to $t_n$ do
    Input all B scequences in $t_i$ into the GRN;
```

Here, β denotes the sigmoid function, and $\mathbf{W}_a$ denote the link weight matrix. Each $\mathbf{w}_j$ represents the scholar attribute category α_k.

3) After calculating $\mathbf{x}_j$, a bidirectional RNN (i.e., long short-memory) and gated recurrent unites are employed to learn the forward hidden sequences $(\overrightarrow{\mathbf{r}_1}, \overrightarrow{\mathbf{r}_2}, \ldots, \overrightarrow{\mathbf{r}_L})$. At the same time, a backward hidden state sequence $(\overleftarrow{\mathbf{r}_1}, \overleftarrow{\mathbf{r}_2}, \ldots, \overleftarrow{\mathbf{r}_L})$ is also learned. Thus, RNN are able to model scholars by interacting with both his/her forward and backward neighbors. Specifically, the forward hidden state sequence $(\overrightarrow{\mathbf{r}_1}, \overrightarrow{\mathbf{r}_2}, \ldots, \overrightarrow{\mathbf{r}_L})$ can be calculated as follows:

$$\mathbf{f}_j = \beta\left(\mathbf{x}_j \mathbf{W}_{xf} + \overrightarrow{\mathbf{r}}_{j-1} \mathbf{W}_{hf} \mathbf{b}_f\right) \tag{11}$$

$$\mathbf{i}_j = \beta\left(\mathbf{x}_i \mathbf{W}_{xf} + \overrightarrow{\mathbf{r}}_{j-1} \mathbf{W}_{hi} \mathbf{b}_i\right) \tag{12}$$

$$\mathbf{o}_j = \beta\left(\mathbf{x}_o \mathbf{W}_{xf} + \overrightarrow{\mathbf{r}}_{j-1} \mathbf{W}_{hf} \mathbf{b}_o\right) \tag{13}$$

$$\mathbf{c}_j = \mathbf{f}_j \circ \mathbf{c}_{j-1} + \mathbf{i}_j \tanh\left(\mathbf{x}_j \mathbf{W}_{xc} + \overrightarrow{\mathbf{r}}_{j-1} \mathbf{W}_{hc} \mathbf{b}_c\right) \tag{14}$$

$$\overrightarrow{\mathbf{r}}_j = \mathbf{o}_j \circ \tanh(\mathbf{c}_j). \tag{15}$$

Here, $\mathbf{f}_j$, $\mathbf{i}_j$, and $\mathbf{o}_j$ are the forget, input, and output's activation vectors, respectively. $\mathbf{c}_j$ denotes the cell state vector. $\circ$ is the Hadamard product. $\tanh$ represents the Hyperbolic tangent function.

4) Finally, we can get the L indices accordingly based on scholar's sequence set τ_i. We can get L embedding vectors $\mathbf{r}_j$ based on the bidirectional RNN layer. Therefore, we can calculated the scholars' final embedding representations $\mathbf{h}_j$ based on the previous proposed architecture [12]. Specifically, the sequences set τ_i is combined into one as $\{\bar{\mathbf{r}}_1, \overrightarrow{\mathbf{r}}_2, \ldots, \overrightarrow{\mathbf{r}}_L\}$ via a pooling strategy. Then, all the $\{\bar{\mathbf{r}}_2, \overrightarrow{\mathbf{r}}_3, \ldots, \overrightarrow{\mathbf{r}}_L\}$ is merged into $\widehat{\mathbf{r}}_i$ via the pooling strategy. Therefore, the final scholar representations $\mathbf{h}_i$ can be calculated as follows:

$$\mathbf{h}_i = [\widehat{\mathbf{r}}_i, \bar{\mathbf{r}}_1]\,. \tag{16}$$

The whole pseudocode is shown in Algorithm (1). Specifically, our key idea is to input scholar attribute into a special multilayer perception, where each scholar's presentation is learned by averaging his/her neighbors' representations in the previous layer. The algorithms run as follows. Given an attributed network $\mathcal{G} = \{\mathbf{V}, \mathbf{G}, \mathbf{A}\}$, the matrix $\mathbf{P}$ is first constructed by normalizing attributed matrix $\mathbf{A}$. The alias method is adopted to sample the attributed random walks based on the discrete probability distribution in $\mathbf{P}$. Then, each scholar i's neighbor sequence is sampled and appended into $\mathcal{T}_i$. The scholar i is regarded as the starting scholar. The GRN model is trained based on the scholar i's label and $\mathcal{T}_i$ with the objective function in (16). After interactions, we can gain the final scholar representation vectors.

D. Optimization

Our proposed model can be trained via various methods, including unsupervised, supervised, or semisupervised approaches because of the good property of GCN. Following previous work [20], we utilized a supervised approach, where the objective function can be calculated as follows:

$$\mathfrak{L} = -\sum_{i \in \mathbf{v}} \mathbf{y}_i^T \log\left(\mathrm{softmax}\left(\beta\left(\mathbf{h}_i \mathbf{W}_h + \mathbf{b}_h\right)\right)\right) \tag{17}$$

where $\mathbf{y}_i$ is the one-shot vector representing the label of scholar i. This objective function can be straightly replaced by other task-specific objective functions such as negative sampling based objective function for optimization.

E. Scholarly Friend Recommendation

Based on the low-dimensional representation matrix $\mathbf{H}$, we can gain the scholars i and j's vector as $\mathbf{h}_i$ and $\mathbf{h}_j$. Therefore, we can calculate the similarity between these two scholars based on the cosine similarity of these two scholar vectors. The detail is as follows:

$$\mathrm{Sim}(\mathbf{h}_i, \mathbf{h}_j) = \frac{\mathbf{h}_i \times \mathbf{h}_j}{||\mathbf{h}_i||||\mathbf{h}_j||}. \tag{18}$$

Finally, the top-K similar scholars are recommended to the target scholar as the potential scholarly friends.

V. Experimental Design

A. Dataset Selection

In the experiments, we utilized two popular academic social networks, including ResearchGate[1] and LinkedIn.[2] We select 100 scholars in the DBLP dataset as the seed scholars and extract their friendships from the online academic social network.

[1] [Online]. Available: https://www.researchgate.net/

[2] [Online]. Available: https://www.linkedin.com

TABLE II
STATISTICS OF THE TWO DATASETS

Datasets	# of scholars	# of edges	Density	Time Duration
ResearchGate	7,271	182,347	0.0069	2010-2015
LinkedIn	8,459	271,961	0.0076	2010-2015

The seed scholars are extracted from the field of data mining area. Specifically, the scholars with top 100 publications in the data mining journals/conference are selected as the seed scholars. These venues include ACM Transactions on Information Systems, IEEE Transactions on Knowledge and Data Engineering, ACM Transactions on Knowledge Discovery from Data, ACM Transactions on Knowledge Discovery from Data, and ACM International Conference on Information and Knowledge Management. Then, we gain the publication information of all the scholars based on the DBLP dataset. It is worth mentioning that we utilize the Aminer[3] dataset to gain the abstract information of all the publications. After processing, we gain 7271 scholars for ResearchGate dataset and 8459 scholars for LinkedIn datasets. The details of the datasets are illustrated in Table II. We can see from this table that the network in LinkedIn dataset is denser than that of ResearchGate dataset.

B. Baselines

Our goal is to recommend scholarly friend by considering both network topology and scholars' attribute information. The mechanism is to measure the similarity between scholars via scholar vector. In order to demonstrate the effectiveness of our proposed model, we mainly compare our method with various baseline methods. Specifically, these methods are as follows:

1) **CN** [7]: CN is the most common measure of how similar two nodes are in social network analysis. If two scholars share more CN, they are more willing to be friends in the future. It recommends potential scholarly friends based on the number of CN.
2) **MVCWalker** [5]: It is a random walk-based scientific collaborator recommendation model, where the similarity between two scholars is measured by a biased random walker. Three kinds of academic factors are utilized to bias the random walk.
3) **DeepWalk** [22]: It employs a stream of truncated random walks to gain the scholar neighbor sequence. The word2vec model is utilized to embed the gained sequences to learn the low-dimensional scholar vectors for recommendation.
4) **AANE** [23]: It learns the scholar representation vectors by considering both the network topology and scholar attributes. It is realized by the decomposition of the scholar attributes affinity and measuring the difference between linked scholars.
5) **CCRec** [4]: It adopts the topic model to calculate scholars' research interest based on the publication information and utilizes the random walk model to calculate the similarity between scholars for recommendation.
6) **CACR** [19]: It jointly represents researchers and research topics as compact vectors based on their co-occurrence relationships and extracts researchers' activeness and conservativeness for context-aware scientific collaborator recommendation.

All baselines are tuned till their best performance for comparison.

C. Model Evaluation

We adopt two widely used evaluation metrics in recommendation systems, including precision@K (P@K), Recall@K (R@K) and MRR@K (M@K). The precision@K is calculated as follows:

$$\mathrm{P@K} = \frac{\text{No. of correct recommended friends}}{\text{No. of recommended friends}}. \tag{19}$$

The R@K is calculated as follows:

$$\mathrm{R@K} = \frac{\text{No. of correct recommended friends}}{\text{No. of true friends}}. \tag{20}$$

The MRR metric denotes the mean reciprocal rank, which can be calculated as follows:

$$\mathrm{M@k} = \frac{1}{|Q| \sum_{i=1}^{|Q|} \frac{1}{\mathrm{rank}_i}}. \tag{21}$$

VI. RESULTS

A. Parameter Selection

There are mainly two parameters that may influence the performance of our proposed recommendation model, i.e., the represented dimension d and the coin probability θ. In this section, we will explore the impact of these two parameters.

Table III presents the experimental results on two datasets with different d. We present the results in terms of Precision@10, Recall@10, and MRR@10. We can see the following from the table.

1) What can be clearly seen from this table is that there is an obvious increase and then decrease with the increase of d. For ResearchGate dataset, the recommendation accuracy is likely to increase before 50, and then decrease. The peak for LinkedIn dataset is 100. Therefore, we will set the d as 50 for ResaerchGate and 100 for LinkedIn in the following experiments.
2) We can notice that the overall performance of our proposed model is better on LinkedIn dataset than that of ResearchGate dataset. The reason may be that the network of LinkedIn dataset is denser than that of DBLP dataset.

Another important parameter is the coin probability θ, which determines how often will the random walker will jump to the scholar attributes. In other word, θ denotes the tradeoff between network topology and scholar attributes.

Fig. 3 reports the recommendation accuracy on two datasets with θ varying from 0.1 to 0.9. We can see the following from this figure.

[3] [Online]. Available: https://aminer.org/

TABLE III
INFLUENCE OF REPRESENTED DIMENSIONS D ON TWO DATASETS

Dataset	RsearchGate			LinkedIn		
d	**Precision@10**	**Recall@10**	**MRR@10**	**Precision@10**	**Recall@10**	**MRR@10**
10	56.1	55	62.23	50.1	45.97	59.23
30	57.23	55.27	62.54	50.23	46.37	60.54
50	**57.34**	**55.9**	**64.72**	51.34	47.2	61.02
100	56.4	55.53	63.2	**52.4**	**47.53**	**61.2**
200	56.32	55.16	62.97	52.32	47.16	60.97

The bold face values represent the optimal values.

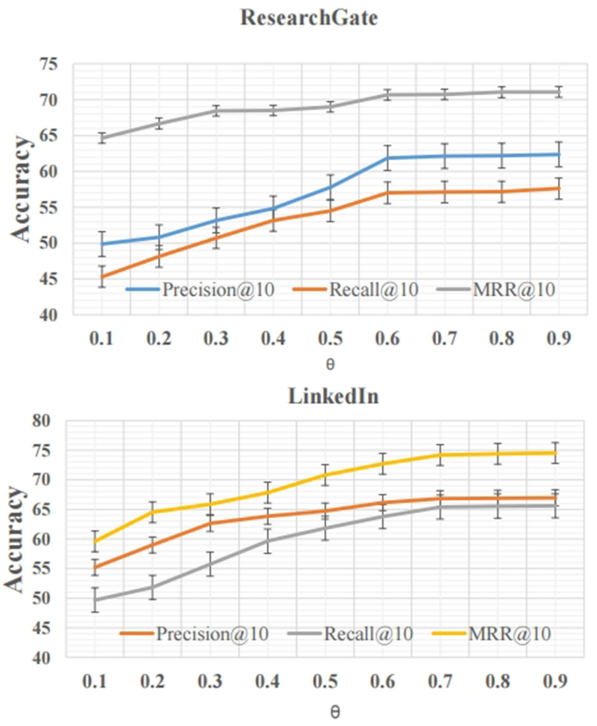


Fig. 3. Influence of θ on two datasets.

1) With the increase of θ, precision@10, Recall@10, and MRR@10 are expected to increase accordingly, and then remain steady after a certain value point. The steady value points for ResearchGate and LinkedIn are 0.6 and 0.7, respectively. Therefore, we set θ as 0.6 and 0.7 for these two datasets, respectively, in the following experiments.
2) The trends of the recommendation accuracy in these two figures are in accordance with common sense. If the θ is small, the impact of scholar attributes is overlooked; if the θ is too large, the impact of the network topology will decrease. Therefore, there is a tradeoff between network topology and scholar attributes.

B. Comparison With Baselines

We compare our proposed model with six baseline recommendation methods. These methods include random walk-based approaches, network embedding-based approaches, and plain network similarity-based approaches. Table IV shows the comparison results on ResearchGate datasets in terms of Precision {5, 10, 20}, Recall{5, 10, 20}, and MRR{5, 10, 20}. We can see the following from this table.

1) Our proposed attributed random walk-based GRN approach achieves the best performance by comparing with all the state-of-the-art approaches. Take Precision@5 for example, our method achieves a 2.3% increase compared with the second best baseline CACR.
2) Baselines considering the scholar attributes may have a better recommendation performance than those methods that merely considers the network topology. Specifically, CACR and AAE performance better than DeepWalk and CN. This indicates that it is beneficial to consider scholar attribute when designing a scholarly friend recommendation system.
3) With the increase of k, i.e., from five to 20, the precision and MRR go up accordingly, while Recall decreases accordingly. This is because of the definition of these evaluation metrics. Take precision as an example. Based on (19), with the increase of K, while more true friends are likely to be recommended, the number of recommended candidates increase faster.

Table V shows the comparison results on LikedIn datasets in terms of Precision{5, 10, 20}, Recall{5, 10, 20}, and MRR {5, 10, 20}. Similarly, we can gain the following observations:

1) The proposed approach achieves the best performance by comparing with all the state-of-the-art approaches. Take Recall@5 for example, our method achieves a 1.3% increase compared with the second best baseline CACR.
2) Baselines considering the scholar attributes such as CACR and AAE have a better recommendation performance than those methods, which merely considers the network topology, such as CN and DeepWalk. This also indicates the advantages of considering scholar attribute in designing scholarly friend recommendation systems.

C. Attribute Contribution Analysis

Based on previous observations, we can see that it is important to consider the scholar attribute for scholarly friend recommendations. In this section, we explore the influence of scholar attributes. We compare our proposed method with a weakened version named variation that dose not incorporate scholar attribute in learning scholar vectors.

Fig. 4 reports the comparison results between our proposed method and its weakened version Proposed in terms of Precision@5, Recall@5, and MRR@5 on two datasets. We can see the following from this figure.

1) Proposed method always performs better than the variation method. This indicates that scholar attributes play an

TABLE IV
PERFORMANCE COMPARISON ON RESEARCHGATE DATASET

Methods	ResearchGate(%)								
	Precision@5	Precision@10	Precision@20	Recall@5	Recall@10	Recall@20	MRR@5	MRR@10	MRR@20
CN	44.2	42.31	40.11	30.27	32.87	36.27	44	46.88	48.24
DeepWalk	48.26	46.26	43.22	34.1	39.25	41.37	47	50.1	51.26
CCRC	50.2	48.1	44.2	35.22	41.1	45.28	50.24	54.31	55.07
MVCWalker	53.25	49.55	45.45	38.25	45.14	47.65	53.25	56.24	59.21
AAE	54.46	50.24	44.53	39	46.1	47.17	54.03	57.16	60.23
CACR	55.26	51.23	45.27	39.55	46.26	47.33	56.19	60.23	63.26
Proposed	**56.57**	**52.42**	**46.13**	**40.24**	**47.55**	**49.2**	**57.55**	**61.22**	**64.08**

The bold face values represent the optimal values.

TABLE V
PERFORMANCE COMPARISON ON LINKEDIN DATASET

Methods	LinkedIn(%)								
	Precision@5	Precision@10	Precision@20	Recall@5	Recall@10	Recall@20	MRR@5	MRR@10	MRR@20
CN	50.11	48.24	45.27	45.67	47.07	48.34	51.07	53.46	56.05
DeepWalk	52.05	51.05	49.26	47.27	48.96	51.76	55.99	58.16	60.54
CCRC	53.04	52.16	50.27	48.27	49.06	51.27	55.27	58.07	59.86
MVCWalker	55.42	54.7	52.04	49.22	51.22	53.55	58.18	61.25	63.22
AAE	56	55.16	52.67	50.13	51.27	53.86	58.38	61.77	64.13
CACR	57.96	57.09	51.23	51.87	54.26	56.85	61.03	63.26	67.67
Proposed	**58.24**	**57.34**	**52.24**	**52.56**	**55.9**	**57.68**	**61.35**	**64.72**	**68.96**

The bold face values represent the optimal values.

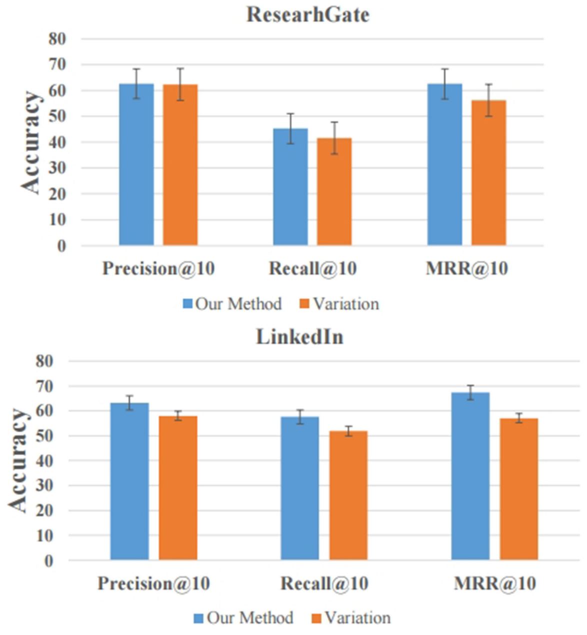


Fig. 4. Comparison results between our proposed method and its weakened version.

important role in measuring scholar similarity. Such an observation is in line with previous findings [4].

2) These two methods have a better performance on the ResearchGate dataset than that on LinkedIn dataset. This may be because of the reason that the network extracted from LinkedIn dataset is much denser than that of ResearchGate dataset.

VII. CONCLUSION

In this article, we proposed to recommend potential friends for scholars using academic social networks. Our proposed method was designed based on the framework network embedding, where the scholar attributes can network topology were jointly embedded via the attributed random walk-based graph recurrent networks. Specifically, we first constructed an attributed social network by extracting scholars' academic attributes from digital libraries. Then, we performed an attributed random walk which can jointly model network structure and scholar attributes. Finally, a novel graph recurrent neural framework was adopted to embed attributed scholar interactions within the model for recommendations. Experimental results on two real-world scholarly datasets demonstrate the effectiveness of our proposed method. Due to the limitation of the datasets, we utilized publication information for scholar attributes. In future work, we would like to consider more academic attributes to improve the performance of our proposed method.

REFERENCES

[1] S. Khan, X. Liu, K. A. Shakil, and M. Alam, "A survey on scholarly data: From big data perspective," *Inf. Process. Manag.*, vol. 53, no. 4, pp. 923–944, 2017.

[2] F. Xia, W. Wang, T. M. Bekele, and H. Liu, "Big scholarly data: A survey," *IEEE Trans. Big Data*, vol. 3, no. 1, pp. 18–35, Mar. 2017.

[3] X. Cai, J. Han, W. Li, R. Zhang, S. Pan, and L. Yang, "A three-layered mutually reinforced model for personalized citation recommendation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 12, pp. 6026–6037, Dec. 2018.

[4] X. Kong, H. Jiang, Z. Yang, Z. Xu, F. Xia, and A. Tolba, "Exploiting publication contents and collaboration networks for collaborator recommendation," *PLoS One*, vol. 11, no. 2, 2016, Art. no. e0148492.

[5] F. Xia, Z. Chen, W. Wang, J. Li, and L. T. Yang, "MVCwalker: Random walk-based most valuable collaborators recommendation exploiting academic factors," *IEEE Trans. Emerg. Topics Comput.*, vol. 2, no. 3, pp. 364–375, Sep. 2014.

[6] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 58, no. 7, pp. 1019–1031, 2007.

[7] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," *Phys. A: Stat. Mech. Its Appl.*, vol. 390, no. 6, pp. 1150–1170, 2011.

[8] R. N. Lichtenwalter, J. T. Lussier, and N. V. Chawla, "New perspectives and methods in link prediction," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2010, pp. 243–252.

[9] P. Goyal and E. Ferrara, "Graph embedding techniques, applications, and performance: A survey," *Knowl.-Based Syst.*, vol. 151, pp. 78–94, 2018.

[10] P. Cui, X. Wang, J. Pei, and W. Zhu, "A survey on network embedding," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 5, pp. 833–852, May 2019.

[11] D. Zhang, J. Yin, X. Zhu, and C. Zhang, "Network representation learning: A survey," *IEEE Trans. Big Data*, to be published, doi: 10.1109/TBDATA.2018.2850013.

[12] Y. Dong, N. V. Chawla, and A. Swami, "metapath2vec: Scalable representation learning for heterogeneous networks," in *Proc. 23 rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2017, pp. 135–144.

[13] X. Kong, Y. Shi, S. Yu, J. Liu, and F. Xia, "Academic social networks: Modeling, analysis, mining and applications," *J. Netw. Comput. Appl.*, vol. 132, pp. 86–103, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1084804519300438

[14] W. Wang *et al.*, "Csteller: Forecasting scientific collaboration sustainability based on extreme gradient boosting," *World Wide Web*, Jul. 2019. [Online]. Available: https://doi.org/10.1007/s11280-019-00703-y

[15] R. Wang *et al.*, "AceKG: A large-scale knowledge graph for academic data mining," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manag.*, 2018, pp. 1487–1490.

[16] W. Wang, J. Liu, Z. Yang, X. Kong, and F. Xia, "Sustainable collaborator recommendation based on conference closure," *IEEE Trans. Comput. Social Syst.*, vol. 6, no. 2, pp. 311–322, Apr. 2019.

[17] J. Tang, "Aminer: Toward understanding big scholar data," in *Proc. 9th ACM Int. Conf. Web Search Data Mining*, 2016, pp. 467–467.

[18] E. Ferrara and A. E. Romero, "Scientific impact evaluation and the effect of self-citations: Mitigating the bias by discounting the h-index," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 64, no. 11, pp. 2332–2339, 2013.

[19] Z. Liu, X. Xie, and L. Chen, "Context-aware academic collaborator recommendation," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2018, pp. 1870–1879.

[20] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

[21] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. IEEE Inf. Conf. Acoust., Speech Signal Process.*, 2013, pp. 6645–6649.

[22] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2014, pp. 701–710.

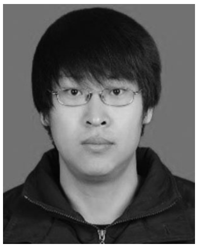
[23] X. Huang, J. Li, and X. Hu, "Accelerated attributed network embedding," in *Proc. SIAM Inf. Conf. Data Mining*, 2017, pp. 633–641.

Chunyou Zhang is currently working toward the Ph.D. degree in mechanical and electrical engineering with Beihang University, Beijing, China.

He is currently a Professor of Mechanical Manufacturing with the College of Mechanical Engineering, Inner Mongolia University for the Nationalities, Tongliao, Inner Mongolia, China. His research interests include energy-saving technology of pumping unit, fluid control, measurement, and control system.

Xiaoqiang Wu was born in Xingtai, Hebei, China, in 1985. He received the B.S. degree in electronic and information engineering from the Agricultural University of Hebei, Baoding, Hebei, China, in 2007, and the M.S. degree in agricultural electrication from Yunnan Agricultural University, Kunming, China, in 2013. He is currently working toward the Ph.D. degree in mechanical engineering with the School of Mechanical Engineering, Tianjin University, Tianjin, China.

Since 2018, he has been an Associate Professor of Electrical Control with the College of Mechanical Engineering, Inner Mongolia University for Nationalities, Tongliao, Inner Mongolia, China. His research interests include advanced manufacturing technology, fundamental study of plasma sources, artificial intelligence, interdisciplinary application of computer, and the internet of things-enabled manufacturing systems.

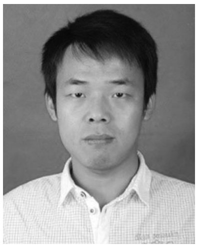


Wei Yan was born in Zhangjiakou City, Hebei Province, China, in 1985. He received the B.S. degree in electronic and information engineering from the Agricultural University of Hebei, Baoding, Hebei, in 2007 and the M.S. degree in circuits and systems, in 2011, from the Hebei University of Technology, Tianjin, China, where he is currently working the Ph.D. degree in electrical engineering with the School of Electrical Engineering.

Since 2013, he has been a Lecturer of Electronic Information with the College of Physics and Electronic Information, Inner Mongolia University for Nationalities. His research interest includes the computer intelligent measurement and control and biomedical signal processing.

Lukun Wang received the B.S. degree in information engineering from the Shandong University of Science and Technology, Shandong, China, in 2006, the M.S. degree in software engineering from the Dalian University of Technology, Dalian, China, in 2009, and the Ph.D. degree in computer science from the Ocean University of China, Qingdao, China, in 2016.

He is currently an Assistant Professor with the Department of Information Engineering, Shandong University of Science and Technology. His current research interests include artificial intelligence, image processing and pattern recognition, machine learning, and big data.



Lei Zhang was born in Shijiazhuang, Hebei, China, in 1987. He received the B.S. degree from the Hebei Normal University of Science and Technology, Qinhuangdao, Hebei, China, in 2010, and the M.S. degree from Yanshan University, Qinhuangdao, China, in 2013, and the Ph.D. degree from Tianjin University, Tianjin, China, in 2018, all in mechanical engineering.

Since 2018, he has been a Lecturer of mechanical and electrical engineering with the School of Mechanical Engineering, Tianjin University of Commerce, Tianjin, China. His research interests include mechanical dynamics, robust control, and intelligent manufacturing.