IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 10, 2020, accepted March 19, 2020, date of publication March 24, 2020, date of current version April 9, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2982986

Deep Neural Network-Embedded Internet of Social Computing Things for Sustainability Prediction

QIAO LI [1], YING SONG [1], BOXIN DU [1], YU SHEN [2,3], AND YUAN TIAN [1]

[1] School of Economics, Chongqing Technology and Business University, Chongqing 400067, China
[2] National Research Base of Intelligent Manufacturing Service, Chongqing Technology and Business University, Chongqing 400067, China
[3] Chongqing South-to-Thais Environmental Protection Technology Research Institute Company Ltd., Chongqing 400069, China

Corresponding author: Yu Shen (shenyu@ctbu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFE0205600, in part by the State Language Commission Research Program of China under Grant YB135-121, in part by the Key Research Project of Chongqing Technology and Business University under Grant ZDPTTD201917, Grant KFJJ2018060, and Grant 1952027, and in part by the Chongqing Natural Science Foundation of China under Grant cstc2019jcyj-msxmX0747.

ABSTRACT Social computing, exploiting utilization of advanced computational techniques to overcome typical problems in social science, has been a more visualized conception in academia. However, existing researches still suffer from two aspects of challenges: 1) lack of reliable multi-source data acquisition and management; 2) absence of high-performance algorithmic approaches. Fortunately, some newly-emerged cross-discipline technologies offer more opportunities to enhance conventional solutions. For the former, characterized by its property of information collection and integration, Internet of Things (IoT) can be introduced to produce a novel architecture named Internet of Social Computing Things (IoSCT). For the latter, specific neural network models can be set up to manipulate complicated calculation. Thus, taking the issue of sustainability prediction as objective situation, deep neural network-embedded Internet of Social Computing Things (NeSoc) is proposed in this paper. Firstly, IoSCT is put forward as bottom support platform, guaranteeing comprehensive resource involvement of social computing. Secondly, a hybrid neural network mechanism is formulated and embedded into IoSCT for centralized modeling. Finally, a series of experiments are conducted on a real-world dataset to evaluate performance of the proposed NeSoc.

INDEX TERMS Internet of Things, social computing, neural network, complex system, sustainability prediction.

I. INTRODUCTION

The Internet of Things (IoT) [1] is a huge network that combines various information with Internet to realize the interconnection among people, machines and things [2]. Accompanied with rapid progress of information technology, almost all of the fields are inevitably committed to promoting the ability of information collection as well as processing [3]. Owing to its excellent capability of information integration [4], IoT has been generalized to a variety of application scenarios [5], deriving different variants under specific situations, such as Internet of Medical Things [6], Internet of Industrial Things [7] and Internet of Financial Things [8].

In recent years, social computing has been a more visualized conception in academia [9]. Generally speaking, it explores utilization of advanced computational theories or approaches to overcome typical research problems in social science [10], in which sustainability prediction acts as a representative one [11]. Nowadays, researches concerning sustainability prediction were mainly implemented upon the basis of conventional sociological approaches like statistical analysis or case study [12]–[18]. However, limited by manual computing power, these sociological mechanism-based methods usually employed simple mathematical theories to model complex social dynamics process filled with uncertainty and nonlinearity. Having recognized this view, some researchers

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

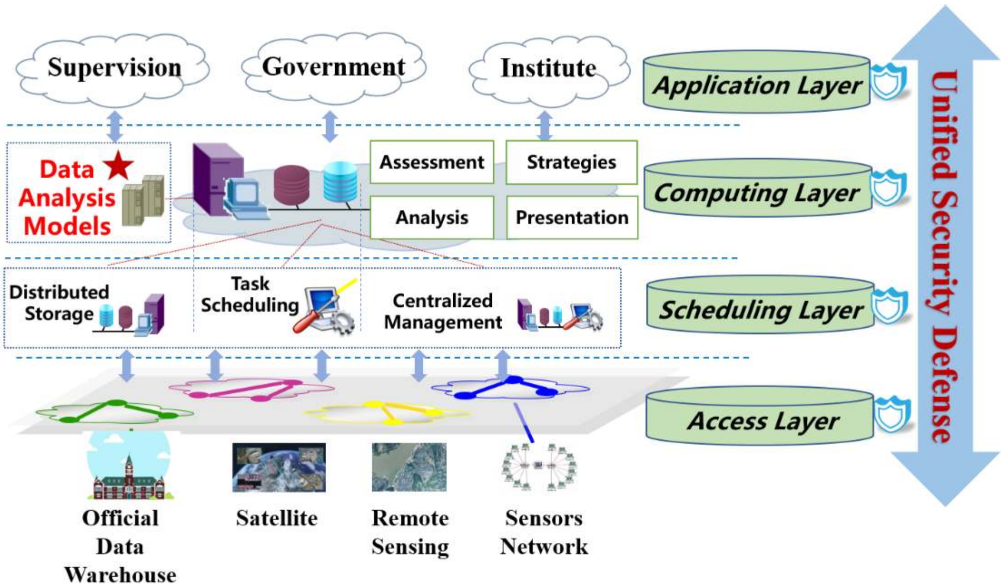


FIGURE 1. Framework design of the Internet of Social Computing Things (IoSCT).

exploited data mining technique to develop improved methods for sustainability prediction issue [19]–[24]. But they still failed to integrate fruitful cross-domain information due to the lack of multi-source information fusion. And from the perspective of system science, process of society development can be viewed as a complicated discrete system driven by joint impact of internal multi-domain factors, e.g. economy, ecology, industry, etc.

Inspired by design of Internet of Medical Things [6], IoT is introduced to serve as fundamental support platform to provide considerable performance of information acquisition and management, establishing Internet of Social Computing Things (IoSCT). Furthermore, in order to manipulate fine-grained feature space in IoSCT, specific neural network models can be set up to express complex dynamics process in society system. Neural network is a type of mathematical algorithm for distributed parallel information processing [29], and imitates neuro system characteristics of animals to realize complicated computation [31], [32]. Specific to the issue of sustainability prediction, major investigation object lies in the coordinated relationships between urbanization and ecological civilization [30] which is major focus of this research. Thus, this paper proposes deep **Ne**ural network-embedded Internet of **So**cial **C**omputing Things (NeSoc) to predict coordination degree between urbanization and ecological civilization in advance for regions. Firstly, architecture of IoSCT is put forward as bottom support platform, ensuring comprehensive resource involvement of social computing. Secondly, a hybrid neural network structure is designed and embedded into IoSCT for centralized modeling. Finally, a series of experiments are conducted on a real-world dataset to evaluate performance of the proposed NeSoc. To the best of our knowledge, we are the first to put up with IoSCT and introduce coupling of deep learning with IoT for sustainability prediction. Obviously, the proposed IoSCT is a deep learning-enhanced Internet of Medical Things, and can also be utilized for medical issues with adjusting some settings.

Main contributions of this paper can be summarized as:

1) We abstract sustainability prediction problem as dynamic process of complex systems, and reveal the importance of multi-source data collection and integration.

2) We propose deep neural network-embedded IoSCT for prediction of coordination degree between urbanization and ecological civilization.

3) We evaluate efficiency and stability of the proposed NeSoc on a real-world dataset.

II. OVERVIEW

A. FRAMEWORK OF IoSCT

This research put forward a novel architecture IoSCT whose framework design is illustrated in FIGURE 1. The architecture contains four layers separately named access layer, scheduling layer, computing layer and application layer. This subsection briefly gives description of each layer to describe the IoSCT:

1) Access layer is the bottom interface of various data sources. The original data in this framework contains two parts: inherent data and sensing data. The former refers to sociological data published by official channels, and is collected through heavyweight crawler programs. The latter is actively acquired through a variety of collection equipment.

2) After acquisition, data is transferred into scheduling layer for distributed storage and centralized management. This layer is responsible for distributed management of data source, in order to facilitate analytical and computational demands. Besides, while receiving requests

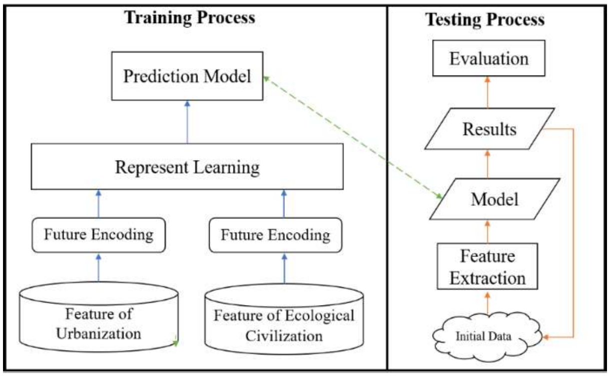


FIGURE 2. Technology roadmap of sustainability prediction.

from computing layer, it is supposed to reasonably schedules data transmission.

3) In computing layer, ensembled data analysis models are implemented to tackle all sorts of social computing issues. In this paper, a hybrid neural network-based prediction model is presented for prediction of coordination degree between urbanization and ecological civilization.

4) Finally, processing results will be accessed into application layer where users are able to obtain results through terminals. Note that IoSCT in this paper is mainly designed for administration departments and research institutes.

B. PROBLEM STATEMENT

This research takes the Urban Agglomeration of Yangtze River Economic Belt in China as research object. And main technology roadmap of sustainability is demonstrated as FIGURE 2. It contains two parts: model training process and model testing process. For the former, its goal is to formulate a hybrid neural network-based prediction model and make it trained through multi-source data from IoSCT. For the latter, it is expected to conduct a series of experiments to evaluate the former part of works.

As society is actually a complex system with complicated internal correlations, graph neural network (GNN) model [26] and convolutional neural network (CNN) model [27] are simultaneously introduced to construct a hybrid neural network mechanism which will be embedded into computing layer of IoSCT further. Then, interaction between urbanization and ecological civilization is exploited by encoding feature spaces into vectorized representations for following algorithmic operations. FIGURE 3 gives a vivid flowchart of the hybrid neural network mechanism. In detail, the whole feature space is divided into two subspaces: urbanization subspace and ecological civilization subspace, both of whom are low-dimensional real-valued vectors.

The former is modeled through GNN model, which is described in Section III.A. The latter is modeled using CNN model, which is described in Section III.B. On the basis of setting up the global feature space with the use of hybrid neural network mechanism, two feature subspaces are concatenated to construct a data mining-driven predictor to output the predicted coordination degree values.

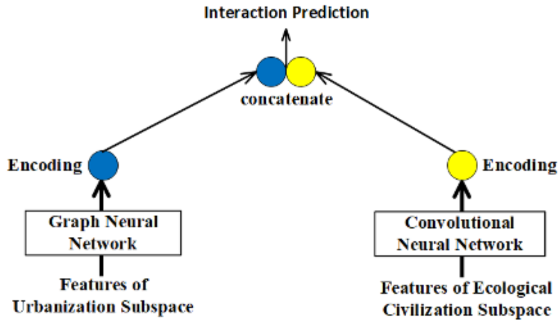


FIGURE 3. Flowchart of the proposed hybrid neural network mechanism.

C. PRELIMINARIES

Feature space of the methodology is separated into two subspaces: urbanization part and ecological civilization part, represented as $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ respectively. Both of the two subspaces are actually vectors, and possess a set of 1st-level indexes and 2nd-level indexes, which is demonstrated in TABLE 1. It can be observed from TABLE 1 that the urbanization subspace contains five 1st-level indexes and twenty-one 2nd-level indexes, and that ecological civilization subspace contains four 1st-level indexes and seventeen 2nd-level indexes.

GNN model is a set of entities and relations in graph network, containing two types of main objects: nodes and edges. In our research, nodes refer to detailed features in the urbanization subspace, corresponding to the twenty-one 2nd-level indexes in TABLE 1 represented as $v_1, v_2, \ldots, v_{21}$. To apply the CNN to ecological civilization feature subspaces, we define the initial input neuros as the seventeen 2nd-level indexes in $\boldsymbol{S_2}$ and represent them as $x_1, x_2, \ldots, x_{21}$. As different attributes possess diverse ranges of values, data normalization operation needs to be conducted in this research, in order to uniform their ranges as [0, 1].

III. METHODOLOGY

A. MODELING OF URBANIZATION

The GNN maps a graph G to a vector y through transition function and output function. It contains a set of vertices $\boldsymbol{V}$ and a set of edges $\boldsymbol{E}$. $\boldsymbol{V}$ denotes the twenty-one features of $\boldsymbol{S_1}$ mentioned in Section 2, while $\boldsymbol{E}$ denotes the pairwise relations among entities in $\boldsymbol{V}$. $\boldsymbol{E}$ is the form of matrix as:

$$\boldsymbol{E} = \begin{bmatrix} e_{1,1} & e_{1,2} & \cdots & e_{1,21} \\ e_{2,1} & e_{2,2} & \cdots & e_{2,21} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1,21} & e_{2,21} & \cdots & e_{21,21} \end{bmatrix} = \{e_{i,j}\} \tag{1}$$

where $i, j = 1, 2, \ldots, 21$.

Give a graph G, we represent a set of all neighboring vertices within radius r from the i-th vertex as $N(i, r)$. Obviously, $N(i, 0)$ denotes the i-th vertex itself. Then, r-radius subgraph for vertex v_i is defined as:

$$G_{v_i}^{(r)} = \left(\boldsymbol{V}_{\boldsymbol{i}}^{(\boldsymbol{r})}, \boldsymbol{E}_{\boldsymbol{i}}^{(\boldsymbol{r})}\right) = \{e_{i,j}\} \tag{2}$$

where $\boldsymbol{V}_i$ and $\boldsymbol{E}_i$ respectively denote all the conforming vertices and edges concerning the i-th vertex. We also define the

TABLE 1. List of multi-source feature space.

Subspaces	1^{st}-level indexes	2^{nd}-level indexes	symbols
Urbanization S_1	Population Urbanization	Population Urbanization Rate (%)	v_1
		Urban Population Density (persons / km^2)	v_2
		Registered Urban Unemployment Rate (%)	v_3
	Spatial Urbanization	Urban area (km^2)	v_4
		Urban Road Area (thousand m^2)	v_5
	Economic Urbanization	Production Value Per Unit Land Area (¥ billion Yuan)	v_6
		Average wage of urban employees (¥ Yuan)	v_7
		GDP per capita (¥ Yuan/person)	v_8
		Ratio of Added Value of Tertiary Industry in GDP (%)	v_9
		Per Capita Disposable Income of Urban Residents (¥ Yuan)	v_{10}
		Per Capita Consumption of Urban Residents (¥ Yuan)	v_{11}
	Social Urbanization	Number of Internet users (million persons)	v_{12}
		Number of Doctors per 10000 people	v_{13}
		Number of public toilets for every 10000 people	v_{14}
		Number of public transport vehicles for every 10000 people	v_{15}
		Number of community health service stations	v_{16}
	Cultural Urbanization	Number of teachers in universities (10000 persons)	v_{17}
		Number of degrees awarded by universities (10000 persons)	v_{18}
		Number of art performance venues	v_{19}
		R & D funds for large-scale enterprises (¥ million Yuan)	v_{20}
		Technical market turnover (¥ billion Yuan)	v_{21}
Ecological Civilization S_2	Resource Bearing Capacity	Coal reserves per capita (10000 tons/person)	x_1
		Iron ore reserves per capita (10000 tons/person)	x_2
		Water reserves per capita (m^3/person)	x_3
		Green area per capita (m^2/person)	x_4
	Environmental Pollution and Ecological Damage	Total wastewater discharge (10000 tons)	x_5
		Sulfur dioxide emissions (ton)	x_6
		Removal volume of garbage (10000 tons)	x_7
		Forest coverage rate (%)	x_8
	Environmental Governance Capacity	Investment in industrial pollution control (¥ 10000 Yuan)	x_9
		Road cleaning area (10000 m^2)	x_{10}
		treatment capacity of urban sewage per day (10000m^3)	x_{11}
		Harmless treatment capacity of garbage (ton/day)	x_{12}
	Strength of Ecological Protection	Proportion of natural reserves area under (%)	x_{13}
		Investment in ecological construction (¥ 10000 Yuan)	x_{14}
		Green coverage（%）	x_{15}
		Total afforestation area (hectares)	x_{16}
		Area of soil erosion control (hectares)	x_{17}

r-radius subgraph for edge $e_{i,j}$ through the following formula:

$$G_{e_{i,j}}^{(r)} = \left(\boldsymbol{V}_i^{(r)} \cup \boldsymbol{V}_j^{(r)}, \boldsymbol{E}_i^{(r)} \cap \boldsymbol{E}_j^{(r)}\right) \quad (3)$$

Then, the transition function for both vertices and edges is defined to set up GNN model. Each feature in subspace $\boldsymbol{S_1}$ is not static constantly, and will evolve over time. Thus, we employ $v_{\boldsymbol{i}}^{(t)}$ to denote the i-th vertex at the t-th timestamp. Timestamps refer to a type of time points, such as year, month, day, etc. Given a graph G, the GNN model updates $v_{\boldsymbol{i}}^{(t)}$ by means of the following transition function:

$$v_i^{(t+1)} = \frac{1}{1+\exp\left[-\left(v_i^{(t)} + \sum c_{i,j}^{(t)}\right)\right]}, \quad v_j \epsilon \in \boldsymbol{V}_{\boldsymbol{i}}^{(t)} \quad (4)$$

where v_j belongs to the set of t-radius neighboring vertices of the i-th node, and v denotes index number of vertices. $c_{i,j}^{(t)}$ denotes the hidden layer vector and is calculated by the following formula:

$$c_{i,j}^{(t)} = \text{ReLU}\left(W_1\begin{bmatrix} v_i^{(t)} \\ e_{i,j}^{(t)} \end{bmatrix} + b_1\right) \tag{5}$$

where ReLU refers to Rectified Linear Unit, acting as a common activation function used in neural network, and can be expressed through the following formula:

$$\text{ReLU}(q) = \max(0, q) \tag{6}$$

where W_1 and b_1 are the model parameters to be estimated. $e_{i,j}^{(t)}$ is the edge between vertex $v_i^{(t)}$ and $v_j^{(t)}$. The vertex transition is helpful for learning more latent features or knowledge in the urbanization subspace.

For edge transition, major thoughts of above procedures can also be employed to construct expression of $e_{i,j}^{(t)}$ through the following formula:

$$e_{i,j}^{(t+1)} = \frac{1}{1+\exp\left[-\left(e_{i,j}^{(t)} + d_{i,j}^{(t)}\right)\right]} \tag{7}$$

$$d_{i,j}^{(t)} = \text{ReLU}\left[W_2\left(v_i^{(t)} + v_j^{(t)}\right) + b_2\right] \tag{8}$$

where W_2 and b_2 are the model parameters to be estimated. The vertex transition is helpful for learning more latent relations in the urbanization subspace.

After setting up vertex transition and edge transition function, the final expression of urbanization subspace can be written as follows:

$$Y_{urban} = \frac{1}{|\boldsymbol{S_1}|}\sum_{i=1}^{|S_1|} v_i^{(t)} \tag{9}$$

where $|\boldsymbol{S_1}|$ is the number of features (vertices) in $\boldsymbol{S_1}$, and equals to 21 in this research.

B. MODELING OF ECOLOGICAL CIVILIZATION

CNN, a typical neural network model, is able to extract features automatically or map explicit features into a new latent feature space (such as Hilbert space). We denote $h_k^{(t)}$ as the hidden layer vector of feature x_k in the t-th timestamp. By introducing CNN model, we define its update rule as follows:

$$h_k^{(t+1)} = \text{ReLU}\left(W_3 h_k^{(t)} + b_3\right) \tag{10}$$

where W_3 and b_3 are the unknown model parameters in CNN.

After setting up recurrent hidden layer function, the final expression of ecological subspace can be written as follows:

$$Y_{eco} = \sum_{k=1}^{|\boldsymbol{S_2}|}\left\{\text{ReLU}\left(\theta\left(h_k^{(t)}\right)\right)\cdot\frac{1}{1+\exp\left[-\delta\left(h_k^{(t)}\right)\right]}\right\} \tag{11}$$

where $|\boldsymbol{S_2}|$ is the number of features in $\boldsymbol{S_2}$, and equals to 17 in this research.

C. PREDICTION OF COORDINATION DEGREE

As the urbanization subspace and ecological civilization subspace have been formulated and modeled, this subsection manages to concatenate them for use of constructing prediction model. The initial hidden layer of the prediction procedure is expressed as:

$$h_1' = \left[\beta\cdot Y_{urban} \oplus' (1-\beta)\cdot Y_{eco}\right] \tag{12}$$

where $\oplus'$ denotes the concatenation operation and β is a non-negative trade-off parameter. Let h_z' $(z = 1, 2, \ldots)$ denote all the hidden layers of prediction part. From $z = 2$, the h_z' can be updated as following formulas:

$$h_z' = W_4 h_{z-1}' + b_4 \tag{13}$$

where W_4 and b_4 are the model parameters to be estimated.

We also let I_P' denote the coordination degree value between urbanization and ecological civilization concerning a time period, such as year, quarter, month, etc. P denotes the designated time period. Therefore, the I_P' can be calculated through the following formula:

$$I_P' = \frac{1}{1+\exp\left(-W_5\cdot h_z' + b_5\right)} \tag{14}$$

where W_5 and b_5 are the model parameters to be estimated. To estimate all the above parameters, we need to specify an objective function for optimization. Let I_P be the real coordination degree value, we compute the objective function as follows:

$$L = (1-\lambda)\left|\sum_{i=1}^{|\boldsymbol{S_1}|}\sum_{k=1}^{|\boldsymbol{S_2}|} ln\frac{I_P'}{I_P}\right| + \lambda\left\|\Theta\right\|^2 \tag{15}$$

where Θ refers to the set of parameters and λ is the parameter of penalty item which is added to prevent overfitting. After all, Dropout [28] is selected as the learning method to solve the above problem. Due to the limitation of textual length, detailed process of the solution for optimization is left out.

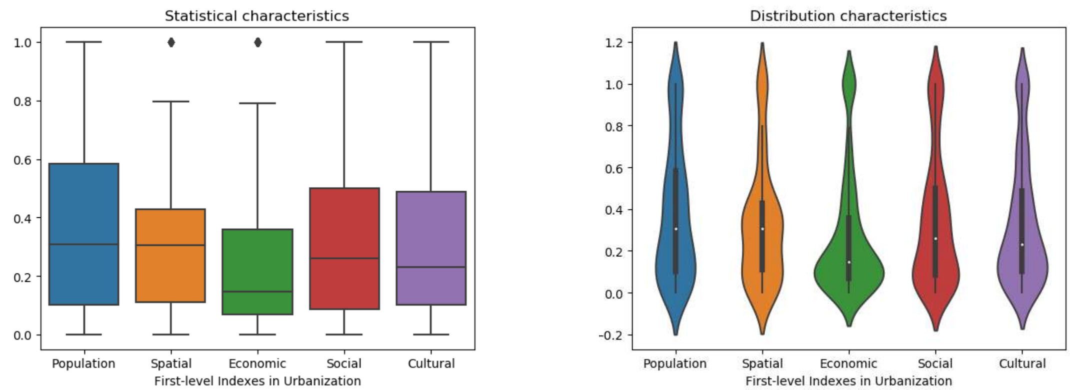
IV. EXPERIMENTS AND ANALYSIS

A. DATASET

To evaluate feasibility of the proposed NeSoc, a series of experiments are carried out on a real-world dataset. This section will introduce the detailed process of our experiments and reveal experimental results.

According to design of IoSCT, experimental dataset of this paper can be divided into two parts: public part and nonpublic part. Public data was collected by the computer technique-based tool—Internet crawler. It is able to automatically grasp information from World Wide Web according to certain rules. And data source of this research contains *1999-2017 China National Statistical Yearbook*,[1] *1999-2017 China Environmental Statistical Yearbook*[2] and other statistical yearbooks of provinces and cities in the Yangtze River economic belt. While the non-public part

[1] http://www.stats.gov.cn/tjsj/ndsj/
[2] http://www.stats.gov.cn/ztjc/ztsj/hjtjzl/#



(a) statistical characteristics (b) distribution characteristics

FIGURE 4. This is statistical distribution characteristics of urbanization subspace in the dataset. Five objects in X-axis separately corresponds to population urbanization, environmental pollution, economic urbanization, social urbanization and cultural urbanization.

was collected through the environmental ecological sensor network which was built by Chongqing Technology and Business University for the National Key Research & Development Program of China (2016YFE0205600). And research contents of this paper are part of the aforementioned national program.

When obtaining the original data, the population urbanization rate is obtained by calculating the proportion of the urban population in each sample province to the permanent population at the end of the year. Coal reserves per capita and iron ore reserves per capita adopt the ratio of coal reserves and iron ore reserves of each sample province to the permanent population at the end of the year. In order to unify the data scale, it is expected to standardize the data of each index, and take the maximum and minimum data of each sample Province as the upper and lower limits. The attributes of final dataset are described in Section II.C.

As the quantity of data acquired from those sources is not large enough, this research adopts Interpolation method to extend the data. For each index, there are a series of data concerning each year. It is supposed to interpolate three pieces of data between each two adjacent data through the principle of gaussian stochastic process, meaning that value evolution from a year to the next year is viewed as a gaussian stochastic process. Besides, the proposed NeSoc is actually a supervised learning scheme, and needs known labels which are coordination degree values of each year in this research. As coordination degree are unknown in the collected dataset, we employ the following steps to generate the coordination degree values. Note that the computed coordination degree values as follows are just reference values. They are just used for evaluation performance of our prediction model, and may not represent real coordination degree values.

Let U_{urban} and U_{eco} denote the contribution degree to coordination respectively from urbanization subspace and ecological subspace. And the two variables are computed as:

$$U_{urban} = \gamma \cdot \sum_{i=1}^{|\boldsymbol{S_1}|} M\left(v_i\right) \tag{16}$$

$$U_{eco} = (1-\gamma) \cdot \sum_{k=1}^{|\boldsymbol{S_2}|} M\left(x_k\right) \tag{17}$$

where γ is the trade-off parameter and is set to 0.5 in this research. $M\left(v_i\right)$ and $M\left(x_k\right)$ denotes satisfaction degree of features in urbanization subspace and ecological subspace respectively. $M\left(v_i\right)$ is computed as follows:

$$M\left(v_i\right) = \begin{cases} \left(v_i - l_i\right) \cdot \left(v_i - n_i\right), & \mathrm{v_i has\ positive\ effect} \\ \left(l_i - v_i\right) \cdot \left(n_i - v_i\right), & \mathrm{v_i has\ negative\ effect} \end{cases} \tag{18}$$

where l_i and n_i are respectively upper and lower limits of the order parameters at the critical point of system stability. $M\left(x_k\right)$ is computed as follows:

$$M\left(x_k\right) = \begin{cases} \left(x_k - l_k\right) \cdot \left(x_k - n_k\right) & \mathrm{x_k has\ positive\ effect} \\ \left(l_k - x_k\right) \cdot \left(n_k - x_k\right) & \mathrm{x_k has\ negative\ effect} \end{cases} \tag{19}$$

On the basis of above computations, coupling degree between urbanization subspace and ecological subspace is calculated as:

$$C_{u,e} = \left\{ \frac{U_{urban} \cdot U_{eco}}{\left[\left(U_{urban} + U_{eco}\right) \times 0.5\right]^2} \right\}^4 \tag{20}$$

and coordination degree between urbanization subspace and ecological subspace is calculated as:

$$D_{u,e} = 0.5 \cdot U_{urban} + 0.5 \cdot U_{eco} \tag{21}$$

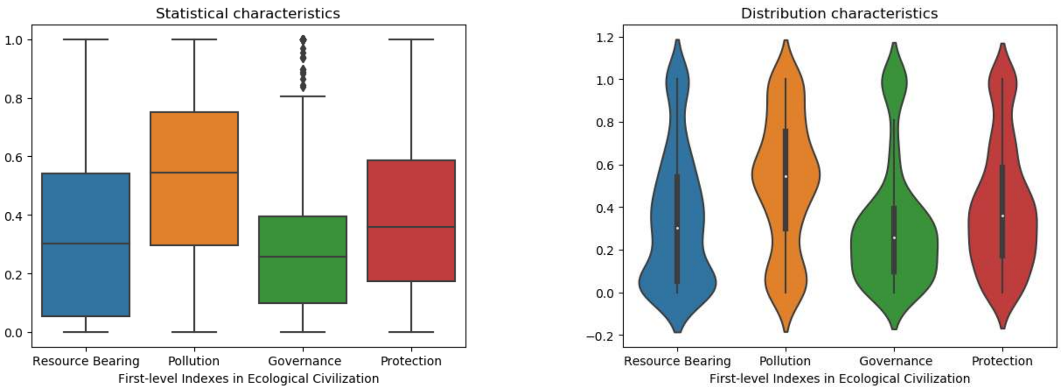
The real coordination degree as golden sample is calculated as:

$$I_P = \sqrt{C_{u,e} \cdot D_{u,e}} \tag{22}$$

B. EXPERIMENTAL SETTINGS

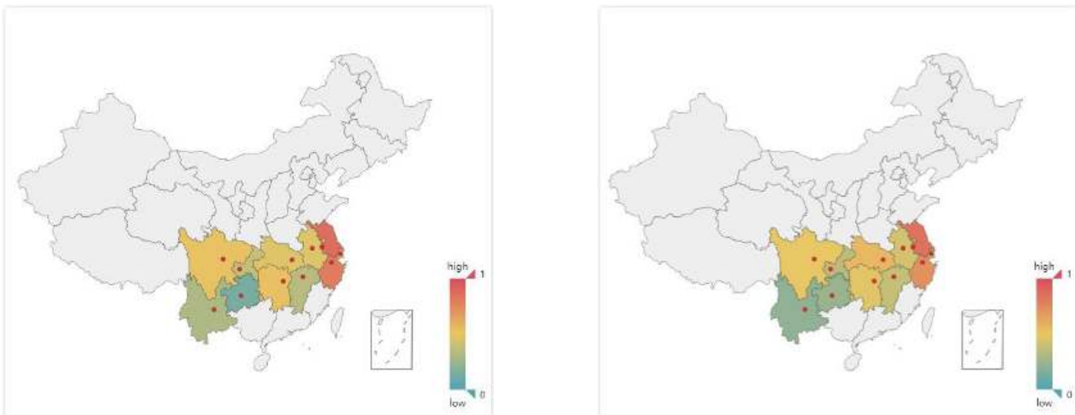
In order to assess data quality before experiments, statistical characteristics of the experimental dataset is visualized through computer programming.

FIGURE 4 illustrates the statistical distribution of urbanization subspace with respect to data in its five first-level indexes, and FIGURE 5 illustrates the statistical distribution



(a) statistical characteristics (b) distribution characteristics

FIGURE 5. **This is statistical distribution characteristics of ecological civilization subspace in the dataset. Four objects in X-axis separately corresponds to resource bearing capacity, environmental governance capacity, economic urbanization and strength of ecological protection.**



(a) Distribution status of coordination degree values in 2010 (b) Distribution status of coordination degree values in 2016

FIGURE 6. **These are average coordination degree values of provinces in the Yangtze river economic belt of China with respect to two typical years: 2010 and 2016.**

of ecological civilization sunspace with respect to data in its four first-level indexes. FIGURE 6 demonstrates the distribution of average coordination degree values concerning provinces in the Yangtze River Economic Belt of China, in which data of two typical years 2010 and 2016 is selected for illustration. This is because the year 2010 is the start of 2010s, and 2016 is the start of 13th five-year plan of China.

It can be observed from the figure that coordination level is decreasing from eastern provinces to western provinces, which is basically in line with the actual situation of China. It can be directly concluded from FIGURE 4 and FIGURE 5 that data in the experimental dataset is balanced, and concluded from FIGURE 6 that the dataset is real and effective. Thus, it is suitable for data mining and analysis.

In order to evaluate performance of the NeSoc, several prediction methods need to be selected as baselines. Here, we compare our approach NeSoc with several universal data mining methods, in order to verify the fact that the proposed NeSoc is able to reflect excellent performance under the situation of coordination degree prediction. The selected baselines are list as follows:

1) MLR: It refers to multi-variate linear regression model which is a kind of regression analysis method using the least squares function to model the relationship between multiple independent variables.

2) MLP: It refers to multi-layer preceptor which is a feed-forward artificial neural network model and whose idea is similar to regression analysis.

3) LSTM: It refers to Long Short-Term Memory which is a kind of sequential neural network model specially designed to solve the long-term dependence problem.

To evaluate feasibility of the NeSoc, the following evaluation metrics are utilized:

$$MAE = \frac{1}{N}\sum_{d=I}^{N}\left|I_P - I'_P\right| \tag{23}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{d=I}^{N}\left(I_P - I'_P\right)^2} \tag{24}$$

where I_P represents the real value of coordination degree at the time period P, I'_P represents the predicted value of coordination degree at the time period P, and N is the total piece number of testing sample. Obviously, the lower values of these metrics represent better performance. As for parameter setting, we initially set the z in Eq. (13) as 10 and the trade-off

TABLE 2. Experimental results of under learning rate of 0.01.

Methods	Training data: 60%		Training data: 70%		Training data: 80%	
	MAE (rank)	RMSE (rank)	MAE (rank)	RMSE (rank)	MAE (rank)	RMSE (rank)
MLR	0.067 (4)	0.093 (4)	0.061 (4)	0.075 (3)	0.066 (4)	0.098 (4)
LSTM	0.058 (2)	0.087 (3)	0.052 (2)	0.077 (4)	0.059 (2)	0.095 (3)
MLP	0.064 (3)	0.084 (2)	0.055 (3)	0.073 (2)	0.060 (3)	0.090 (2)
NeSoc	**0.055 (1)**	**0.081 (1)**	**0.046 (1)**	**0.069 (1)**	**0.052 (1)**	**0.084 (1)**

TABLE 3. Experimental results of under learning rate of 0.008.

Methods	Training data: 60%		Training data: 70%		Training data: 80%	
	MAE (rank)	RMSE (rank)	MAE (rank)	RMSE (rank)	MAE (rank)	RMSE (rank)
MLR	0.069 (4)	0.096 (3)	0.064 (4)	0.091 (4)	0.067 (4)	0.097 (4)
LSTM	0.059 (2)	0.091 (2)	0.053 (2)	0.088 (3)	0.061 (3)	0.094 (3)
MLP	0.063 (3)	0.098 (4)	0.055 (3)	0.083 (2)	**0.053 (1)**	0.089 (2)
NeSoc	**0.051 (1)**	**0.085 (1)**	**0.042 (1)**	**0.077 (1)**	0.056 (2)	**0.082 (1)**

TABLE 4. Experimental results of under learning rate of 0.003.

Methods	Training data: 60%		Training data: 70%		Training data: 80%	
	MAE (rank)	RMSE (rank)	MAE (rank)	RMSE (rank)	MAE (rank)	RMSE (rank)
MLR	0.051 (3)	0.075 (3)	0.067 (4)	0.090 (4)	0.044 (2)	0.078 (3)
LSTM	0.052 (4)	**0.068 (1)**	0.060 (2)	0.089 (3)	0.045 (3)	0.075 (2)
MLP	0.048 (2)	0.079 (4)	0.062 (3)	0.086 (2)	0.051 (4)	0.082 (4)
NeSoc	**0.040 (1)**	0.072 (2)	**0.048 (1)**	**0.079 (1)**	**0.039 (1)**	**0.074 (1)**

parameter λ in Eq. (15) as 0.5. Learning rate in optimization method Dropout for solving Eq. (15) is initially set to 0.01. As for dataset, we split it to training set and testing set with the multiple divisions during the experiments, and the default proportion rate is 70% training to 30% testing.

C. RESULTS AND ANALYSIS

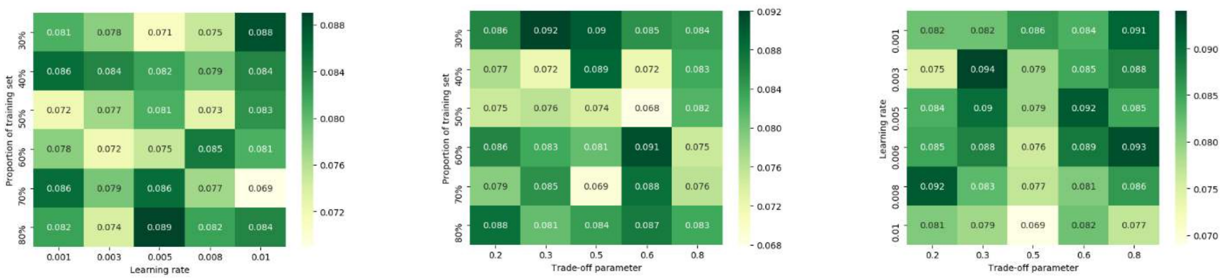
With parameters z and λ setting to initial values, and proportion of training data setting to 60%, 70% and 80% in order, we evaluate NeSoc and baselines with the learning rate in Droupout changing to three values: 0.01, 0.008 and 0.003. Table 2, Table 3 and Table 4 list results of NeSoc and baselines under learning rate of 0.01, 0.008 and 0.003 separately. In area of data display in these tables, each cell has two parts: data value and the rank of such data value in that column. It can be directly observed from these tables that the proposed NeSoc performs better than baselines in almost all of situations, and that data mining methods can be well suited for predicting coordinate degree of urbanization and ecological civilization. Although the NeSoc cannot obtain the best performance in a few situations, the distance is not too large.

This phenomenon can be attributed to two main aspects of reasons. Firstly, the proposed NeSoc utilizes CNN model to extract abstract features for ecological civilization subspace, which proposes a more fine-grained feature representation than other methods. Secondly, the proposed NeSoc employs GNN model to express deep interval relationships in urbanization subspace, which is able to fully capture feature structures of this subspace. Thirdly, the proposed NeSoc is embedded a hybrid neural network mechanism combining CNN with GNN together. Such a design produces a quite flexible method that models the coordinate process between urbanization and ecological civilization. In all, this set of experiments are able to well verify efficiency of the proposed NeSoc.

Besides, another set of experiments are conducted to evaluate stability of the proposed NeSoc. The scheme of stability testing is to make a combination of two parameters change into different groups of values, and test whether the experimental results will remain stable as a whole. During this set of experiments, the NeSoc is just assessed the parameter sensitivity singly, will not be compared with baselines. FIGURE 7 and FIGURE 8 are parameter sensitivity results of NeSoc with respect to MAE and RMSE. They both have three subfigures, respectively corresponding to parameter combination of: (a) learning rate and proportion of training set, (b) parameter combination of trade-off parameter and proportion of training set, and (c) trade-off parameter and learning rate. It can be observed from the two figures that the

(a) (b) (c)

FIGURE 7. Parameter sensitivity results of NeSoc with respect to MAE.



(a) (b) (c)

FIGURE 8. Parameter sensitivity results of NeSoc with respect to RMSE.

proposed NeSoc reflects a relatively stable performance with the dynamic parameter combinations. A possible reason can be deduced to explain this phenomenon. The proposed NeSoc simultaneously introduces CNN and GNN to capture two subspaces of the coordination process. Such a comprehensive insight which makes itself not susceptible to changing of parameter situations. Therefore, no matter how the parameter groups change, experimental results never heavily fluctuate and remain relatively stable.

To sum up, a large amount of experiments are conducted in this research to evaluate the NeSoc method from the perspectives of efficiency and stability. Results reveal that our proposals are effective and feasible for the problem of coordination degree prediction.

V. CONCLUSION

Two main obstacles exist in general problems of social computing: 1) lack of multi-source data acquisition and management; 2) absence of high-performance algorithmic approaches. The IoT is a huge network characterized by significant ability of data integration and management. And deep learning has been proved an effective computational scheme for many traditional scientific problems. The combination of them two seems to be a promising insight to enhance many classical methods in various fields. Inspired by this view, this paper deals with the first one by introducing cross-discipline technique IoT and designing a novel framework named IoSCT. In addition, the second one can be solved by embedding a hybrid neural network mechanism composed of CNN and GNN into IoSCT. Thus, taking sustainability prediction as an example, deep neural network-embedded Internet of Social Computing Things (NeSoc) is presented in this paper. And a series of experiments are conducted on a real-world dataset to evaluate performance of the proposed NeSoc.

REFERENCES

[1] M. Ben-Daya, E. Hassini, and Z. Bahroun, "Internet of Things and supply chain management: A literature review," *Int. J. Prod. Res.*, vol. 57, nos. 15–16, pp. 4719–4742, Aug. 2019.

[2] Z. Cui, Y. Cao, X. Cai, J. Cai, and J. Chen, "Optimal LEACH protocol with modified bat algorithm for big data sensing systems in Internet of Things," *J. Parallel Distrib. Comput.*, vol. 132, pp. 217–229, Oct. 2019.

[3] A. Darwish, A. E. Hassanien, M. Elhoseny, A. K. Sangaiah, and K. Muhammad, "The impact of the hybrid platform of Internet of Things and cloud computing on healthcare systems: Opportunities, challenges, and open problems," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 10, pp. 4151–4166, Oct. 2019.

[4] M. Noura, M. Atiquzzaman, and M. Gaedke, "Interoperability in Internet of Things: Taxonomies and open challenges," *Mobile Netw. Appl.*, vol. 24, no. 3, pp. 796–809, Jun. 2019.

[5] J. Dizdarević, F. Carpio, A. Jukan, and X. Masip-Bruin, "A survey of communication protocols for Internet of Things and related challenges of fog and cloud computing integration," *ACM Comput. Surveys*, vol. 51, no. 6, pp. 1–29, Jan. 2019.

[6] A. Gatouillat, Y. Badr, B. Massot, and E. Sejdic, "Internet of medical things: A review of recent contributions dealing with cyber-physical systems in medicine," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3810–3822, Oct. 2018.

[7] D. Packwood, M. Sharma, D. Ding, H. Park, Z. Salcic, A. Malik, and K. I.-K. Wang, "FPGA-based mixed-criticality execution platform for SystemJ and the Internet of industrial things," in *Proc. IEEE 18th Int. Symp. Real-Time Distrib. Comput.*, Apr. 2015, pp. 161–174.

[8] R. M. Dijkman, "Business models for the Internet of Things," *Int. J. Inf. Manage.*, vol. 35, no. 6, pp. 672–678, 2015.

[9] S. E. Middleton, S. Papadopoulos, and Y. Kompatsiaris, "Social computing for verifying social media content in breaking news," *IEEE Internet Comput.*, vol. 22, no. 2, pp. 83–89, Mar. 2018.

[10] J. Vitak, N. Proferes, K. Shilton, and Z. Ashktorab, "Ethics regulation in social computing research: Examining the role of institutional review boards," *J. Empirical Res. Human Res. Ethics*, vol. 12, no. 5, pp. 372–382, Dec. 2017.

[11] P. Chamoso, A. González-Briones, A. Rivas, F. De La Prieta, and J. M. Corchado, "Social computing in currency exchange," *Knowl. Inf. Syst.*, vol. 61, no. 2, pp. 733–753, Nov. 2019.
[12] S. Zinatizadeh, A. Azmi, S. M. Monavari, and S. Sobhanardakani, "Evaluation and prediction of sustainability of urban areas: A case study for Kermanshah city, Iran," *Cities*, vol. 66, pp. 1–9, Jun. 2017.
[13] Z. Wang, L. Yang, J. Yin, and B. Zhang, "Assessment and prediction of environmental sustainability in China based on a modified ecological footprint model," *Resour., Conservation Recycling*, vol. 132, pp. 301–313, May 2018.
[14] P. Cannavo, "Prediction of long-term sustainability of constructed urban soil: Impact of high amounts of organic matter on soil physical properties and water transfer," *Soil use Manage.*, vol. 30, no. 2, pp. 272–284, 2014.
[15] W. Wand, "Is scientific collaboration sustainability predictable?" in *Proc. 26th Int. Conf. World Wide Web Companion*, Perth, WA, Australia, 2017, pp. 853–854.
[16] E. Lee, J.-W. Lim, H. S. Moon, and K.-K. Lee, "Assessment of seawater intrusion into underground oil storage cavern and prediction of its sustainability," *Environ. Earth Sci.*, vol. 73, no. 3, pp. 1179–1190, Feb. 2015.
[17] L. P. Chung, "Application of CFD in prediction of indoor building thermal performance as an effective pre-design tool towards sustainability," *World Appl. Sci. J.*, vol. 30, no. 30, pp. 269–279, 2014.
[18] S. Jackson, "Prediction, explanation and big(GER) data: A middle way to measuring and modelling the perceived success of a volunteer tourism sustainability campaign based on 'nudging,'" *Current Issues Tourism*, vol. 19, no. 7, pp. 643–658, Jun. 2016.
[19] A. Capozzoli, D. Grassi, M. S. Piscitelli, and G. Serale, "Discovering knowledge from a residential building stock through data mining analysis for engineering sustainability," *Energy Procedia*, vol. 83, pp. 370–379, Dec. 2015.
[20] R. S. Middleton, R. Gupta, J. D. Hyman, and H. S. Viswanathan, "The shale gas revolution: Barriers, sustainability, and emerging opportunities," *Appl. Energy*, vol. 199, pp. 88–95, Aug. 2017.
[21] Z. Li, J. Han, B. Ding, and R. Kays, "Mining periodic behaviors of object movements for animal and biological sustainability studies," *Data Mining Knowl. Discovery*, vol. 24, no. 2, pp. 355–386, Mar. 2012.
[22] N. Ramakrishnan, M. Marwah, A. Shah, D. Patnaik, M. S. Hossain, N. Sundaravaradan, and C. Patel, "Data mining solutions for sustainability problems," *IEEE Potentials*, vol. 31, no. 6, pp. 28–34, Nov. 2012.
[23] J. Quinn, V. Frias-Martinez, and L. Subramanian, "Computational sustainability and artificial intelligence in the developing world," *AI Mag.*, vol. 35, no. 3, pp. 36–47, 2014.
[24] M. F. Hassan, M. Z. M. Saman, S. Sharif, and B. Omar, "Sustainability evaluation of alternative part configurations in product design: Weighted decision matrix and artificial neural network approach," *Clean Technol. Environ. Policy*, vol. 18, no. 1, pp. 63–79, Jan. 2016.
[25] M. Babaee, D. T. Dinh, and G. Rigoll, "A deep convolutional neural network for video sequence background subtraction," *Pattern Recognit.*, vol. 76, pp. 635–649, Apr. 2018.
[26] F. Gama, A. G. Marques, G. Leus, and A. Ribeiro, "Convolutional neural network architectures for signals supported on graphs," *IEEE Trans. Signal Process.*, vol. 67, no. 4, pp. 1034–1049, Feb. 2019.
[27] L. Lin, Z. He, and S. Peeta, "Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach," *Transp. Res. C, Emerg. Technol.*, vol. 97, pp. 258–276, Dec. 2018.
[28] S. Wang and C. Manning, "Fast dropout training," in *Proc. 30th Int. Conf. Mach. Learn.*, Atlanta, GA, USA, 2013, pp. 118–126.
[29] J. Su, Y. Yu, and Y. Tao, "Measuring knowledge diffusion efficiency in R&D networks," *Knowl. Manage. Res. Pract.*, vol. 16, no. 2, pp. 208–219, 2018.
[30] J. Su, C. Li, Q. Zeng, J. Yang, and J. Zhang, "A green closed-loop supply chain coordination mechanism based on third-party recycling," *Sustainability*, vol. 11, no. 19, p. 5335, 2019.
[31] D. Xu, P. Ren, J. A. Ritcey, and Y. Wang, "Code-frequency block group coding for anti-spoofing pilot authentication in multi-antenna OFDM systems," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 7, pp. 1778–1793, Jul. 2018.
[32] D. Xu, P. Ren, and J. A. Ritcey, "Independence-checking coding for OFDM channel training authentication: Protocol design, security, stability, and tradeoff analysis," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 2, pp. 387–402, Feb. 2019.

QIAO LI received the B.S. degree from Chongqing Technology and Business University, in 2013, and the M.S. degree in industrial economics from Chongqing University, in 2016. She is currently responsible for works related to management and scientific research with Chongqing Technology and Business University. Her recent research interests include social computing and ecological economics.

YING SONG received the Ph.D. degree from Southwest University, in 2014. Since 2005, she has been working with Chongqing Technology and Business University. She was promoted as an Associate Professor, in 2011. She has led a number of scientific research projects, including the National Social Science Foundation of China. Her current research interests include sociological economics and sustainable engineering.

BOXIN DU is currently pursuing the B.S. degree with Chongqing Technology and Business University. She will begin to pursue the M.S. and Ph.D. degrees, in 2020. She has implemented survey tasks of research projects in Hong Kong SAR, China. Her main research interests are related to data analysis and sociological economics. She received the Second Prize in the National Business Negotiation Competition of China, in 2018.

YU SHEN received the B.S. degree from Huazhong Agricultural University, in 2004, and the Ph.D. degree from the Dalian University of Technology, in 2009. He has led or is leading the three National Projects of China. He is currently a Researcher with Chongqing Technology and Business University. His research interests include ecological engineering and the Internet of Things.

YUAN TIAN received the M.S. and Ph.D. degrees from Chongqing University, in 2015 and 2019, respectively. She is currently a Lecturer with Chongqing Technology and Business University. She has been devoted to research fields of ecological economics and computational sustainability for at least five years. Her research interests are in regional economics and financial data analysis.