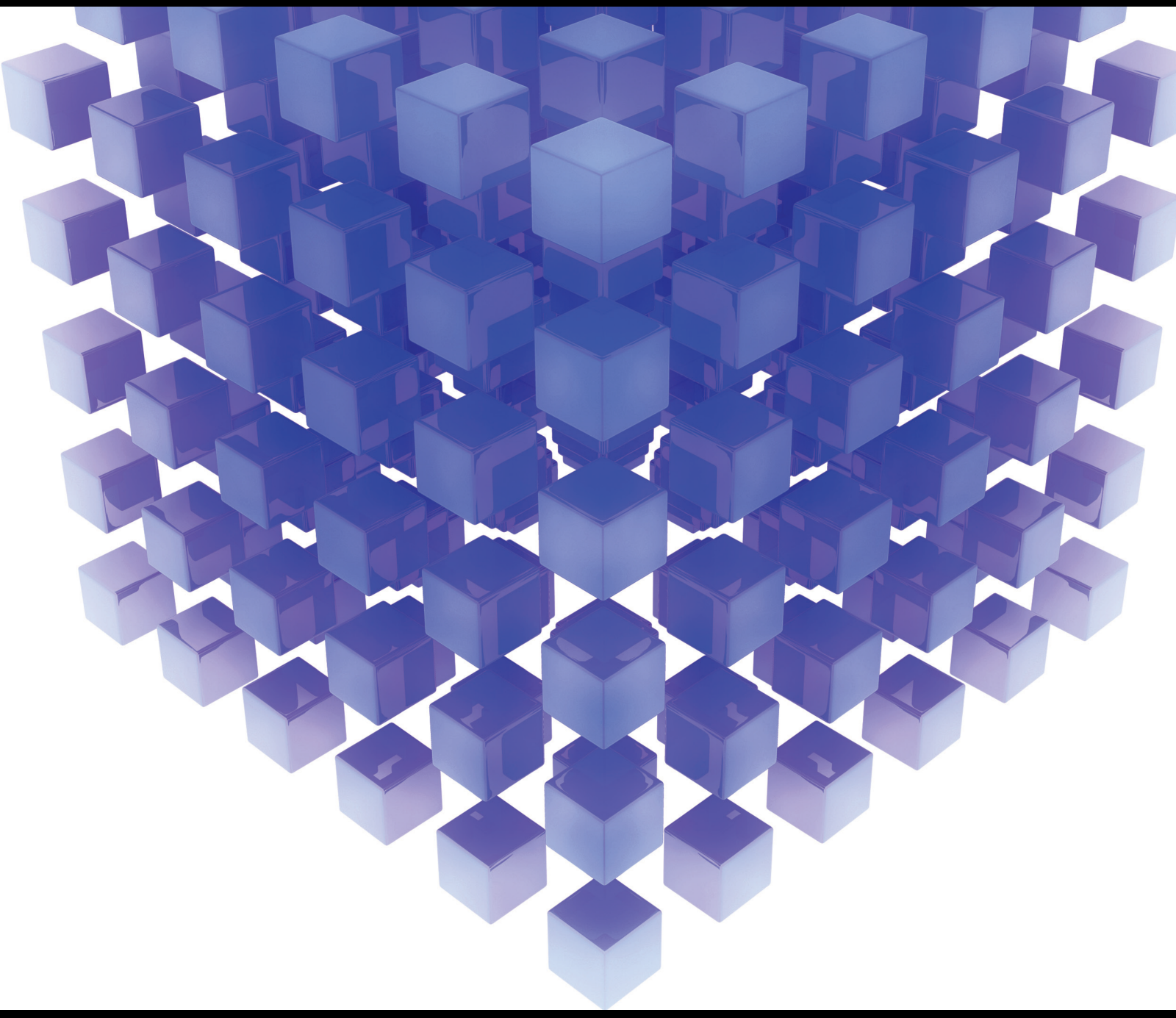
Mathematical Problems in Engineering

Big Data Modelling of Engineering and Management

Lead Guest Editor: Wen-Tsao Pan
Guest Editors: Shianghau Wu, Yungho Leu, and Weilin Xiao



Big Data Modelling of Engineering and Management

Mathematical Problems in Engineering

Big Data Modelling of Engineering and Management

Lead Guest Editor: Wen-Tsao Pan
Guest Editors: Shianghau Wu, Yungho Leu, and Weilin Xiao

Chief Editor

Academic Editors

Contents

Contents

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 2349716, 17 pages
https://doi.org/10.1155/2021/2349716

Research Article

Integrating Kano Model, AHP, and QFD Methods for New Product Development Based on Text Mining, Intuitionistic Fuzzy Sets, and Customers Satisfaction

Ming Li and Jie Zhang

East China University of Science and Technology, 130 Meilong Road, Xuhui District, Shanghai 200237, China

Correspondence should be addressed to Jie Zhang; hdlghd@163.com

Received 6 August 2020; Revised 2 February 2021; Accepted 15 March 2021; Published 31 March 2021

Academic Editor: Wen Tsao Pan

Online reviews are crucial to any online business that wants to increase sales on the Internet. Customer reviews have information about product attributes, customer requirements (CRs), and shopping experience; mining reviews provide the direction of decision-making for new product development and design (NPDD). Besides, the information of customer preference has vagueness and uncertainty, and the accuracy of decision-making information directly affects the success of NPDD. This paper proposed a methodology that integrates the Kano model (KM), analytic hierarchy process (AHP), and quality function deployment (QFD) methods with intuitionistic fuzzy set (IFS) to solve decision-making problems in NPDD. By the new method, the web crawler technology was first applied to e-commerce web sites to collect raw data, and the representative CRs were extracted through combining LDA model with Apriori algorithm. Second, the intuitionistic fuzzy Kano model (IFKM) is proposed to evaluate adjustment coefficient of CRs and Kano categories via customer preference membership functions. Thirdly, overall weights which contained emotional needs (ENs) and functional needs (FNs) are obtained via intuitionistic fuzzy analytic hierarchy process (IFAHP); thus, the adjusted weights are calculated from IFKM and IFAHP. Next, the intuitionistic fuzzy quality function deployment (IFQFD) is proposed to acquire engineering characteristics (ECs) of weights through combining competition benchmarks and based on technical benchmarks to make goals for a company's NPDD. Finally, the method was applied to study vertical-configured air conditioner (VAC) as an example. The results showed that the application of text mining and IFS to improve CS is both reliable and scientific.

1. Introduction

As market competition intensifies and product life cycles constantly get shorten, new product design and development (NPDD) methodology has become an effective way for enterprises to respond to market competition. With the rapid development of science and technology, NPDD also becomes a comprehensive process of diversification and complication. Anderson and Mittal [1] found that, for every 1% increase in customer satisfaction (CS) with products, the corresponding average return on investment (ROI) is increased by 2.37%, whereas, for every 1% decrease in satisfaction, the average ROI is decreased by 5.08%. Therefore, improving CS in product development is the key to enterprises' survival and success in the fierce market competition. At the same time, the rapid development of the Internet has provided unprecedented development opportunities for e-commerce platforms; more and more customers purchase products from online and generate huge comments, which contain information like customer needs and product attributes [2]. So, mining customer reviews can provide a scientific direction for NPDD. Moreover, Newman and Patel [3] mentioned three types of product attributes that are associated with customer requirements (CRs), namely, the attributes of basic function, the attributes of convenient function, and the attributes of psychological satisfaction. Both attributes of basic function and convenient function belong to customer functional requirements, and

the attributes of psychological satisfaction are the higher psychological response of customer experience; that is, the product forms will further promote CS after functions meet customer's basic requirements, thereby resonating with the customer's psychology. In today's environment in which CRs are diversified and personalized, emotional needs (ENs) become increasingly important in customer decision-making to purchase products. This requires NPDD not only to fulfill the customers' basic function needs (FNs) for products, but also to satisfy customers ENs. Therefore, enterprises should pay attention to both ENs and FNs in NPDD.

In recent years, many scholars use the Kansei engineering methodology to explore customer emotional factors and to help companies better understand customer ENs [4]. The Kansei engineering is a product development model originally proposed by Nagamachi. The model is used to quantify Kansei impression based on CRs, thus allowing ones to analyze the relationship between human impression and products forms to determine the right direction of product development [5]. However, Kansei engineering model primarily focuses on directing product form design and lacks an effective approach to assist engineering characteristics (ECs) development. In addition, it is essential to obtain accurate CRs and their categories in the process of CS-oriented NPDD, which determines the success/failure of new product development. The Kano model proposed by Professor Noriaki Kano can solve the problem of CRs' category classification and prioritization. The analysis of the impact of CRs on CS indicates a nonlinear relationship between product performance and CS [6]. But information on human perception and preference is often vague and uncertain. Some scholars have proposed combining fuzzy mathematic and Kano model to transform qualitative analysis into a quantitative one to obtain the preference information of interviewees. However, fuzzy Kano model can only express customer's positive and negative information and cannot deal with customer's uncertainty and hesitation. At last, the quality function deployment (QFD) technique establishes the relationship between the voice of customer (VOC) and the voice of technology (VOT); ENs and FNs are crucial to the improvement of CS [7], because the product form is the first avenue for customer perception to form emotional cognition, then continuously deepening the experience through the use of product functions to fuse the final CS. But enterprises tend to pay more attention to functionality and utility in NPDD, so CRs should be divided into ENs and FNs. Thus, those qualities can be embedded into a product at an early design stage which help companies save time and money, shorten product development cycles, and improve product market competitiveness.

In summary, to improve CS in NPDD process, one will face three key issues: how to get CRs from customer reviews, how to accurately investigate and obtain CRs' preference information, and how to efficiently and accurately improve CRs. In our study, we deal with those issues in several steps. First, we applied Kansei image to express customer's emotional factors (adjective words), combined with topic model and Apriori association algorithm to mine web reviews to obtain CRs including ENs and FNs. Secondly, we conducted questionnaire survey of two-dimensional CS through a combination of the intuitionistic fuzzy set (IFS) and Kano model to obtain customer intuition preference information (positive, negative, and hesitant) and thus to determine CRs categories and adjustment coefficient. Thirdly, we applied intuitionistic fuzzy analytic hierarchy process (IFAHP) to determine the intuition preference relations via the pairwise comparison between each CRs criterion and subcriterion obtained through expert interview. And then, each CRs adjusted importance can be determined by intuitionistic fuzzy Kano model (IFKM) and IFAHP. Lastly, we combined IFS with QFD to gain strong and weak relationships of CRs with ECs and surveyed competition benchmarks to obtain absolute importance of CRs and technical benchmarks to establish EC goals that could be used to guide companies and designers in product development and design optimization.

2. Literature Review

2.1. Kansei Engineering. Kansei is a Japanese word, it means "consumer's psychological feeling and image" when translated into English. Kansei engineering is a model for translating feelings and impressions into product parameters. The model was invented in the 1970s by Professor Nagamachi at Kure University. He recognized that companies usually want to quantify the customer's impression of their products. Kansei engineering can measure the feelings and show the correlation to certain product features. Therefore, products can be designed in a way that responds to the intended feeling. Many applications of Kansei engineering focus on the physical design of products appearance. For example, Yang et al. [8] applied rule-based inference model to design telephone's form and style under the framework of Kansei engineering. Yang [9] used a hybrid Kansei engineering system of multiple images for design form of mobile phone. Shieh and Yeh [10] established the relationship between the shape of running shoe and the emotional response of consumers using Kansei engineering methodology, showing that the appearance of running shoe has a significant impact on consumers' overall perception, and their results provided effective guidance for designer in shoe design. Wang [11] combined Kansei engineering with conjoint analysis to achieve product customization of digital cameras and provided decision support for the development of next-generation digital cameras. However, Kansei engineering is rarely used to guide product functions design; most of its applications focus on the psychological impact of product appearance on customers' psychological needs. In addition, it only reflects the corresponding customer's feelings in the process of extracting customer image features and cannot really reflect the customer's demand preferences. Therefore, the current Kansei engineering method has limitations in the analysis of CS with product design.

2.2. Topic Model. The topic model is a statistical model for discovering abstract "topics" that occur in a collection of documents in machine leaning and natural language

processing. In 1998, Papadimitriou first proposed an early topic model named latent semantic indexing (LSI) [12]. Then, Hofmann created probabilistic latent semantic indexing (PLSI) in 1999 [13]. After continuous research and improvement, latent Dirichlet allocation (LDA) becomes the most common topic model currently in use that is a generalization of PLSI. Developed by David Blei and his colleagues [14] in 2003, LDA model is a Bayesian probability model, which is an unsupervised learning algorithm. Its basic points are words that often show different probability distribution characteristics in different potential topic distributions and the probability of words is relatively close under the same topic. Now, LDA model is usually applied in mining topics, text classification, and text similarity. For instance, Walter et al.[15] used LDA to extract topics from Twitter comments and distinguish man-made accounts. Maier et al.[16] proposed a reliable and effective method with LDA in communication field. Wei et al.[17] created improved LDA to cluster texts. However, there may be a single review involving multiple CRs, and different customers have different idiomaticity that means the same CR was described by different expressions. In addition, a single document often contains multiple topics and topic distributions will change depending on the document. Therefore, based on the LDA model, this paper introduced Apriori association algorithm to solve the above problems.

2.3. Apriori Association Algorithm. Apriori was proposed by Agrawal in 1994 to solve market basket problem. Its purpose is to discover the association rules between different commodities in transaction database [18]. The algorithm uses iterative method of layer-by-layer search to find the relationship between item set to form rules in database, and the concept of item set is the set of items (commodities). If candidate item set meets the minimum support degree, then it is called frequent item set. Now, Apriori algorithm is widely used in business, mobile communication, network security, and management; for example, Guo et al.[19] applied Apriori algorithm to improve the recommendation system of e-commerce platform. Panjaitan et al.[20] used Apriori algorithm to explore consumption pattern. Liu et al.[21] applied Apriori algorithm to solve enterprise stock management problem. Therefore, we attempt to mine the list of topic words as transaction which is got by LDA model and analyze the frequently matched vocabulary with attribute words, thus finding out the content related to these attributes, and finally locate CRs information.

2.4. Kano Model. In 1984, Professor Noriaki Kano and his colleague Fumio Takahashi [22], inspired by Herzberg's motivator-hygiene theory, proposed attractive quality and must-be quality theory. Then, Kano divides quality attributes into five categories: attractive quality, one-dimensional quality, must-be quality, indifferent quality, and reverse quality (Figure 1). The pair of questionnaire surveys, one assumed to be with and another one without quality (CRs) attributes, is conducted to provide insight into which quality attributes fall into which quality categories (Table 1), and it

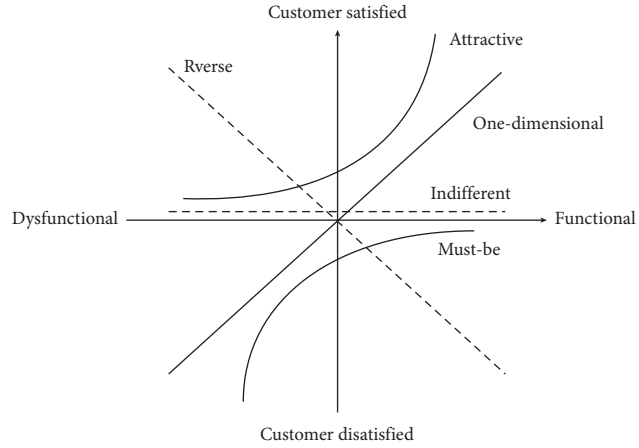


Figure 1: Kano model.

Table 1: The Kano evaluation table.

Pref.	Dysfunctional				
Functional	1. Like	2. Must-be	3. Neutral	4. Live with	5. Dislike
1. Like	Q	A	A	A	O
2. Must-be	R	I	I	I	M
3. Neutral	R	I	I	I	M
4. Live with	R	I	I	I	M
5. Dislike	R	R	R	R	Q

Note: A: attractive quality; O: one-dimensional quality; M: must-be quality; I: indifferent quality; R: reverse quality; Q: questionable quality.

offers a better understanding of how customers evaluate a product; thus this can be used to assist enterprises to focus on the improvement of most important attributes. After that, the Kano model is widely used to do research by many scholars in NPDD, such as Matzler and Hinterhuber [23], who discussed the strategic importance of the Kano model in NPDD, and proposed the satisfaction coefficient to identify contributions to CS. Tan and Shen [24] incorporated the Kano model into the QFD planning matrix to help understand VOC accurately and deeply and proposed a method of adjusting the importance of CRs by approximating transformation function. Xu et al. [25] attempted to quantify the Kano two-dimensional questionnaire scores in the gradient range of -0.5 to 1 and then converted their values to polar coordinate to get the classification of Kano. The above research on the improvement of the Kano model has moved from qualitative to quantitative level.

Some scholars have tried to study the ambiguity and uncertainty in the process of evaluating human information. For example, Foldesi et al. [26] tried to apply the fuzzy and Kano model to obtain the maximum CRs with financial constraints. Lee and Huang [27] developed the fuzzy Kano two-dimensional questionnaire method, which converts one answer of the traditional Kano questionnaire into multiple answers in percentages. Each percentage represents the membership degree of the standard answer. Although the above-mentioned Kano models with incorporation of fuzzy theory have led to more accurate VOC in expressing uncertainty and ambiguity of CRs, these two methods are still

of qualitative analysis and do not consider continuous problems. To solve the continuous problems, Wu and Wang [28] developed a continuous fuzzy Kano model (CFKM) to analyze the ambiguity of CRs more accurately and proposed the influence value matrix to adjust the positive and negative emotional differences in the survey. However, CFKM only describes positive and negative information of customer preference and cannot express hesitation information. Consumers' perception differs according to their experience, knowledge, and environment, which results in different degrees of decision hesitation among attributes. IFs not only describe the degree of membership and nonmembership of a certain preference, but also reflect the degree of hesitation.

Therefore, in this paper, we attempt to obtain the relationship between CRs and CS based on IFs and then help companies analyze consumer preferences more accurately and efficiently in NPDD.

2.5. Intuitionistic Fuzzy Analytic Hierarchy Process. Analytic hierarchy process (AHP) is a decision-making tool combining qualitative and quantitative methods, was first proposed by Professor Saaty [29], and has been widely used to solve multicriteria decision-making problems. AHP uses five steps to calculate the combined weight of all constituent elements of each level to the goal: defining goal or problem, establishing hierarchy structure, constructing judgment matrix, sorting single hierarchy, and sorting total hierarchy, thus obtaining the comprehensive evaluation value of different feasible schemes, which provides a basis for choosing the best. However, the key of AHP is to establish a judgment matrix. However, whether the judgment matrix is scientific and reasonable will directly affect the analysis results, and subjectivity and ambiguity of decision-makers in the scoring process always exist. Van Laarhoven and Pedrycz [30] combined fuzzy and AHP to express the ambiguity of decision-makers as triangular fuzzy numbers to solve vagueness problem. Kwiesielewicz [31] mentioned that van Laarhoven and Pedrycz's fuzzy AHP has issues of normal equation having one or more degrees of freedom and applied the generalized pseudoinverse approach to find a structure of general solution. Boender et al. [32] modified van Laarhovern and Pedrycz's fuzzy method and proposed a more robust approach to the normalization of the local priorities. However, these studies have not solved the problem of affirmation, negation, and hesitation of complex cognition, because the membership function used is only single-valued function which cannot be applied to express the support and objection information. Xu and Liao [33] proposed intuitionistic fuzzy analytic hierarchy process (IFAHP) to solve positive, negative, and uncertain preference; besides AHP and FAHP need to reinvestigate the experts when judgment matrix is inconsistent. IFAHP only needs to achieve consistency through the conversion of the judgment matrix. This method also saves time and resources. In addition, the multiattribute decision-making tool represented by AHP has been widely used in NPDD [34–36]. In consequence, this paper used IFAHP to take CS as the target and construct subgoal of customers' ENs and FNs and then obtain the overall weights of CRs, thus multiplying intuitionistic fuzzy Kano evaluation value. In this way, CRs overall adjusted evaluation weights and priority can be obtained.

2.6. Quality Function Deployment. The concept of quality function deployment (QFD) was proposed by Yoji Akao in Japan in the 1960s [37], and this process was integrated with other improvement tools to generate a lot of opportunities for product designers. QFD is a customer-oriented product development tool that achieves higher CS by translating CRs into design requirements, part features, or production plans. The intents of applying QFD are to incorporate VOC into the product development cycle through customer surveys and interviews and to assume the achievement of CS quality. Within the elements of QFD, House of Quality (HOQ) is the core framework of QFD to organize VOC with VOT, and the VOC is expressed in the customer's terms which can be in the form of linguistic or crisp variables. According to research by Hauser and Clausing [38], QFD can reduce product development time by 50% and cost by 30%. Lai et al. [39] took advantage of Kano model in CRs and combined it with QFD to direct new product development. Tu et al. [40] combined AHP with QFD to develop a sport headphone. Gül Bayraktaroğ and Özge Özgen [41] integrated Kano model, AHP, and QFD to analyze CRs of library services to improve service quality in market strategy development and provide guidance library managers on how to allocate budget, arrange services, and develop their marketing strategies. In the face of vagueness and uncertainty of VOC, Khoo and Ho [42] proposed a fuzzy QFD model, which expresses VOC as a triangular fuzzy number and applied it to the design and development. Wu [43] developed fuzzy HOQ and fuzzy QFD for fuzzy regression evaluation. Similarly, there are characteristics of affirmation, negation, and hesitation in linguistic relation (weak correlation, medium correlation, and strong correlation) in fuzzy HOQ. The membership function cannot fully describe the essential information. So, our study tried to combine IFS and QFD to cover experts' preference information; thus we can achieve engineering characteristics ranking and weights, to finally get goals in NPDD according to the technical benchmarks.

3. Research Methodology

3.1. Intuitionistic Fuzzy Set Background. Fuzzy set theory (FS) is an effective technical method for modeling inaccuracy and uncertainty in the real world. FS and IFS are often used to express uncertain information in multiattributes decision-making processes, which has resulted in a great deal of research achievement [44–46]. IFS is an extended fuzzy information concept proposed by Atanassov [47] in 1983 on the basis of FS theory. It extends FS from a single scale to two scales to describe the ambiguity of a physical phenomenon, which possesses three fuzzy states of support, neutrality, and opposition. It has the advantage over FS in accuracy, flexibility, and comprehension.

Definition 1. Let a fixed universe crisp set Xbe fixed and let $A \subset X$ be a fixed set [47–49]. An IFS $\widetilde{A}$ in $X$is an object of the following form:

$$\widetilde{A} = \{(x, \mu_A(x), \nu_A(x)) | x \in X\}, \tag{1}$$

where the functions μ_A: $X \longrightarrow [0,1]$ and ν_A: $X \longrightarrow [0,1]$ represent, respectively, the degree of membership and nonmembership of elements$x \in X$ to the set A. And, for every $x \in X$, $0 \le \mu_A + \nu_A \le 1$ holds.

Definition 2 [50–52] Triangular intuitionistic fuzzy number (TIFN) $\widetilde{a} = ((a, a, \overline{a}); w_{\widetilde{a}}, u_{\widetilde{a}})$ obviously is a special case of IFS on a real number set R, whose membership function and nonmembership function are defined as follows [47–49]:

$$\mu_{\widetilde{a}}(x) = \begin{cases} \left(x - \underline{a}\right) w_{\widetilde{a}} / (a - \underline{a}), & \text{if } a \le x < a, \\ w_{\widetilde{a}}, & \text{if } x = a, \\ (\overline{a} - x) w_{\widetilde{a}} / (\overline{a} - a), & \text{if } a < x \le \overline{a}, \\ 0, & \text{if } x < \underline{a} \text{ or } x > \overline{a}, \end{cases}$$

$$\nu_{\widetilde{a}}(x) = \begin{cases} \left[a - x + u_{\widetilde{a}}(x - \underline{a})\right] / (a - \underline{a}), & \text{if } a \le x < a, \\ u_{\widetilde{a}}, & \text{if } x = a, \\ \left[x - a + u_{\widetilde{a}}(\overline{a} - x)\right] / (\overline{a} - a), & \text{if } a < x \le \overline{a}, \\ 1, & \text{if } x < \underline{a} \text{ or } x > \overline{a}. \end{cases} \tag{2}$$

The values $w_{\widetilde{a}}$and $u_{\widetilde{a}}$represent the maximum and the minimum degree of nonmembership, respectively, and satisfy the conditions $0 \le w_{\widetilde{a}} \le 1$, $0 \le u_{\widetilde{a}} \le 1$, and $0 \le w_{\widetilde{a}} + u_{\widetilde{a}} \le 1$.

Definition 3. For each IFS $\widetilde{A}$ on X[47–49], then

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x), \tag{3}$$

is called the degree of nondeterminacy (uncertainty) of the membership of the element $x \in X$ to the set A. In the case of ordinary fuzzy sets, $\pi_A(x) = 0$ for every $x \in X$.

Szmidt and Kacprzyk [53] pointed out that $\pi_A(x)$ is a hesitancy degree of x to A and cannot be omitted when calculating the distance between two IFSs. For convenience, Xu [54] called $\alpha = (\mu_\alpha, \upsilon_\alpha, \pi_\alpha)$ an intuitionistic fuzzy value (IFV), where $\mu_\alpha \in [0,1]$, $\nu_\alpha \in [0,1]$, $\pi_\alpha \in [0,1]$, and $0 \le \mu_\alpha + \nu_\alpha \le 1$, and assume that Ψ be the set of all IFVs.

In order to rank the IFVs, Szmidt and Kacprzyk [53] developed a function, defined mathematically as follows:

$$\rho(\alpha) = 0.5\left(1 + \pi_\alpha\right)\left(1 - \mu_\alpha\right). \tag{4}$$

The smaller the value of $\rho(\alpha)$ is, the greater the IFV α is in the sense of the amount of positive information included and reliability of information.

Definition 4. Let $\widetilde{A}$ and $\widetilde{B}$ be two IFSs in universe X and let λ be a real number and $\lambda > 0$[55]. The arithmetical operations are stipulated as follows:

$$\widetilde{A} \oplus \widetilde{B} = \{(x, \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x), \nu_A(x)\nu_B(x)) | x \in X\}, \tag{5}$$

$$\widetilde{A} \otimes \widetilde{B} = \{(x, \mu_A(x)\mu_B(x), \nu_A(x) + \nu_B(x) - \nu_A(x)\nu_B(x)) | x \in X\}, \tag{6}$$

$$\lambda\widetilde{A} = \left\{\left(x, 1 - \left(1 - \mu_A(x)\right)^\lambda, \left(\nu_A(x)\right)^\lambda\right) | x \in X\right\}, \tag{7}$$

$$\widetilde{A}^\lambda = \left\{\left(x, \left(\mu_A(x)\right)^\lambda, 1 - \left(1 - \nu_A(x)\right)^\lambda\right) | x \in X\right\}. \tag{8}$$

3.2. Selecting Representative Customer Needs. CRs attributes determine the direction of product development process; thus selecting and using improper CRs information could waste companies' large amount of time and money. Thus, to accurately obtain CRs is the first step to successfully develop products. In recent years, many e-commerce platforms such as Amazon and JD.com have launched online review systems for their products or services. Those platforms allow consumers to share and acquire more product information through online reviews. A huge amount of customer comments can be extracted by web crawler technology in e-commerce platform, and ENs and FNs are mined with LDA model and Apriori association algorithm.

The LDA model is represented as probabilistic graphical model in Figure 2 [14]; α is the Dirichlet prior parameter of the topic distribution in each document, reflecting relative strength of the hidden topics in the corpus; β is the Dirichlet prior parameter of the word distribution in each topic, reflecting the probability distribution of all hidden topics. And LDA model has two basic assumptions: each document is a set of series of topics and each word belongs to a topic in the document.

The analysis process of LDA model mainly consists of two parts: generative process and parameter estimation. In the generative process part, LDA assumes a corpus D consisting of M documents each of length N_i and K denotes topic in D.

(1) Choose $\theta_i \sim$ Dirichlet(α), where $i \in \{1, 2, 3, \ldots, M\}$, θ_i is the topic distribution for document i, and Dirichlet(α) is a Dirichlet distribution with a symmetric parameter α which is typically sparse ($\alpha < 1$).

(2) Choose $\varphi_k \sim$ Dirichlet(β), where $k \in \{1, 2, 3, \ldots, K\}$, φ_k is the word distribution for topic k, and β is typically sparse.

(3) For each of the word positions i and j, where $i \in \{1, 2, 3, \ldots, M\}$ and $j \in \{1, 2, 3, \ldots, N_i\}$, we het the following. Then, choose a topic $z_{ij} \sim$ Multinomial$(\theta_{i,k})$, where z_{ij} is the topic for jth word in document i and choose a word $w_{ij} \sim$ Multinomial$(\varphi_{k,t})$, where w_{ij} is the specific word; multinomial distribution refers to the multinomial with only one trial, which is also known as the categorical distribution.

In the parameter estimation part, through the estimation of φ and θ, we can obtain the relevant information of the

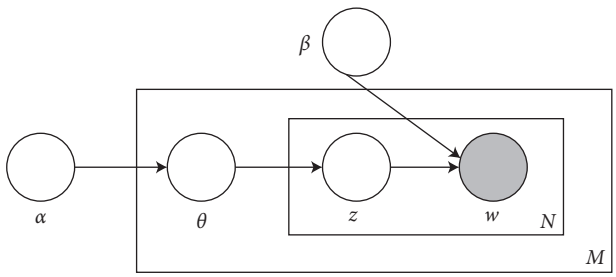


Figure 2: Graphical model representation of LDA.

topics contained in the corpus and its proportion in each document. These parameters using Gibbs sampling [56] are estimated as follows:

(1) Estimate the probability value of a document.

$$p(w|\alpha,\beta) = \int p(\theta|\alpha)\left(\prod_{j=1}^{N}\sum_{z_j} p\left(z_j|\theta\right)p\left(w_j|z_j,\beta\right)\right)d\theta. \tag{9}$$

Collapsed Gibbs sampling first estimates the relationship between each word and topic and then calculates these parameters through frequency statistics.

(2) Calculate the conditional probability of the topic sequence under the word sequence.

$$p\left(z_j = k|z_{-j}, w\right) = \frac{p(w,z)}{p\left(w,z_{-j}\right)} \propto \frac{n_{k,-j}^{t} + \beta_t}{\sum_{t=1}^{V} n_{k,-j}^{t} + \beta_t} \frac{n_{i,-j}^{k} + \alpha_k}{\sum_{i=1}^{K} n_{i,-j}^{k} + \alpha_k}, \tag{10}$$

where z_j represents the assignments of jth word in a document to topic k, z_{-j} represents all topic assignments not including jth word, n_k^t denotes the number of times in which word t is assigned to topic z_k, and n_i^t denotes the number of times in which topic z_k is assigned to document m_i.

(3) Calculate φ and θparameters.

$$\varphi_{k,t} = \frac{n_k^t + \beta_t}{\sum_{t=1}^{V} n_k^t + \beta_t},$$
$$\theta_{i,k} = \frac{n_i^k + \alpha_t}{\sum_{t=1}^{K} n_i^k + \alpha_t}, \tag{11}$$

where $\varphi_{k,t}$ represents the probability of word t in topic z_k and $\theta_{i,k}$ represents the probability of topic z_k in document i. So, we can get the relations topic-word and document-topic by probability value.

In LDA model process, the perplexity [14] is used to determine the number of topics, and the smaller the perplexity, the better the model performance; see the following formula:

$$\text{perplexity}(w) = exp\left\{-\frac{\log(p(w))}{\sum_{i=1}^{N}\sum_{t=1}^{V} n_i^t}\right\}, \tag{12}$$

where V is the number of words in corpus D and n_i^t denotes the number of times in which word t is in ith document.

However, there are two issues in mining CRs. Firstly, a single customer review (document) may include multiple CRs; secondly, due to different idiomaticity, the same need is expressed in different ways. Hence, mining words frequently collocate with attribute words; we can find out the content related to these attributes so that CRs are located. In the example, we used the topic words of multiple iterations of the LDA model as transactions.

The Apriori algorithm uses breadth-first search and Hash tree structure to count candidate item sets efficiently [18]. It generates candidate item set of length k from sets of length k-1, then it prunes the candidates which have an infrequent subpattern, and these sets of items have at least the given minimum support. After that, it scans the transaction database to determine frequent items sets among the candidates. The minimum support and minimum confidence are measured association rules in transaction, which represents certainty and usefulness, respectively, as in (13)-(14). And lift is a supplement to association rules, which reflects the worth of the rules as in (15):

$$\text{support}(A \Rightarrow B) = P(A \cup B), \tag{13}$$

$$\text{confidence}(A \Rightarrow B) = P(B|A) = \frac{\text{support}(A \cup B)}{\text{support}(A)} = \frac{\text{support\_count}(A \cup B)}{\text{support\_count}(A)}, \tag{14}$$

$$\text{lift}(A \Rightarrow B) = \frac{\text{confidence}(A \Rightarrow B)}{\text{support}(B)}, \tag{15}$$

where A and B represent the item sets (topic words), $A \Rightarrow B$denote an association rule, greater than minimal support value S, and S represents the ratio of all topics that include both topic words A and B, which denote$P(A \cup B)$. If S is smaller, it means that the relationship between A and B is not great; otherwise, A and B are related. The confidence degree C represents the probability that the topic of item sets A and B appears at the same time, which denote $P(B|A)$. If C is smaller, it means that both A and B have little chance to appear. The lift degree L reflects the correlation between A and B in the association rules. If L is greater than 1, thus this means higher positive correlation; if L is less than 1, thus this means higher negative correlation; otherwise, A and B are not relevant.

3.3. Intuitionistic Fuzzy Kano Model Acquires CRs Category and Adjustment Coefficient. The Kano model has been widely used in industry and scientific research. More specifically, it is used to analyze and classify CRs. In addition, the Kano model is a two-dimensional quality model and is used to present asymmetry and nonlinear relationship between product performance and CS [27]. The questionnaire survey (Table 2) includes a pair of questions (function/dysfunction CRs attributes), which is used to calculate Kano categorization [22]. As mentioned before, information for evaluation, which is solely based on crisp value (classical Kano model) or positive/negative value (FKM and CFKM), is not complete, because people often express their preference in the scoring process with uncertainty and vagueness.

The triangular IFNs (TIFNs) are a special case of fuzzy sets and are important for fuzzy multiattributes decision-making problems [51]. Intuitionistic fuzzy Kano model (IFKM) is different from FKM, and it only uses one point in the survey to express customers' preference, and it is also necessary when investigating the scores of satisfaction and dissatisfaction and the need for obtaining positive, negative, and hesitant preference in the questions (Figure 2). Besides, the degree of membership of one's preference to a standard answer (enjoy, expect, neutral, live with, or dislike) can be considered as a decreasing function of the distance between the chosen point and the standard answer. The closer the chosen point to the standard answer, the greater the membership degree of one's preference to it. In this paper, we use triangular membership functions to represent the memberships of one's preference to the five standard answers as follows:

(1) Dislike (1, 1, 3)

$$f_R(x) = \begin{cases} w_{\tilde{a}}, & x = 1, \\ \frac{1}{2}(3-x)w_{\tilde{a}}, & 1 < x \le 3, \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

(2)Live with (1, 3, 5)

$$f_R(x) = \begin{cases} \frac{1}{2}w_{\tilde{a}}(x-1), & 1 \le x < 3, \\ w_{\tilde{a}}, & x = 3, \\ \frac{1}{2}(5-x)w_{\tilde{a}}, & 3 < x \le 5, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

(3)Neutral (3, 5, 7)

$$f_R(x) = \begin{cases} \frac{1}{2}w_{\tilde{a}}(x-3) & 3 \le x < 5 \\ w_{\tilde{a}} & x = 5 \\ \frac{1}{2}(7-x)w_{\tilde{a}} & 5 < x \le 7 \\ 0 \text{ otherwise.} \end{cases} \tag{18}$$

(4)Expect (5, 7, 9)

$$f_R(x) = \begin{cases} \frac{1}{2}w_{\tilde{a}}(x-5), & 5 \le x < 7, \\ w_{\tilde{a}}, & x = 7, \\ \frac{1}{2}(9-x)w_{\tilde{a}}, & 7 < x \le 9, \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

(5)Enjoy (7, 9, 9)

$$f_R(x) = \begin{cases} \frac{1}{2}w_{\tilde{a}}(x-7), & 7 \le x < 9, \\ w_{\tilde{a}}, & x = 9. \end{cases} \tag{20}$$

where $w_{\tilde{a}}$ is the maximum membership of one's attribute and the membership degrees of one's preference to the standard answers can be obtained through the membership functions. For example, the maximum membership and minimum nonmembership are 0.9 and 0, respectively. And an interviewee chooses the point 3.3 as his/her answer to a certain question; the membership degrees of his/her preference to 'like with' and 'neutral' are 0.765 and 0.135, respectively (Figure 3).

IFKM uses Kano evaluation table to combine the membership of the function with that of the dysfunction of a certain attribute (Table 1). For instance, one respondent's preference may be assigned to cells (2, 4), (2, 5), (1, 4), and (1, 5) in Table 1. The membership degree μ_{nij} of his/her preference combination (i, j) is determined by the following form [27]:

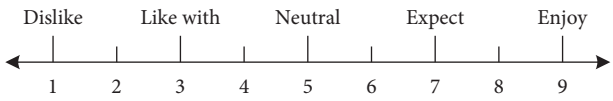
$$\mu_{nij} = m(F_i)_n \times m(D_j)_n, \tag{21}$$

where the subscript n represents nth participant; $m(F_i)_n$ and$m(D_j)_n (i, j = 1, 2, \ldots, 5)$ are the membership degrees of his/her preference to ith standard answer to the functional questions and jth standard answer to the dysfunctional questions, respectively. Then, the membership degree μ_{nij} of nth participant's preference can be identified as the degree of "attractive quality," "must-be quality," "one-dimensional quality," "indifferent quality," or "reverse quality."

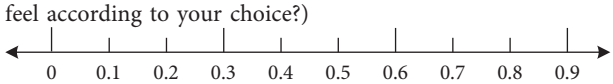
According to the Kano evaluation in Table 1, 25 combinations contribute different values to CS, because "like" has more intense emotion, and "must-be" has only weaker expression of satisfaction. Similarly, "dislike" is a very strong dislike, whereas "live with" is only a weak expression of dissatisfaction [57]. In addition, Tan and Shen [24] and Chen and Ko [58] suggested assigning values of 2, 1, and 0.5 to the corresponding categories of must-be, one-dimensional, and attractive. Therefore, in the upper-right quadrant of Table 1, the influence values of pure "must-be," "one-dimensional," and "attractive" are 2, 1, and 0.5, respectively. In lower-left quadrant of Table 1, the answers of the function are stronger responses than dysfunction ones, so the negative influence values are half as large as the corresponding positive ones. Thus, a standard matrix of the influence value was given as follows:

Table 2: The IFKM questionnaire.

Function and dysfunction questions (e.g., if the VAC has/does not have simple and liberal attribute, how do you feel?)



Degree of score satisfaction and dissatisfaction (e.g., what do you feel according to your choice?)



Note. Satisfaction + dissatisfaction is less than 1.

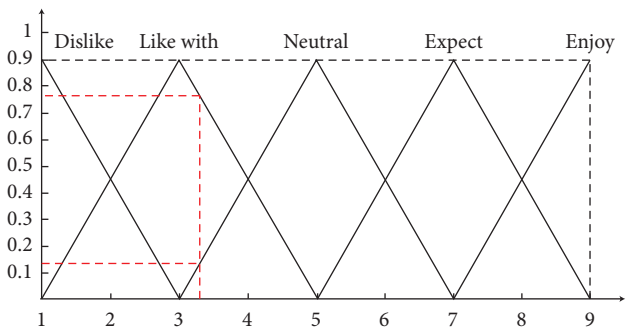


Figure 3: Membership functions of IFNs.

$$v_{ij} = \begin{bmatrix} 0 & 0.200 & 0.250 & 0.300 & 0.500 \\ -0.100 & 0 & 0.050 & 0.075 & 0.900 \\ -0.125 & -0.025 & 0 & 0.100 & 1.000 \\ -0.150 & -0.038 & -0.050 & 0 & 0.8 \\ -0.250 & -0.450 & -0.500 & -0.400 & 0 \end{bmatrix}. \quad (22)$$

At last, nth participant's satisfaction value of the quality SF_n can be obtained by a mathematical aggregation of all influence values weighted by the corresponding membership degrees.

$$SF_n = \sum_{i=1}^{5}\sum_{j=1}^{5} v_{ij} \times \mu_{nij}. \quad (23)$$

3.4. Intuitionistic Fuzzy Analytic Hierarchy Process to Obtain ENs and FNs Importance. IFKM can be used to obtain priorities and Kano categories of customers' ENs and FNs, respectively, but the weights of ENs and FNs cannot be compared directly. Because Kansei image and function are different abstract concepts, experts are more comprehensive and professional than customers to distinguish the weight. IFAHP can be considered as an analysis method of extending AHP and FAHP [33], where FAHP can analyze and process the participants' membership information better than AHP, and IFAHP can obtain nonmembership and neutral information better than FAHP. IFAHP analysis includes the following 7 steps:

(1) Identify the objective, criteria, and subcriteria of the decision-making problem and then construct the hierarchy of the considered problem. This step is similar to the classical AHP.

(2) Determine the intuitionistic preference relation via the pairwise comparison between each criterion and subcriterion. And then, the intuitionistic preference relations are constructed. The scale regarding the relative importance degrees is denoted by IFVs.

(3) If all the intuitionistic preference matrices have acceptable consistency, go to the next step. The consistency of each intuitionistic preference matrix is given as follows:

$$\overline{\mu}_{ik} = \frac{\sqrt[k-i-1]{\prod_{t=i+1}^{k-1}\mu_{it}\mu_{tk}}}{\sqrt[k-i-1]{\prod_{t=i+1}^{k-1}\mu_{it}\mu_{tk}} + \sqrt[k-i-1]{\prod_{t=i+1}^{k-1}(1-\mu_{it})(1-\mu_{tk})}}, \quad k > i+1, \quad (24)$$

$$\overline{v}_{ik} = \frac{\sqrt[k-i-1]{\prod_{t=i+1}^{k-1}v_{it}v_{tk}}}{\sqrt[k-i-1]{\prod_{t=i+1}^{k-1}v_{it}v_{tk}} + \sqrt[k-i-1]{\prod_{t=i+1}^{k-1}(1-v_{it})(1-v_{tk})}}, \quad k > i+1, \quad (25)$$

$$d(R,\overline{R}) = \frac{1}{2(n-1)(n-2)}\sum_{i=1}^{n}\sum_{k=1}^{n}\left(\left|\overline{\mu}_{ik}-\mu_{ik}\right| + \left|\overline{v}_{ik}-v_{ik}\right| + \left|\overline{\pi}_{ik}-\pi_{ik}\right|\right), \quad (26)$$

$$d(R,\overline{R}) < \tau, \quad (27)$$

where $\overline{\mu}_{ik}$ is the consistency preference membership; $\overline{v}_{ik}$ is consistency preference nonmembership, and $i < k < t\,(i,k,t = 1,2,3,\ldots,n)$ are elements of matrix; μ_{ik} is original preference membership; v_{ik} is original preference membership; R is the expert judgment preference matrix; $\overline{R}$ is consistency judgment matrix; τ is consistency threshold value and is generally set to 0.1.

(4) Recalculate the inconsistent intuitionistic preference relations as follows:

$$\widetilde{\mu}_{ik}^{(p)} = \frac{\left(\mu_{ik}^{(p)}\right)^{1-\sigma}\left(\overline{\mu}_{ik}^{(p)}\right)^{\sigma}}{\left(\mu_{ik}^{(p)}\right)^{1-\sigma}\left(\bar{\mu}_{ik}^{(p)}\right)^{\sigma} + \left(1-\mu_{ik}^{(p)}\right)^{1-\sigma}\left(1-\bar{\mu}_{ik}^{(p)}\right)^{\sigma}}, \quad i,k = 1,2,\ldots,n, \tag{28}$$

$$\widetilde{\nu}_{ik}^{(p)} = \frac{\left(\nu_{ik}^{(p)}\right)^{1-\sigma}\left(\overline{\nu}_{ik}^{(p)}\right)^{\sigma}}{\left(\nu_{ik}^{(p)}\right)^{1-\sigma}\left(\bar{\nu}_{ik}^{(p)}\right)^{\sigma} + \left(1-\nu_{ik}^{(p)}\right)^{1-\sigma}\left(1-\bar{\nu}_{ik}^{(p)}\right)^{\sigma}}, \quad i,k = 1,2,\ldots,n, \tag{29}$$

$$d\left(\overline{R}, R^{(p)}\right) = \frac{1}{2(n-1)(n-2)}\sum_{i=1}^{n}\sum_{k=1}^{n}\left(\left|\overline{\mu}_{ik} - \mu_{ik}^{(p)}\right| + \left|\overline{\nu}_{ik} - \nu_{ik}^{(p)}\right| + \left|\overline{\pi}_{ik} - \pi_{ik}^{(p)}\right|\right), \tag{30}$$

where p denotes the number of iterations and $p \geq 2$; $\widetilde{\mu}_{ik}^{(p)}$ is iterative membership; $\widetilde{\nu}_{ik}^{(p)}$ is iterative nonmembership; σ denotes controlling parameter determined by the decision-maker, and $0 \leq \sigma \leq 1$. The smaller the value of σ, the closer the $\widetilde{R}^{(p)}$ to $R^{(p)}$.

(5) Calculate the weight vector $K = (K_1, K_2, \ldots K_n)$ of each intuitionistic preference relation as follows:

$$\begin{aligned} w_i &= \frac{\sum_{k=1}^{n}[\mu_{ik}, 1-\nu_{ik}]}{\sum_{i=1}^{n}\sum_{k=1}^{n}[\mu_{ik}, 1-\nu_{ik}]} \\ &= \frac{\left[\sum_{k=1}^{n}\mu_{ik}, \sum_{k=1}^{n}(1-\nu_{ik})\right]}{\left[\sum_{i=1}^{n}\sum_{k=1}^{n}\mu_{ik}, \sum_{i=1}^{n}\sum_{k=1}^{n}(1-\nu_{ik})\right]}, \\ &= \left[\frac{\sum_{k=1}^{n}\mu_{ik}}{\sum_{i=1}^{n}\sum_{k=1}^{n}(1-\nu_{ik})}, \frac{\sum_{k=1}^{n}(1-\nu_{ik})}{\sum_{i=1}^{n}\sum_{k=1}^{n}\mu_{ik}}\right], \quad i = 1,2,\ldots,n. \end{aligned} \tag{31}$$

Then, we can transform each interval into a corresponding IFV.

$$W_i = \left(\frac{\sum_{k=1}^{n}\mu_{ik}}{\sum_{i=1}^{n}\sum_{k=1}^{n}(1-\nu_{ik})}, 1 - \frac{\sum_{k=1}^{n}(1-\nu_{ik})}{\sum_{i=1}^{n}\sum_{k=1}^{n}\mu_{ik}}\right). \tag{32}$$

(6) Combine all the ratio of importance from the lowest level to the highest level by the operations of IFVs, to obtain the overall weightsWK_iusing formula (6).

3.5. Intuitionistic Fuzzy Quality Function Deployment to Obtain Engineering Characters Priorities. The impact of ENs on CS is becoming more and more important, but enterprise usually ignores the ENs in the NPDD. And the CRs are divided into ENs and FNs in a way that could effectively improve CS. Furthermore, in the NPDD processing, ENs directly affect product form design, and FNs impact greatly on functional parameter design, because product function matches customer usage needs and product form satisfies customer aesthetic needs. Thus, application QFD finds out the relations of CRs and ECs and establishes the goals of NPDD.

The construction of HOQ is the most important process in implementing QFD (Figure 4), and it consists of CRs input left wall (WHAT), ECs input ceiling (HOW), competition benchmarks input right wall (WHO), CRs and ECs relation matrix input ROOM, ECs correlation matrix input roof (WHY), ECs weights output floor (HOW MANY), technical benchmarks input basement (WHERE), and strategy of NPDD output goals (METHOD). CRs are obtained by LDA model and Apriori algorithm, and the importance of CRs is gained by IFKM and IFAHP (32). HOWs is acquired via expert interview; ROOM is gained through the relationship between VOC and VOT. By the relation symbols proposed by Cohen [59], △ denotes weak correlation, ○denotes moderate correlation, and ◎denotes strong correlation. Every EC has to be paired with at least one relation symbol of CRs; otherwise, it should be abandoned, and every corresponding symbol is converted into IFV (Table 3); WHO consists of market competition evaluation, target satisfaction, rate of satisfaction improvement, and absolute weights of CRs, where market competition evaluation relies on five-Likert scale (extreme importance/extreme unimportance) to evaluate CRs; target satisfaction depends on importance of CRs and developers to determine subjectively; rate of satisfaction improvement is obtain using (34); absolute weights of CRs use formula (35); WHY is gained by expert interview, + denotes positive correlation, and - denotes negative correlation; according to (36)–(38); HOW MANY can be obtained; WHERE is acquired by technical developer number through web crawler; METHOD is obtained by developers based on technical competition benchmarks.

$$W_i = WK_i SF_i, \tag{33}$$

$$Rsi_i = Ts_i / Cs_i, \tag{34}$$

$$AW_i = W_i Rsi_i, \tag{35}$$

where Rsi_i denotes rate of satisfaction improvement, Ts_i denotes competitor target satisfaction of CRs, Cs_i is development company satisfaction of CRs, W_i denotes adjusted importance, WK_i denotes ratio of importance, SF_i denotes adjustment coefficient, AW_i denotes absolute weight, and $i = 1, 2, \ldots n$is the number of CRs.

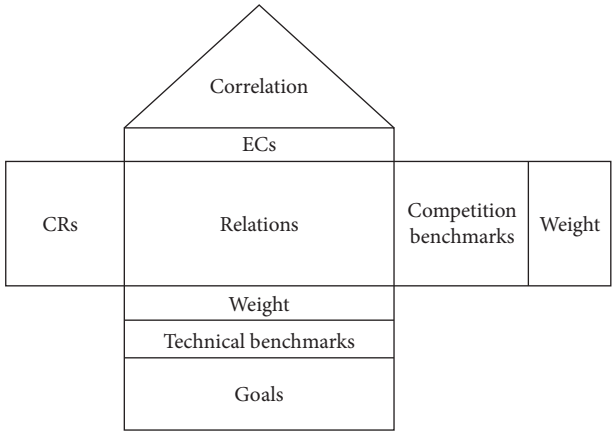


Figure 4: Model of QFD procedures.

Table 3: Symbols and IFVs in QFD.

Symbols	Definition	IFVs
Blank	No relationship	(0.1, 0.9)
△	Weak relationship	(0.3, 0.6)
○	Medium relationship	(0.6, 0.3)
◎	Strong relationship	(0.9, 0.1)

Then, based on HOQ model, we can obtain the standard weight value ECN_j of ECs as follows:

$$EC_j = \sum_{i=1}^{n} AW_i \otimes S_{ij}, \tag{36}$$

$$\rho\left(EC_j\right) = 0.5\left(1 + \pi_{EC_j}\right)\left(1 - \mu_{EC_j}\right), \tag{37}$$

$$ECN_j = \frac{\rho\left(EC\right)}{\sum_{j=1}^{m} \rho\left(EC\right)_j}, \tag{38}$$

where S_{ij} is CRs and ECs intuitionistic fuzzy relation matrix, EC_j is intuitionistic importance of jth WHAT, $\rho(EC_j)$ is the real number of ECs, ECN_j is the standard weight, and $i = 1, 2, \ldots, n$ denotes CRs number, and $j = 1, 2, \ldots m$ denotes ECs number.

4. Case Study

For this section, we use vertical-configured air conditioner (VAC) as a research example. VAC has become essential electrical device for homes and workplaces where they create a comfortable environment for living and work, increasing people's feeling of happiness in their daily life. When purchasing a VAC, customers not only consider the product's basic functions, but also concern about its external form and appearance. Thus, Kansei factors play a critical role in customer's purchasing decision. Since VAC becomes widely available throughout homes and workplaces, people have accumulated lots of experience with the VAC and developed good understanding of the various aspects of the products. Therefore, customers could provide valuable information on the VAC products, which could facilitate our research study. Thus, we could acquire the information on which CRs are more important for influencing customers' decisions of buying the VAC products and how and where companies innovate and prioritize their product design and provide assistance to designers and engineers in product development. In the next section, we will apply the proposed method to the VAC. It should be noted that the method is also applicable to other products.

4.1. Step 1. Mining CRs. We collected customer reviews by Octopus and saved in xls format from e-commerce platform (JD.com), with a result of total 5391 comments in Chinese. The original data was preprocessed, including removing invalid and duplicate comments and deleting punctuation. The Python language is applied throughout the retrieval process. Then, we called the Jieba library for Chinese segmentation and obtained 5310 valid comments. According to perplexity (14) to determine the topic number K, we do not estimate the hyperparameters α and β—instead the smoothing parameters are fixed at $50/K$ and $200/M$, respectively, where K is the topic number of LDA models, M is training samples containing 4248 (80%) documents in LDA, and the LDA model tested samples are 1062 (20%). The result of perplexity is shown in Figure 5; the topic number is 90 and the model has better generalization ability.

In consequence, the topic number is 90; we used Gensim library to construct LDA model which iterated 6000 through corpus and every 100 iteration to record the topic words. The results (Table 4) do not quite focus on product attributes by LDA model. So, we applied Apriori association algorithm to analyze 60 transactions; each transaction contains 90 topics, and each topic includes 10 topic words.

We obtained 7 strong association rules about sound in Table 5, and the rules are described in 7 ways meaning the same CR; we can adjust the order of words in these strong association rules to gain easy-to-understand sentence that is low noise.

Similarly, we obtained all the CRs in Table 6 when minimum support set is 0.005 and minimum confidence set is 0.5, and the CRs of VAC contain 3 ENs and 6 FNs from topic words. But there are massive invalid rules. There may be three reasons. Firstly, since the frequency of words about dehumidification and air purification is quite low in corpus, the LDA model cannot cover these words in priority of 10 topic words. Secondly, the Apriori algorithm prunes frequency item sets (transactions) which is noun and does not consider the relationship with different parts of speech. Thirdly, Chinese segmentation may be inaccurate, because some Chinese words have multiple parts of speech and different order has different meaning.

4.2. Step 2. Obtaining CRs Category and Adjustment Coefficient. According to Table 2, we made Kano questionnaire including 9 CRs attributes. 200 interviewees including housewives, white collar workers, and college students (100 males and 100 females), from 20 to 55 years old, were invited. Every participant was required to answer two-dimensional function and dysfunction quality questions and fill with their scores of satisfaction and dissatisfaction. 195 questionnaires were returned, a response rate of 97.5 percent, and 182 questionnaires were reasonable, a

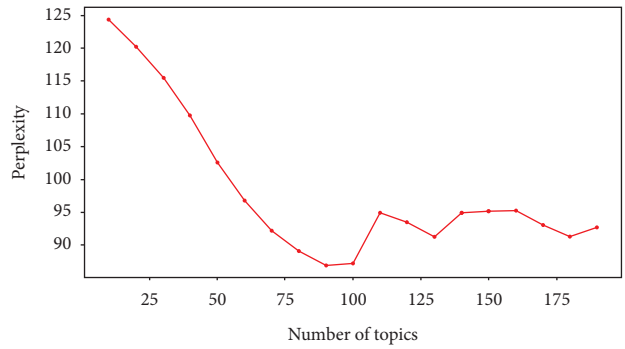


Figure 5: Perplexity on different topics.

Table 4: Inductive topics.

No.	Summary	Main topics
1	Function	Sound, energy, cooling, heating, mobile, and wind
2	Form	Beautiful appearance
3	Service	Installation, delivery, and after sale
4	Experience	Shopping and working
5	Price	Double 11, 618
6	Brand	Midea, Gree, and Aux

Table 5: Example of strong rules.

Strong association rules	Chinese/English sentence	Support	Confidence	Lift
[听不到] => [声音] [cannot hear] => [sound]	听不到声音/no sound	0.0081	0.7333	12.4138
[一点] => [声音] [a little] => [sound]	一点声音/a little sound	0.0061	0.5593	9.4681
[几乎, 听不到] => [声音] [almost, cannot hear] => [sound]	几乎听不到声音/almost no sound	0.0081	0.7333	12.4138
[一点, 风] => [声音] [a little, wind] => [sound]	一点风的声音/a little wind	0.0057	0.7209	12.2038
[一点, 听不到] => [声音] [a little, cannot hear] => [sound]	一点听不到声音/no sound at all	0.0061	0.7333	12.4138
[制暖, 听不到] => [声音] [heating, cannot hear] => [sound]	听不到制暖声音/no heating sound	0.0069	0.7115	12.0449
[风, 几乎, 听不到] => [声音] [wind, almost, cannot hear] => [sound]	几乎听不到风的声音/hardly hear the wind	0.0076	0.7321	12.3936

Table 6: Representative CRs.

No.	CRs	Support
C_1	Being simple and liberal	0.0513
C_2	High-tech and intelligence	0.0391
C_3	Being fashionable and novel	0.0291
C_4	Low noise	0.0487
C_5	Energy conservation	0.0559
C_6	Fast cooling and heating	0.0828
C_7	Strong mode and soft wind	0.0270
C_8	Safety protection	0.0457
C_9	Mobile APP control	0.0446

reasonable response rate of 91 percent. Based on Definition 2, the maximum of membership (satisfaction) is 0.9 and minimum of membership (dissatisfaction) is 0; thus, the CRs membership function can be obtained (Figure 2) with every interviewee's membership to standard answers via functions (16)–(20) and CR value SF_i (Table 7) via functions (21)–(23). Besides, according to Chen [60], the adjustment coefficient of attractive quality, one-dimensional quality, must-be quality, and indifferent quality set is 4, 2, 1, and 0 in KM and FKM, respectively.

4.3. Step 3. Obtaining CRs Weights. According to Table 6, we constructed CRs hierarchy process (Figure 6). 10 experts with more than 5-year development experience and 5-year design experience were selected to act as interviewees; thus relation matrices were obtained via pairwise comparison, and calculated intuitionistic preference matrices are of acceptable consistency via functions (24)–(27). Then, nonconsistency relation matrices were calculated via (28)–(30). The consistency threshold value $\tau$was selected as 0.1 (Table 8) and controlling parameterσ was set to 0.8. The 1st, 6th, and

Table 7: The comparative results of the case study.

CRs	Classical KM					FKM					CFKM		IFKM	
	M	O	A	I	Cat.	M	O	A	I	Cat.	Eval.	Pri.	Eval.	Pri.
C_1	0	49	80	90	I (0)	2	57	116	102	A (4)	0.2173	1	0.1973	1
C_2	6	15	54	111	I (0)	7	25	50	138	I (0)	0.0912	3	0.1393	3
C_3	0	11	77	97	I (0)	54	4	103	116	I (0)	0.1425	2	0.1500	2
C_4	1	12	79	94	I (0)	77	15	43	46	M (1)	0.2736	3	0.2696	1
C_5	2	12	64	101	I (0)	80	52	51	31	M (1)	0.3312	1	0.2481	2
C_6	1	4	48	124	I (0)	77	24	58	49	M (1)	0.3130	2	0.2088	3
C_7	0	2	42	144	I (0)	57	49	36	92	I (0)	0.1376	5	0.1371	5
C_8	13	2	37	138	I (0)	41	55	29	71	I (0)	0.2084	4	0.1652	4
C_9	0	11	62	113	I (0)	21	91	37	58	O (2)	0.1359	6	0.1296	6

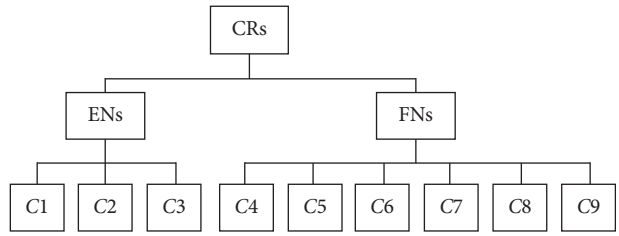


Figure 6: Hierarchical structure of the case study.

Table 8: Consistency test of IFAHP.

Experts	First		Second	
	ENs threshold	FNs threshold	ENs threshold	FNs threshold
1	0.0000	0.0767	—	—
2	0.3087	0.0915	0.0628	—
3	0.3087	0.0563	0.0628	—
4	0.2923	0.0422	0.0541	—
5	0.2632	0.0716	0.0508	—
6	0.0000	0.0965	—	—
7	0.0429	0.0966	—	—
8	0.1032	0.0278	0.0159	—
9	0.1077	0.0216	0.0240	—
10	0.1448	0.0433	0.0235	—

7th experts make acceptable consistency in ENs, and all experts make acceptable consistency in FNs. At last, we fused all the weights from the lowest level to the highest level via functions (31-32) (Table 9).

4.4. Step 4. Obtaining ECs Weights and Goals. Most of VACs consist of refrigeration cycle system, air circulation system, electrical control system, and shell system (Table 10). We calculated the adjusted weights obtained by formula (33) and collected relation matrixes of HOWs and WHATs from 10 experts and converted them into IFVs via Table 3. Then, the target satisfaction was set through interview in 3000–4000 RMB VAC; the rate of satisfaction improvement is shown via (34); the absolute weights are gained by formula (35). Thus, we obtained the standard weights of ECN_j via functions (36)-(38) and HOQ of VAC (Figure 7). At last, according to technical benchmarks, goals were established in NPDD.

5. Comparative Analysis

Comparing the result weights of CRs in AHP, FAHP, and IFAHP (Table 11), the weights are similar; FNs are higher than ENs, and the results of FAHP are quite different between AHP and IFAHP. However, in the IFAHP, if we do not consider the nonmembership degree of WK_j when ranking the priority, the result of the IFAHP method will be ENs > FNs as well because $0.0469 > 0.0438 > 0.0338 > 0.0094 \ldots > 0.0052$, which is consistent with the result of the FAHP method. The different results imply that the IFVs can represent the preferences of the pairwise comparison more comprehensively.

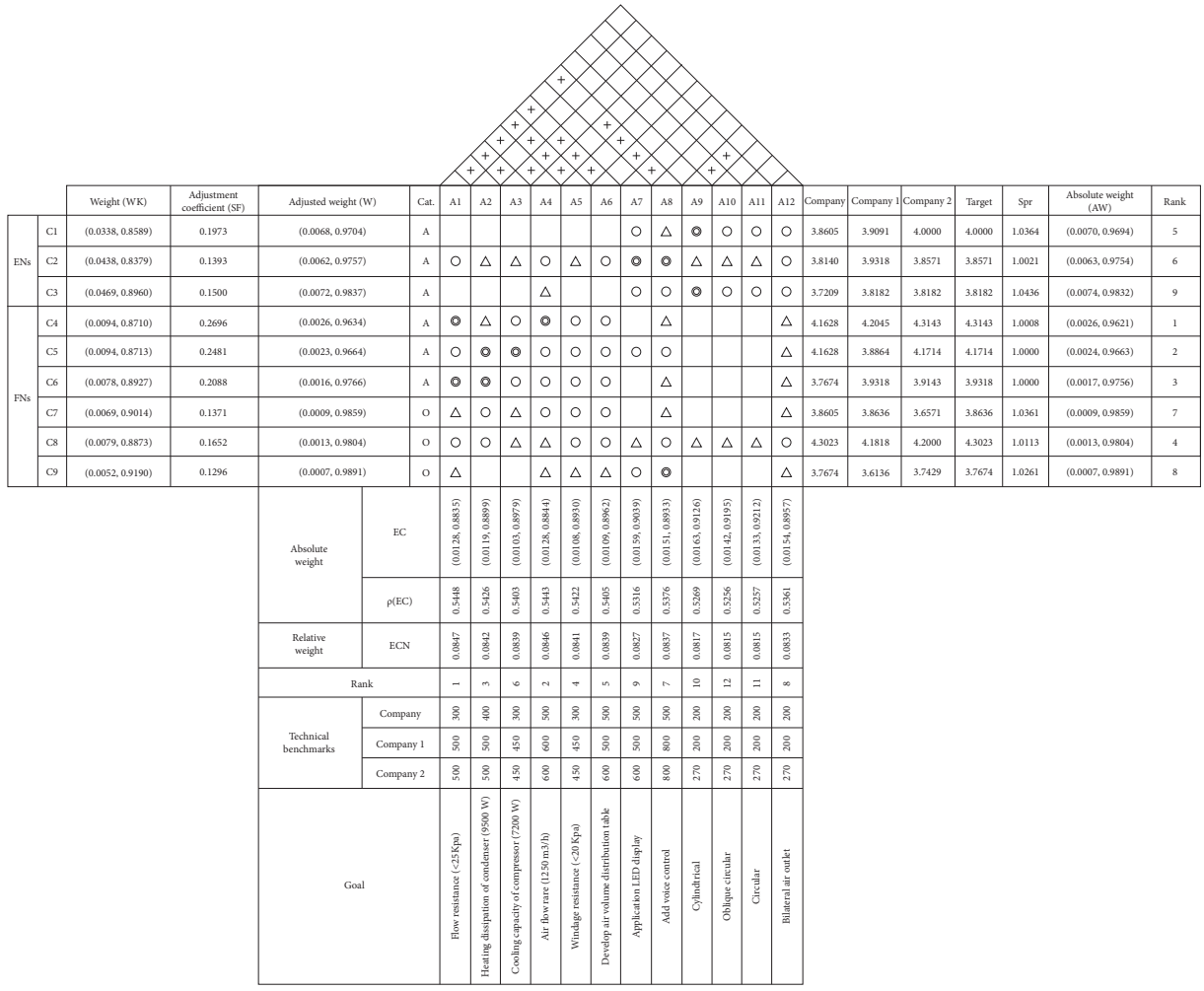
Comparing adjustment coefficient in KM, FKM, CFKM, and IFKM (Table 7), the classical Kano model does not work well in this case study, and FKM is worse than CFKM and IFKM in rationality of the classification results of CRs. The CFKM and IFKM have similar results, and the

Table 9: Weights of IFAHP.

Criterion	W_i	Subcriterion	W_{ij}	WK_i
ENs	(0.3529, 0.6494)	C_1	(0.0957, 0.5974)	(0.0338, 0.8589)
		C_2	(0.1241, 0.5375)	(0.0438, 0.8379)
		C_3	(0.1328, 0.7034)	(0.0469, 0.8960)
FNs	(0.5158, 0.4643)	C_4	(0.0183, 0.7598)	(0.0094, 0.8710)
		C_5	(0.0183, 0.7598)	(0.0094, 0.8713)
		C_6	(0.0151, 0.7998)	(0.0078, 0.8927)
		C_7	(0.0133, 0.8159)	(0.0069, 0.9014)
		C_8	(0.0153, 0.7895)	(0.0079, 0.8873)
		C_9	(0.0100, 0.8489)	(0.0052, 0.9190)

Table 10: ECs of VAC.

Refrigeration cycle system	Air circulation system	Electrical control system	Shell system
Refrigeration cycle power parts (A_1)	Forming airflow parts (A_4)	Information display (A_7)	Body (A_9)
Thermal exchange parts (A_2)	Air charge parts (A_5)	Intelligent control parts (A_8)	Top (A_{10})
Refrigerant flow and pressure parts (A_3)	Air flow direction controller parts (A_6)		Base (A_{11})
			Air outlet grille (A_{12})



		Weight (WK)	Adjustment coefficient (SF)	Adjusted weight (W)	Cat.	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	Company	Company 1	Company 2	Target	Spr	Absolute weight (AW)	Rank
ENs	C1	(0.0338, 0.8589)	0.1973	(0.0068, 0.9704)	A							○	△	◎	○	○	○	3.8605	3.9091	4.0000	4.0000	1.0364	(0.0070, 0.9694)	5
	C2	(0.0438, 0.8379)	0.1393	(0.0062, 0.9757)	A	○	△	△	○	△	○	◎	◎	△	△	△	○	3.8140	3.9318	3.8571	3.8571	1.0021	(0.0063, 0.9754)	6
	C3	(0.0469, 0.8960)	0.1500	(0.0072, 0.9837)	A				△			○	○	◎	○	○	○	3.7209	3.8182	3.8182	3.8182	1.0436	(0.0074, 0.9832)	9
FNs	C4	(0.0094, 0.8710)	0.2696	(0.0026, 0.9634)	A	◎	△	○	◎	○	○		△				△	4.1628	4.2045	4.3143	4.3143	1.0008	(0.0026, 0.9621)	1
	C5	(0.0094, 0.8713)	0.2481	(0.0023, 0.9664)	A	○	◎	◎	○	○	○	○	○				△	4.1628	3.8864	4.1714	4.1714	1.0000	(0.0024, 0.9663)	2
	C6	(0.0078, 0.8927)	0.2088	(0.0016, 0.9766)	A	◎	◎	○	○	○	○		△				△	3.7674	3.9318	3.9143	3.9318	1.0000	(0.0017, 0.9756)	3
	C7	(0.0069, 0.9014)	0.1371	(0.0009, 0.9859)	O	△	○	△	○	○	○		△				△	3.8605	3.8636	3.6571	3.8636	1.0361	(0.0009, 0.9859)	7
	C8	(0.0079, 0.8873)	0.1652	(0.0013, 0.9804)	O	○	○	△	△	○	○	△	○	△	△	△	○	4.3023	4.1818	4.2000	4.3023	1.0113	(0.0013, 0.9804)	4
	C9	(0.0052, 0.9190)	0.1296	(0.0007, 0.9891)	O	△			△	△	△	○	◎				△	3.7674	3.6136	3.7429	3.7674	1.0261	(0.0007, 0.9891)	8
				Absolute weight	EC	(0.0128, 0.8835)	(0.0119, 0.8899)	(0.0103, 0.8979)	(0.0128, 0.8844)	(0.0108, 0.8930)	(0.0109, 0.8962)	(0.0159, 0.9039)	(0.0151, 0.8933)	(0.0163, 0.9126)	(0.0142, 0.9195)	(0.0133, 0.9212)	(0.0154, 0.8957)							
					ρ(EC)	0.5448	0.5426	0.5403	0.5443	0.5422	0.5405	0.5316	0.5376	0.5269	0.5256	0.5257	0.5361							
				Relative weight	ECN	0.0847	0.0842	0.0839	0.0846	0.0841	0.0839	0.0827	0.0837	0.0817	0.0815	0.0815	0.0833							
				Rank		1	3	6	2	4	5	9	7	10	12	11	8							
				Technical benchmarks	Company	300	400	300	500	300	500	500	500	200	200	200	200							
					Company 1	500	500	450	600	450	500	500	800	200	200	200	200							
					Company 2	500	500	450	600	450	600	600	800	270	270	270	270							
				Goal		Flow resistance (<25Kpa)	Heating dissipation of condenser (9500 W)	Cooling capacity of compressor (7200 W)	Air flow rare (1250 m3/h)	Windage resistance (<20 Kpa)	Develop air volume distribution table	Application LED display	Add voice control	Cylindrical	Oblique circular	Circular	Bilateral air outlet							

Figure 7: HOQ case study.

Table 11: Weight comparison.

	AHP	Pri.	FAHP	Pri.	IFAHP	Pri.
C_1	0.0502	7	0.1563	2	(0.0338, 0.8589)	7
C_2	0.1250	4	0.1686	1	(0.0438, 0.8379)	8
C_3	0.0247	9	0.1363	3	(0.0469, 0.8960)	9
C_4	0.1863	2	0.0975	4	(0.0094, 0.8710)	1
C_5	0.2328	1	0.0969	5	(0.0094, 0.8713)	2
C_6	0.1127	5	0.0894	6	(0.0078, 0.8927)	4
C_7	0.0728	6	0.0861	8	(0.0069, 0.9014)	5
C_8	0.1498	3	0.0891	7	(0.0079, 0.8873)	3
C_9	0.0455	8	0.0798	9	(0.0052, 0.9110)	6

Table 12: Absolute weight comparison.

	Rsi	AHP *KM *Rsi	FAHP *FKM *Rsi	IFAHP *IFKM *Rsi	Pri.
C_1	1.0361	0.0000	0.6478	(0.0070, 0.9694)	5
C_2	1.0113	0.0000	0.0000	(0.0063, 0.9754)	6
C_3	1.0261	0.0000	0.0000	(0.0074, 0.9832)	9
C_4	1.0364	0.0000	0.1010	(0.0026, 0.9621)	1
C_5	1.0021	0.0000	0.0971	(0.0024, 0.9663)	2
C_6	1.0436	0.0000	0.0933	(0.0017, 0.9756)	3
C_7	1.0008	0.0000	0.0000	(0.0009, 0.9859)	7
C_8	1.0000	0.0000	0.0000	(0.0013, 0.9804)	4
C_9	1.0000	0.0000	0.1596	(0.0007, 0.9891)	8

difference between the two methods is mainly in the attractive and one-dimensional qualities which are located in the cells (1, 4) and (1, 5), respectively, because the IFS is more comprehensive and flexible than FS in describing uncertainty.

According to Chen [60], the adjustment coefficient of attractive quality, one-dimensional quality, must-be quality, and indifferent quality set is 4, 2, 1, and 0 in KM and FKM, respectively. The absolute weight of CRs is done by comparison in Table 12; there are limitations which ignore the difference of the satisfaction coefficient between the consumer demand items and the same quality attributes. Besides, the absolute weight cannot obtain good performance in HOQ which used KM and FKM as adjustment coefficient.

6. Discussion

In the classical Kano model, designers should follow the quality ranking for NPDD, which is must-be > attractive > one-dimensional > indifferent. One strategy is to integrate all the must-be, one-dimensional, and attractive quality into the product. However, the resource is quite limited; thus it is hard for them to decide which four of the six attractive qualities should be included with the addition of other qualities. In consequence, we proposed that IFKM can obtain CR preference, which includes affirmation, negation, hesitation, and priority, which effectively express VOC. In this paper, we chose VAC as an example. The ENs' priority of KM category (Table 7) is simple and liberal > fashionable and novel > high-tech and intelligence, and all of them belong to attractive quality. The FNs' priority of category is low noise > energy conservation > fast cooling and heating, which are attractive quality; the order of one-dimensional quality is safety protection > strong mode and soft wind > mobile APP control.

In order to obtain and compare the weight of the ENs and FNs, we conducted IFAHP process that analyzed overall weights about CRs and multiplies the adjustment coefficient via IFKM. Thus, we got the adjusted weights in the field of VAC's NPDD. The absolute weights should contain adjusted weights and ratios of satisfaction, which based on competition benchmarks (Figure 7). Thus, the result of absolute weights pirority is low noise > energy saving > fast cooling and heating > safety protection > simple and liberal > high-tech and intelligence > strong mode and soft wind > mobile APP control > fashion and novel. According to IFKM category and resource constraint, the strategy of this company should choose low noise, energy conservation, fast cooling and heating, and being simple and liberal as the attractive quality, finally with the addition of other qualities. Because the attractive quality massively improves CS (Figure 1) and the one-dimensional quality has linear relation to CS.

The refrigeration cycle power parts and forming airflow parts have the highest priority via IFQFD, because of low noise and energy conservation, which are the main sources of sound and power consumption, being the top two of CRs. And the thermal exchange part is the third importance ECs, because it has strong relation to energy conservation and fast cooling and heating. In additional, the ranking of the ECs of refrigeration cycle system, air circulation system, and electrical control system is higher than that of shell system; thus it is consistent with the FNs being larger than the ENs. Thus, this research has high credibility in CS analysis and can provide effective support for decision-making. Finally, the company established development goals based on technical benchmarks.

7. Conclusions

In this study, we proposed a new method for NPDD based on text mining, CS, and IFS. We applied text mining technique to retrieve CRs from big data and combined this with advantage of IFS which can better process uncertain and hesitant information, to integrate Kano model, AHP, and QFD methods. Thus, we can establish goals about the customer-oriented NPDD. In additional, text mining can directly obtain expected information from unstructured data and provide a scientific basis for NPDD. The IFSs have more flexible and comprehensive features in describing objective phenomena than classical FS. And the evaluation of various preference types is unified into IFS based on the heterogeneous characteristics of participants that can minimize the loss of decision information.

Since certain limitations still exist, this research needs to be continually improved. First, it is necessary to develop a supervised algorithm for the extraction of CRs with low frequency in text mining, because lower frequency of CRs like dehumidification and air purification will be affected by geographical and environment factors. Second, the impact of multiple Kansei images of product on customers satisfaction remains to be studied, which means that a product may contain factors like being fashionable, elegant, simple, beautiful, etc. and generally have nonlinear relationship to product form. Third, the threshold (ϕ) of the *SF* is determined by practitioners based on experience. There is a need to develop an approach to determine the suitable values for decision-making.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Shanghai Summit Discipline in Design of China and Shanghai Philosophy and Social Science Planning Project (Grant nos. DC17014 and 2019ZJX002).

References

[1] E. W. Anderson and V. Mittal, "Strengthening the satisfaction-profit chain," *Journal of Service Research*, vol. 3, no. 2, pp. 107–120, 2000.

[2] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177, Seattle, WA, USA, August 2004.

[3] A. J. Newman and D. Patel, "The marketing directions of two fashion retailers," *European Journal of Marketing*, vol. 38, no. 7, pp. 770–789, 2004.

[4] M. Nagamachi, "Kansei engineering as a powerful consumer-oriented technology for product development," *Applied Ergonomics*, vol. 33, no. 3, pp. 289–294, 2002.

[5] M. Nagamachi, "Kansei Engineering: a new ergonomic consumer-oriented technology for product development," *International Journal of Industrial Ergonomics*, vol. 15, no. 1, pp. 3–11, 1995.

[6] S.-M. Yang, M. Nagamachi, and S.-Y. Lee, "Rule-based inference model for the Kansei engineering system," *International Journal of Industrial Ergonomics*, vol. 24, no. 5, pp. 459–471, 1999.

[7] H. M. Khalid and M. G. Helander, "Customer emotional needs in product design," *Concurrent Engineering*, vol. 14, no. 3, pp. 197–206, 2006.

[8] S. Yang, M. Nagamachi, and S. Lee, "Rule-based inference model for the Kansei engineering system," *International Journal of Industrial Ergonomics*, vol. 24, no. 5, pp. 459–471, 2011.

[9] C.-C. Yang, "Constructing a hybrid Kansei engineering system based on multiple affective responses: application to product form design," *Computers and Industrial Engineering*, vol. 60, no. 4, pp. 760–768, 2011.

[10] M.-D. Shieh and Y.-E. Yeh, "Developing a design support system for the exterior form of running shoes using partial least squares and neural networks," *Computers and Industrial Engineering*, vol. 65, no. 4, pp. 704–718, 2013.

[11] C.-H. Wang, "Integrating Kansei engineering with conjoint analysis to fulfil market segmentation and product customisation for digital cameras," *International Journal of Production Research*, vol. 53, no. 8, pp. 2427–2438, 2015.

[12] C. H. Papadimitriou, P. Raghavan, H. Tamaki et al., "Latent semantic indexing: a probabilistic analysis," *Journal of Computer and System Sciences*, vol. 61, no. 2, pp. 217–235, 1998.

[13] T. Hofmann, "Probabilistic latent semantic indexing," in *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 50–57, Berkeley, CA, USA, August 1999.

[14] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[15] D. Walter, Y. Ophir, and K. H. Jamieson, "Russian twitter accounts and the partisan polarization of vaccine discourse, 2015–2017," *American Journal of Public Health*, vol. 110, no. 5, pp. 718–724, 2020.

[16] D. Maier, A. Waldherr, P. Miltner et al., "Applying LDA topic modeling in communication research: toward a valid and reliable methodology," *Communication Methods and Measures*, vol. 12, no. 2-3, pp. 93–118, 2018.

[17] X. Wei X and W. Croft, "LDA-based document models for Ad-hoc retrieval," in *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 178–185, Seattle, WA, USA, August 2006.

[18] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proceedings of the 20th International Conference on Very Large Data Bases*, vol. 1215, pp. 487–499, Santiago, SM, Chile, September 1994.

[19] Y. Guo, M. Wang, and X. Li, "Application of an improved Apriori algorithm in a mobile e-commerce recommendation system," *Industrial Management & Data Systems*, vol. 117, no. 2, pp. 287–303, 2017.

[20] S. Panjaitan, M. Amin, S. Lindawati et al., "Implementation of apriori algorithm for analysis of consumer purchase patterns,"

Journal of Physics: Conference Series, IOP Publishing, vol. 1255, no. 1, Article ID 012057, 2019.

[21] Y. Liu, "Study on application of apriori algorithm in data mining," in *Proceedings of the 2010 Second International Conference on Computer Modeling and Simulation*, vol. 3, pp. 111–114, Sanya, HN, China, January 2010.

[22] N. Kano, N. Seraku, and F. Takahashi, "Attractive quality and must-be quality," *Journal of the Japanese Society for Quality Control*, vol. 14, no. 2, pp. 39–48, 1984.

[23] K. Matzler and H. H. Hinterhuber, "How to make product development projects more successful by integrating Kano's model of customer satisfaction into quality function deployment," *Technovation*, vol. 18, no. 1, pp. 25–38, 1998.

[24] K. C. Tan and X. X. Shen, "Integrating Kano's model in the planning matrix of quality function deployment," *Total Quality Management*, vol. 11, no. 8, pp. 1141–1151, 2000.

[25] Q. Xu, R. J. Jiao, X. Yang, M. Helander, H. M. Khalid, and A. Opperud, "An analytical Kano model for customer need analysis," *Design Studies*, vol. 30, no. 1, pp. 87–110, 2009.

[26] P. Földesi, L. T. Kóczy, and J. Botzheim, "Fuzzy extension for Kano's model using bacterial evolutionary algorithm," in *Proceedings of the 2007 International Symposium on Computational Intelligence and Intelligent Informatics*, pp. 147–151, Agadir, SM, Morocco, March 2007.

[27] Y.-C. Lee and S.-Y. Huang, "A new fuzzy concept approach for Kano's model," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4479–4484, 2009.

[28] M. Wu and L. Wang, "A continuous fuzzy Kano's model for customer requirements analysis in product development," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 226, no. 3, pp. 535–546, 2012.

[29] T. L. Satty, *The Analysis Hierarchy Process*, McGraw-Hill, New York, NY, USA, 1980.

[30] P. J. M. van. Laarhoven and W. Pedrycz, "A fuzzy extension of Saaty's priority theory," *Fuzzy Sets and Systems*, vol. 11, no. 1, pp. 229–241, 1983.

[31] M. Kwiesielewicz, "A note on the fuzzy extension of Saaty's priority theory," *Fuzzy Sets and Systems*, vol. 95, no. 2, pp. 161–172, 1998.

[32] C. G. E. Boender, J. G. de Graan, and F. A. Lootsma, "Multicriteria decision analysis with fuzzy pairwise comparisons," *Fuzzy Sets and Systems*, vol. 29, no. 2, pp. 133–143, 1989.

[33] Z. Xu and H. Liao, "Intuitionistic fuzzy analytic hierarchy process," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 4, pp. 749–761, 2014.

[34] S.-W. Hsiao and Y.-C. Ko, "A study on bicycle appearance preference by using FCE and FAHP," *International Journal of Industrial Ergonomics*, vol. 43, no. 4, pp. 264–273, 2013.

[35] K. Kamvysi, K. Gotzamani, A. Andronikidis, and A. C. Georgiou, "Capturing and prioritizing students' requirements for course design by embedding Fuzzy-AHP and linear programming in QFD," *European Journal of Operational Research*, vol. 237, no. 3, pp. 1083–1094, 2014.

[36] K. Chakraborty, S. Mondal, and K. Mukherjee, "Analysis of product design characteristics for remanufacturing using Fuzzy AHP and axiomatic design," *Journal of Engineering Design*, vol. 28, no. 5, pp. 338–368, 2017.

[37] Y. Akao, "Development history of quality function deployment," *The Customer Driven Approach to Quality Planning and Deployment*, vol. 339, p. 90, 1994.

[38] J. R. Hauser, "The house of quality," *Harvard Business Review*, vol. 66, no. 3, pp. 63–73, 1988.

[39] X. Lai, M. Xie, and T. C. Tan, "Optimizing product design using the Kano model and QFD," in *Proceedings of the 2004 IEEE International Engineering Management Conference*, vol. 3, pp. 1085–1089, Singapore, October 2004.

[40] N. Tu, T. Zhang, Q.-Y. He, H.-F. Zhang, and Y.-H. Li, "Applying combined AHP-QFD method in new product development: a case study in developing new sports earphone," in *Proceedings of the 2011 International Conference on Management Science & Industrial Engineering*, Harbin, HL, China, January 2011.

[41] G. Bayraktaroğlu and O. Özge, "Integrating the Kano model, AHP and Planning Matrix: QFD application in library services," *Library Management*, vol. 29, no. 45, pp. 327–351, 2008.

[42] L. P. Khoo and N. C. Ho, "Framework of a fuzzy quality function deployment system," *International Journal of Production Research*, vol. 34, no. 2, pp. 299–311, 1996.

[43] Q. Wu, "Fuzzy measurable house of quality and quality function deployment for fuzzy regression estimation problem," *Expert Systems with Applications*, vol. 38, no. 12, pp. 14398–14406, 2011.

[44] I. Mahdavi, N. Mahdavi-Amiri, A. Heidarzade, and R. Nourifar, "Designing a model of fuzzy TOPSIS in multiple criteria decision making," *Applied Mathematics and Computation*, vol. 206, no. 2, pp. 607–617, 2008.

[45] T. Aouam, S. I. Chang, and E. S. Lee, "Fuzzy MADM: an outranking method," *European Journal of Operational Research*, vol. 145, no. 2, pp. 317–328, 2003.

[46] S. J. Chen and C. L. Hwang, *Fuzzy Multiple Attribute Decision Making*, Springer, Berlin, Heidelberg, Germany, 1992.

[47] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.

[48] K. T. Atanassov, *Intuitionistic Fuzzy Sets: Theory and Applications*, Springer, Berlin, Heidelberg, Germany, 1999.

[49] K. T. Atanassov, *On Intuitionistic Fuzzy Sets Theory*, Springer, Germany, 2012.

[50] D. F. Li, J. X. Nan, and M. J. Zhang, "A ranking method of triangular intuitionistic fuzzy numbers and application to decision making," *International Journal of Computational Intelligence Systems*, vol. 3, no. 5, pp. 522–530, 2010.

[51] D.-F. Li, "A ratio ranking method of triangular intuitionistic fuzzy numbers and its application to MADM problems," *Computers and Mathematics with Applications*, vol. 60, no. 6, pp. 1557–1570, 2010.

[52] D.-F. Li, "Multiattribute decision making models and methods using intuitionistic fuzzy sets," *Journal of Computer and System Sciences*, vol. 70, no. 1, pp. 73–85, 2005.

[53] E. Szmidt and J. Kacprzyk, "Amount of information and its reliability in the ranking of atanassov's intuitionistic fuzzy alternatives," *Studies in Computational Intelligence*, vol. 222, pp. 7–19, 2009.

[54] Z. S. Xu, "Intuitionistic fuzzy aggregation operators," *Fuzzy System*, vol. 15, no. 6, pp. 1179–1187, 2007.

[55] Z. Xu, "Intuitionistic preference relations and their application in group decision making☆," *Information Sciences*, vol. 177, no. 11, pp. 2363–2379, 2007.

[56] M. Rosen-Zvi, T. L. Griffiths, M. Steyvers et al., "The author-topic model for authors and documents," in *Proceedings of the 20th Conference in Uncertainty in Artificial Intelligence*, Banff, AB, Canada, July, 2004.

[57] C. Berger, R. E. Blauth, and D. Boger, "Kano's methods for understanding customer-defined quality," *Computer Science*, vol. 2, no. 4, pp. 3–35, 1995.

[58] L.-H. Chen and W.-C. Ko, "A fuzzy nonlinear model for quality function deployment considering Kano's concept," *Mathematical and Computer Modelling*, vol. 48, no. 3-4, pp. 581–593, 2008.

[59] L. Cohen, *Quality Function Deployment: How to Make QFD Work for You*, Addison-Wesley, Boston, MA, USA, 1995.

[60] C.-C. Chen and M.-C. Chuang, "Integrating the Kano model into a robust design approach to enhance customer satisfaction with product design," *International Journal of Production Economics*, vol. 114, no. 2, pp. 667–681, 2008.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 6654461, 12 pages
https://doi.org/10.1155/2021/6654461

Research Article

Dependence-Cognizant Locking Improvement for the Main Memory Database Systems

Ouya Pei,[1,2] **Zhanhuai Li**,[1,2] **Hongtao Du**,[1,2] **Wenjie Liu**,[1,2] **and Jintao Gao**[1,2]

[1]*School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China*
[2]*Key Laboratory of Big Data Storage and Management, Northwestern Polytechnical University, Xi'an 710072, China*

Correspondence should be addressed to Ouya Pei; peiouya2013@mail.nwpu.edu.cn and Hongtao Du; duhongtao@nwpu.edu.cn

Received 29 October 2020; Revised 9 December 2020; Accepted 15 January 2021; Published 22 February 2021

Academic Editor: Shianghau Wu

The traditional lock manager (LM) seriously limits the transaction throughput of the main memory database systems (MMDB). In this paper, we introduce dependence-cognizant locking (DCLP), an efficient improvement to the traditional LM, which dramatically reduces the locking space while offering efficiency. With DCLP, one transaction and its direct successors are collocated in its context. Whenever a transaction is committed, it wakes up its direct successors immediately avoiding the expensive operations, such as lock detection and latch contention. We also propose virtual transaction which has better time and space complexity by compressing continuous read-only transactions/operations. We implement DCLP in Calvin and carry out experiments in both multicore and shared-nothing distributed databases. Experiments demonstrate that, in contrast with existing algorithms, DCLP can achieve better performance in many workloads, especially high-contention workloads.

1. Introduction

Although have arisen since the 1980s, but until recent decade, with the emergence of larger capacity and cheaper memory in which most or even all of data can be resident, MMDB began to achieve some success in commercial fields.

However, for MMDB, many of their workloads only consist of short transactions that access only a few records each. Therefore, the traditional locking mechanism has become the primary performance bottleneck. Based on the stand-alone single-core architecture, Harizopoulos et al. reported that 16% to 25% of the transaction overhead is spent on the lock manager [1]. As for the multicore processor architecture, some research studies similarly show that concurrently accessing the lock manager needs larger amounts of overhead [2–4].

As described in [5], the most common way to implement LM is as a hash table that maps each record's primary key to a linked list of lock requests. For lock acquiring, LM (1) probes the internal hash table to find the corresponding entry and then latches it, (2) appends a new request object to the entry list and blocks if it is incompatible with current holders, and (3) unlatches the entry and returns the (granted) request to the transaction. For lock releasing, it has to go through the steps of the lock request, and the only difference is removing the request from the requests' list.

These operations are still expensive. Firstly, row locking consumes too many computations and memory resources. Secondly, too many latches are applied to ensure correctness. Thirdly, list operations impose high overhead as the number of active transactions increases.

For reducing these overheads, some studies have reduced the number of lock acquisitions by redesigning the processing logic of LM [4, 6]. Others focus on optimizing the design and implementation of LM for improving LM scalability by reducing the latch and critical sections [7, 8], but these research studies still use the basic two-phase locking design.

With the vigorous development of deterministic transaction execution strategy, very lightweight locking (VLL) and selective contention analysis (SCA) were proposed by Kun Ren et al. [5, 9].

VLL's core idea is that use two simple semaphores containing the number of outstanding requests for that lock

(C_x for write requests and C_s for read requests). In this way, linked list operations are completely eliminated. However, tracking less information about contention results in poor throughput under high contention. SCA simulates the standard lock manager's ability to discover unblocked transactions by rescan TxnQueue that stores all active transactions in the order of arrival.

However, there are still some drawbacks which are as follows:

Unable to timely detect unblocked transactions. SCA only just makes VLL have the ability to detect unblocked transactions.

Easily blocked. For efficiency, low threshold is set for the number of blocked transactions. Thus, encountering long chain dependent or high contention workloads, the low threshold is easily exceeded so that subsequent transactions make no progress even if they are conflict-free.

1.1. Brief Introduction about DCLP. On the basis of the above analysis, it is necessary to revisit LM's design in MMDB. In this study, DCLP is introduced to reduce the expensive overhead of LM and ensure the ability to detect unblocked transactions without delay. Its core idea is tracking contention through co-locating one transaction with its direct successors. These following strategies are adopted to ensure efficiency:

Fragmented storage of lock requests means the lock request lists are partially removed, and the direct conflict information between transactions is collocated with the corresponding direct predecessors

Virtual transaction: using virtual transaction compresses continuous read requests by reference counting, and list operations are almost eliminated

With these two strategies, DCLP almost eliminates latch acquisitions and list operations and has the ability to timely detect which transactions should inherit released locks. Comparative experiments show that compared with VLL\SCA, DCLP achieves better (maximum 50% higher in some workloads) performance in high-contention workloads, long-dependent workloads, and workloads with high multipartition transactions, while it has approximately equal or slightly better performance in low-contention workloads.

1.2. Contributions and Paper Organization. This paper makes the following contributions:

We propose fragmented storage of lock requests. It completely eliminates the expensive overhead of list operations in terms of write-only workloads.

We propose the virtual transaction strategy. It effectively reduces the overhead of list operations by compressing continuous read requests.

We design a large number of experiments to evaluate DCLP overall. Four locking mechanisms are implemented in Calvin, and two benchmarks are used. Experiments show that DCLP is effective and efficient.

2. Principles and Algorithms

It is known that the major advantage of hash or hash-based data structures is they are efficient to search efficiency and more apparent when processing massive data. Especially, if the hash buckets are relatively large enough, a hash-based data structure that is a low collision or even collision-free can be easily designed and provides O(1) time complexity.

Hence, hash lock, a coarse-grained lock, which locks the primary keys with the same hash value at once, is the best choice. An ingenious hash lock is presented, and its format is as follows (for details, refer to Table 1):

$$(T_{lw}, V_{\text{ref}}, V_{\text{wait}}, V)$$

2.1. Fragmented Storage of Lock Requests. As the above analysis, lock acquire and release have to traverse the list. As the number of active transactions, the traversing list imposes higher overhead.

For reducing this overhead, fragmented storage of lock requests (fragmented storage for short) is proposed. For ease of interpretation, direct predecessors and direct successors are introduced, and their formal definitions are as follows.

Direct successor: here is from the perspective of the read-write lock conflict. Especially, if $T_i \prec T_k \prec T_j$ and $R_i(a) \prec R_k(a) \prec W_j(a)$, T_j is both a direct successor of T_i and a direct successor of T_k**.** Assume that T_i and T_j are any two transactions that appear in the same lock request list at the same time. If T_i precedes T_j, then $T_i \prec T_j$. For any two transactions T_i and T_j, T_j is the direct successor of T_i if and only if any one of the following three conditions is met:

(1) If key $\in T_i$.readSet $\cap$ T_j.writeSet, $T_i \prec T_j$, then (Tex translation failed), and $T_i \prec T_k \prec T_j$
(2) If key $\in T_i$.readSet $\cap$ T_j.writeSet, $T_i \prec T_j$, then (Tex translation failed), and $T_i \prec T_k \prec T_j$
(3) If key $\in T_i$.readSet $\cap$ T_j.writeSet, $T_i \prec T_j$, then (Tex translation failed), and $T_i \prec T_k \prec T_j$

Direct predecessor: for any two transactions T_i and T_j, T_i is the direct predecessor of T_j if T_j is the direct successor of T_i.

The core idea of fragmented storage is that a transaction is collocated with its direct successors. Each transaction has a list named successor List (preallocation and variable-length) to store its direct successors. Therefore, the lock request list is simplified, and its regular form is $W\{0, 1\}R^*$ (W for the last write request and R for the read request following W without other write transactions between them).

Hence, operations of granting to lock requests are significantly simplified. A write request can be granted if its corresponding request list is empty. A read request

Table 1: Notations.

Symbol	Meaning
T_{lw}	The last write transaction pointer
V_{ref}	The number of read transactions following one write transaction
V_{wait}	The number of transactions waiting to be granted
V	The virtual transaction pointer

can be granted if the head of its corresponding request list is not a write request. Releasing a write request is to remove the head of its request list if the head is itself. Releasing a read request (suppose R_{release}) is slightly complex, and the following two situations need to be considered:

(1) If the head of the request list is write, there is nothing to do because of compressing by the new write request
(2) If the head of the request list is read, traverse the request list, and then remove R_{release} if found

Write compression is essential for the efficiency of fragmented storage. The higher the write ratio, the better the efficiency of fragmented storage. For write-only workloads, the length of the request list is strictly no more than 1, and thereby, fragmented storage gains the best efficiency. For read and write mixed workloads, fragmented storage gains the second-highest efficiency. Why? The reason is that the length of the request list under such workloads is just shortened but generally not less than 2 (possibly larger). As the read ratio increases, the efficiency of fragmented storage gets worse and worse.

For more detailed processing logic, refer to Algorithms 1 and 2. These two algorithms are relatively simple, and so, we ignore additional supplementary instructions.

2.2. Virtual Transaction. For performing well in a continuous reading scenario, the virtual transaction is proposed to effectively reduce the overhead of list operations by compressing continuous read requests.

Virtual transaction: a special transaction which has only the right to lock and unlock primary keys and only operates shared read lock.

Continuous read compression: it is used to compress continuous reads per hash(key) to one which is represented by a virtual transaction. A simple semaphore containing the number of continuous outstanding read requests for that lock with no other write requests interrupted (actually, V_{ref} for continuous read requests) is used, and each successive read operation corresponds to a virtual transaction. From the above analysis, the regular form of the lock request list per lock under the combination of hash lock and fragmented storage is $W\{0,1\}R^*$, and the subpattern R^* is the essential reason for poor effectiveness of fragmented storage in both high-ratio-read workloads and read-only workloads. With read compression, the regular form of the lock request list per lock is further strictly limited to $W\{0,1\}R\{0,1\}$.

```
Input: T: a transaction pointer
(1) foreach key ∈ T ⟶ writeSet do
(2)    L = data[key].lockRequestList;
(3)    if T ≡ L ⟶ front.ownTrans then
(4)       L ⟶ popFront;
(5)    end
(6) end
(7) foreach key ∈ T ⟶ readSet do
(8)    L = data[key].lockRequestList;
(9)    if W ≡ L ⟶ front.lockMode then
(10)      L ⟶ removeIfExist(T);
(11)   end
(12) end
(13) foreach E ∈ T ⟶ successorList do
(14)   E ⟶ prednum − −;
(15)   if 0 ≡ E ⟶ prednum then
(16)      ReadyQueue←E;
(17)   end
(18) end
```

Algorithm 1: Lock release with fragmented storage.

This brings the following two benefits:

The memory cost and CPU cost of the lock request list are constant and minimized. $W\{0,1\}R\{0,1\}$ means that the length of the lock request list is no greater than 2, and thereby, the fixed-length list can be preallocated to reduce the high cost of operations (i.e., create and destroy). Two transaction pointers are placed in each hash lock, respectively, pointing to the last write transaction and the virtual transaction.

The cost of write lock request processing is significantly reduced. Processing a write lock request (assume W) may traverse the corresponding lock request list (assume L), and the transaction that issues W is added to successorList of each element of L. The longer L, the more times the additions. However, after introducing reference counting, W will be used as a direct successor to only two transactions (the last write transaction and the virtual transaction).

How does it work? Consider, for example, the sequence of transactions A, B, C, and D. Among them, A and D only write x, while B and C only read x. At first, x.T_{lw}, x.V, $x.V_{\text{ref}}$, and $x.V_{\text{wait}}$ are initialized to NULL, NULL, 0, and 0. Transaction A arrives. Because of conflict free, $x.T_{lw}$ is set to A and then grants A. Transaction B follows. Because of read-write collision, the following three steps are executed sequentially. (1) Add B to successorList of A. (2) Create a new virtual transaction T_{virtual}. (3) Increase prednum of T_{virtual} by one ($x.V_{\text{ref}}$: 0 ⟶ 1). When C arrives, steps 1 and 3 are selected to execute sequentially. Now, x.T_{lw}, x.V, $x.V_{\text{ref}}$, and $x.V_{\text{wait}}$ are A, T_{virtual}, 2, and 0, respectively. Finally, transaction D comes. Because of read-write collision, D is only added to successorList of T_{virtual} rather than to that of B and C.

Unique virtual transaction for each hash key: one virtual transaction was used to represent all reads for each key. In other words, there cannot be several virtual transactions on the same primary key. Any continuous read operations are

```
    Input: T: a transaction pointer
    Output: Granted: 1 if granted, otherwise 0
(1) Granted:  = 1;
(2) foreach key ∈ T ⟶ writeSet do
(3)   L = data[key].lockRequestList;
(4)   if L ⟶ Empty then
(5)     L ⟵ ⟨W, T⟩;
(6)   else
(7)     if T ≡ L ⟶ back.own Trans then
(8)       foreach E ∈ L do
(9)         l = E.ownTrans ⟶ successorList;
(10)        if T ≡ l ⟶ back then
(11)          Granted = 0;
(12)          l ⟵ T;
(13)          T ⟶ prednum + +;
(14)        end
(15)      end
(16)      L ⟶ clear;
(17)      L ⟵ ⟨W, T⟩;
(18)    end
(19)  end
(20) end
(21) foreach key ∈ T ⟶ readSet do
(22)  L = data[key].lockRequestList;
(23)  if L ⟶ Empty then
(24)    L ⟵ ⟨R, T⟩;
(25)  else
(26)    if T ≡ L ⟶ back.own Trans then
(27)      E = L ⟶ front;
(28)      if W ≡ E.lockMode then
(29)        l = E.ownTrans ⟶ successor List;
(30)        if T ≡ l ⟶ back then
(31)          Granted = 0;
(32)          l ⟵ T;
(33)          T ⟶ prednum + +;
(34)        end
(35)      end
(36)      L ⟵ ⟨R, T⟩;
(37)    end
(38)  end
(39) end
```

Algorithm 2: Lock acquire with fragmented storage.

passed to the write operation immediately following them by reference counting.

How to implement it? The only to do is to save multiple successor transactions for the virtual transaction rather than only to save its direct successor transactions. For saving multiple successor transactions, new $V_{\text{successorList}}$ is attached to virtual transactions each (for details, refer to Algorithms 3 and 4).

Assume that transaction is T and its corresponding hash lock is Lock.

For write lock acquire, see Algorithm 3, 2 to 27 lines. The basic processing logic is as follows:

(1) Check whether Lock.V_{ref} is greater than zero:

If true, prednum of T is increased by LockInfo.V_{ref}. And then, tid of T is added to successorList of Lock.V.

If false, T is added to successorList of Lock.T_{lw} if Lock.T_{lw} is not NULL.

(2) Lock.T_{lw} is assigned T.

For write lock release, see Algorithm 4, 1 to 6 lines. The basic processing logic is as follows:

(1) If Lock.T_{lw} and T are the same, Lock.T_{lw} is assigned NULL.

For read lock release, see Algorithm 4, 7 to 38 lines. The basic processing logic is as follows:

If Lock.T_{lw} is NULL, Lock.V_{ref} is decreased by one, and Lock.V_{wait} is assigned zero.

If Lock.T_{lw} is not NULL, successorList of $L.V$ is traversed sequentially to find the first transaction (assume T_{found}) whose prednum is greater than zero. If T_{found} is found, prednum of T_{found} is decreased by one. Finally, T_{found} is added to ReadyQ ueue immediately if its prednum is zero.

For read lock acquire, see Algorithm 3, 28 to 41 lines. The basic processing logic is as follows:

(1) If Lock.T_{lw} is not NULL, prednum of T is increased by one, and T is added to successorList of Lock.T_{lw}
(2) Lock.V_{ref} is increased by one

Why does successor List of virtual transaction store tid? When one transaction is completed, its related resources are released, and thereby, its pointer is invalid. When a virtual transaction (assume V_1) wakes the first successor transaction (assume T_1) whose prednum is greater than 0, T_1 may not be removed from V successor List of V_1 if prednum is still greater than 0 after decreased. When V_1 performs the wake-up operations again, V_1 will acquire the first successor transaction whose prednum is greater than 0 again. So, if V successor List stores pointers to transactions, V_1 needs to operate these pointers (of course, includes T_1), while the pointer of T_1 may be invalid because of other transaction's waking operations. In addition, TransMap is used to track all unfinished normal transactions, and its format is $\langle$tid, $p_t\rangle$ (tid for transaction ID and p_t for transaction pointer). When a virtual transaction performs wake operations, it will use TransMap to detect whether its successor transaction is active. Hence, the virtual transaction mechanism is running correctly and smoothly.

In a word, these techniques incur far less overhead than maintaining a traditional lock manager and timely detect which transactions should inherit released locks.

3. Dependence-Cognizant Locking

In DCLP, lock is a four-tuple: $(T_{lw}, V_{\text{ref}}, V_{\text{wait}}, V)$. A global map of transaction requests (called TransMap) is maintained, tracking all active transactions with unorder. Because of only detecting whether one transaction is still active, std: : unorder_map is adopted for the possibly best performance.

In order to facilitate the explanation of DCLP, this paper assumes that each partition has one lock thread to run DCLP.

```
     Input: T: a transaction pointer
     Output: Granted: 1 if granted, otherwise 0
 (1) Granted: = 1;
 (2) foreach key ∈ T ⟶ writeSet do
 (3)   LockInfo = data[hash(key)];
 (4)   if T ≡ LockInfo.T_lw then
 (5)     if 0 < LockInfo.V_ref then
 (6)       l: = LockInfo.V ⟶ successor List;
 (7)       if 0 == LockInfo.V_wait then
 (8)         l ⟶ reuse;
 (9)       end
(10)       l←T ⟶ tid;
(11)       LockInfo.V_wait + +;
(12)       T ⟶ prednum+ = LockInfo.V_ref;
(13)       LockInfo.V_ref = 0;
(14)       Granted = 0;
(15)     else
(16)       if NULL ≡ LockInfo.T_lw then
(17)       l: = LockInfo.T_lw ⟶ successorList;
(18)       if l ⟶ empty‖T ≡ l ⟶ back then
(19)         Granted = 0;
(20)         l←T;
(21)         T ⟶ prednum + +;
(22)       end
(23)     end
(24)   end
(25)   LockInfo.T_lw = T;
(26)   end
(27) end
(28) foreach key ∈ T ⟶ readSet do
(29)     LockInfo = data[hash(key)];
(30)     if T ≡ LockInfo.T_lw then
(31)       if NULL ≡ LockInfo.T_lw then
(32)         l: = LockInfo.T_lw ⟶ successorList;
(33)         if l ⟶ empty‖T ≡ l ⟶ back then
(34)           T ⟶ prednum + +;
(35)           l←T;
(36)           Granted = 0;
(37)         end
(38)       end
(39)     end
(40)     LockInfo.V_ref + +;
(41)   end
(42) TransMap←⟨T ⟶ tid, T⟩;
```

Algorithm 3: Lock acquire with DCLP.

```
     Input: T: a transaction pointer
 (1) foreach key ∈ T ⟶ writeSet do
 (2)   LockInfo = data[hash(key)];
 (3)   if T ≡ LockInfo.T_lw then
 (4)     LockInfo.T_lw: = NULL;
 (5)   end
 (6) end
 (7) foreach key ∈ T ⟶ readSet do
 (8)     LockInfo = data[hash(key)];
 (9)     if NULL ≡ LockInfo.T_lw then
(10)       LockInfo.V_ref − −;
(11)       LockInfo.V_wait = 0;
(12)     else
(13)       L: = LockInfo.V ⟶ successorList;
(14)       grantedLock: = 0;
(15)       while !L ⟶ empty do
(16)         got: = TransMap.find(L.front);
(17)         if got ≡ TransMap.end then
(18)           E: = got ⟶ second;
(19)           if 0 ≡ E ⟶ prednum then
(20)             L ⟶ popFront;
(21)           else
(22)         E ⟶ prednum − −;
(23)         grantedLock = 1;
(24)         if 0 ≡ E ⟶ prednum then
(25)           ReadyQueue←E;
(26)         end
(27)         break;
(28)       end
(29)     else
(30)       L.popFront;
(31)     end
(32)   end
(33)   if 0 ≡ granted then
(34)   LockInfo.V_ref − −;
(35)   end
(36)     LockInfo.V_wait = L ⟶ size;
(37)   end
(38) end
(39) foreach E ∈ T ⟶ successorList do
(40)   E ⟶ prednum − −;
(41)   if 0 ≡ E ⟶ prednum then
(42)     ReadyQueue←E;
(43)   end
(44) end
(45) TransMap.erase(T ⟶ tid);
```

Algorithm 4: Lock release with DCLP.

When a transaction arrives at a partition, it requests locks on records belonging to the partition. According to the type of lock request, the processing logic of lock requests is different. In Algorithm 3, 2 to 27 lines correspond to the write lock processing, while 28 to 41 lines correspond to the read lock processing. Write locks are granted to the requesting transaction if $V_{\text{ref}} = 0$ and $T_{lw} = \text{NULL}$ (or is itself). Similarly, read locks are granted if $T_{lw} = \text{NULL}$ (or is itself).

For correctness, one transaction must be added to TransMap after requesting all lock requests. Because of the sequential processing mode, the requesting of the locks and the adding of the transaction to the map can be treated as an atomic operation. As for predetermining the read set and write set of transactions in advance, Thomson [9] had proposed a solution method by allowing a transaction to prefetch whatever reads it needs to (at no isolation) for it to figure out what data it will access before it enters the critical section.

Similar to VLL, when one transaction completes all lock requests, it can be tagged with only one of the three states: free, blocked, and waiting. The specific meanings of these three states are as follows:

> Free means that a transaction has acquired all locks whether it is a single-partition or multi-partition transaction

Blocked means that a transaction has not acquired all of its lock requests immediately upon requesting them

Waiting means that a transaction could not be completely executed without the result of an outstanding remote read request

For a single-partition transaction, its state must be either free or blocked, and thereby, only one state transition is from blocked to free. For a multipartition transaction, its state can be any one of the three states. So, there are two state transition paths. One is from waiting to free if all remote readings are successfully returned. The other is first from blocked to waiting if all ungranted locks are granted and then to free. Only free transactions can begin execution immediately.

DCLP records the dependencies between transactions and thereby has the ability to timely detect unblocked transactions. When one transaction is completed, it is easy to hand over its locks directly to its direct successors by decreasing prednum of its direct successors (see Algorithm 4 for more details).

For VLL, with an increase of active transactions, the probability of a new transaction being tagged free decreases. For good performance, the number of active transactions should be limited. However, it is difficult to tune the number of active transactions since different levels of contention demand different thresholds. So, a threshold is set by the number of active but blocked transactions since high-contention workloads will reach this threshold sooner than low-contention workloads.

However, the fixed threshold brings two defects, and they are as follows:

(1) For most of the real workloads, the contention is not fixed, but is changing from time. Therefore, it is difficult to select a fixed value of this threshold.

(2) When encountering the workloads that include numerous fragmented continuous conflict transactions, the performance is greatly affected. It is because that once the threshold is exceeded, new transactions make no progress even if they are free.

For DCLP, the above two defects have been completely overcome. First, there is no need to set a threshold for active but blocked transactions. DCLP tracks the direct conflict between transactions, and thus, blocked transactions can be unblocked and executed immediately. A preallocated and variable-length list attaches each transaction to store all its direct successors. Hence, compared with VLL, DCLP can nonblocking handle more transactions without being restricted by the threshold of blocked transactions.

However, in consideration of memory and performance, the number of active transactions should be set a threshold, especially when encountering workloads with long transaction execution time.

Figure 1 depicts an example execution trace for a sequence of transactions. Free transactions (here A) are pushed directly into ReadyQueue. Transactions B and A have RAW dependencies on key x, and thus, B is termed blocked. Meanwhile, B is added to successorList of A. The end of A will unblock B and then push B to ReadyQueue. The processing logic of C is similar to that of transaction B and will not be expanded in detail.

4. Evaluation

To evaluate DCLP's performance, a large number of experiments are designed to compare DCLP against deadlock-free 2PL (Calvin's deterministic locking protocol), VLL, and VLL\SCA in a number of contexts. These experiments are divided into two groups: single-machine multicore experiments and experiments in which data are partitioned across many commodity machines in a shared-nothing cluster. The single-machine multicore experiments were conducted on a Linux (2.6.32–220.el6) of two 2.3 GHz six-core Intel Xeon E5-2630 machine with 24 CPU threads and 64 GB RAM. The experiments in a distributed database were conducted on the same server configuration as the single-machine multicore experiments used, connected by a single 10 gigabit Ethernet switch.

As a comparison point, DCLP is implemented inside Calvin. This allows for an apples-to-apples comparison in which the only difference is the locking strategy.

The same configuration as [5] was used: devote 3 out of 8 cores on every machine to those components that are completely independent of the locking scheme, and devote the remaining 5 cores to worker threads and lock management threads. For all experiments, we stationary use fixed 4 work threads and one lock thread.

Although OCC (or MVOCC) shows greater advantages than PCC in short-transaction workloads for MMDB, PCC is still widely adopted by many databases. This paper focuses on PCC and single data version (not MVCC). Therefore, OCC and MVCC are not within the scope of this paper to discuss.

4.1. Multicore, Single-Server Experiments

4.1.1. Standard Microbenchmark Experiments. Two experiments are presented in this section:

The first experiment is called short microbenchmark. Each microbenchmark transaction reads 10 records and updates a value at each record. Of the 10 records accessed, one is chosen from a small set of "hot" records, and the rest are chosen from a larger set of 'cold' records. The contention levels are tuned by varying the size of the set of hot records. The set of cold records is always large enough so that transactions are extremely unlikely to conflict. In addition, the term contention index is used to represent contention levels. For example, if the number of hot records is 1000, the contention index would be 0.001.

The second experiment is called long microbenchmark. The only difference between long microbenchmark and short microbenchmark is that the former consumes a certain amount time to calculate after reading each record (default provided by Calvin).

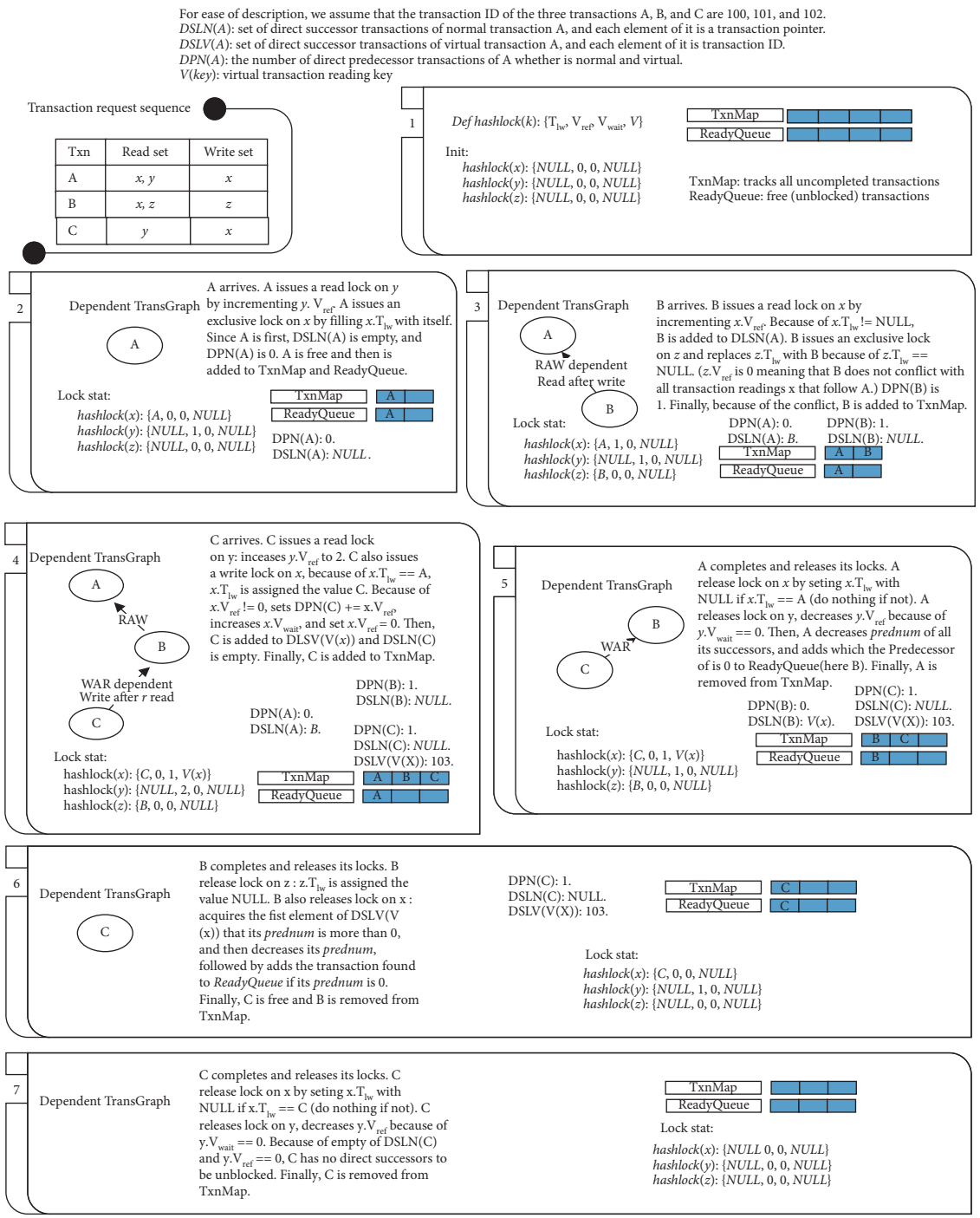


Figure 1: Example execution of a sequence of transactions A, B, and C using DCLP. Each transaction's read and write set is shown in the top left box.

As shown in Figure 2, when contention is low, DCLP, VLL, and VLL\SCA yield near-optimal transaction throughput, while only almost 50% of their transaction throughput can be provided by deadlock-free 2PL. The reason is that VLL and VLL\SCA almost completely eliminate the overhead of latch acquisitions and linked list operations.

Why does DCLP gain such high transaction throughput in low contention? There are three reasons. First, DCLP also adopts acquiring all locks at once for reducing the overhead of latch acquisitions. Second, the hash lock almost completely eliminates the overhead of primary key storage. Third, fragmented storage and virtual transaction effectively eliminate the overhead of linked list operations. Low contention means that the overhead to maintain the conflict information between transactions is negligible.

As contention increases, all decrease, while VLL presents the trend of the first decline and then is almost stable.

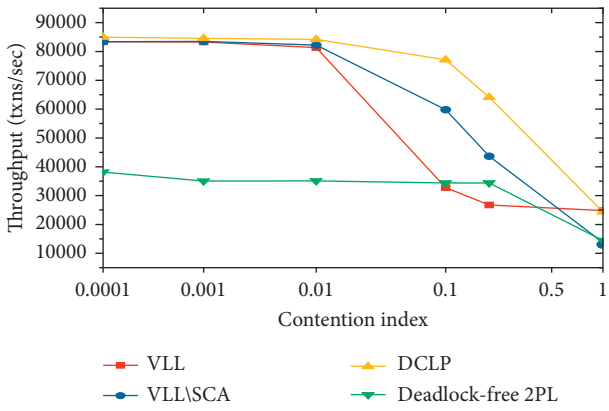


Figure 2: Transaction throughput with short microbenchmark vs. contention under a deadlock-free workload.

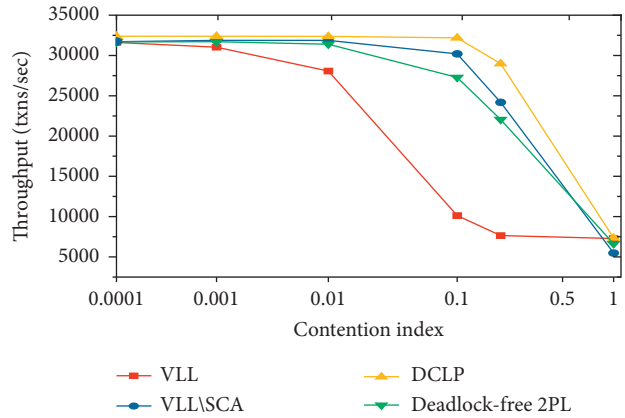


Figure 3: Transaction throughput with long microbenchmark vs. contention under a deadlock-free workload.

Without tracking conflict information, blocked transactions can only be executed serially, and then VLL's performance falls quickly. With tracking conflict information, the higher the contention, the more useful the conflict information. So, the performance of the remaining three mechanisms falls slowly.

Figure 3 shows the trend of the transaction throughput of the four locking mechanisms is very similar to the short microbenchmark results, and DCLP still performs well. However, the transaction throughput of all the four locking mechanisms receives a significant impact and is only about 40% of their throughput in short microbenchmark. The reason is that the executing time of long transactions becomes the primary performance bottleneck rather than the lock overhead.

4.1.2. Modified Microbenchmark Experiments. The standard microbenchmark is write-only, and therefore, these four locking mechanisms cannot be fully evaluated. For a comprehensive and complete evaluation, some modified microbenchmarks by combining different ratios of read to write and different degrees of conflict were designed.

The first set of modified microbenchmarks only just replaces write-only with read-only, with the purpose to prove that DCLP is efficient and has the same high performance as VLL and VLL\SCA on read-only workloads.

As expected, Figure 4 confirms that DCLP has almost the same high performance as VLL and VLL\SCA with respect to read-only workloads no matter how high or low the contention is.

The second set of modified microbenchmarks varies contentions with a fixed ratio of read to write. As we all know, for some real transaction processing systems, reading more and writing less is a very important feature. Considering some of the default workloads of YCSB [10], a special workload with 95% read and 5% write is chosen for this set of experiments.

As shown in Figure 5, when contention is low (actually only 5% write means that factual contention is lower), DCLP, VLL\SCA, and VLL have almost the same high performance. Low contention does not fully demonstrate the advantage of these three locking mechanisms in detecting unblocked transactions. As contention increases, DCLP

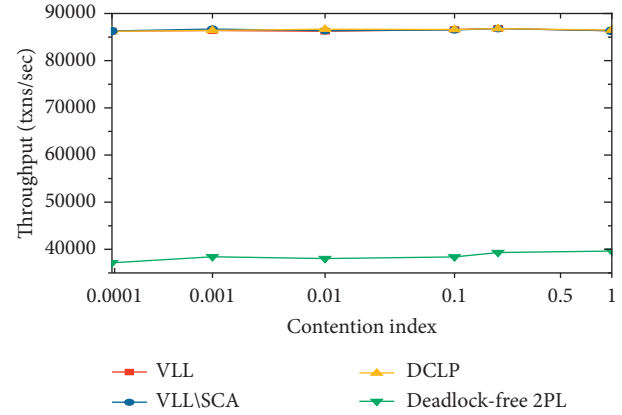


Figure 4: Transaction throughput with read-only short microbenchmark vs. contention under a deadlock-free workload.

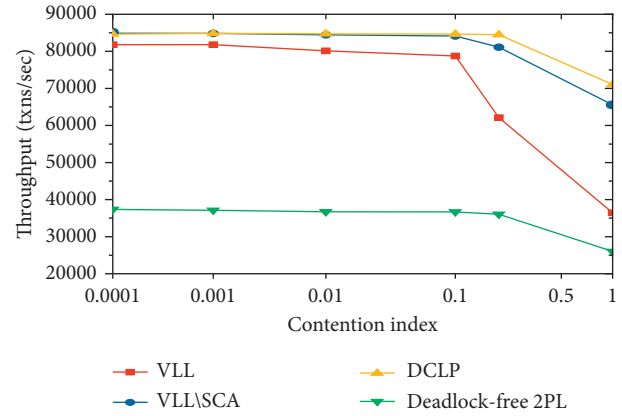


Figure 5: Transaction throughput with modified short microbenchmark (5% write and 95% read) vs. contention under a deadlock-free workload.

gradually demonstrates this advantage, and thereby, its performance is the highest. Due to the loss of the ability to detect unblocked transactions, VLL is lower than VLL\SCA. As for deadlock-free 2PL, its performance is still the lowest because of its high overhead of locking processing.

The third set of modified microbenchmarks varies both contention and write percent. Two representative contention points are selected: 0.1 for the high-contention and 0.0001 for the low-contention point (as [5]). Due to similarity and simplification, short microbenchmark with low contention was executed, but no longer presented.

For short transactions, the high overhead of locking processing is the primary bottleneck that affects transaction throughput. Therefore, DCLP and VLL\SCA overtake deadlock-free 2PL. As for VLL, without the ability to detect unblocked transactions, its performance dramatically decreases. Figure 6 presents these conclusions.

In real transaction processing systems, long chain dependence between transactions is also an important feature, for example, exchange rate transactions in which the exchange rate is being read continuously along with regular update with a small update interval. So, we design the last set of experiments to test such workloads. Similar to the standard short microbenchmark, each transaction has one hot record and nine cold records, and any of the cold records of each transaction is read and updated. The difference is that the size of the hot dataset is fixed to 1. And then, we tune the ratio of read to write to simulate the depth of dependence between transactions.

Table 2 shows the results of the last set of experiments. Every transaction operates the only one hot record meaning that one transaction blocks all its subsequent transactions. So, for the four locking mechanisms, the efficiency in detecting unblocking transactions can improve their transaction throughput. The conclusion is that DCLP is the most efficient in detecting unblocked transactions, and the second is VLL\SCA. Without the ability in detecting unblocked transactions, VLL achieves poor transaction throughput. In addition, although it has the ability in detecting unblocked transactions, deadlock's performance is even lower than that of VLL. It also verifies from another perspective that deadlock-free 2PL is expensive.

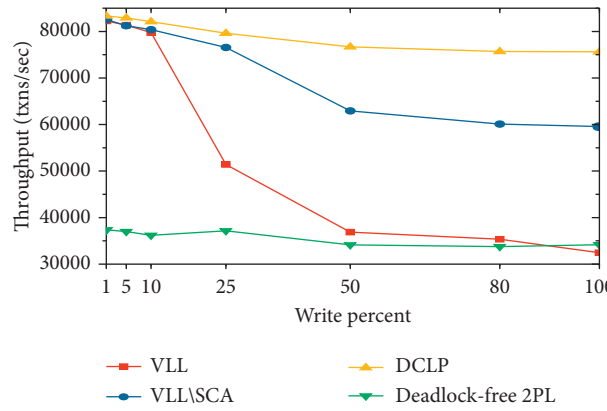


Figure 6: Transaction throughput with the modified short microbenchmark (high contention) under a deadlock-free workload, varying the write percent.

Table 2: Transaction throughput with the modified microbenchmark under a deadlock-free workload, varying read and write interdependent depth.

Throughput (txns/sec)	9 depth	19 depth	99 depth
VLL	18315.4	23381.7	56825.4
VLL\SCA	20016.2	33874.2	65905.2
DCLP	27632.5	51659.9	95499.4
Deadlock-free 2PL	15261.1	21644.7	38026.6

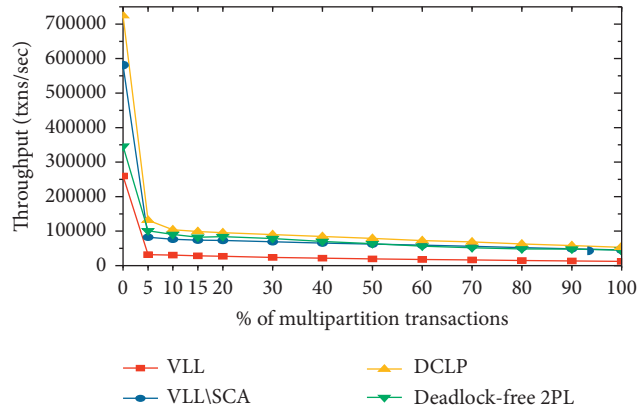


Figure 7: Microbenchmark throughput with high contention under a deadlock-free workload, varying how many transactions span multiple partitions.

4.2. Distributed Database Experiments

4.2.1. Standard Microbenchmark Experiments. The same standard short microbenchmark as in Section 4.1.1 is used. Both contention and percentage of multipartition transactions are varied. Two representative contention points are also selected: 0.1 for the high-contention and 0.0001 for the low-contention point (same as [5]). In addition, these experiments were run on 8 machines.

Although the data partition strategy is different from that of [5], there is no doubt that SCA is extremely important no matter contention is high or low. Under high contention, SCA improves the performance of VLL by at least 55% (maximum of up to 73%). Under low contention, SCA improves the performance of VLL by at least 40% (maximum of up to 67%) when the percentage of multipartition transactions is greater than 20%.

Figure 7 shows that DCLP is the best one and has about 15% performance advantage over VLL\SCA. It is easy to understand that DCLP is more efficient in detecting unblocked transactions. Because of the long executing time, the performance of all four locking mechanisms presents a downward cliff from non-multipartition transactions to 5%. When the percentage is greater than 5%, all transactions almost can only be executed serially so that the performance is almost stable.

Figure 8 also shows that DCLP is the best one. Compared with VLL\SCA, DCLP has about 2% to 21% performance advantage. Although contention is low, with the increase of multipartition transactions and the only one lock thread, more and more transactions are blocked. The performance gap between DCLP, VLL, and VLL\SCA is mainly due to the efficiency of detecting unblocked transactions.

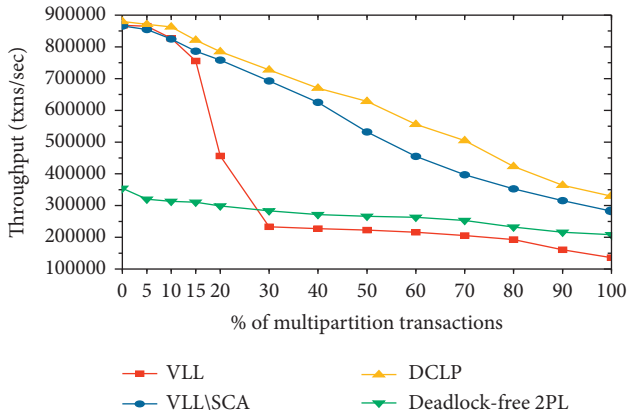


Figure 8: Microbenchmark throughput with low contention under a deadlock-free workload, varying how many transactions span multiple partitions.

The scalability of three of the four locking mechanisms is also tested at low contention when there are 10% and 20% multipartition transactions. These locking mechanisms are deadlock-free 2PL, VLL\SCA, and DCLP. We scale from 2 to 8 machines in the cluster. Figure 9 shows that DCLP achieves the same linear scalability as VLL\SCA and deadlock-free 2PL and still has the best performance at scale.

4.2.2. TPC-C Experiments. The same TPC-C benchmark as in [5] is used. In order to vary the percentage of multipartition transactions in TPC-C, the percentage of new order transactions that access a remote warehouse is varied. 96 TPC-C warehouses were divided across the same 8-machine cluster described in the previous sections, but there is a subtle difference that all the four lock mechanisms partitioned the TPC-C data across 8 twelve-warehouse partitions (one per machine and one partition has twelve warehouses). Each partition corresponds to one machine. In the end, we would expect to achieve similar performance if we were to run the complete TPC-C benchmark.

Figure 10 shows the transaction throughput results of the four locking mechanisms. Overall, the performance of different locking mechanisms is very similar to the result of the high-contention microbenchmark with multipartition transactions. DCLP still has the best performance and has maximum of about 5% performance advantage over VLL\SCA.

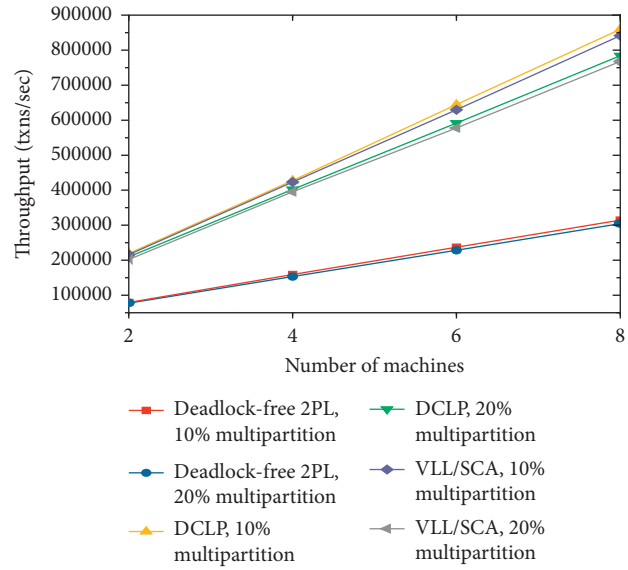


Figure 9: Scalability under a deadlock-free workload.

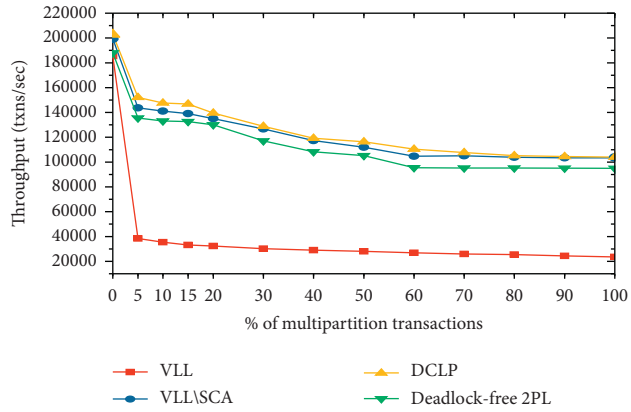


Figure 10: Transaction throughput with TPC-C under a deadlock-free workload, varying how many transactions span multiple partitions.

5. Related Work

5.1. Reducing Lock Acquisitions. These research works mainly reduce the number of lock acquisitions by redesigning the processing logic of LM, such as [4, 6, 7, 11]. Speculative lock inheritance (SLI) [4] allows a completing transaction to pass on some locks (hot locks, frequently acquired in a short time) which it acquired to transactions which follow. This successfully avoids a pair of release and acquire calls to the lock manager for each such lock. However, SLI only performs well on hot locks because SLI does not optimize LM itself and only just reduces the number of calls to LM. RHS [7] adopts the RAW pattern and barrier synchronization to greatly reduce the use of latches and thus generally improves the performance. However, it is still in the category of traditional lock manager (optimized).

5.2. Job Scheduling. In other communities except for the database community, there has been extensive research on scheduling problems in general. These research works are to minimize time costs from four different perspectives and have achieved better performance. These four aspects are, respectively, to minimize the sum of completion time [12], the latest completion time [13], the completion time variance [14, 15], and the waiting time variance [16]. However, the assumption of these research works that each processor/worker can be used by only one job at a time is not applicable to the database community, while locks can be held in shared and exclusive modes in the database.

5.3. Dependency-Based Scheduling. Scheduling transactions with dependencies among them has been studied for many years. Tian et al. [17] proposed a contention-aware transaction scheduling algorithm, which captures the contention and the dependencies among concurrent transactions. It is different from most existing systems which rely on a FIFO (first in, first out) strategy to decide which transactions to grant the lock to. Although it achieves better performance than FIFO, it may cause some problems to the applications based on FIFO because [17] granted the lock to the transaction with the largest dependency set, rather than to the first requested transaction. However, in real applications, in some situations, applications want databases to execute transactions as their sending order.

5.4. Optimistic Concurrency Control. Optimistic concurrency control (OCC) is a popular concurrency control due to its low overhead in low-contention settings [18–26]. Huang et al. [18] presented two optimization techniques, commit-time updates and timestamp splitting, which can dramatically improve the high-contention performance of OCC. TicToc [25] used a technique called data-driven timestamp management to completely eliminate the timestamp allocation bottleneck. AOCC [19] adaptively chose an appropriate tracking mechanism and validation method to reduce the validation cost according to the number of records read by a query and the size of write sets from concurrent update transactions.

5.5. Multiversion Concurrency Control. The main advantage of MVCC is that it potentially allows for greater concurrency by permitting parallel accesses on different versions. Many MMDBs are in favor of MVCC (e.g., Hekaton and MemSQL). They utilize mechanisms commonly known as MVOCC or MV2PL [27, 28]. Wu et al. [27] made a comprehensive empirical evaluation and identified the limitations of different designs and implementation choices.

5.6. Lightweight Locking. Lightweight locking also attracts many researchers. In some situations, [5, 9] almost eliminate all the expensive overhead of traditional lock managers. In VLL, LM is extremely simplified by replacing the lock request list with two simple semaphores containing the number of outstanding requests for that lock (C_x for write requests and C_s for read requests). However, VLL performs badly on high-contention workloads, while VLL\SCA does not have the ability to schedule unblocked transactions without delay.

From previous literature about the optimization and redesign of LM, we can conclude that designing optimal LM for databases has remained an open problem. Considering that almost all existing systems rely on a FIFO (first in, first out) strategy to decide which transactions to grant the lock to, we decide to follow the FIFO strategy. In addition, acquiring all locks at once can avoid deadlock and then achieve good performance. In this paper, we blend the advantages of dependency-based scheduling and lightweight locking to redesign LM for reducing the locking space and providing higher performance.

6. Conclusion

In this paper, we presented dependence-cognizant locking (DCLP) combined with refined scheduling with lightweight locking that provides a fast and efficient concurrency control for database systems. For MMDB, the lock and latch cost of the traditional lock mechanism is expensive. In DCLP, we manage transactions by dependency chains. It eliminates most of the cost of latch, and more importantly, transactions can be awakened immediately when their required data are unlocked. Furthermore, for better lock transfer performance, we proposed two optimization techniques. One is named fragment storage of lock mechanism which ensures transactions getting successors in place. The other is the virtual transaction mechanism (VT for short), compressing continuous read requests into a special read request as VT, which reduces the scheduling complexity significantly, especially in heavy workloads. Experiments show that DCLP achieves better performance than deadlock-free 2PL and VLL\SCA, without inhibiting scalability. In the future, we will intend to integrate hierarchical locking approaches, column locking, and row/column hybrid locking mechanism into DCLP and investigate multiversion variants of the DCLP strategy.

Data Availability

All experiments are based on Calvin, which is an open-sourced transactional database, and its URL is https://github.com/yaledb/calvin. Except DCLP, the benchmarks and the other three lock mechanisms used in this paper have been implemented in Calvin. This paper has given the pseudocode of the algorithm DCLP. Further help is available from the first author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 61672434, 61472321, and 61732014.

References

[1] S. Harizopoulos, D. J. Abadi, S. Madden et al., "Oltp through the looking glass, and what we found there making databases work: the pragmatic wisdom of michael stonebraker," 2018.

[2] P. Larson, "High-performance concurrency control mechanisms for main-memory databases," *Proceedings of Vldb Endowment*, vol. 24, 2011.

[3] I. Pandis, R. Johnson, N. Hardavellas, and A. Ailamaki, "Data-oriented transaction execution," *Proceedings of the VLDB Endowment*, vol. 3, no. 1-2, pp. 928–939, 2010.

[4] R. Johnson, I. Pandis, A. Ailamaki et al., "Improving OLTP scalability using speculative lock inheritance," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 479–489, 2009.

[5] K. Ren, A. Thomson, and D. J. Abadi, "Lightweight locking for main memory database systems," *Proceedings of the Vldb Endowment*, vol. 6, no. 2, pp. 145–156, 2012.

[6] A. M. Joshi, G. M. Lohman, A. Sernadas et al., "Adaptive locking strategies in a multi-node data sharing environment," in *Proceedings of the International Conference on Very Large Data Bases*, Catalonia, Spain, February 1991.

[7] H. Jung, H. Han, A. Fekete, G. Heiser, and H. Y. Yeom, "A scalable lock manager for multicores," *Acm Transactions on Database Systems*, vol. 39, no. 4, pp. 1–29, 2014.

[8] T. Horikawa, "Latch-free data structures for DBMS: design, implementation, and evaluation," in *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data. ACM*, New York, NY, USA, June 2013.

[9] A. Thomson, "VLL: a lock manager redesign for main memory database systems," *Vldb Journal the International Journal of Very Large Data Bases*, 2015.

[10] B. F. Cooper, A. Silberstein, E. Tam et al., "Benchmarking cloud serving systems with YCSB," in *Proceedings of the Symposium on Cloud Computing*, pp. 143–154, Indianapolis, IN, USA, June 2010.

[11] X. Yu, G. Bezerra, A. Pavlo et al., "Staring into the abyss: an evaluation of concurrency control with one thousand cores," *Proceedings of the VLDB Endowment*, vol. 7, 2014.

[12] C. He, J. Y.-T. Leung, K. Lee, and M. L. Pinedo, "Improved algorithms for single machine scheduling with release dates and rejections," *4OR-Q J Operational Research*, vol. 14, no. 1, pp. 41–55, 2016.

[13] B.-C. Choi, S.-H. Yoon, S.-J. Chung et al., "Minimizing maximum completion time in a proportionate flow shop with one machine of different speed," *European Journal of Operational Research*, vol. 176, no. 2, pp. 964–974, 2007.

[14] A. M. Krieger and M. Raghavachari, "V-shape property for optimal schedules with monotone penalty functions," *Computers & Operations Research*, vol. 19, no. 6, pp. 533-534, 1992.

[15] X. Cai, "V-shape property for job sequences that minimize the expected completion time variance," *European Journal of Operational Research*, vol. 91, no. 1, pp. 118–123, 1996.

[16] S. Eilon and I. G. Chowdhury, "Minimising waiting time variance in the single machine problem," *Management Science*, vol. 23, no. 6, pp. 567–575, 1977.

[17] B. Tian, J. Huang, B. Mozafari et al., "Contention-aware lock scheduling for transactional databases," *Proceedings of the VLDB Endowment*, vol. 11, no. 5, 2018.

[18] Y. Huang, W. Qian, E. W. Kohler et al., "Opportunities for optimism in contended main-memory multicore transactions," *Proceedings of the VLDB Endowment*, vol. 12, no. 4, 2020.

[19] J. Guo, P. Cai, J. Wang, W. Qian, and A. Zhou, "Adaptive optimistic concurrency control for heterogeneous workloads," *Proceedings of the VLDB Endowment*, vol. 12, no. 5, pp. 584–596, 2019.

[20] M. E. Schule, L. Karnowski, J. Schmeißer et al., "Versioning in main-memory database systems: from musaeusdb to tardisdb," in *Proceedings of the 31st International Conference on Scientific and Statistical Database Management*, pp. 169–180, Santa Cruz, CA, USA, July 2019.

[21] T. Zhu, Z. Zhao, F. Li et al., "SolarDB," *ACM Transactions on Storage*, vol. 15, no. 2, pp. 1–26, 2019.

[22] B. Ding, L. Kot, and J. Gehrke, "Improving optimistic concurrency control through transaction batching and operation reordering," *Proceedings of the VLDB Endowment*, vol. 12, no. 2, pp. 169–182, 2018.

[23] H. Lim, M. Kaminsky, and G. David, "Andersen. cicada: dependably fast multi-core in-memory transactions," in *Proceedings of the ACM international conference on management of data. ACM*, Chicago IL, USA, May 2017.

[24] T. Wang and H. Kimura, "Mostly-optimistic concurrency control for highly contended dynamic workloads on a thousand cores," *Proceedings of the VLDB Endowment*, vol. 10, no. 2, pp. 49–60, 2016.

[25] X. Yu, A. Pavlo, D. Sanchez et al., "Tictoc: time traveling optimistic concurrency control," in *Proceedings of the 2016 International Conference on Management of Data*, pp. 1629–1642, San Francisco, CA, USA, June 2016.

[26] H. T. Kung and J. T. Robinson, "On optimistic methods for concurrency control," *ACM Transactions on Database Systems*, vol. 6, no. 2, pp. 213–226, 1981.

[27] Y. Wu, J. Arulraj, J. Lin, R. Xian, and A. Pavlo, "An empirical evaluation of in-memory multi-version concurrency control," *Proceedings of the VLDB Endowment*, vol. 10, no. 7, pp. 781–792, 2017.

[28] G. Weikum and Gottfried Vossen, *Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control and Recovery*, ACM, New York NY, USA, 2001.

Hindawi
Mathematical Problems in Engineering
Volume 2021, Article ID 3421478, 9 pages
https://doi.org/10.1155/2021/3421478

Research Article

Prediction Modelling of Cold Chain Logistics Demand Based on Data Mining Algorithm

Bo He and Lvjiang Yin

School of Economics and Management, Hubei University of Automotive Technology, Shiyan 442002, China

Correspondence should be addressed to Bo He; zmestudy@163.com

Received 21 August 2020; Revised 4 September 2020; Accepted 5 February 2021; Published 16 February 2021

Academic Editor: Ernesto Zambrano-Serrano

Modern information technologies such as big data and cloud computing are increasingly important and widely applied in engineering and management. In terms of cold chain logistics, data mining also exerts positive effects on it. Specifically, accurate prediction of cold chain logistics demand is conducive to optimizing management processes as well as improving management efficiency, which is the main purpose of this research. In this paper, we analyze the existing problems related to cold chain logistics in the context of Chinese market, especially the aspect of demand prediction. Then, we conduct the mathematical calculation based on the neural network algorithm and grey prediction. Two forecasting models are constructed with the data from 2013 to 2019 by R program 4.0.2, aiming to explore the cold chain logistics demand. According to the results estimated by the two models, we find that both of models show high accuracy. In particular, the prediction of neural network algorithm model is closer to the actual value with smaller errors. Therefore, it is better to consider the neural network algorithm as the first choice when constructing the mathematical forecasting model to predict the demand of cold chain logistic, which provides a more accurate reference for the strategic deployment of logistics management such as optimizing automation and innovation in cold chain processes to adapt to the trend.

1. Introduction

With the economic growth and social development, people's living standard has been steadily improved these years. Nowadays, there are increasingly consumers that have updated view of health and consumption. They tend to focus on a higher level of life quality such as the requirements of freshness and varieties when selecting and purchasing product, which greatly promoted the vigorous development of cold chain logistics. Actually, cold chain logistics is a special category of supply chain logistics, depending on the refrigeration technology to keep products in the specified temperature environment during the process from production period, circulation period, to sale period. Such professional logistics activities aim to ensure the quality of fresh products and satisfy the needs and requirements of consumers. Cold chain is vital in preserving the integrity and freshness of transported temperature-sensitive products. Generally, the subdivision product of fresh agricultural products includes meat, aquatic product, mill, fruit vegetable, etc. These fresh products are prone to spoilage and deterioration, which is the reason why they have to be transferred, stored, and distributed in a relatively complicated environmental where the temperature and humidity are specially designed and controlled during the logistics process. In order to preserve the freshness of these agricultural products to the greatest extent and deliver them to consumers perfectly, it is necessary to treat these products in a suitable low temperature environment. Besides, the newest data show that China's cold chain logistics market scale reached 303.5 billion yuan in 2019. Such a large scale of cold chain logistics is due to the continuous enrichment of fresh products, which objectively increases the demand for low-temperature preservation and transportation of products. According to the data revealed in National Bureau of Statistics of China, the production of various fresh agricultural

products is rising year by year, except for the production of meat. From an overall perspective, the total production of China's fresh products is going up (see Table 1). The data shown in Table 1 obviously illustrate that the annual total output of products has steadily increased, which solves the problem of supply shortage and provides a solid material foundation for the development of cold chain logistics.

On the contrary, the increasingly extensive and continuous improvement of China's cold chain infrastructure also provides equipment support and technical support for the development of cold chain logistics. In 2019, the total number of cold storages in China reached 60.53 million tons, and the number of refrigerated trucks reached 21,147 million tons. The increasing popularity of refrigeration equipment has greatly alleviated the problem of "availability of cold" and further released the development potential of cold chain logistics.

From the overall point of view, although the development of cold chain logistics in China has made remarkable progress, the logistics demand has not been fully met due to the relatively latter reaction to the cold chain market, leading to the mismatch between supply and demand. What is worse, the outbreak of COVID-19 causes all kinds of effects on all walks of life, and the cold chain logistic is no exception. As the social distance became one of the important factors of consumer behaviour, "Contactless" distribution is more and more popular and adopted widely to the logistic development under this special situation, which made many electronic e-business platforms providing fresh products emerge blowout growth in the epidemic situation. Unfortunately, faced with such a huge flow of market demand, the problems of cold chain logistics industry are also more severe, such as the instability of product supply, shortage of personnel, lack of transport capacity, system collapse, and other issues that need to be solved. The rapid increase in cold chain logistics demand and the lag of effective supply in the short term can account for the emergence of these problems basically. Due to the high equipment cost of cold chain logistics, most enterprises are not willing to purchase a large number of refrigeration equipment in advance when facing the fuzzy market demand but tend to add corresponding supporting facilities only when the demand arises in order to avoid waste and save expenses. At the same time, compared with the general logistics supply chain, cold chain logistics has more difficulties in management, higher technical requirements, and longer equipment construction cycle, so it needs to make corresponding preparations in advance. All the analysis proves the importance of demand forecasting of cold chain logistics. Consequently, it is very difficult to supplement quickly in a short period of time if there is a shortage of supply, which also explains why the cold chain logistics is in short supply when the outbreak of the epidemic.

To solve the above problems, a more reasonable way is to forecast the demand of cold chain logistics scientifically by conducting mathematical model. Accurate prediction of cold chain logistics development trend is of great strategic significance for the sustainable development of the overall national economy. From the macrolevel, the cold chain logistics infrastructure construction, cold chain-related support policies, etc., can have better effect only if the layout is planned according to the demand [1]. From the microlevel, if the demand scale and development trend of cold chain logistics can be predicted in advance, it can play a role of wind vane for microindividuals to a certain extent. Therefore, this paper innovatively uses data mining to conduct the mathematical analysis and forecast the demand of cold chain logistics. According to the mathematical logic of grey prediction and neural network algorithm, relevant program codes with R language are created and run to establish prediction models. Then, we compare the prediction effect of the two and find out the relatively better prediction model. This paper provides more accurate and effective prediction methods by constructing the neural network algorithm model based on data mining, so as to promote the progress of society and the development of cold chain logistics engineering and management.

2. Literature Review

China's cold chain logistics industry started late, and its foundation is relatively weak. The relevant data statistics and research work are not comprehensive enough so far. From the existing literature, the current research on cold chain logistics mainly focuses on the following aspects. The first kind is qualitative research, stressing on the theoretical analysis of cold chain logistics status and countermeasures. Although China's cold chain logistics is booming, there are still many problems related to the cold chain logistic, which cannot be ignored. The problems of cold chain logistics mainly include the backward of infrastructure construction, insufficient investment in construction funds, the incomplete formation of cold chain logistics system, and the low level of informatization. From the perspective of the development of commercial circulation industry, China's cold chain logistics has deficiencies in transportation cost, cold chain continuity, enterprise development factor investment, agricultural product standardization, and cold chain awareness, which restrict the development of cold chain logistics [2]. In view of these problems, some scholars believe that it is necessary to promote the development of cold chain logistics from the aspects of macromanagement, standardization, brand building, and consumer awareness training of cold chain logistics. At the same time, we should increase the investment in cold chain logistics infrastructure and information construction, accelerate the construction and improvement of agricultural products cold chain network system, and strengthen the research and development and application of new technologies of agricultural products cold chain logistics [3]. Through strengthening coordination and supervision, strengthening policy support, increasing capital investment, and accelerating personnel training, the development of cold chain logistics can be promoted [4].

On the literature of cold chain logistics, the second hotpot of research direction mainly focuses on empirical analysis from different influencing factors and evaluates the effect and efficiency of cold chain logistics, which emphasizes data mining and processing. Guo Mingde et al.

Table 1: China's fresh food production from 2012 to 2019 (unit: 10,000 tons).

Category	2012	2013	2014	2015	2016	2017	2018	2019
Aquatic products	5502	5744	6002	6211	6379	6445	6458	6480
Meat	8471	8633	8818	8750	8628	8654	8625	7759
Milk	3175	3001	3160	3180	3064	3039	3075	3201
Fruit	22092	22748	23303	24525	24405	25242	25688	27401
Vegetable	61624	63198	64949	66425	67434	69193	70347	72103
Total	100864	103324	106231	109090	109911	112573	114192	116944

calculated and analyzed the development level of cold chain logistics in 12 provinces and cities by using a factor analysis method, multilayer perception, and cluster analysis method and pointed out that there are significant differences in the cold chain logistics level of agricultural products in China. The development level of cold chain logistics in the eastern region is the highest, followed by the middle part, and then the western region. Generally speaking, the development level of cold chain logistics in China presents a ladder shape [5]. Gao Fan analyzed the influencing factors of deterioration loss of cold chain agricultural products and pointed out the development trend of controlling the deterioration loss of cold chain agricultural products from the perspectives of cold chain inventory, cold chain storage, transportation and distribution time, cold chain investment, and temperature control [6]. Tian Yujie et al. adopted the comprehensive weighting method combining AHP method and entropy weight method to determine the weight of the comprehensive evaluation index and comprehensively evaluated the logistics safety of fresh products from the aspects of storage, transportation, packaging, distribution, and handling [7]. Cao Wujun et al. analyzed the distribution efficiency of cold chain logistics based on system dynamics and found that when enterprises pay attention to the dynamic market demand, actively meet the needs of consumers, and increase the supply-demand ratio of cold and fresh meat products, the distribution efficiency can be improved to a certain extent, thus improving customer satisfaction and increasing the market share of enterprises [8]. Based on the perspective of time and space, Wang Jun and Li Hongchang discussed the role of the intermediate layer organization in the agricultural product cold chain logistics and pointed out that the role of the intermediate layer of the agricultural product cold chain is affected by the external environment such as social economy and policy and also affected by the internal governance performance [9].

The third aspect of current literature is the combination of qualitative and quantitative, focusing on the study of practical problems in cold chain logistics, such as location of distribution points, quality and safety monitoring system, and path optimization. Aiming at the network layout and transportation problems of cold chain logistics network, Zhang Wenfeng proposed a nonlinear mixed integer programming model with the construction cost and operation cost of cold chain logistics network as the optimization objectives and solved the nonlinear mixed integer planning model of cold chain logistics network by using a quantum particle swarm optimization algorithm and the layout problem of precooling station and cold chain logistics distribution center in a cold chain logistics network [10]. Zhao Zhixue et al. considered the economic cost and environmental cost and integrated the road congestion factor into the cold chain logistics green vehicle routing optimization mathematical model [11]. Zhou Qiang et al. established the mathematical model of distribution route optimization with minimum total cost based on the Internet and cloud computing technology. They established an intelligent comprehensive technology prevention and control system, with core functions including real-time monitoring, safety early warning, and food traceability [12]. Based on real-time traffic information, Yao Yuanguo et al. analyzed the fixed cost, transportation cost, refrigeration cost, damage cost, and penalty cost of agricultural products cold chain logistics distribution and established a mathematical model of distribution path optimization with the total cost minimization [13]. Aiming at the problem of slow convergence speed caused by insufficient pheromone in the initial stage of ant colony algorithm, Fang Wenting et al. constructed a hybrid ant colony algorithm and established a mathematical model of cold chain logistics path optimization with the minimum total cost as the research objective and carried out simulation optimization and comparative analysis on an example to verify the effectiveness of the model and algorithm [14]. Similarly, according to the technical advantages of the Internet and the characteristics of cold chain logistics, Yao Zhen and Zhang Yi used the improved genetic algorithm to solve the mathematical model and proposed the logistics distribution path optimization model with the minimum total cost as the objective function, and the effectiveness and rationality of the model and algorithm were demonstrated by an example [15].

In the related research of cold chain logistics, demand forecasting is also a research hotspot. Demand forecasting can help avoid over or undersupply of cold chain logistics and play a certain role in indicating the direction of investment. Therefore, it has attracted more and more scholars' attention in recent years. Based on a combination of weight distribution methods, Wang Xiumei ensemble partial least squares method, time series ARIMA method, and quadratic exponential smoothing method; three single forecast methods, respectively, predicted China's aquatic products, meat, poultry, eggs and milk products, and fruit and vegetable products. The demand trend of cold chain logistics for large-scale agricultural products pointed out that the forecast accuracy of the weight distribution combination method is better than that of the three single forecast methods [16]. Yuan Jing compared the results of one-way forecasts of agricultural cold chain logistics and

positive weight combination forecasts using extended trend, exponential smoothing, neural network algorithms, regression methods, and grey forecasting methods and found that the positive weight combination forecasting method is closer to the true value [17]. Based on the grey model, support vector machine, BP neural network, RBF neural network, and genetic neural network, Wang Xiaoping and Yan Fei established a forecasting model of cold chain logistics demand for agricultural products. It is found that the ability of the five models to analyze agricultural product cold chain logistics demand problems is ranked as follows: genetic neural network model > RBF neural network model > BP neural network model > support vector machine model > grey model. Chain logistics demand analysis has advantages [18]. In the follow-up research, they constructed their index system from the five perspectives of agricultural product supply, social economy, cold chain development, humanities, and logistics demand scale. Using the global search capability of genetic algorithm, they constructed BP neural network to forecast agricultural cold chain logistics demand of urban agricultural products [19]. Yang Yangwei and Cao Wei integrated traceability information and monitoring information in the supply chain and also used BP neural network to establish an early warning indicator system for fruits and vegetables [20].

The academic community attaches great importance to the research of cold chain logistics technology. Many scholars are keen on the exploration of regional cold chain logistics development strategies and pay attention to the analysis of cold chain demand. However, there are still very few quantitative researches on cold chain logistics demand forecasting, especially the analysis of demand influencing factors. Considering that cold chain logistics demand is easily affected by many factors and the statistical data of the year are limited, the research intends to find out a cold chain logistics demand forecasting model.

3. Research Methods

3.1. Backpropagation Network (BPN). The first model constructed in this research is the backpropagation neural network prediction model. Specifically, the neuralnet function in the R program is used to run the backward pass neural network algorithm to construct a predictive model.

Artificial neural network is a famous strategy to solve the problem of multitarget prediction. Neural networks contain a large number of models and learning methods, which are mainly divided into feedforward neural networks and feedback neural networks [21]. By imitating the behavioural characteristics of animal neural networks, distributed and parallel information dissemination processing is carried out, thereby constructing algorithmic mathematical models. This kind of network relies on the complexity of the system and achieves the purpose of processing information by adjusting the weight value and threshold value of the interconnection between a large number of internal nodes. The backward pass neural network is the most widely used neural network. It is a multilayer feedforward network trained by error backpropagation. The basic idea is to use the gradient descent method to search technology to make the predicted output value and expected output of the network. The mean square error of the value is the smallest. The calculation structure of neural network nodes is shown below (see Figure 1).

As shown in Figure 1, the input layer includes two input nodes (X_1, X_2), the middle layer also includes two input nodes (H_1, H_2), and the output layer has an output node (Y_1). If the node receives the input, the vector is represented by X, the network weight value vector of the node and the upper layer is represented by W, and the threshold value of the node is represented by θ, and then, the adder U_j of the j^{th} node is defined as follows:

$$U_j = \sum_{i=1}^{n} W_{ij} X_i + \theta_j. \tag{1}$$

In (1), n is the number of nodes in the upper layer, X_i is the output of the i^{th} node in the upper layer, and W_{ij} is the link weight value of the i^{th} node in the upper layer and the j^{th} node in this layer. The startup function $y = F(U)$, using the sigmoid function, is defined as

$$F\left(U_j\right) = \frac{1}{1 + e^{-U_j}}. \tag{2}$$

In the backward pass phase, the link weights between nodes will be corrected in the reverse direction according to the prediction error. After repeated corrections, the predicted output of the BPN will approach the target value. The error function formula is as follows:

$$\text{error} = \frac{1}{2} \sum_{i=1}^{n} \left(T_i - Y_i\right)^2. \tag{3}$$

In (3), T_i is the expected output value of the i^{th} output neuron in the output layer and Y_i is the predicted output value of the i^{th} output neuron in the output layer. Also, the partial differential of the error function to the weight value is as follows:

$$\Delta\omega = -n \frac{\partial \text{error}}{\partial \omega}. \tag{4}$$

In (4), $\Delta\omega$ is the correction amount of the link weight value of each layer of neurons and N is the learning rate parameter of the neural network, which mainly controls the learning speed of the BPN neural network.

3.2. Grey Prediction. Grey prediction is to identify the degree of difference in the development trend between system factors and process the original data to find the law of system changes, generate a data sequence with strong regularity, and then establish the corresponding differential equation model to predict the future development trend of things [22]. It constructs a grey prediction model with a series of quantitative values that reflect the characteristics of the predicted object observed at the same time interval to predict the feature quantity at a certain time in the future or the time to reach a certain feature quantity [23]. In order to weaken the randomness of the original time series, before establishing

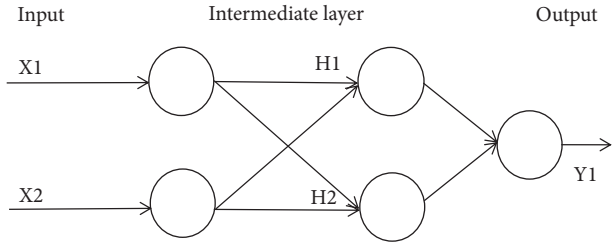


Figure 1: Neural network node calculation structure diagram.

the grey forecast model, the original time series must be processed with data. The time series after data processing is called the generated column. The data processing methods commonly used in grey systems include accumulation and accumulation. The steps required to run the grey prediction model are as follows.

Generate original modelling sequence as follows:

$$X^{(0)} = \left[X^{(0)}(1), X^{(0)}(2), \ldots, X^{(0)}(n)\right]. \tag{5}$$

Calculate the accumulation to generate the next-to-average sequence as follows:

$$X^{(1)} = \left[X^{(1)}(1), X^{(1)}(2), \ldots, X^{(1)}(n)\right]. \tag{6}$$

Calculate the mean as follows:

$$Z^{(1)}(k) = \frac{1}{2}\left[X^{(1)}(k-1) + X^{(1)}(k)\right]. \tag{7}$$

Construct matrix B and matrix Y as follows:

$$B = \begin{pmatrix} -Z^{(1)}(2) \\ -Z^{(1)}(3) \\ \vdots \\ -Z^{(1)}(n) \end{pmatrix}, \quad Y = \begin{pmatrix} X^{(0)}(2) \\ X^{(0)}(3) \\ \vdots \\ X^{(0)}(n) \end{pmatrix}. \tag{8}$$

Solve the model accuracy (a, b), where a is the development coefficient and b is the grey effect as follows:

$$\begin{pmatrix} a \\ b \end{pmatrix} = \left(B^T B\right)^{-1} B^T Y. \tag{9}$$

Bring (a, b) into the improved grey prediction model formula and calculate $k = 1, 2, 3, \ldots, n$ as follows:

$$\widehat{X}^{(2)}(k+1) = \left(1 - e^{-a}\right)\left[X^{(0)}(1) - \frac{b}{a}\right]e^{-ak}. \tag{10}$$

The relative error of the grey prediction model is as follows:

$$e(\widehat{k}) = \left|\frac{X^{(0)}(k) - \widehat{X}(k)}{X^{(0)}(k)}\right|. \tag{11}$$

The average relative error of the grey prediction model is as follows:

$$\varepsilon = \frac{1}{n}k = \sum_{k=1}^{n} |\varepsilon(k)||. \tag{12}$$

Generally, the average prediction accuracy should be higher than 80%. If the prediction accuracy is above 90%, it means that the prediction effect of the model is better.

4. Prediction Modelling

4.1. Data Source. The data for this study come from the official website of the National Bureau of Statistics of China, the Cold Chain Committee of China Logistics Alliance, and iiMedia Data Center (see Table 2). The time span of the data is from 2013 to 2019. When establishing the neural network model, the input indicators include per capita disposable income of residents and per capita consumption expenditure of residents, and the output indicator is the scale of China's cold chain market.

In the past literature, the production of fresh agricultural products was usually regarded as the demand for cold chain logistics directly. Although this definition has its rationality and operability, it also has certain defects because it ignores the situation that the target market and the original produce place are the same. The sales destination of fresh agricultural products can be mainly divided into urban and rural areas. Fresh agricultural products sold in urban areas do require cold chain logistics, but some agricultural products sold in rural areas are sold to surrounding rural areas near the place of origin. The transportation distance is not far, and the transit time is very short. In that case, the demand for the cold chain logistics of these fresh agricultural products is not so strong. In addition, some fresh agricultural products are even sold directly at the where they were grown, and this part has no effective demand for cold chain logistics in fact. Moreover, there is no clear correlation between the total product changes in fresh agricultural products and the regional distribution ratio of product sales. That is to say, when the output of fresh agricultural products increases, we cannot judge that the new output is transported more to cities and towns, transported to the surrounding rural areas, or retained in the place of origin for self-selling. If fresh agricultural products are used as the demand for cold chain logistics, it cannot fully reflect the real situation of cold chain demand and is not accurate enough. Therefore, this article takes the variable of China's cold chain market size as a proxy indicator of cold chain demand. There is a significant positive correlation between demand and market size, and the main purpose of predicting future demand is to accurately understand the development trend of cold chain logistics. The market size is also an important measurement factor reflecting the development trend of the industry. Therefore, the scale of China's cold chain market is more reasonable as a proxy indicator of cold chain demand. In the construction of the two models, the variable of China's cold chain market size was used. However, in addition to the expected output variables, the neural network algorithm also

Table 2: Related indicators of cold chain logistics demand forecasting model.

Category	2013	2014	2015	2016	2017	2018	2019
Per capita disposable income (ten thousand yuan)	1.83	2.02	2.20	2.38	2.60	2.82	3.07
Per capita consumption expenditure (ten thousand yuan)	1.32	1.45	1.57	1.71	1.83	1.98	2.16
Cold chain market size (100 billion yuan)	1.26	1.5	1.8	2.25	2.55	3.035	3.78

needs other exogenous influencing factors as input variables, so this article selects the per capita disposable income of residents and per capita consumption expenditure of residents and they are used as proxy indicators. This operation is because it is considered that the role of cold chain logistics is mainly to ensure the quality of fresh food, which reflects the residents' pursuit of quality life, and at the same time, these two indicators reflect the income level and consumption concept of residents to a large extent. It is the economic foundation and endogenous driving force for residents to pursue a quality life. Therefore, the per capita disposable income and per capita consumption expenditure of residents can be used as important factors affecting the demand for the cold chain.

4.2. Model Based on Neural Network Algorithm. According to the mathematical logic of the neural network and the rules of the R language, we write the neural network model algorithm program statement for data mining and substitute the relevant index values into the program to run. Among them, the data from 2013 to 2018 are used to train and revise the model, and the latest data of 2019 are used to test the difference and accuracy between the predicted value and the actual value. The complete R language program algorithm is as follows:

```
install.packages ("neuralnet")
library (neuralnet)
traininginput1 < -as.data.frame (c (1.83, 2.02, 2.20, 2.38, 2.60, 2.82))
traininginput2 < -as.data.frame (c (1.32, 1.45, 1.57, 1.71, 1.83, 1.98))
trainingoutput < -as.data.frame (c (1.26, 1.5, 1.8, 2.25, 2.55, 3.035))
trainingdata < -cbind (traininginput1, traininginput2, trainingoutput)
trainingdata
colnames (trainingdata) < -c ("Input1", "Input2", "Output")
trainingdata
f = Output ~ Input1 + Input2
net.sqrt < -neuralnet (f, trainingdata, hidden = c (5, 3), rep = 50, algorithm = 'rprop+', threshold = 0.001, linear.output = T)
print (net.sqrt)
plot (net.sqrt)
testdata1 < -as.data.frame (c (3.07))
testdata2 < -as.data.frame (c (2.16))
testdataout < -as.data.frame (c (3.78))
testdata < -cbind (testdata1, testdata2)
colnames (testdata) < -c ("testdata1", "testdata2")
testdata
net.results < -compute (net.sqrt, testdata)
cleanoutput < -cbind (testdata1, testdata2, testdataout, as.data.frame (net.results$net.result))
colnames (cleanoutput) < -c ("testdata1", "testdata2", "Expected Output", "Neural Net Output")
print (cleanoutput)
```

The above algorithm codes are set to establish the neural network. The parameter "Hidden" determines the number of intermediate layers of the neural network. This study controls a multilayer neural network structure, so the *c* function "hidden = c (5, 3)" is used, which means that the middle layer has two layers: the first layer has 5 neurons, and the second layer has 3 neurons. The parameter "threshold" sets the stop threshold of the neural network error function. This study sets it to 0.001, which means the algorithm stops running when the neural network error drops to 0.001. The parameter "rep" neural network training is repeated 50 times. Generally speaking, the more the repetitions, the higher the prediction accuracy. However, since the upper limit of the image processing ability of the R 4.0.2 is 63, that is, rep <63, and after testing, it is found that the prediction results of training repetitions higher than 50 are not ideal, the prediction accuracy in the interval [50, 63] does not increase but decreases, so this study sets the training repetition accuracy to 50. The parameter "algorithm" calculates the algorithm of the neural network. The algorithm includes backprop, rprop+, rprop-, etc. However, due to the poor performance of the backprop test in the neuralnet function, this study uses elastic backprop to construct the neural network prediction model. The parameter "linear.output" sets whether the result is a continuous numeric output, if it is set to "TRUE", the output is numeric; otherwise, it is binary output. In this study, the numerical output is set, so it is set to "TRUE." After the analysis of the above R language program, the neural network structure diagram (see Figure 2) is finally obtained as follows.

In the neural network structure diagram, the value on black line represents the link weight value between neurons, and value on blue line represents the error term added to the blue line at each step during the neural network fitting process. It can be seen from the figure that a total of 21198 steps were taken to construct the model, and the average error term was only 0.0019. After running the codes in R for more than 50 times, we find that the output result of the neural network algorithm once reached 3.60, which has a

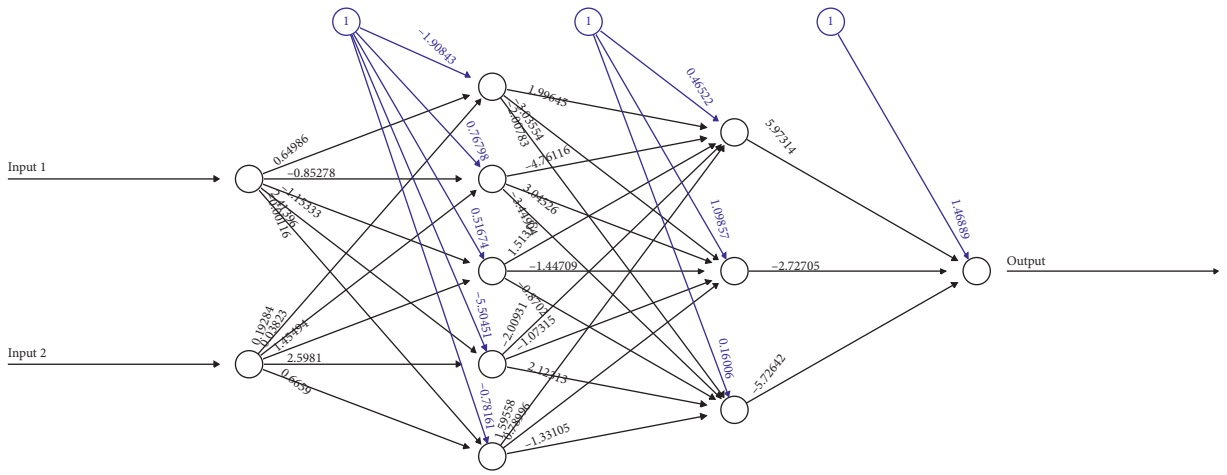


Figure 2: Neural network algorithm structure.

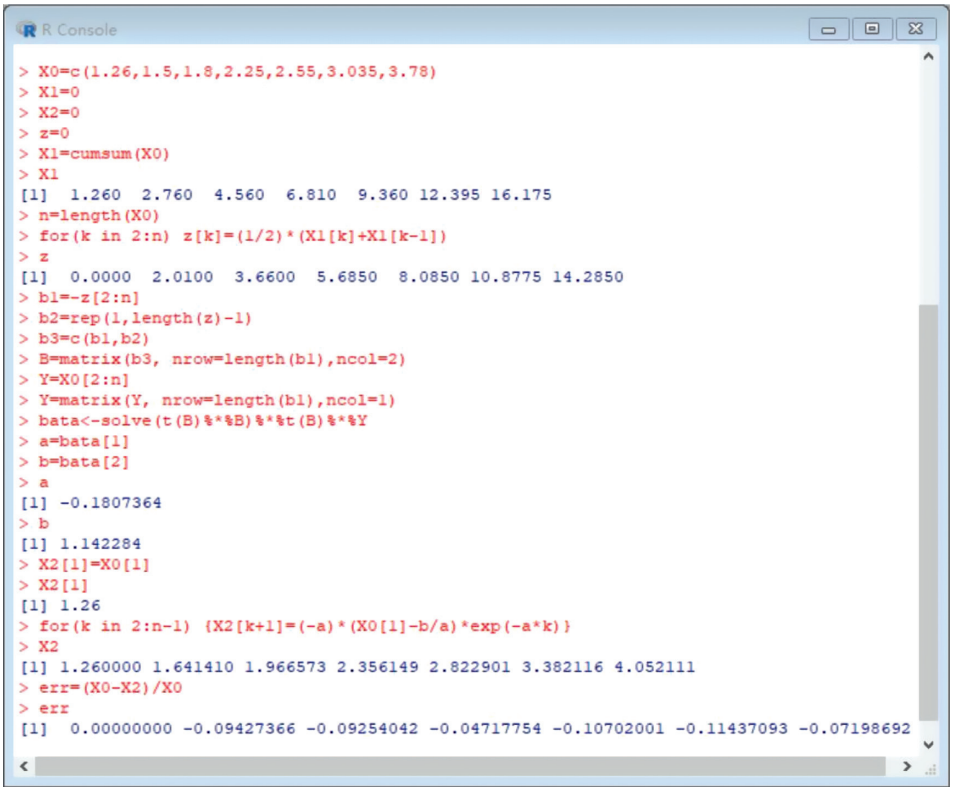


Figure 3: Grey prediction modelling process in R.

very small error value from the expected input; that is, the true value is 3.78. It can be seen that the backward pass neural network has an effective predictive ability.

4.3. Model Based on Grey Prediction. According to the R language rules and the mathematical logic of grey prediction, a grey prediction program is written, and the required data are substituted into the program to run. The specific program and process screenshots are as follows:

$X_0 = \mathrm{c}\ (1.26, 1.5, 1.8, 2.25, 2.55, 3.035, 3.78)$

$X_1 = 0$

$X_2 = 0$

$z = 0$

$X_1 = \mathrm{cumsum}\ (X_0)$

$n = \mathrm{length}\ (X_0)$

$\mathrm{for}\ (k\ \mathrm{in}\ 2:n)\ z\ [k] = (1/2)^*(X_1\ [k] + X_1\ [k-1])$

$b_1 = -z\ [2:n]$

$b_2 = \text{rep}\ (1,\ \text{length}\ (z)-1)$

$b_3 = \text{c}\ (b_1,\ b_2)$

$\text{B} = \text{matrix}\ (b_3,\ \text{nrow} = \text{length}\ (b_1),\ \text{ncol} = 2)$

$Y = X_0\ [2:n]$

$Y = \text{matrix}\ (Y,\ \text{nrow} = \text{length}\ (\text{b1}),\ \text{ncol} = 1)$

bata < -solve (t (B) %*% B)%*% t (B) %*% Y

$a = \text{bata}\ [1]$

$b = \text{bata}\ [2]$

$X_2\ [1] = X_0\ [1]$

for $(k$ in $2:n-1)$ $\{X_2\ [k+1] = (-a)^*(X_0\ [1]-b/a)^*\exp(-a^*k)\}$

$\text{err} = (X_0-X_2)/X_0$

As shown in Figure 3, the demand for cold chain logistics in 2019 obtained by the grey forecast model is about 4.052. Compared with the real value of 3.78, the error is −0.072. Although the absolute value of the error is very small, it is still slightly larger than that of neural network. The error of the prediction result of the network algorithm model is −0.048. In other words, the neural network algorithm model has a higher accuracy and better forecasting results for the demand of cold chain logistics.

5. Conclusions

This paper analyses the problems existing in cold chain logistic in Chinese market and constructs the demand prediction models based on data mining. Neural network algorithm and grey prediction are applied in the R language program to study the cold chain logistics demand. The results show that the accuracy of grey prediction model and the neural network algorithm is very high, and the prediction effect is significant, while the neural network prediction model has smaller errors and slightly higher accuracy than the grey prediction model. Therefore, neural network algorithm model is more suitable for predicting demand of cold chain logistic management.

As for the positive contribution of this paper, we find out the better one between the two mathematical model based on data mining. According to the findings, in cold chain logistics and supply chain management, the use of neural network algorithm models for demand forecasting should be more popularized, and more appropriate mathematical models should be used for data mining to provide forward-looking guidance for production. First, it can guide the production scale of farmers who grow fresh agricultural products and ensure that the supply of goods is abundant and does not cause too much waste; the second is to help logistics companies clarify the budget for frozen preservation equipment and staffing, such as the number of refrigerated trucks and cold storage, and guarantee that the cold chain logistics and transportation capabilities will not cause too much financial burden to the enterprise; the third is to promote the maintenance and upgrade of the relevant information management system of the enterprise and let the enterprise clarify the investment in data management. In short, the forecast of cold chain demand is very important. It can have a relatively clearer understanding of the future cold chain logistics demand scale and development trend, provide forward-looking guidance for logistics management, and better ensure the balance of supply and demand.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Social Science Foundation of China (NSSFC) under grant no. 17BGL238.

References

[1] X. Zhang, "Study on economic characteristics, difficulties and countermeasures of cold chain logistics of agricultural products in China," *Modern Economic Research*, vol. 12, pp. 100–105, 2019.

[2] Y.-P. Liu, "Research on the problems in the agricultural cold chain logistics development of China and its countermeasure," *Prices Monthly*, vol. 9, pp. 60–63, 2018.

[3] L. Dong, "The development trend, problems and countermeasures of China's cold chain logistics-from the perspective of the development of the commercial," *Circulation Industry*, vol. 12, pp. 94–96, 2019.

[4] K. Leng, L. Jing, and L. Pan, "Study on the status quo and problems of cold chain logistics development in Hubei province under the new normal economy," *Theories Monthly*, vol. 5, pp. 125–131, 2017.

[5] M. Guo and Li Hong, "Evaluation of the development level of agricultural cold chain logistics based on data analysis of 12 typical provinces and cities," *Business Economics Research*, vol. 1, pp. 125–127, 2019.

[6] G. Fan, "Analysis on the influencing factors of the development of agricultural products cold chain green logistics," *Agricultural Economy*, vol. 4, pp. 143-144, 2019.

[7] Y.-J. Tian, Q.-H. Xie, and Z.-H. Wang, "Safety assessment of fresh agricultural products cold chain logistics based on AHP-entropy weight method," *Storage and Process*, vol. 19, no. 5, pp. 185–190, 2019.

[8] W.-J. Cao and H. X. Hao, "Analysis and promotion strategy of cold chain logistics distribution efficiency factors based on system dynamics," *Science and Technology Management Research*, vol. 38, no. 14, pp. 217–223, 2018.

[9] J. Wang and H. Li, "Research on the role of intermediate organizations in the cold chain logistics of agricultural products from the perspective of time and space," *Journal of Beijing Jiaotong University*, vol. 18, no. 2, pp. 119–128, 2019.

[10] W.-F. Zhang and K.-H. Liang, "Optimization of cold-chain network nodes and delivey for fresh agricultural products," *Systems Engineering*, vol. 35, no. 1, pp. 119–123, 2017.

[11] Z. Zhao, X. Li, X. Zhou, and C. Liu, "Research on green vehicle routing problem of cold chain distribution: considering traffic congestion," *Computer Engineering and Applications*, vol. 56, no. 1, pp. 224–231, 2020.

[12] Q. Zhou and S. Fu, "Research on intelligent cold chain logistics integrated prevention and control technology," *System*, vol. 40, no. 13, pp. 196–201, 2020.
[13] Y. Yao and S. He, "Research on optimization of distribution route for cold chain logistics of agricultural products based on traffic big data," *Management Review*, vol. 31, no. 4, pp. 240–253, 2019.
[14] F. Wen-Ting, A. I. Shi-Zhong, Q. Wang, and J.-B. Fan, "Research on cold chain logistics distribution path optimization based on hybrid ant colony algorithm," *Chinese Journal of Management Science*, vol. 27, no. 11, pp. 107–115, 2019.
[15] Z. Yao and Z. Yi, "Research on the optimization of cold chain logistics distribution route under Internet of things and low carbon," *Ecological Economy*, vol. 36, no. 2, pp. 61–66, 2020.
[16] X. Wang, "Prediction of demand trend of agricultural products cold chain logistics based on weight distribution combination method," *Statistics and Decision*, vol. 34, no. 9, pp. 55–58, 2018.
[17] J. Yuan, "Forecast of agricultural products cold chain logistics demand under the forward weight combination forecasting mechanism," *Jiangsu Agricultural Sciences*, vol. 45, no. 19, pp. 341–346, 2017.
[18] X.-P. Wang and F. Yan, "Logistic demands and forecasting of agriculture cold chain serving Beijing, Tianjin and Hebei province," *Fujian Journal of Agricultural Sciences*, vol. 33, no. 8, pp. 870–878, 2018.
[19] X.-P. Wang and F. Yan, "The demand forecast of Beijing urban agricultural products cold chain logistics based on GA-BP model," *Mathematics in Practice and Knowledge*, vol. 49, no. 21, pp. 17–27, 2019.
[20] W. Yang and W. Cao, "The application of BP neural network in the early warning of fruit and vegetable cold chain logistics," *Computer Engineering and Science*, vol. 37, no. 9, pp. 1707–1711, 2015.
[21] S. Han and Na Liu, "Evaluation of the competitiveness level of my country's logistics enterprises," *Business Economics Research*, vol. 15, pp. 101–104, 2019.
[22] L. I. Xia-pei, "Forecast of agricultural products logistics demand based on grey linear combination model," *Journal of Beijing Jiaotong University*, vol. 16, no. 1, pp. 120–126, 2017.
[23] X. I. E. Huang-Ming, L. Sun, and W. Liu, "The layout planning of cold storage based on the origin type of cold chain logistics market," *Storage and Process*, vol. 17, no. 5, pp. 120–128, 2017.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 2343860, 12 pages
https://doi.org/10.1155/2020/2343860

Research Article

Study of the Influencing Factors on Development of Ports in Guangdong, Hong Kong, and Macao from the Perspective of Spatial Economics

Lianhua Liu and Hai Ping

Department of Business Administration, Huashang College Guangdong University of Finance & Economics, Guangzhou 510000, China

Correspondence should be addressed to Lianhua Liu; lianhua16@126.com

Received 16 August 2020; Revised 14 October 2020; Accepted 7 November 2020; Published 27 November 2020

Academic Editor: Wen Tsao Pan

Under the background of "one belt, one road initiative" and the "Outline Development Plan for the Guangdong-Hong Kong-Macao Greater Bay Area," Guangdong, Hong Kong, and Macao port group development is facing new opportunities. The port group has a large throughput, covering the hinterland with dense population and high economic density, excellent transportation infrastructure, and broad consumption market. However, the port group internal competition is fierce, and the development level is different in space. Based on this, this paper uses the HHI index and spatial economics as research method and aims to study the spatial evolution characteristics and influencing factors of Guangdong-Hong Kong-Macao port group. Firstly, the HHI index is used to describe the aggregation status. Secondly, the development level index according to port throughput and container throughput is constructed. The spatial development and evolution process of the port group of Guangdong, Hong Kong, and Macao is analyzed by combining with spatial econometrics and economic geography. In summary, the influencing factor "diamond model" is constructed and took empirical research to verify its rationality and scientificity. The empirical results show a strong spatial correlation between the development of the Guangdong, Hong Kong, and Macao port group. The government intervention and the development of port industry have a negative correlation, and this impact will be weakened over time. There is a negative relationship between the level of marketization and the development of the port industry. There is a multiple and complex relationship between the port auxiliary industry and the development of the port industry, and it shows the short-term influence is smaller than the long-term impact. The level of port transportation infrastructure, the port industry competition, and the economic openness have significant and positive effects on the development of the port group.

1. Introduction

Under the background of "one belt, one road initiative" and the "Outline Development Plan for the Guangdong-Hong Kong-Macao Greater Bay Area," Guangdong, Hong Kong, and Macao port group development is facing new opportunities. The port group of Guangdong, Hong Kong, and Macao is an important distribution point for the import and export of goods in the South Asia international logistics corridor of Sichuan, Guizhou, Guangxi, and Guangdong and also an important water node connecting Guangdong, Jiangsu, Anhui, and Jiangxi provinces. The port group of Guangdong, Hong Kong, and Macao has a large throughput. There are three top ten ports in the world in terms of container throughput in 2018, including the Hong Kong port, Shenzhen port, and Guangzhou port. The port group of Guangdong, Hong Kong, and Macao covers the hinterland with high population and economic density and has excellent transportation infrastructure and broad consumption market. However, there are great differences in the internal development level in space, and the competition within the port group is fierce.

Since the documentation of "development planning outline of the Guangdong-Hong Kong-Macao Bay Area" in

2019, this paper studies the spatial and temporal evolution characteristics and influencing factors of the spatial pattern of the Guangdong-Hong Kong-Macao port group, explores the effective measures to alleviate the spatial development differences of the Guangdong-Hong Kong-Macao port group, studies the influencing factors of its development spatial evolution, and puts forward effective development suggestions, which is conducive to promoting the coordinated development of ports in the region and improving the Guangdong-Hong Kong-Macao Greater Bay Area development level.

2. Literature Review

2.1. Theory of Port Development Stage. The formation and development of ports is the result of the development of regional economy and world economy. Based on the theory of spatial economics, this paper discusses the spatial distribution of the port industry, which provides an action theory perspective for comprehensively and scientifically explaining the development of the port industry. Scholars have studied the spatial distribution, network structure, spatial evolution mechanism, and driving factors of ports and put forward a six-stage model of port spatial evolution which covers six stages: isolated port development; route penetration stage and port concentration stage; branch line interconnection stage; hinterland traffic sustainable development stage; hinterland node concentration stage; national trunk line formation stage [1]. The port system has to go through five stages: preparation period, adoption period, concentration period, hub center period, and marginal challenge period. In the first four stages, the basic direction of spatial structure evolution of the port system is centralization [2].

2.2. Research Theory of Port Development Spatial Pattern. Spatial economics theory, Gini coefficient, Herfindahl-Hirschman index, and Theil Index are widely used in the study of spatial characteristics and structural evolution of ports [3–8]. For example, the Gini coefficient and HerfindahlHirschman index are used to study the spatial structure and competition pattern of China's container port system [3]. The Theil Index is used to study the spatial differences and evolution of port logistics economy in the Yangtze River Delta. Traffic location, related industries, and development policies are the main influencing factors of spatial differences and evolution [4]. The spatial econometric model is introduced to study the spatial spillover and threshold effect of port efficiency in the Yangtze River Delta region [5]. From the overall and local perspectives, the distribution degree and spatial structure evolution trend of the main cargo types of the port group in the Yangtze River Delta are analyzed [6]. The spatial distribution characteristics of the ports along the maritime Silk Road by are also analyzed using global spatial autocorrelation and local spatial autocorrelation [7]. Based on HHI index and spatial autocorrelation theory, the spatial structure and evolution of China's coastal container port system are studied [8].

The influencing factors of port development are few of the important contents of port development research. Relevant research shows that the main influencing factors of port development include the economic development degree of the port hinterland, port location, logistics infrastructure, aggregation of enterprises in the logistics industry, and external policy environment. The spatial linkage of port and hinterland is mainly affected by the adsorption function of port, the "water gathering" function of traffic network, the sea driving function of hinterland, and the market and policy guidance function [9]. Port logistics and regional economy coordinated the development relevance and spatial differences [10]. The better the economic development, the higher the development degree of logistics, port input-output synergies, and industrial ecology synergies [11–13]. Different regional economic development degrees form the port spatial economic pattern [14]. Logistics infrastructure affects the formation and cost of logistics network [15, 16], forming different spatial distribution of ports, and the aggregation degree of logistics enterprises is one of the main factors affecting the development of ports [17]. The main factors of the spatial evolution of port hinterland include geographical location, comprehensive scale of port city, foreign trade environment, port infrastructure, and efficiency [18, 19].

To sum up, the current research on the spatial characteristics and agglomeration of port industry mainly focuses on the quantitative analysis of agglomeration phenomenon. According to the situation of spatial development characteristics of ports in some river basins, the aggregation degree and reasons are analyzed and discussed. In terms of research methods, there is a lack of in-depth quantitative research. The existing research does not take the spatial correlation into the empirical study of port industry agglomeration. The agglomeration of the port industry is a geographical spatial phenomenon in the process of regional economic evolution. The spatial spillover effect will be ignored if the traditional OLS regression method is used to explain the agglomeration difference of the port industry. Based on the study of the spatial correlation of port clusters in Guangdong, Hong Kong, and Macao, this paper establishes an analysis framework of influencing factors of the port industry agglomeration from the perspective of regional evolution. Based on the port panel data from 2007 to 2017, this paper analyzes the impact on the agglomeration of port clusters in Guangdong, Hong Kong, and Macao.

3. Research Objects and Methods

3.1. Research Objects and Data Sources. Based on the port statistical yearbook, this paper breaks through the limitation of administrative region, and based on the concept of port group, this paper takes the port group of Guangdong Province, Hong Kong, and Macao as the research object. Based on the theory of port group, the Guangdong, Hong Kong, and Macao port group is divided into four groups, named the Pearl River Delta port group, including Guangzhou port, Shenzhen port, Hong Kong port, Macao port, Dongguan port, Huizhou Port, Jiangmen port, Zhongshan port, and Zhaoqing port. The Eastern

Guangdong port group includes Shantou port, Chaozhou port, Jieyang port, and Shanwei port. The Western Guangdong port group includes Zhanjiang port, Maoming port, Yangjiang port, and Yunfu port. The Inland river ports include Qingyuan port, Shaoguan port, Heyuan port, and Meizhou port. The relevant data from 2008 to 2017 are obtained from China port yearbook and statistical yearbook by city as the basic data for the study. The development level of ports in Guangdong, Hong Kong, and Macao is reviewed, and the spatial structure and evolution mechanism of the development and the influencing factors of the spatial evolution of development are studied.

3.2. Research Methods

(1) *HerfindahlHirschman index*. There are many methods to measure the industrial agglomeration, such as spatial Gini coefficient, location entropy, HerfindahlHirschman index, and Theil Index. Taking into account the characteristics of the port industry, this paper selects the HHI as the measurement of the regional concentration of ports in Guangdong, Hong Kong, and Macao. HHI is an important index to measure the specialization of industrial market, which can fully compare the agglomeration degree of port industry in different regions. The formula as follows:

$$\mathrm{HHI} = \sum_{i=1}^{n} \frac{C_i}{C}, \tag{1}$$

where "n" is the number of ports, "C_i" is the cargo throughput of port i, "C" is the total cargo throughput of the port system, and "HHI" is the HerfindahlHirschman index of the port system ($-1 < \mathrm{HHI} < 1$). The higher the HHI value, the more unbalanced the cargo throughput in the port system is, and the spatial structure of the port system tends to be centralized. On the contrary, it indicates that the port industry lacks agglomeration and the spatial structure of the port system tends to be decentralized.

(2) *Spatial Correlation Measurement Method* Spatial correlation refers to the spatial interdependence, mutual restriction, and interaction between things and phenomena in different regions, that is, the dependence between observation value and location.

The Moran index is used to measure the spatial correlation, which reflects the similarity of attribute values of spatial adjacent or spatial adjacent regional decision-making units and measures the agglomeration effect of regional decision-making units. Its calculation formula is as follows:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \overline{x})}{S^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}, \tag{2}$$

where $S^2 = (\sum_{i=1}^{n} (x_i - \overline{x}) / n)$ is the sample variance, W_{ij} is the (i, j) of the spatial weight matrix, which is used to measure the distance between port i and port j, and $\sum_{i=1}^{n} \sum_{j=0}^{n} W_{ij}$ is the sum of the spatial weights among all ports.

The value range of the Moran index is between $[-1, 1]$. If the economic activity of different regions shows spatial positive correlation, its value will be larger; otherwise, it will be smaller.

According to the calculation results of the Moran index, the standard normal distribution method is used to test whether there is spatial correlation among regions. Its calculation formula is as follows:

$$Z(I) = \frac{\mathrm{Moran}'s - E(I)}{\sqrt{\mathrm{VAR}(I)}}. \tag{3}$$

According to the value of "Z", we can judge whether there is spatial autocorrelation between Guangdong, Hong Kong, and Macao ports. When $Z \geq 0$ and significant, it indicates that there is a positive spatial correlation between the ports of Guangdong, Hong Kong, and Macao, which is characterized by spatial homogeneity and positive spatial spillover effect. When $Z \leq 0$ and significant, it indicates that there is a spatial negative correlation between ports in Guangdong, Hong Kong, and Macao, which shows spatial heterogeneity. When $Z = 0$, the ports of Guangdong, Hong Kong, and Macao present a random distribution.

(3) *Spatial Lag Model* (SLM). It is given as follows:

$$Y = \rho WY + X\beta + \varepsilon, \tag{4}$$

$$\begin{aligned} Y &= X\beta + \varepsilon, \\ \varepsilon &= \lambda W\varepsilon + \mu, \end{aligned} \tag{5}$$

where "Y" is the dependent variable, "X" is the matrix of exogenous explanatory variables, "ρ" is the spatial autoregressive coefficient, which reflects the influence degree and direction of the adjacent observed values on the local observation values, "WY" is the spatial lag dependent variable, and "ε" is the random error term.

(4) *Spatial Error Model (SEM)*. it is given as follows:

where "ε" is the random error term, "λ" is the spatial error coefficient of the cross section dependent variable vector of $N + 1$, and "μ" is the random error vector obeying normal distribution. The spatial error model mainly investigates the degree of spatial dependence existing in the error disturbance term, and the degree to which the random impact of the corresponding variables in the adjacent area affects the local observation value.

(5) *Selection of SLM or SEM Models* through the Moran index test, lmerr, and lmlag and their robust forms r-lmerr and r-lmlag, we can judge whether there is a spatial correlation of the port industry agglomeration in Guangdong, Hong Kong, and Macao. On the

basis of scholars' research, the following rules are used as the basis of SLM and SEM model selection. Lmlag is more significant than lmerr, r-lmlag is significant, r-lmerr is not significant, and so the SLM model is suitable. lmerr is more significant than lmlag, r-lmerr is significant, and r-lmlag is not significant, so SEM model is suitable.

4. Change of the Concentration Degree of Ports in Guangdong, Hong Kong, and Macao

Based on the port cargo throughput and the formula (1), the HerfindahlHirschman index of the Guangdong, Hong Kong, and Macao port group is shown in Table 1. The HHI index value of Guangdong, Hong Kong, and Macao ports decreased from 0.1719 in 2008 to 0.1269 in 2017. This change trend shows that the development of the port group system in Guangdong, Hong Kong, and Macao is increasingly dispersed, and there is spatial spillover effect.

As shown in Figure 1, the HHI values of the four port systems in Guangdong, Hong Kong, and Macao have different spatial distribution trends. The overall HHI value of the Guangdong, Hong Kong, and Macao port group is lower than the internal major groups, which indicates that the concentration degree of spatial distribution tends to be decentralized, but the local area tends to be concentrated. The HHI index of eastern Guangdong port group has a very obvious downward trend, while that of Pearl River Delta port group decreases slowly. The HHI index of the western Guangdong port group fluctuates and shows an upward trend. The HHI index of inland river port group in mountainous area rises rapidly.

Combined with the market share, the Guangzhou port, Shenzhen port and Hong Kong port accounts for a high proportion of the market share in the whole Guangdong-Hong Kong-Macao port group. However, the total market share of the three ports has decreased year by year. The market share of the three ports was 67.14% in 2008 and 49.1% in 2017. The market share of Guangzhou port's cargo throughput was 29.54% in 2008 and decreased to 26.4% in 2017. The market share of Shenzhen port's cargo throughput was 16.9% in 2008 and decreased to 10% in 2017. Hong Kong's 20.7% fell to 12.7% from 2008 to 2017. In the port group of Guangdong, Hong Kong, and Macao, the source of goods gradually diverted from the three major ports to the surrounding ports, which led to the development of Zhuhai, Zhaoqing, Huizhou, Dongguan, and other ports, resulting in spatial spillover effect.

The HHI index of the Pearl River Delta port group decreased year by year from 0.2242 in 2008 to 0.1823 in 2017, which shows that under the background of the construction of Guangdong-Hong Kong-Macao Greater Bay Area, the division and cooperation of ports in the bay area are promoted, and the phenomenon of space spillover exists.

The HHI index of the port group in western Guangdong is relatively stable, which fluctuates between 0.5814 and 0.6498. The reason is that the market share of Zhanjiang port remains stable, ranging from 74.8% to 79.2%, which forms the spatial distribution of the port group in western Guangdong.

Among the port group in eastern Guangdong, the market share of Shantou Port decreased year by year from 73.9% in 2008 to 47.6% in 2017. The HHI index of the eastern Guangdong port group also reflects the spatial change characteristics of this cargo source from concentration to dispersion. The HHI index decreases from 0.4937 to 0.3375, which reflects the spatial trend of the eastern Guangdong port group system tends to be decentralized, and the competition among ports is intensified.

The HHI index of inland river ports in mountainous areas increased from 0.4536 to 0.9225, reflecting the spatial characteristics of highly concentrated distribution of inland river ports in mountainous areas. The reason is the rapid development of the Qingyuan port. The Qingyuan port is an important port in the Beijiang River Basin. It has rich natural resources and faces the vast market of the Pearl River Delta. Besides, it is close to Guangzhou and has certain geographical advantages. The Qingyuan port market share of inland river port group increased from 70.3% in 2008 to 96% in 2017, forming the phenomenon of high spatial agglomeration of the inland river port group in inland area.

5. Spatial Evolution Characteristics of Port Group in Guangdong, Hong Kong, and Macao

5.1. Development Level Index of Port Group in Guangdong, Hong Kong, and Macao. In order to further explore the spatial characteristics and evolution characteristics of the port development in Guangdong, Hong Kong, and Macao, this paper constructs the port development capacity index system to study the spatial distribution of the port group development in Guangdong, Hong Kong, and Macao. The agglomeration level index of port development is a comprehensive study on the development of port logistics. From the perspective of comprehensiveness and availability, based on the two indicators which are port container throughput X_1 (teu) and port cargo throughput of X_2 (ton) as the port development capacity index of ports in Guangdong, Hong Kong, and Macao. The port development capacity index is constructed to reflect the development capacity of Guangdong, Hong Kong, and Macao ports. The mean-square difference method is used to determine the port development capacity index of Guangdong, Hong Kong, and Macao. Firstly, normalization is used to normalize the original data of the indicators, and the weights of the two indicators are obtained by means of mean-square deviation normalization, indicating the dispersion degree of the data samples under a certain index. Finally, the port development level index of Guangdong, Hong Kong, and Macao is obtained by weighted sum, as shown in Table 2.

5.2. Evolution of Spatial Pattern of Port Clusters in Guangdong, Hong Kong, and Macao. In order to study the evolution process of the spatial pattern of ports in Guangdong, Hong Kong, and Macao, this paper selects the port development level indexes of 2008, 2012, and 2017 for spatial processing.

Table 1: HHI values of the Guangdong, Hong Kong, and Macao port group and subport group.

Port system	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Total ports	0.1719	0.1672	0.1523	0.1441	0.1374	0.1381	0.1267	0.1263	0.1265	0.1269
Pearl River Delta port group	0.2242	0.2240	0.2046	0.1960	0.1900	0.1955	0.1787	0.1815	0.1815	0.1823
Eestern Guangdong port group	0.6498	0.6474	0.6127	0.6131	0.6001	0.5814	0.6006	0.6019	0.6329	0.6342
Eastern Guangdong port group	0.4937	0.4851	0.4167	0.3947	0.4000	0.3898	0.3830	0.3676	0.3480	0.3375
Inland Guangdong river port	0.4536	0.5810	0.6344	0.6338	0.5596	0.6639	0.8289	0.8555	0.8752	0.9225

Source: according to the calculation criteria.

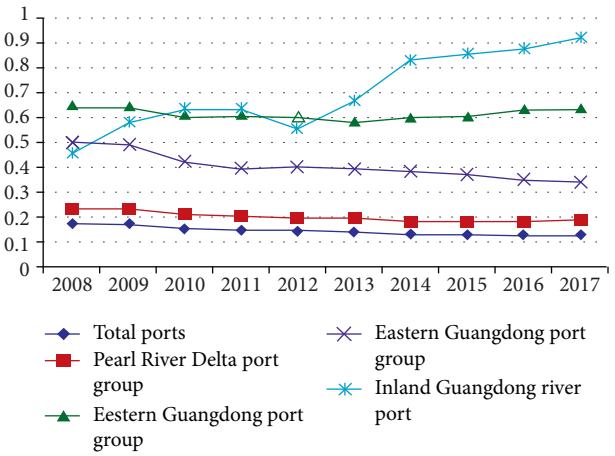


Figure 1: HHI value change chart.

Table 2: Guangdong-Hong Kong-Macao port development level index (2008–2017).

Port	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Guangzhou	0.2263	0.2381	0.2224	0.2120	0.2310	0.2360	0.2329	0.2449	0.2511	0.2597
Shenzhen	0.2614	0.2496	0.2520	0.2056	0.2490	0.2472	0.2378	0.2377	0.2281	0.2026
Zhuhai	0.0193	0.0201	0.0222	0.0242	0.0250	0.0274	0.0307	0.0330	0.0354	0.0417
Zhaoqin	0.0063	0.0072	0.0081	0.0106	0.0124	0.0125	0.0118	0.0115	0.0117	0.0191
Foshan	0.0418	0.0465	0.0407	0.0333	0.0350	0.0353	0.0350	0.0364	0.0377	0.0437
Huizhou	0.0109	0.0144	0.0148	0.0162	0.0153	0.0195	0.0146	0.0161	0.0167	0.0125
Dongguang	0.0123	0.0153	0.0194	0.1314	0.0339	0.0418	0.0488	0.0531	0.0563	0.0484
Zhongshan	0.0192	0.0118	0.0235	0.0229	0.0226	0.0266	0.0266	0.0254	0.0235	0.0254
Jianmen	0.0198	0.0194	0.0193	0.0213	0.0216	0.0230	0.0237	0.0237	0.0245	0.0243
Qingyuan	0.0015	0.0016	0.0020	0.0021	0.0020	0.0024	0.0054	0.0068	0.0070	0.0078
Maoming	0.0064	0.0074	0.0065	0.0065	0.0063	0.0062	0.0062	0.0062	0.0057	0.0053
Zhanjiang	0.0367	0.0410	0.0391	0.0435	0.0433	0.0444	0.0447	0.0485	0.0545	0.0582
Shantou	0.0157	0.0159	0.0174	0.0177	0.0213	0.0221	0.0207	0.0197	0.0192	0.0186
Chaozhou	0.0012	0.0013	0.0020	0.0025	0.0022	0.0024	0.0023	0.0023	0.0022	0.0023
Jieyang	0.0020	0.0024	0.0034	0.0041	0.0037	0.0057	0.0054	0.0056	0.0057	0.0056
Shanwei	0.0015	0.0017	0.0017	0.0016	0.0018	0.0016	0.0014	0.0018	0.0023	0.0021
Yuanfu	0.0025	0.0025	0.0032	0.0036	0.0039	0.0046	0.0047	0.0050	0.0055	0.0061
Shaoguan	0.0002	0.0001	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001
Heyuan	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006	0.0006	0.0005	0.0000	0.0000
Meizhou	0.0008	0.0005	0.0004	0.0004	0.0003	0.0003	0.0003	0.0002	0.0002	0.0002
Yangjiang	0.0008	0.0011	0.0021	0.0030	0.0037	0.0046	0.0035	0.0042	0.0042	0.0051
Mocao	0.0021	0.0015	0.0013	0.0007	0.0010	0.0010	0.0011	0.0012	0.0011	0.0010
Hong kong	0.3049	0.2942	0.2744	0.2318	0.2599	0.2483	0.2384	0.2126	0.2037	0.2064

According to the data processing principle of the natural breakpoint method, the same type of ports is divided into two groups.

According to the principle of the biggest difference among different types of ports, the classification should be based on the classification criteria of less than 0.012, 0.012–0.034, and 0.034–0.231 and greater than 0.231. It is divided into four levels: low-level development, medium-level development, medium- and high-level development, and high-level development. The port group of Guangdong,

Hong Kong, and Macao in 2008, 2012, and 2017 shows obvious spatial development differences. The spatial pattern of ports in Guangdong, Hong Kong, and Macao is characterized by polarization. As shown in Table 3, the high-level ports covered Shenzhen and Hong Kong in 2008 and 2012, and the Guangzhou port became the only high-level port in 2017. According to the data of three time sections in 2008, 2012, and 2017, the number of ports with low development level was 14, 11, and 11, respectively; the number of ports with medium development level was 4, 6, and 5, respectively; there are 3 ports, 4 ports, and 6 ports with medium and high development levels.

In 2008 and 2012, Guangzhou port was a medium and high development level port, and in 2017, it was upgraded from medium high development level to the only high development level port. In 2012, Dongguan port directly jumped from low development level to medium and high development level and maintained a good situation of medium and high development level in 2017. Zhuhai port was a port of general development level in 2012 and rose to medium and high levels in 2017. Under the fierce port competition, Shenzhen port and Hong Kong port were downgraded from high development level to medium high development level in 2017.

The development level of ports in Guangdong, Hong Kong, and Macao is highly concentrated. The development level of ports in Guangdong, Hong Kong, and Macao is mainly distributed around the Pearl River estuary waters and evenly distributed along the Xijiang River, Beijiang River, and Dongjiang River. The Pearl River estuary is the center, and Hong Kong, Shenzhen, and Guangzhou are the port centers of the whole Guangdong, Hong Kong, and Macao bay area. The eastern and western wings of Guangdong present regional centralization, forming the port group in western Guangdong with Zhanjiang as the center and Shantou as the center in eastern Guangdong.

5.3. Spatial Autocorrelation Analysis of Port Group Development in Guangdong, Hong Kong, and Macao. Spatial correlation of the port group development in Guangdong, Hong Kong, and Macao. This paper uses the development capacity index of Guangdong, Hong Kong, and Macao ports from 2008 to 2017 to calculate Moran's I index by using Geoda software, in which the spatial weight matrix adopts to the adjacent space matrix. Basis on the formula (2) and formula (3), the calculation results are shown in Table 4.

According to Table 4, Moran's I is positive and Z values of normal statistics pass the significance level test of 5%. It shows that the port development in the Guangdong, Hong Kong, and Macao port group has obvious spatial positive correlation, that is, there is a tendency of spatial dependence and homogenization. This shows that the development of ports in Guangdong, Hong Kong, and Macao ports is not randomly distributed in space and has certain similarity. Specifically, ports with high development level tend to cluster, while those with low development level tend to gather, which has obvious spatial spillover effect. Ports with high development level tend to be close to ports with high development level, and ports with low development level are close to ports with low development level. From 2008 to 2017, overall, Moran's I value showed an increasing trend, and the Z value of normal statistics showed an increasing trend, which was greater than 1.96 from 2013 to 2017. The research results show that the development of ports in Guangdong, Hong Kong, and Macao has obvious spatial autocorrelation characteristics. The Lisa clustering maps of 2008, 2012, and 2017 are shown in Figure 2.

According to the Lisa cluster diagram, this paper classifies the spatial relationship of ports in Guangdong, Hong Kong, and Macao, as shown in Table 5. High-high (HH) means that the high growth region of the port industry is surrounded by other regions with high growth; low-low (LL) means that the low growth area of the port industry is surrounded by other regions with low growth; low-high (LH) means that the low growth area of the logistics industry is surrounded by other regions with high growth; and high-low (HL) means that the high growth area of the port industry is surrounded by other regions with low growth. Not significant means there is no obvious correlation between the development space of ports. In 2008 and 2012, the aggregation of ports in Guangdong, Hong Kong, and Macao was consistent. The port areas with no significant correlation in spatial spillover included 14 ports including the western Guangdong port group and mountainous port group. Ports with certain spatial development correlation and spatial spillover effect were mainly concentrated in Pearl River Delta port group and eastern Guangdong port group.

The 2017 Lisa cluster map shows the port space spillover effect in the Pearl River Delta region is obvious. The reason is, under the policy guidance of the development of Guangdong, Hong Kong, and Macao Bay area, the coordinated development advantages of the Pearl River Delta port group are reflected, which promotes the cooperation and division of labor of the Pearl River Delta port group and forms a positive correlation relationship between the spatial development of Guangzhou, Shenzhen, Hong Kong, Foshan, and Dongguan. That is a regional synergy effect. Huizhou, Zhongshan, and Zhuhai have formed the relationship of regional competition, heterogeneity of regional development, and negative correlation of low and high spatial development. The results show that there is a certain spatial correlation in the development of Guangdong, Hong Kong, and Macao port group. Therefore, it is necessary to incorporate the spatial correlation into the influencing factors of port industry development and explore the development characteristics of Guangdong, Hong Kong, and Macao port group from the perspective of spatial economics.

6. Influencing Factor Model of the Port Development Spatial Pattern

6.1. Theoretical Framework of Influencing Factors on Spatial Pattern of Port Cluster Development in Guangdong, Hong Kong, and Macao. At present, the research on the spatial pattern of the development of ports in Guangdong, Hong Kong, and Macao has not formed a complete theoretical analysis framework. Combining with the theory of spatial

Table 3: Classification of development level of the port group of Guangdong, Hong Kong, and Macao in different periods.

Year	High level	Medium and high levels	Medium level	Low level
2008	Shenzhen, Hong Kong	Guangzhou, Foshan, Zhanjiang	Jiangmen, Zhongshan, Zhuhai, Macao, Shantou	Shaoguan, Qingyuan, Zhaoqing, Yunfu, Maoming, Yangjiang, Heyuan, Meizhou, Chaozhou, Jieyang, Shanwei, Huizhou, Dongguan
2012	Shenzhen, Hong Kong	Guangzhou, Foshan, Zhanjiang, Dongguan	Zhaoqing, Jiangmen, Huizhou, Zhongshan, Zhuhai, Macao, Shantou	Shaoguan, Qingyuan, Yunfu, Maoming, Yangjiang, Heyuan, Meizhou, Chaozhou, Jieyang, Shanwei
2017	Guangzhou	Foshan, Zhanjiang, Dongguan, Zhuhai, Shenzhen, Hong Kong	Zhaoqing, Jiangmen, Huizhou, Zhongshan, Macao, Shantou	Shaoguan, Qingyuan, Yunfu, Maoming, Yangjiang, Heyuan, Meizhou, Chaozhou, Jieyang, Shanwei

Table 4: Moran index and Z value of port group development in Guangdong, Hong Kong, and Macao.

Year	Moran's I	Sd	Z value	P value
2008	0.1798	0.1254	1.7874	0.0481
2009	0.181	0.1266	1.7781	0.0487
2010	0.1964	0.1266	1.9009	0.0452
2011	0.2872	0.1325	2.4993	0.0274
2012	0.2055	0.1277	1.9561	0.0434
2013	0.221	0.1282	2.0686	0.0417
2014	0.228	0.1286	2.1158	0.0406
2015	0.2322	0.1284	2.1522	0.0398
2016	0.2321	0.1283	2.1519	0.0397
2017	0.2232	0.1277	2.0914	0.04

Source: according to the calculation criteria.

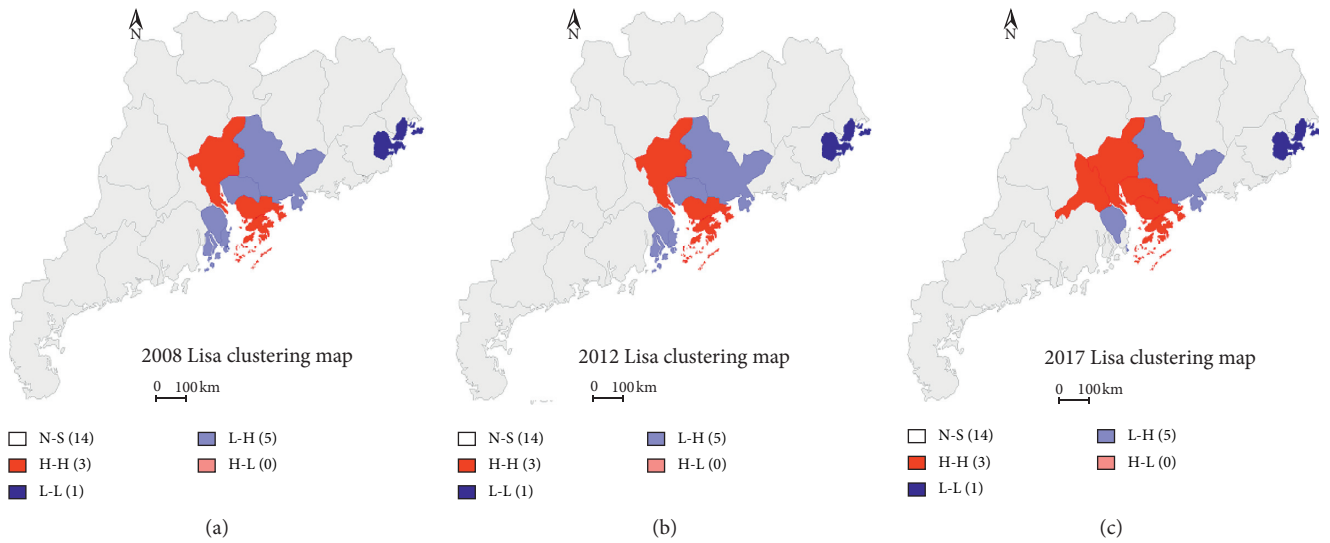


Figure 2: Lisa clustering map of the spatial pattern of the development of the Guangdong-Hong Kong-Macao port group.

Table 5: Spatial correlation models of development of ports in Guangdong, Hong Kong, and Macao in different periods.

Year	HH	LL	LH	HL	Not significant
2008	Guangzhou,	Shantou	Huizhou, Dongguan, Zhongshan, Zhuhai, Macao	None	Shaoguan, Qingyuan, Zhaoqing, Yunfu, Maoming, Yangjiang,
2012	Shenzhen, Hong Kong				Jiangmen, Heyuan, Meizhou, Chaozhou, Jieyang, Shanwei, Zhanjiang, Foshan
2017	Guangzhou, Shenzhen, Hong Kong, Foshan, Dongguan	Shantou	Huizhou, Zhongshan Zhuhai	None	Shaoguan, Qingyuan, Zhaoqing, Yunfu, Maoming, Yangjiang, Jiangmen, Heyuan, Meizhou, Chaozhou, Jieyang, Shanwei, Zhanjiang, Macao

economics and the analysis results of agglomeration characteristics, the affecting factors of the spatial agglomeration in the Guangdong, Hong Kong, and Macao port group are studied. Referring to Michael Porter's diamond model, the theoretical framework model of influencing factors of the spatial pattern of port clusters in Guangdong, Hong Kong, and Macao is constructed, which is shown in Figure 3. This model covers the port production factors like transportation infrastructure; port market development level such as social retail, GDP, and secondary industry development; the port auxiliary industry which includes the development of related industries and upstream industries; the port industry formats include the strategy, structure, and horizontal competition of port enterprises; and the port industry development opportunities and the business environment.

(1) Port transportation infrastructure level: the level of port traffic infrastructure is the basic element and resource condition of port development. It includes the physical and geographical environment, the infrastructure and equipment, and the collection and distribution system of the port. Convenient and good port production infrastructure is conducive to the cluster development of ports. The level of port transportation infrastructure can promote the efficiency of port resource allocation and improve the efficiency of regional scale economy.

Hypothesis 1: there is a positive correlation between port infrastructure level and port industry agglomeration.

(2) Development level of port market: as a logistics transportation node, the development level of the port market affects the development of the port industry. The level of port market development is mainly affected by the level of economic development and the degree of socialization of transportation services. The development and distribution of hinterland economy and manufacturing industry affect the market distribution and demand levels.

Hypothesis 2: there is a positive correlation between the level of market development and the agglomeration of the port industry.

(3) Port industry structure: the structure of the port industry includes the development cycle, development strategy, and competition situation.

Hypothesis 3: there is a positive correlation between the competitive situation of port industry and the agglomeration of the port industry.

(4) Auxiliary industry: to a certain extent, supporting industries of port industry affects the development level and agglomeration degree of ports. It mainly includes the development of related industries of port logistics, such as the cargo freight market and transportation industry. The higher the development degree of the transportation industry, the more conducive it is to the agglomeration of the port industry.

Hypothesis 4: there is a positive correlation between the development degree of port supporting industries and the agglomeration of the port industry.

(5) The opening environment of the port: the opening environment of the port refers to foreign trade development situation. The higher the degree of development of foreign trade, the stronger the demand of foreign trade logistics, and the more conducive it is to the development of the port industry, the more stable it is to the effect of port industry agglomeration.

Hypothesis 5: the opening environment is positively related to the agglomeration of port industry.

(6) Government: the policy and service of the government affect the development of the port. The policy environment of the government refers to the government's policies and systems to support the development and construction of ports. Government service environment refers to the influence of government's work efficiency, customs clearance policy, and port service ability.

Hypothesis 6: the government intervention on the port industry development presents agglomeration in the initial stage, showing a positive correlation in short-term and a long-term negative correlation.

6.2. Econometric Model of Influencing Factors of Port Cluster Development Spatial Pattern in Guangdong, Hong Kong, and Macao Port Group. According to the above theories and assumptions, this paper sets up the linear model of the influencing factors of the port cluster development spatial pattern in the Guangdong, Hong Kong, and Macao port group as follows:

$$F(Y_i) = \beta_0 + \beta_1 \mathrm{Fac} + \beta_2 \mathrm{Dev} + \beta_3 \mathrm{Com} + \beta_4 \mathrm{Transport} + \beta_5 \mathrm{Trade} + \beta_6 \mathrm{Gov} + \mu, \tag{6}$$

where "β" is the regression parameter, "i" is the port, "Y" is the dependent variable that is the port development agglomeration level index, "Fac" is the transportation infrastructure, expressed by the highway mileage and the number of port berths in the land area, "Dev" is the market development level, mainly expressed by per capita GDP, secondary industry output value, and total social retail sales, "Com" is the competitive of the port industry which is expressed by the number of transportation enterprises, "Transport" is the auxiliary industry of port industry, expressed by the total freight volume, "Trade" is the degree of opening, expressed by the total amount of foreign trade, "Gov" is the government environment, expressed by the proportion of financial expenditure in GDP, and "μ" is the error item.

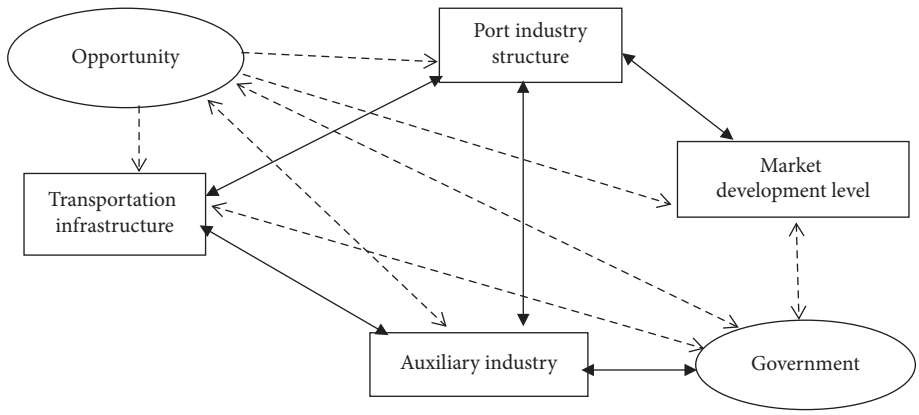


Figure 3: Theoretical model of influencing factors of spatial pattern of port development.

6.3. Empirical Analysis on Influencing Factors of Port Cluster Development Spatial Pattern in Guangdong, Hong Kong, and Macao Port Group. In order to explore the influence of the above factors in different periods on the aggregation level of port development, model I is set as the current model, the value of the dependent variable is the port level development index in 2017, and the independent variable is the value of each influencing factor in 2017, reflecting the influence degree of the current independent variable on the current dependent variable and the direction of spatial evolution. Model II is an intertemporal model. The dependent variable value is the port development level index in 2017, and the independent variable value is the value of various influencing factors in 2007, reflecting the intertemporal influence of initial variables on the current variables.

6.3.1. Estimation Results and Analysis Based on OLS Regression of Model I and Model II. Firstly, OLS regression was used to analyze model I and model II. The results are shown in Tables 6 and 7. The goodness of fit's R-squared values of model I and II were 0.9657 and 0.9511, which indicated that the regression line fitted the observed values well. The results show that the above five variables have a positive impact on the agglomeration level of the port industry, which is consistent with the theoretical hypothesis. However, the coefficient of market development level is negative, which is inconsistent with the hypothesis, which means the market development level variable can not explain the regional differences of port industry development in Guangdong, Hong Kong, and Macao. There may be omissions of important explanatory variables, or problems related to spatial analysis are not considered in the model. Therefore, we introduce the spatial econometric model to correct the problems of the OLS model.

6.3.2. Estimation Results and Analysis Based on SLM Model I and Model II. In order to further test whether the spatial autocorrelation exists in the Guangdong, Hong Kong, and Macao port group, we use the spatial lag model to further study model I and model II. Based on formula (4) and formula (5), the spatial dependence results of model I in Table 8 show that Moran index is -0.2452 which shows that the classical regression error has spatial correlation. LM (lag) and R-LM (lag) pass the significance level test of 5%; the significance level of R-LM (err) is 10%, LM(err). The err value did not pass the significance level test. It shows that SLM is more suitable for model I. Similarly, model II is more suitable for the SLM model.

The spatial effect is added to models I and II, and the current and intertemporal SLM models of the theoretical model for the development of the Guangdong, Hong Kong, and Macao port group is established. The calculation results are shown in Tables 9 and 10. Compared with the OLS model, the goodness of fit values of SLM model I and II were 0.9713 and 0.9488, which were both improved. The logarithm likelihood function (LOGL) value was greatly improved. AIC and SC values were smaller than those of the OLS model. It shows that the model considering the spatial lag effect can more scientifically explain the relationship between the development of port industry spatial evolution. Most of the parameters in the model pass the significance level test of 5%, which shows that, with the integration and development of regional transportation and the improvement of opening-up level, the development of auxiliary industries of the port industry, and the concentration of port competition, the spatial dependence of port industry is strengthened and the spatial spillover effect is obvious.

Based on the SLM model, this paper analyzes the influencing factors of port cluster development spatial pattern in the Guangdong, Hong Kong, and Macao port group.

(1) The improvement of port transportation infrastructure promotes the development of the port industry in Guangdong, Hong Kong, and Macao. The Fac coefficients of model I and model II are all positive, which indicates that the improvement of the transportation infrastructure level is conducive to the development of the port industry and can contribute to the economic spillover of port space. The Fac correlation coefficient of model I is smaller than that of model II, which indicates that the level of transportation infrastructure has a long-term effect on the agglomeration of the port industry, and the

Table 6: OLS estimation results of models I and II

Model	C	FAC	Dev	Com	Transport	Trade	GOV
Model I	−0.0125* (−1.7236)	0.0166 (0.8284)	−0.0901** (−2.5742)	8.5533** (2.2355)	1.5387*** (6.3983)	5.8404 (1.4504)	1.4456 (0.9694)
Model II	−0.0149* (−0.9452)	0.0484** (2.1726)	−0.1450** (−3.6382)	5.8031*** (7.4780)	4.1558*** (5.4361)	3.4461*** (2.3601)	−0.0001** (−2.5819)

Remarks: *, **, and ** denote 10% significance level, 5% significance level, and 1% significance level, and the figures in brackets are t-statistical value.

Table 7: OLS estimation results of models I and II

Model	R-squared	Ad R-squared	LOGL	AIC	SC
Model I	0.9657	0.9657	67.1263	−120.253	−112.615
Model II	0.9511	0.9511	63.2261	−112.152	−104.815

Table 8: Spatial dependence test of models I and II

Model	Moran I	LM (lag)	R-LM (lag)	LM (err)	R-LM (err)
Model I	−0.2452	9.8056**	9.5609**	0.3429	0.0982
Model II	0.5593	8.1953**	9.3265**	0.0155	1.1467

Remarks: *, **, and ** denote 10% significance level, 5% significance level, and 1% significance level, and the figures in brackets are t-statistical value.

Table 9: SLM estimation results of models I and II

Model	C	FAC	Dev	Com	Transport	Trade	GOV
Model I	−0.0062 (−1.4727)	0.02799** (2.437)	−0.0565** (−2.7596)	1.1981** (5.3236)	1.3104*** (9.2340)	1.6500*** (5.4089)	−8.0131 (−0.851)
Model II	−0.0029 (−0.5244)	0.04329** (3.1283)	−0.1585*** (−6.3897)	5.8132*** (12.1547)	4.7384*** (9.6236)	5.8036*** (5.4424)	−0.0002*** (−5.4554)

Remarks: *, **, and ** denote 10% significance level, 5% significance level, and 1% significance level, and the figures in brackets are t-statistical value.

Table 10: SLM estimation results of models I and II

Model	R-squared	LOGL	AIC	SC
Model I	0.9887	75.4016	−134.803	−126.075
Model II	0.9806	69.4392	−122.878	−114.151

investment in transportation infrastructure lags behind the development of the port industry, which means the impact of transportation infrastructure on the development of port industry is long term.

(2) In models I and II, there is a negative correlation between the development level of port market and the development of the port industry in Guangdong, Hong Kong, and Macao. Generally speaking, the higher the market level, the more concentrated the development of the port industry. The emergence of negative correlation is caused by the change of market development level logistics mode. With the development of manufacturing industry agility, the poor timeliness of marine transportation makes it difficult to meet the requirements of timeliness.

(3) The development cycle of the port group in Guangdong, Hong Kong, and Macao is in a concentrated period. In the models I and II, the correlation coefficient of the index is positive, and both pass the significance level test of 5%. It reflects the development trend of the port industry, with the increase of the number of enterprises in the industry; the port industry shows the characteristics of specialization, socialization, networking, and large-scale development.

(4) The relationship between port auxiliary industry and the development of port industry in Guangdong, Hong Kong, and Macao is complex, showing the short-term influence is smaller than the long-term impact. The correlation coefficient of the index is 1.1981 in model I and 5.8132 in model II, both of which pass the significance level test of 1%. Model I shows that the increasing demand of auxiliary industries related to port development, such as land transportation, will promote the development trend of port spatial agglomeration. From the intertemporal model II, because the infrastructure construction of other logistics modes has been improved, the business volume of the port industry has been greatly impacted, which is greater than that of the current period model I.

(5) The opening environment has a positive correlation with the development of the port industry. In models I and II, the correlation coefficient of the index was positive, which passed the significance level test of 5%. It shows that the strength of opening to the outside world promotes the development of

international trade, and the development of international trade promotes the demand of international freight transport, so as to improve the demand of port service and promote the further development of the port industry.

(6) Excessive government intervention will result in low efficiency of port development. The short-term impact of government intervention is greater than the long-term impact. The correlation coefficients of GOV index in models I and II are negative, the short-term correlation coefficient is −8.0131, and the long-term correlation coefficient is −0.0002.x. The negative coefficient of the index in model I is not significant, which indicates that excessive government intervention will lead to low efficiency of port development. In model II, the index is negative and passes the 1% significance test, which shows that the negative impact of government intervention has a certain lag. The reason is the investment cycle of the port industry is long, especially the construction period of infrastructure. Therefore, the negative effect of government intervention on port construction and operation is reflected in the long-term market operation.

7. Conclusions and Suggestions

In this paper, the HHI index is used to describe the aggregation status of the Guangdong, Hong Kong, and Macao port group, and the development level index is constructed according to the port throughput and container throughput. Combined with spatial econometrics and economic geography, the influencing factors of the development of the Guangdong, Hong Kong, and Macao port group are analyzed. Basic on this, the diamond model of influencing factors on the spatial pattern of the port group development in Guangdong, Hong Kong, and Macao is modeled. For the lack of spatial factors in classical economics and to explain the regional differences in the development level of the Guangdong, Hong Kong, and Macao port group, this paper introduced the partial measurement method to further study the scientific of the influencing factor model. The empirical results show that there are strong spatial autocorrelation and heterogeneity in the development of the Guangdong, Hong Kong, and Macao port group. There is a negative correlation between government intervention and the development of the port industry, and the influence of government intervention will be weakened as time goes on. The level of marketization weakens the efficiency of port development. The port auxiliary industry and the development of the port industry have many repetitions and miscellaneous relations, showing the short-term influence is smaller than the long-term impact. The reason is that the supply capacity of the auxiliary industry is limited in the near future, and the port industry plays a role in diverting the service demand of the auxiliary industry. There is a positive correlation between the level of port transportation infrastructure, the level of port industry competition, and the environment of opening to the outside world.

Based on the analysis of the empirical results, some suggestions are put forward to promote the development of the Guangdong, Hong Kong, and Macao port group.

Firstly, the development of the port industry in Guangdong, Hong Kong, and Macao has obvious spatial correlation. The suggestion is the local government should improve its own transportation infrastructure, increase the intensity of opening to the outside world, and coordinate different transportation modes from the perspective of the overall economic development of Guangdong, Hong Kong, and Macao. The integration of the regional port industry should be actively promoted, and the overall development level of the Guangdong, Hong Kong, and Macao port industry is improved.

Secondly, we should reduce the administrative intervention on the port industry, relax the control over the development of the port industry, actively change the role of the government, especially the intervention of the port and shipping administrations in the development of the port industry, do a good job in guiding the role, focus on the construction of infrastructure, improve the business environment, and provide convenient customs clearance conditions and promotional policy for the port industry.

Thirdly, we should guide and support port enterprises to actively participate in market competition, strengthen the cooperative operation mechanism between port industry and other transportation modes, and promote the restructuring of port enterprises and participate in market-oriented competition. We should do a good job in the division of labor and cooperation between different ports to avoid homogeneous competition. The overall development of the port industry should be promoted and the development of transportation infrastructure, bonded insurance, finance, legal consultation, port warehousing and logistics should be improved to build a good ecological environment for the development of the port industry.

Data Availability

The data used to support the findings of this study are available from all the corresponding author upon request.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was financially supported by the 2018 Social Science Planning Project of Guangzhou "Research on the Construction and Development of Guangzhou Smart International Shipping Center Based on the One Belt One Road Strategy" (Grant no. 2018GZGJ169); 2019 Construction of Provincial Characteristic Specialty in Guangdong Province (Grant no. HS2019CXQX04); and 2016 Humanities and Social Sciences Research Projects of Universities in Guangdong Province "Construction of key disciplines in business administration" (Grant no. 2015WTSCX126).

References

[1] E. J. Taaffe, R. L. Morrill, and P. R. Gould, "Transport expansion in underdeveloped countries: a comparative analysis," *Geographical Review*, vol. 53, no. 4, pp. 503–529, 1963.

[2] Y. Hayuth, "Rationalization and concentrati on of the U.S. container port system," *The Professional Geographer*, vol. 40, no. 3, pp. 279–288, 1988.

[3] Y. Cao, H. Li, and C. Wen, "Spatial structure and competition pattern of container port system in China," *Acta Geographica Sinica*, vol. 6, pp. 1020–1027, 2004.

[4] S. Liang, Y. Cao, W. Cao, and W. Wu, "Analysis on spatial difference and evolution of port logistics economy in Yangtze River Delta," *Economic Geography*, vol. 29, no. 7, pp. 1081–1086, 2009.

[5] M. Song and H. Huang, "Research on spatial spillover and threshold effect of port efficiency in Yangtze River Delta," *Shanghai Economy*, vol. 2, pp. 33–44, 2018.

[6] G. Dong and X. Chen, "Spatial structure evolution and collaborative strategy of Yangtze River Delta port group from the perspective of regional expansion," *Journal of Dalian Maritime University (Social Science Edition)*, vol. 16, no. 4, pp. 61–66, 2017.

[7] Q. Zeng, K. Wu, and T. Teng, "Spatial distribution characteristics of ports along the maritime Silk Road," *Journal of Dalian University of Technology (Social Science Edition)*, vol. 37, no. 1, pp. 25–30, 2016.

[8] shaob Wang, "Spatial structure and evolution of China's coastal container port system," *Economic Geography*, vol. 36, no. 8, pp. 93–98, 2016.

[9] X. Dong, R. Wang, and Z. Han, "Analysis of spatial structure evolution of port hinterland system: a case study of Dalian Port Liaoning economic hinterland system," *Economic Geography*, vol. 30, no. 11, pp. 1761–1766, 2010.

[10] L. Jiang, S. Wang, and Z. Liu, "Study on spatial structure and regional role of port city zone in Liaoning Province," *Regional Research and Development*, vol. 29, no. 3, pp. 25–30, 2010.

[11] X. Zou, S. Somenahalli, and D. Scrafton, "Evaluation and analysis of urban logistics competitiveness and spatial evolution," *International Journal of Logistics Research and Applications*, vol. 23, no. 5, 2020.

[12] M. Zhao, W. Wu, and M. Zhang, "The spatial and temporal evolution of Chinese coastal rural logistics: fractal development and empirical simulation," *Journal of Coastal Research*, vol. 98, no. 1, 2019.

[13] P. W. de Langen, H. Sornn-Friese, and J. Hallworth, "The role of port development companies in transitioning the port business ecosystem; the case of port of Amsterdam's circular activities," *Sustainability*, vol. 12, no. 11, 2020.

[14] Y. Luo and B. Peng, "Coordinated development of port logistics and regional economy: measurement of correlation and spatial difference—an Empirical Study Based on Guangdong Province," *Journal of Chongqing Jiaotong University (Social Science Edition)*, vol. 16, no. 2, pp. 43–49, 2016.

[15] X. Gong and F. Zammori, "Coupling coordinated development model of urban-rural logistics and empirical study," *Mathematical Problems in Engineering*, vol. 2019, Article ID 9026795, 12 pages, 2019.

[16] Y. Su, Q. Jin, P. Yang, Q. Jiang, and G.-W. Weber, "A supply chain-logistics super-network equilibrium model for urban logistics facility network optimization," *Mathematical Problems in Engineering*, vol. 2019, Article ID 5375282, 12 pages, 2019.

[17] M. He, L. Zeng, X. Wu, and J. Luo, "The spatial and temporal evolution of logistics enterprises in the Yangtze River Delta," *Sustainability*, vol. 11, no. 19, 2019.

[18] W. Xu and Y. Xu, "Comprehensive strength evaluation of China's coastal ports and evolution of main port hinterland space," *Economic Geography*, vol. 38, no. 5, pp. 26–35, 2018.

[19] L.-Z. Cong, D. Zhang, M.-L. Wang, H.-F. Xu, and L. Li, "The role of ports in the economic development of port cities: panel evidence from China," *Transport Policy*, vol. 90, pp. 13–21, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 4020752, 9 pages
https://doi.org/10.1155/2020/4020752

Research Article

A Study on the Health Output Effect of Chinese Medical Service Industry Agglomeration Based on Big Data Analysis

Yan Shu,[1] **Longxin Lin**,[2] **and Yueqian Hu**[3]

[1]*School of Economics and Management, Guangzhou University of Chinese Medicine, Guangzhou, China*
[2]*College of Information Science and Technology, Jinan University, Guangzhou, China*
[3]*School of Business, Guangdong University of Foreign Studies, Guangzhou, China*

Correspondence should be addressed to Yueqian Hu; 20170401772@gdufs.edu.cn

Received 20 August 2020; Revised 13 September 2020; Accepted 21 September 2020; Published 27 October 2020

Academic Editor: Weilin Xiao

The agglomeration health output effect of the medical service industry in the era of big data is an important part of the agglomeration innovation of medical resources. This paper used the regression model of data mining to set up the fixed effect model and system GMM model to study the relationship between the agglomeration of medical service industry and resident's health level, based on the panel data of 31 provinces of China from 2003 to 2017. The results show that the health outcome of the medical industrial agglomeration is positive and different in provinces. The influence of medical service cluster on residents' health level in the eastern region fails the significance test, while the medical service cluster in the central and western regions can significantly improve residents' health level. And, this effect is also related to the characteristics of medical resources, economic development, demographic characteristics, and other heterogeneous factors. On this basis, the paper puts forward policy suggestions to promote the market structure of the medical industry from the aspects of strengthening synergies and policy guidance.

1. Introduction

In 2018, Paul Zane Pilzer, an American economist, proposed that the healthcare industry is a star industry with a global scale of trillions after the electronic information industry. From the development practice of all countries in the world, the health care industry has become a strong driving force for the growth of national economy. In OECD countries, the added value of the health care industry accounts for 10% of GDP, while in the United States it has exceeded 17%, while in China it is only 5% [1]. With the advancement of a new round of health care reform, the development of health service in China has risen as a national strategy. In October 2016, the "healthy China 2030" planning outline, issued by the State Council, explicitly pointed out that the total scale of China's health service industry will reach 16 trillion yuan by 2030, and the medical and health industry will become the pillar industry of China's national economy. With the development of big data, the most important innovative resources in the healthcare industry will focus on the allocation and utilization of information resources, so the economic benefits of big data in healthcare are becoming increasingly prominent. McKinsey's (2013) The Big Data Revolution in Health Care estimates that big data has reduced health spending in the United States by $300 billion to $450 billion, equivalent to 12 percent to 17 percent of US health spending in 2011 [2].

As an important branch of medical and health industry, the health output effect of medical service industry agglomeration is an important content of the innovation of health and medical resources agglomeration. There has been fast development of medical services in China since the reform and opening up. According to statistics, the total number of national medical and health institutions increased from 1978 in 170000 to 2017 in 987000, practicing physician per thousand population rose from 1.08 in 1978 to 2.44 in 2017, and medical and health institutions of beds per thousand population from 1.94 in 1978 rose to 5.72 in 2017. However, because the regional imbalance in the development of China's medical service industry still exists, the

phenomenon of cross-regional medical treatment is increasing, and the medical industry cluster in different regions is significantly different. Will the level of medical industry cluster affect the health level of residents? Does it promote or inhibit the health of residents? What factors influence its performance? The regression analysis method based on data mining technology to study the health output effect of the agglomeration of China's medical service industry will be helpful to promote the agglomeration innovation of China's medical and health resources, which effectively solve the heterogeneous differences in the distribution of regional health resources.

2. Literature Review

Industrial cluster is a geographical phenomenon in the dynamic evolution of industry, so scholars at home and abroad have conducted many fruitful researches on this economic phenomenon. However, there are few studies on cluster of medical service industry, especially the literature on the influence of cluster of medical service industry on health. After reviewing the existing literature, it can be summarized into the following aspects.

2.1. Research on the Essence and Development Mode of Medical Service Industry. Scholars mainly study the connotation of medical service industry from the aspects of economic impact and medical reform policy. In terms of economic impact, Relman proposed that the medical service industry was a high-profit industry, which not only affected the implementation effect of the national health policy, but also promoted economic growth [3]. DeVol and Koepp measured the cluster degree of health services in the US, and the empirical evidence showed that health services in the US played a significant role in promoting economic development [4]. Zhou and Liu put forward that the medical service industry is an important channel to achieve economic transformation and upgrading and put forward suggestions on developing the medical service industry from the aspects of top-level design, reform, and innovation [5]. Gao pointed out that, in the stage of high-quality economic development, promoting the supply-side reform of the medical service industry is the key to achieving the sick having access to medical care [6].

In terms of medical reform policies, Serrano Ibis analyzed the current situation and characteristics of American medical service industry and discussed the impact of American medical reform on the pharmaceutical industry [7]. Devlin-Foltz et al. pointed out that the social security system plays a leading role in the income level of most American retirees based on the perspective of distribution [8]. Schoen et al. proposed policy recommendations to improve access to medical services and reduce the economic burden of high cost of health services, regarding low-income and disadvantaged groups as the research subjects [9]. Domestic scholars have conducted extensive studies on medical reform policy. For example, Xing focused on medical system reform on public hospitals and proposed that public hospitals are the core carrier of the development of medical service industry [10]. Song et al. studied the allocation of public resources between different hierarchies and regions and pointed out the importance and urgency of hierarchical medical reform [11].

2.2. Research on the Mechanism of Service Industry Cluster. Scholars at home and abroad have studied the cluster mechanism of service industry mainly from endogenous mechanism and exogenous mechanism. In endogenous mechanism, firstly, the cluster of service industry reduces the information cost and transaction cost of enterprises, thus gaining advantages in competitive cost (Keeble & Nachum) [12]. Secondly, the complementary symbiosis is the driving force of service industry cluster; that is, agglomeration between industries within the service industry and between the service industry and other industries can effectively achieve its ultimate goal (Pandit et al.; Bathelt) [13, 14].Thirdly, the service industry cluster aims to acquire tacit knowledge. Tacit knowledge cannot be transmitted by traditional means and can only be acquired through close interaction between industries (Zhao et al.) [15]. Fourthly, the aggregation motivation of service industry is the demand for industrial innovation. Through agglomeration, a large-scale innovation network is formed, and the communication between industries can stimulate the application of new technologies and methods (Keeble & Nachum; Han) [12, 16]. In exogenous mechanism, most scholars believe that the main reason for service industry cluster is externality, which means that, due to external economies of scale, enterprises can reduce transaction costs by concentrating in geographical regions to improve their competitiveness (Pandit et al.; Breandan & Timothy; Wang; Zhang et al.; Liu et al.) [13, 17–20]. In addition, some scholars hold that the motivation of aggregation is the reservoir effect of human resources. The high-quality talents can reduce the search cost of human capital of enterprises and achieve common benefits in the area.

2.3. Research on the Influence of Medical Service Industry Cluster on Health. Health was originally a medical problem, but since the 1980s, scholars at home and abroad began to study the relationship between health and economy. Most studies have focused on the impact on health such as medical insurance and health expenditure. Literature studies have focused on the impact of health insurance and health expenditure on the resident health.

In terms of the impact of medical insurance on health, most scholars believe that medical insurance lowers the relative price of medical services to a certain extent and improves the utilization rate of medical services, thus reducing the incidence of some diseases and significantly improving the health conditions of residents (Goldman et al.; Card et al.; Huang and Gan) [21–23]. However, some scholars believe that medical insurance significantly reduces the marginal cost of medical services, but the moral hazard will lead to the waste of medical resources, and the insured's health conditions do not improve significantly (Finkelstein

& Mcknight; King et al.; Chen & Jin) [24–26]. In terms of the impact of health expenditure on health, scholars believe that the allocation of health resources is highly dependent on the government's health expenditure, which can not only guarantee the public welfare of health services, but also improve the effective of health services and the health level of residents (Grossman; Farag et al.; Sun et al.) [27–29]. In the context of rapid supply-side reform of health service industry, social health expenditure develops rapidly and improves the effective supply of health service. As an important part of health financing, social health expenditure, like government health expenditure, has significant health output effect (Hu) [30].

2.4. Research on the Impact of Big Data on the Medical Service Industry. At present, domestic and foreign researches on the effect of big data on the health output of the medical service industry are rarely involved, and existing relevant researches mainly focus on the following two aspects. The first one is about the impact of big data on the medical industry. Kayyali et al. studied the impact of big data on the medical industry in the United States and pointed out that many potential values of medical big data are gradually being discovered [31]. Kaggal et al. proposed that the lack of big data analysis technology and the openness of medical big data will seriously hinder the development of medical service industry [32]. Wang et al., taking Jiangsu Province as an example, pointed out that the establishment of big data sharing mechanism for health care is the key [33]. The other one is about the processing technology and risk of medical big data. Raghupathi and Raghupathi pointed out that the biggest difference between the application architecture of medical big data and the traditional medical information architecture lies in distributed computing and information storage [34]. Guo et al. proposed that the service based on medical big data involves the benefit distribution and responsibility sharing of all participants, which may lead to new problems such as the definition of medical responsibility and the prevention of medical accidents [35]. Huang studied the impact of big data technology on medical service system and expounded the innovation and risks that big data technology may bring from the aspects of knowledge, organization, rules, and culture [36].

Scholars at home and abroad mainly focus on the connotation, development model, agglomeration mechanism of the service industry, and the application of big data in the medical field and have made many fruitful research results. However, there are few literatures on the effect of big data on health output of medical service industry agglomeration. Based on the theory of agglomeration economy, this paper uses provincial panel data of China's medical service industry to estimate the degree of agglomeration of medical service industry and conducts an empirical study on the influence mechanism of agglomeration of medical service industry on residents' health under the background of big data.

3. Research Hypothesis

3.1. Influence of Medical Service Industry Cluster on Residents' Health. With accelerating trend of urbanization and population aging, the health of residents is threatened of infectious and chronic diseases, so the demand for medical and health services is increasing. Due to the particularity of the medical service market, information asymmetry, and the contradiction between the supply and demand of medical service, the Chinese government implements a relatively strict supervision policy on the medical service market. In the stage of high-quality economic development, the supply of medical services is difficult to meet the diversified and multilevel medical needs. It is too difficult and expensive to see a doctor in China. With the deepening of marketization and reform in the property rights of public hospital, the development of China's medical service industry is characterized by insufficiency and imbalance, and urban-rural gap and regional imbalance are increasingly significant. At the same time, China's medical service industry shows agglomeration phenomenon, which is reflected in medical funds that continue to gather in the eastern coastal areas. The medical service industry cluster has brought about lots of high-quality medical resources, which is conducive to exerting scale effect, improving the accessibility of high-quality medical services for residents, and thus improving the health level of residents.

H1: medical service industry cluster has a positive impact on residents' health level.

3.2. Influence of Medical Service Industry Cluster on Residents' Health

3.2.1. Influence of Medical Resources on the Agglomeration of the Medical Service Industry. According to the theory of industrial cluster, industrial economic activities tend to be concentrated in areas with rich sensitive resources. The medical service industry cluster is less affected by natural resource endowment but more affected by high-quality talents. If the medical service industry enterprises in a region attract more high-quality medical technical talents, the transaction cost of the enterprises will decrease and the market scale will expand rapidly, which will further promote the medical service industry cluster and form the accumulation cycle of agglomeration. At the same time, high-quality medical technical personnel represent the innovation ability of the region and can produce significant knowledge spillover effect, which will further promote the agglomeration of medical service industry. In China's first-tier cities, medical service enterprises often have the country's first-class medical technical personnel and the most advanced medical service facilities. Lv pointed out that China's medical resources show a polarization, on the one hand, between rural and urban areas and, on the other hand, between major hospitals in cities and grass-roots health organizations [37]. Major hospitals in big cities have far more medical resources than those in second-tier cities and other regions.

H2: medical resources have a positive impact on the medical service industry cluster and the health of residents.

3.2.2. Influence of Resident Income on the Medical Service Industry Cluster. Residents' income directly determines their material living standard. In general, the higher the disposable income is, the more the residents spend on nutrition, exercise, physical examination, and health care, which means that the increase of disposable income can bring more expenditure on medical service and more demand for medical service, thus affecting the agglomeration of medical service industry. Zhang proposed the income health effect and the income effect [38]. Income health effect means the health status of high-income people may be relatively good, and the medical expenditure of high-income people is less when other conditions are similar, while income effect means medical consumption is a normal, and the higher the income is, the more medical services and the more medical expenditure they spend.

The income gap also affects the medical services cluster. On the one hand, the income gap leads to the difference in the demand for medical services between the rich and the poor. The decrease in the investment and expenditure of public goods affects the medical service industry cluster. Krugman pointed out that large differences in people's preferences would lead to the undervaluation of public goods, which would lead to the reduction of public expenditure and increase the difficulty in the implementation of public policies [39]. On the other hand, according to the demand theorem, the increase in income leads to the increase in the demand for medical services, but high-income people take up more medical resources, which will not only reduce the accessibility of medical services for low-income people, but also allocate medical resources to the place with the greatest profit in the market price system. Wei and Gustafsson pointed out that the insufficient public medical expenditure would lead to the higher-income class finding it easier to access medical services. The widening income gap eventually leads to better medical services concentrate in big cities [40].

H3: residents' income has a positive impact on medical service industry cluster and has a positive impact on residents' health.

4. Empirical Study

4.1. Empirical Model. According to Grossman's theory of healthy production, health is a commodity jointly produced by a series of factors such as education, income, medical service, and lifestyle. In order to examine the impact of the market structure of the medical service industry on health, this paper takes the agglomeration degree of the medical service industry as an input factor and sets the static panel model as follows:

$$\text{Health}_{i,t} = \alpha + \beta_1 LQ_{i,t} + \beta_2 X_{i,t} + \mu_i + \varepsilon_{i,t}, \quad (1)$$

where $\text{Health}_{i,t}$, an explanatory variable, is the health level of population. $LQ_{i,t}$, a core explanatory variable, is the degree of agglomeration of medical services. $X_{i,t}$ is a control variable. α is constant term. μ_i is region effect. $\varepsilon_{i,t}$ is random perturbed variable.

The relationship between residents' health and the concentration of medical service industry is complex, and the current health level of residents may be affected by the health level of the previous period. In this paper, residents' health level with a lag of one stage is taken as an explanatory variable, and the dynamic panel model is established as follows:

$$\text{Health}_{i,t} = \alpha + \beta_1 \text{Health}_{i,t-1} + \beta_2 LQ_{i,t} + \beta_3 X_{i,t} + \mu_i + \varepsilon_{i,t}, \quad (2)$$

where, $\text{Health}_{i,t-1}$ is 1 period lag in the health level of population. The other variables are the same.

4.2. Variable Declaration

4.2.1. Explained Variable. The explained variable is the level of health of residents ($\text{Health}_{i,t}$).The UN's Millennium Development Goals are to reduce child mortality, improve maternal health, and combat HIV and other diseases. Due to the availability of data, Maternal Mortality Rate (Mmr) is chosen to reflect the health level of the residents. The maternal mortality rate is the number of maternal deaths per 100,000 live births during the year. The higher the value of this negative indicator is, the lower the health level of residents will be.

4.2.2. Explaining Variable. We considered the degree of medical service industry cluster ($LQ_{i,t}$). This paper applies the location entropy method to measure the agglomeration degree of medical service industry in China. Location entropy was first proposed by Hargate, which is used in location analysis. Also known as the regional concentration index of production, it is the ratio of ratios. The calculation formula is as follows:

$$LQ_{i,j,t} = \frac{q_{i,j,t}/q_{i,t}}{q_{i,t}/q_t}, \quad (3)$$

where. $LQ_{i,j,t}$ represents the concentration degree of industry i in region j during period t. $q_{i,j,t}$ represents the employment number of industry i in region j during period t. $q_{i,t}$ represents the number of people employed in nationwide i industry during t period. q_t represents the number of employed people in the whole country during period t. In this paper, i represents the employment in medical service industry, which consists of professional physicians, registered nurses, pharmacist, rural doctors, licensed (assistant) physicians, licensed physicians, registered nurses, pharmacists, and health workers.

The characteristic variable of medical resources is represented by the number of beds in medical institutions per 1,000 population (Bper1000), which reflects the level of medical facilities and services in different regions.

Economic development variables include per capita disposable income and urbanization rate. The population

proportion index method is generally adopted to measure the urbanization level, that is, the proportion of urban population in the total population.

The demographic characteristics variable refers to the level of education. In this paper, the length of schooling (Edu) is used to indicate that illiteracy is 0 years, primary school education is 6 years, junior high school education is 9 years, senior high school education is 12 years, and junior college education and above are 16 years. In order to weaken the heteroscedasticity of the data, all variables were logarithmically processed.

The balance panel data of 31 provincial regions in China from 2003 to 2017 were adopted in this paper. The data were obtained from China Statistical Yearbook and China Statistical Yearbook of Health and Family Planning.

5. Quantitative Analysis

5.1. Descriptive Statistical Analysis. Table 1 shows the variables and descriptive statistical results in the empirical model. The minimum value and maximum value of Mmr of residents' health level are 1.1 and 399.1, indicating that there are large differences in residents' health level in the sample. The minimum value of LQ, the indicator variable of the agglomeration degree of medical service industry, is 0.53 and the maximum value is 3.19, indicating that there are significant differences in the agglomeration degree of medical service industry in various provinces of China. At the same time, per capita disposable income varies in a larger range, and the standard deviation is as high as 8226, which shows the regional economic development and its imbalance in China. The above problems also exist in the urbanization rate and the years of schooling (Edu), and the gap between the minimum and maximum values is also large.

5.2. Full Sample Regression of the Influence of the Medical Service Industry Cluster on the Health Level of Residents. Regression methods of panel data generally include mixed regression, fixed effect regression, and random effect regression. Generally, F test is used to judge whether the empirical model should use mixed regression or fixed effect regression, while Hausman test is used to judge whether the empirical model should use fixed effect regression or random effect regression. The empirical test of the sample data shows that both the F value and the Hausman value show that the static panel model (1) should adopt the fixed effect regression method.

In order to avoid the result bias caused by the possible endogeneity problems between variables, instrumental variable method is further introduced for estimation. The estimation methods commonly used in the instrumental variable method are generalized moment estimation (GMM) and two-stage least square method (2SLS). The assumptions of generalized moment estimation (GMM) are more lenient, so it is not necessary to assume the same variance, but to use instrumental variables to estimate parameters, which can effectively solve the endogenous problem. System-GMM estimation method introduces the horizontal equation on the difference-GMM estimation method and takes the lag difference variable as the instrumental variable, which improves the effectiveness of the estimation results. In this paper, residents' health level lagged by one stage was added into the static panel model, and the dynamic panel model (2) above was constructed, and the robustness test was carried out by applying the system GMM estimation method. Specific regression results are shown in Table 2:

The validity test of system GMM estimation mainly includes constraint test of overidentification (Sargan test or Hansen test) and sequence correlation test (AR(1) and AR(2)). The former is used to test the validity of instrumental variables in sample estimation. The latter is used to verify the existence of sequence correlation. The system GMM estimation method allows the existence of first-order difference sequence correlation, but it does not allow the existence of second-order difference sequence correlation. The regression results in Table 2 show that the Sargan test P value of the dynamic panel model (2) is 0.669, indicating that the selection of instrumental variables is effective, and the P value of the AR(2) test is 0.269, indicating that there is no second-order sequence correlation between the residuals after the first-order difference. Therefore, it is reasonable to judge that the system GMM estimation method is adopted.

The fixed effect regression in Table 2 is consistent with the sign of each coefficient in the system GMM regression, and individual coefficient estimates become more significant, which also indicates the robustness of the estimated results. According to the GMM estimation results of the system, the estimated coefficient of the agglomeration degree of the medical service industry is -0.056, which is significant at the 10% level. This result confirms hypothesis H1; the higher the agglomeration degree of the medical service industry, the lower the maternal mortality rate, indicating the higher the health level of residents. This is because the agglomeration of medical service industry has brought a large number of high-quality medical resources and advanced medical equipment, which contributes to a more effective allocation of medical resources and the realization of scale effect. The estimated coefficient of the number of beds in medical institutions per 1,000 population is −0.168, which is significant at the 1% level, indicating that the higher the number of beds in medical institutions per 1,000 population, the lower the maternal mortality rate. This result confirms hypothesis H2. The reasons are as follows. The more medical resources, the better the agglomeration of medical service industry and the better the accessibility of medical services and the lower the probability of residents not seeking medical treatment due to illness, which is conducive to improving the health level of residents.

The estimated coefficient of per capita disposable income is −0.200, which is significant at the level of 5%, indicating that the higher the per capita disposable income is, the higher the residents' health level is, confirming hypothesis H3. This result shows that the higher the level of disposable income of residents, the greater the demand for medical services and the greater the expenditure on medical services, which promotes the agglomeration of medical services to a

Table 1: Descriptive statistics of variables.

	Variable	Mean	SD	Mini	Max
Explained variable					
	Mmr	31.796	41.373	1.1	399.1
Explaining variable					
	LQ	1.124	0.411	0.531	3.194
	Bper1000	3.935	1.372	1.48	7.55
	Income	17227.83	8226.242	6530.48	58988
	Urban	50.815	14.958	20.21	89.6
	Edu	8.519	1.222	3.738	12.5

Table 2: Full sample regression of the influence of medical service agglomeration on residents' health level.

Variables	(1)FE	(2)RE	(3)GMM
L. health			0.843∗∗∗
			(9.84)
LQ	−0.096	−0.076	−0.056∗
	(−1.05)	(−0.88)	(−1.82)
Bper1000	−0.006	−0.156∗	−0.168∗∗∗
	(−0.05)	(−1.84)	(−2.67)
Income	−0.441∗∗∗	−0.435∗∗∗	−0.200∗∗
	(−6.53)	(−6.46)	(−2.37)
Urban	−2.389∗∗∗	−1.886∗∗∗	−0.102
	(−7.44)	(−8.06)	(−0.47)
Edu	−0.715∗	−0.634∗	−0.199∗∗
	(−1.94)	(−1.95)	(−2.38)
_cons	18.102∗∗∗	16.119∗∗∗	2.935∗∗
	(16.24)	(22.47)	(2.01)
R^2	0.778	0.774	
F	25.67∗∗∗		
Hausman		30.18∗∗∗	
AR(1) P value			0.000
AR(2)的 P value			0.269
Sargan value for inspection			0.669
Obs	465	465	434

Note: (1) FE represents fixed effect model, FE represents random effect model, and GMM represents system GMM estimation of dynamic panel model; (2) FE estimation and GMM estimation are both T value in brackets, and Re estimation is Z value in brackets; (2) symbols ∗, ∗∗, and ∗∗∗, respectively, represent the significance level of 10%, 5%, and 1%.

certain extent, which confirms the above income effect. Urbanization rate has a positive influence on residents' health level, but it fails to pass the significance test. This may be due to the fact that, on the one hand, the acceleration of urbanization has brought an increase in the number of urban residents, and medical resources are gradually inclined to populated areas. The improvement of residents' access to medical services contributes to the improvement of their health level. On the other hand, urban work pressure and life pressure are relatively high, and residents may have a significantly higher incidence of disease and a lower level of health due to higher stress. The improvement of education level has significantly improved the health level of residents. This is because the higher the education level of residents, the more initiative to obtain health information. At the same time, the improvement of health awareness is conducive to increasing the number of health checks, early detection, and early treatment of diseases. Therefore, the higher the level of education, the greater the effect of healthy production function.

5.3. Regional Difference Analysis of the Influence of Medical Service Cluster on Residents' Health Level. To investigate whether there are regional differences in the health output effect of medical service agglomeration, the whole sample was divided into eastern, central, and western regions. Table 3 shows the regression results of the influence effect of the agglomeration of medical service industry on the health level of residents in the subregion. F test value and Hausman test say the eastern, central, and western regions are suitable for the fixed effects regression model; to estimate the robustness of results, further to the eastern, central, and western regions, respectively, system GMM estimation, the sample data of AR (1) and AR (2), and Sargan test P values show that a systematic GMM estimation method is effective and feasible.

It can be seen from Table 3 that the agglomeration degree of medical service industry in eastern China has a positive impact on the health level of residents, but it fails the significance test. The agglomeration degree of medical service industry in central and western China significantly improves the health level of residents. As the economic level of eastern region is more developed, with leading medical services and diminishing the marginal effect of health, medical service industry cluster is not the key to improve the health level of residents. The economic level of central and western region is less developed, which the medical service level is relatively backward, so the medical service industry cluster brings an increasing marginal effect in health. The effect of the number of beds per 1,000 people in the eastern region on the health of residents did not pass the significance test, but in the central and western regions, the more the beds per 1,000 people in the medical institutions, the better the health of residents. The reason is that the eastern region has abundant medical resources, while the central and western regions are in short supply. The agglomeration of medical service industry in the central and western regions brings a larger scale effect of medical resources and plays a more important role in improving the health level of residents. The impact of per capita disposable income on residents' health in eastern, central, and western

Table 3: Regional differences of the influence of medical service agglomeration on residents' health level.

Variable	Eastern region		Central region		West region	
	FE	GMM	FE	GMM	FE	GMM
L. health		0.639***		0.793***		0.825***
		(8.99)		(11.71)		(17.71)
LQ	−0.209	−0.027	−0.004	−0.036*	−0.137	−0.062*
	(−0.92)	(−0.13)	(−0.02)	(−1.85)	(−1.31)	(−1.84)
Bper1000	0.301	−0.073	−0.122	−0.173*	−0.551***	−0.077**
	(1.09)	(−0.42)	(−0.62)	(−1.82)	(−3.28)	(−2.21)
Income	−0.271	−0.335*	−0.471***	−0.357***	−0.468***	−0.160**
	(−1.51)	(−1.91)	(−3.96)	(−3.10)	(−6.25)	(−2.07)
Urban	−3.064***	−0.746*	−2.484***	−0.087	−1.435***	−0.022
	(−3.23)	(−1.96)	(−5.83)	(−0.35)	(−3.53)	(−0.17)
Edu	−2.280**	−1.323*	−0.503	−0.174**	−0.142	−0.554***
	(−2.17)	(−1.79)	(−0.68)	(−2.30)	(−0.40)	(−2.67)
_cons	22.426***	4.400***	17.932***	3.773**	13.799***	3.171***
	(7.20)	(3.34)	(8.61)	(2.19)	(10.08)	(3.40)
R^2	0.591		0.856		0.903	
F	14.73***		13.85***		35.97***	
Hausman	14.25**		14.46**		26.21***	
AR(1)'s value		0.069		0.003		0.012
AR(2)'s value		0.287		0.185		0.424
Sargan P value for insection		0.455		0.740		0.758
Obs	165	154	120	112	180	168

Note: (1) FE estimation and GMM estimation are both T values in brackets; (2) symbols *, **, and ***, respectively, represent the significance level of 10%, 5%, and 1%.

regions has a significant promoting effect, and high income contributes to the improvement of residents' health level, for the same reasons as in the full sample test. The promotion effect of urbanization rate on the residents' health in eastern regions is more significant than the central and western regions because in the eastern region, urbanization rate is higher, attracting a large number of population and increasing higher medical service demand. At the same time, the improvement of urbanization rate has brought more medical resources and the increase of the medical service accessibility to improve the health level of the residents. In eastern, central, and western regions, years of education have a significant effect on the health of residents, which is consistent with the full sample regression.

6. Conclusion

An outline for the "Healthy China 2030" initiative proposed that health is the basic condition for economic and social development. Previous studies have focused on the health output effect and comparative analysis of medical input, but few have discussed the health output effect of medical industrial structure. Based on Grossman's healthy production function, using China's provincial data from 2003 to 2017, this paper studies the health output effect of medical service industry agglomeration with full sample regression and regional difference, drawing the following conclusions: (1) using fixed effects, random effects, and system GMM method to analyze the sample, the result showed that medical service significantly promoted the residents to improve the level of health. However, there are significant regional differences in the health output effect of medical service industry cluster. The medical industry cluster in central and western regions has significantly improved the health level of residents, while the effect in eastern regions has not passed the significance test. (2) Considering other control variables, this paper found that the medical institution beds per thousand population to promote the health level of residents in the sample, central, and western regions of the country's samples passed the test of significance; personal disposable income and education level of the improvement of residents' health passed the test of significance in all samples and regional samples. While the health effects of urbanization rate in the national sample, central, and western regions sample are not by significance test, the eastern part of the effect is significant.

Based on the empirical research results, this paper proposes the following policy recommendations.

First, strengthen the construction of medical and health big data application system and promote the aggregation of advantageous resources in the medical service industry. To remove data barriers, government should promote the data sharing, data mining, and application of health big data based on the regional population health information platform. Standardize the management of medical and health big data, give full play to the resource optimization function of big data technology, improve the coordination and comprehensiveness of the medical service industry, enhance the accessibility of high-quality medical resources, and promote the aggregation of advantageous resources of the medical service industry.

Second, use regional synergy and promote structure optimization of the medical industry. In central and western regions where the development of medical service industry is

not comprehensive, it is necessary to play to the ability of medical industry cluster to optimize the allocation of medical resources and improve the efficiency, so as to strengthen the scale effect of medical service industry cluster. In eastern region, where medical resources are relatively abundant, the technical level should be improved, and the regional gradient transfer and radiation range should be promoted, so as to improve the health of the residents.

Third, the government needs to strengthen policy guidance to adapt to local conditions and to intensify policy support in central and western areas, through preferential tax credit and financial policies to strengthen the supervision function of government, based on the goal of balanced development of medical resources among regions in China. At the same time, aided by private capital to improve the construction of medical facilities, we can improve the health of the residents and gradually narrow the development gap with the eastern region.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors acknowledge the Philosophy and Social Science Planning Project of Guangdong Province (Grants: GD18XYJ24 and 2019GXJK088), Humanity and Social Science foundation of Guangzhou University of Chinese Medicine (Grant: 2020SKZD07), Lancang-Mekong Cooperation Special Fund in 2020, and the Philosophy and Social Science Researching Colleges and Universities in Guangdong Province.

References

[1] Y.-f. Liu and Zhen Wang, "The development model of health management in USA and its experiences for China," *Asia-pacific Economic Review*, vol. 3, pp. 75–81, 2016.

[2] McKinsey Consulting, *The Big Data Revolution in Health Care*, McKinsey Consulting, New York, NY, USA, 2013.

[3] A. S. Relman, "The new medical-industrial complex," *New England Journal of Medicine*, vol. 303, no. 17, pp. 963–970, 1980.

[4] R. C. DeVol and R. Koepp, *America's Health Care Economy*, Milken Institute, Santa Monica, CA, USA, 2003.

[5] X. Zhou and J. Liu, "Progress, problems and prospects of China's medical service industry reform," *Price:Theory & Practice*, vol. 5, pp. 5–11, 2018.

[6] C. Gao, "The path choice of the development and governance of medical services in the stage of high-quality development," *Social Science Front*, vol. 5, pp. 222–229, 2019.

[7] S. Serrano Ibis, *The World's Health Care Crisis-1:the united states' Leadership*, pp. 3–28, World's Health Care Crisis, Geneva, Switzerland, 2011.

[8] S. Devlin-Foltz, A. M. Henriques, and J. E. Sabelhaus, "Is the U.S. retirement system contributing to rising wealth inequality?," *Journal of the Social Sciences*, vol. 6, pp. 59–85, 2016.

[9] C. Schoen, K. Davis, A. Willink, and C. Buttorf, *A Policy Option to Enhance Access and Affordability for Medicare's Low-Income Beneficiaries*, pp. 1–15, Commonwealth Found, London, UK, 2018.

[10] W. Xing, "Practice exploration and policy thinking on the development of health service industry," *Macroeconomic Management*, vol. 6, pp. 29–31, 2014.

[11] X. Song, W. Deng, P. Zhou, S. Zhang, J. Wan, and Y. Liu, "Spatial equity and influences of two-level public healthcare resources:A background to hierarchical diagnosis and treatment reform in China," *Acta Geographica Sinica*, vol. 6, pp. 1178–1189, 2019.

[12] D. Keeble and L. Nachum, "Why do business ser vice firms cluster? small consultancies, clustering and decentralization in London and southern England," *Transactions of the Institute of British Geographers*, vol. 1, pp. 67–90, 2001.

[13] N. R. Pandit, G. A. S. Cook, P. G. M. Swann, and G. M. Peter Swann, "The dynamics of industrial clustering in British financial services," *The Service Industries Journal*, vol. 21, no. 4, pp. 33–61, 2001.

[14] H. Bathelt, "The Re-emergence of a media industry cluster in leipzig," *European Planning Studies*, vol. 10, no. 5, pp. 583–611, 2002.

[15] S. X. B. Zhao, L. Zhang, and D. T. Wang, "Determining factors of the development of a national financial center: the case of China," *Geoforum*, vol. 35, no. 5, pp. 577–592, 2004.

[16] M. Han, *Research on the Clustering Development of Producer Services Based on the Perspective of Interaction with Manufacturing*, Science-Technology and Management, Palaiseau, France, 2009.

[17] B. Ó hUallacháin and T. F. Leslie, "Producer services in the urban core and suburbs of phoenix, Arizona," *Urban Studies*, vol. 44, no. 8, pp. 1581–1601, 2007.

[18] X. Wang, "Employment absorption, industry agglomeration and the development of producer service sector," *Collected Essays on Finance and Economics*, vol. 1, pp. 15–19, 2011.

[19] Y. Zhang, Y.-J. Pu, and L.-T. Chen, "Urbanization and service industry agglomeration—a view based on system coupling interaction," *China Industrial Economics*, vol. 6, pp. 57–69, 2013.

[20] J.-y. Liu, M. Wang, J.-f. Li, and W.-z. Wang, "Literature review of agglomeration on producer services," *Journal of Chongqing University of Technology(Social Science)*, vol. 7, pp. 34–39, 2014.

[21] D. P. Goldman, J. Bhattacharya, D. F. McCaffrey et al., "Effect of insurance on mortality in an HIV-positive population in care," *Journal of the American Statistical Association*, vol. 96, no. 455, pp. 883–894, 2001.

[22] D. Card, C. Dobkin, and N. Maestas, "Does medicare save lives?∗," *Quarterly Journal of Economics*, vol. 124, no. 2, pp. 597–636, 2009.

[23] F. Huang and L. Gan, "Excess demand or appropriate demand? —health insurance, medical care and mortality of the elderly in urban China," *Economic Research Journal*, vol. 45, no. 06, pp. 105–119, 2010.

[24] A. Finkelstein and R. McKnight, "What did medicare do? the initial impact of medicare on mortality and out of pocket medical spending," *Journal of Public Economics*, vol. 92, no. 7, pp. 1644–1668, 2008.

[25] G. King, E. Gakidou, K. Imai et al., "Public policy for the poor? a randomised assessment of the Mexican universal health insurance programme," *The Lancet*, vol. 373, no. 9673, pp. 1447–1454, 2009.

[26] Y. Chen and G. Z. Jin, "Does health insurance coverage lead to better health and educational outcomes? evidence from rural China," *Journal of Health Economics*, vol. 31, no. 1, pp. 1–14, 2012.

[27] M. Grossman, "On the concept of health capital and the demand for health," *Journal of Political Economy*, vol. 80, no. 2, pp. 223–255, 1972.

[28] M. Farag, A. K. Nandakumar, S. Wallack, D. Hodgkin, G. Gaumer, and C. Erbil, "Health expenditures, health outcomes and the role of good governance," *International Journal of Health Care Finance and Economics*, vol. 13, no. 1, pp. 33–52, 2013.

[29] Y.-lin Sun, Hu-wei Wen, and H.-feng Xu, "Analysis on health production effectiveness of government health expenditure: based on the perspective of international comparison," *Chinese Health Economics*, vol. 35, no. 3, pp. 66-67, 2016.

[30] C. Hu, "The impact of social health expenditure on macro health level and its regional differences," *Chinese Health Economics*, vol. 6, pp. 77-78, 2016.

[31] B. Kayyali, D. Knott, and S. V. Kuiken, *The Big-Data Revolution in US Health Care: Accelerating Value and Innovation*, Mc Kinsey & Company, New York, NY, USA, 2013.

[32] V. C. Kaggal, R. K. Elayvilli, and S. Mehrabi, "Toward a learning health-care system–knowledge delivery at the point of care empowered by big data and NLP," *Biomed Inform Insights*, vol. 8, pp. 13–22, 2016.

[33] J. Wang, K. Leng, and H. Lu, "Study on the innovation of WIT120 medical service mode based on "Internet+"," *Journal of Nanjing Medical University(Social Sciences)*, vol. 2, pp. 84–87, 2020.

[34] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," *Health Information Ence & Systems*, vol. 2, no. 1, p. 3, 2014.

[35] X. Guo, X. Zhang, X. Liu, and V. Doug, "eHealth service management research in the big data era:challenges and future directions," *Journal of Management Science*, vol. 30, no. 1, pp. 3–14, 2017.

[36] N. Huang, "The elemnt modality of medical service system under the background of the big data: evolution, anomie and construction," *Social Sciences in Ningxia*, vol. 219, no. 1, pp. 113–119, 2020.

[37] G. Lv, "From polarization to equilibrium allocation-a path to integrate urban and rural health care resources," *Economic Management Journal*, vol. 31, no. 12, pp. 155–159, 2009.

[38] C. Zhang, "Empirical analysis on impact of health change on labor supply and income," *Economic Review*, vol. 4, pp. 79–88, 2011.

[39] P. Krugman, "The spiral of inequality," pp. 44–49, Mother Jones, San Francisco, CA, USA, 1996.

[40] Z. Wei and B. Gustafsson, "Inequity in financing China's healthcare," *Economic Research Journal*, vol. 12, pp. 26–34, 2005.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9462048, 14 pages
https://doi.org/10.1155/2020/9462048

Research Article

Energy Saving in Flow-Shop Scheduling Management: An Improved Multiobjective Model Based on Grey Wolf Optimization Algorithm

Lvjiang Yin[1], **Meier Zhuang**[2], **Jing Jia**[1] and **Huan Wang**[1]

[1]*School of Economics and Management, Hubei University of Automotive Technology, Shiyan 442002, China*
[2]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510006, China*

Correspondence should be addressed to Meier Zhuang; 990883130@qq.com

Received 16 June 2020; Accepted 21 August 2020; Published 14 October 2020

Guest Editor: Weilin Xiao

Currently, energy saving is increasingly important. During the production procedure, energy saving can be achieved if the operational method and machine infrastructure are improved, but it also increases the complexity of flow-shop scheduling. Actually, as one of the data mining technologies, Grey Wolf Optimization Algorithm is widely applied to various mathematical problems in engineering. However, due to the immaturity of this algorithm, it still has some defects. Therefore, we propose an improved multiobjective model based on Grey Wolf Optimization Algorithm related to Kalman filter and reinforcement learning operator, where Kalman filter is introduced to make the solution set closer to the Pareto optimal front end. By means of reinforcement learning operator, the convergence speed and solving ability of the algorithm can be improved. After testing six benchmark functions, the results show that it is better than that of the original algorithm and other comparison algorithms in terms of search accuracy and solution set diversity. The improved multiobjective model based on Grey Wolf Optimization Algorithm proposed in this paper is conducive to solving energy saving problems in flow-shop scheduling problem, and it is of great practical value in engineering and management.

1. Introduction

Many mathematical problems in scientific research and practical engineering essentially belong to multiobjective optimization problem. The analysis of constrained multiobjective optimization algorithm has become a research hotspot in recent years.

Different theories exist in the literature regarding optimization algorithm such as the Improved Multiobjective Grey Wolf Optimizer (IMOGWO) that hybridize with the fast nondominated sorting strategy [1]. In fact, significant progress has been made in solving constrained multiobjective optimization problems at home and abroad due to its efficiency and simplicity [2], but there is still much room for improvement in terms of the diversity and convergence of solution sets.

In previous research, some scholars proposed a differential evolution algorithm based on two-population search mechanism, which randomly deletes one of the two individuals with the smallest Euclidean distance [3]. In this way, the boundary solution may be lost, and the diversity of solution set may be affected. At the same time, when updating the infeasible solution set, the individuals with a small degree of constraint violation are preferred, but the individual objective function value selected in this way may be poor, thus affecting the convergence speed of the algorithm.

Several lines of evidence suggest that a number of penalty terms were applied to modify the value of individual objective function. In the process of evolution, feasible nondominant solutions were retained, and infeasible solutions with a small degree of constraint violation were also retained

[4]. This algorithm also has the condition of losing boundary solution, which indicates that it has some defects in the diversity maintenance of solution set. When updating feasible solution sets and infeasible solution sets, individuals located in the sparse region are given priority. However, when updating the infeasible solution sets, individuals with a large degree of constraint violation will be retained, thus reducing the convergence speed of the algorithm [5].

As for the improved elite selection strategy, it can make the solution set more widely distributed by setting preference points and expand the application of constrained multiobjective optimization algorithm to high-dimensional problems by combining with Deb criterion [6], but it still has some drawbacks.

Up to now, plenty of differential evolution algorithms have been proposed, which minimize the value of the objective function for the feasible solution and minimize the degree of constraint violation for the infeasible solution [7]. However, the information interaction between the feasible solution set and the infeasible solution set is insufficient, and the population diversity needs to be improved. On the other hand, some scholars proposed a constrained multiobjective optimization algorithm based on adaptive e truncation strategy, which can improve the diversity of solution sets and give consideration to the convergence [8]. But the selection of e parameters is difficult, which needs to be adjusted according to different problems [9]. In brief, it is difficult for most algorithms to achieve enough balance for the key performance indexes in the constrained multiobjective optimization problems in terms of diversity and convergence [10–12].

Therefore, we propose an improved Multiobjective Grey Wolf Optimizer related to Kalman filtering and reinforcement learning (MKGWO) in this paper. The main innovation of the algorithm is that Kalman filter facilitates the convergence from solution set to Pareto optimal front end introduced into the static multiobjective algorithm. It combines the characteristics of Kalman filter with the robustness, reliability, and high efficiency of the reinforcement learning system when solving problems [13]. In the process of evolution, the algorithm uses an elite population to store feasible nondominant solutions and retains the nondominant solutions generated by historical iteration. The problem of scheduling research is to allocate scarce resources to different tasks within a certain period of time [14, 15]. As the production with the continuous expansion of scale, the importance of scheduling and decision-making to enterprise management and production is increasingly prominent [16].

The scheduling problem is an interdisciplinary field of research, which involves operations research, computer science, control theory, industrial engineering, and many other disciplines [17]. A good scheduling scheme can greatly improve the production level of enterprises, making rational use of resources and enhancing the market competitiveness [18–21].

From a different perspective, combining the data mining technology and mathematical logic, we establish an improved multiobjective operation model based on Grey Wolf Optimization Algorithm to give consideration to energy-saving problems in engineering. The results show that the algorithm can solve the Pareto front end problem in flow-shop scheduling successfully, and it is of great practical value in engineering and management.

2. Multiobjective Grey Wolf Optimizer

2.1. Multiobjective Problem. As briefly mentioned in the introduction, multiobjective optimization refers to the optimization of a problem with more than one objective function [22, 23]. Without loss of generality, it can be formulated as a maximization problem as follows:

$$\begin{aligned} &\text{maximize:} && F(x) = f_1(x),\ f_2(x), \ldots, f_o(x), \\ &\text{subject to:} && g_i(x) \geq 0, i = 1, 2, \ldots, m, \\ & && h_i(x) = 0, \quad i = 1, 2, \ldots, p, \\ & && L_i \leq X_i \leq U_i, \quad i = 1, 2, \ldots, n, \end{aligned} \tag{1}$$

where n is the number of variables, o is the number of objective functions, m is the number of inequality constraints, p is the number of equality constraints, g_i is the ith inequality constraints, h_i indicates the ith equality constraints, and $[L_i, U_i]$ are the boundaries of ith variable [24–26].

In single-objective optimization, solutions can be compared easily due to the unary objective function. For maximization problems, solution X is better than Y if and only if $X > Y$. However, the solutions in a multiobjective space cannot be compared by the relational operators due to multicriterion comparison metrics. In this case, a solution is better than (dominates) another solution if and only if it shows better or equal objective value on all of the objectives and provides a better value in at least one of the objective functions [27]. The concepts of comparison of two solutions in multiobjective problems were first proposed by Khamis et al. [28] and then extended by Khamis et al. [29]. Without loss of generality, the mathematical definition of Pareto dominance for a maximization problem is as follows [30].

Definition 1 (Pareto dominance).

Suppose that there are two vectors such as $x = (x\_1, x\_2, \ldots, x\_k)$ and $y = (y\_1, y\_2, \ldots, y\_k)$.

Vector x dominates vector y (denoted as $x \succ$ y) if

$$[f(x_i) \geq f(y_i)] \wedge [\exists i \in 1, 2, \ldots, k : f(x_i)], \quad \forall i \in \{1, 2, \ldots, k\}. \tag{2}$$

The definition of Pareto optimality is as follows.

Definition 2 (Pareto optimality).

A solution $x \in X$ is called Pareto-optimal if

$$\exists y \in X | F(y) \succ F(x)|. \tag{3}$$

A set including all the nondominated solutions of a problem is called Pareto-optimal set, and it is defined as follows.

Definition 3 (Pareto optimality set).

The set of all Pareto-optimal solutions is called Pareto set as follows:

$$P_s := \{x, y \in X \mid \exists F(y) \succ F(x)\}. \tag{4}$$

A set containing the corresponding objective values of Pareto-optimal solution in Pareto-optimal set is called Pareto-optimal front [31]. The definition of the Pareto-optimal front is as follows.

Definition 4 (Pareto-optimal front).

A set containing the values of objective functions for Pareto solutions set is

$$P_f := \{F(x) \mid x \in P_s\}. \tag{5}$$

2.2. MOGWO. MOGWO algorithm was proposed by Holland [32]. The social leadership and hunting technique of grey wolves were the main inspiration of this algorithm. In order to mathematically model the social hierarchy of wolves when designing MOGWO, the fittest solution is considered as the alpha (α) wolf. Consequently, the second and third best solutions are named beta (β) and delta (δ) wolves, respectively. The rest of the candidate solutions are assumed to be omega (ω) wolves. In the GWO algorithm, the hunting (optimization) is guided by α, β, and δ. The ω wolves follow these three wolves in the search for the global optimum [33–35]:

$$D = \left|C * X_P(t) - X(t)\right|, \tag{6}$$

$$X(t+1) = X_P(t) - A * D, \tag{7}$$

where t indicates the current iteration, A and C are coefficient vectors, X_p is the position vector of the prey, and X indicates the position vector of a grey wolf [36].

The vectors A and C are calculated as follows:

$$A = 2a * r_1 - a, \tag{8}$$

$$C = 2r_2, \tag{9}$$

where elements of a linear decrease from 2 to 0 over the course of iterations and r_1, r_2 are random vectors in [0, 1]. Position updating mechanism of search agents and effects A is indicated in Figure 1 [37]. In this figure, we can see that the three top wolves (namely, the fittest solutions) guide the directions of other wolves (namely, the candidate solutions).

The following formulas are run constantly for each search agent during optimization in order to simulate the hunting and find promising regions of search space:

$$D_\alpha = \left|C_1 * X_\alpha - X\right|, \tag{10}$$

$$D_\beta = \left|C_2 * X_\beta - X\right|, \tag{11}$$

$$D_\delta = \left|C_3 * X_\delta - X\right|, \tag{12}$$

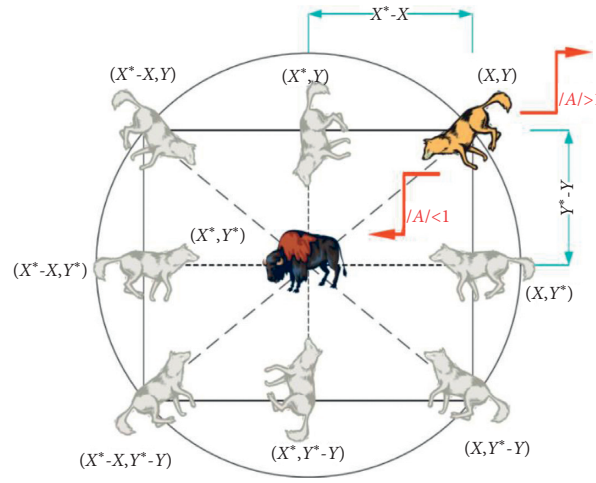


Figure 1: Position updating mechanism of search agents and effects.

$$X_1 = X_\alpha - A_1 * D_\alpha, \tag{13}$$

$$X_2 = X_\beta - A_2 * D_\beta, \tag{14}$$

$$X_3 = X_\delta - A_3 * D_\delta, \tag{15}$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3}. \tag{16}$$

C is a random value that is generated in [0, 2]. The storage mechanism of the nondominant solution is grid in MOGWO. When the archive is full, each hypercube is taken out by roulette according to the probability:

$$P_i = \frac{c}{N_i}, \tag{17}$$

where c is a constant number greater than one and N is the number of obtained Pareto-optimal solutions in the segment [38].

2.3. Defect in MOGWO. Traditional Multiobjective Grey Wolf Optimizer is a grey wolf group predation was inspired by multiobjective optimization algorithm, using a fixed external file to store nondominated solution, simple multiobjective grey wolves optimizer in solving static multiobjective problem, because without a good promotion strategy, lead to being not close to the Pareto-optimal front end, and the diversity of solution set is not high [39]. To solve the above problems, an improved algorithm KMGWO is proposed in the next section.

3. An Improved MOGWO Based on Kalman Filter and Reinforcement Learning

3.1. Kalman Filter. In 1960, R. E. Kalman published a paper describing a method which can process a time series of measurements and predict unknown variables more precisely than that based on a single measurement alone. This is

referred to as the Kalman filter. Kalman filter maintains state vectors, which describe the system state, along with its uncertainties. The equations for the Kalman filter fall into two groups, time update and measurement update equations, which are performed recursively for the Kalman filter to make prediction. Here, the Kalman filter is used to directly predict for future generations in the decision space, and the two major steps are described below [40].

3.1.1. Measurement Update. The measurement update equations are responsible for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate [41]. The individual solutions just before the change occurs are taken as the actual measurements of the previous predictions. This information is used to update the Kalman filter prediction model [42].

3.1.2. Time Update. The time update equations are responsible for projecting forward the current state and error estimate covariance estimates to obtain the a priori estimates for the next step. New solutions are predicted based on the corrected Kalman filter associated with each individual in the decision space [43]. These are a priority estimates of the future.

Pareto-optimal solutions will then be used to update the reference points and subproblems. The specific equations for the two steps are presented in the following [44]:

Time update step:

$$x_k^- = Ax_{k-1} + bu_{k-1}, \tag{18}$$

$$P_k^- = AP_{k-1}A^T + Q. \tag{19}$$

Measurement update step:

$$K_k = P_k^- H^T \left(HP_K^- H^T + R\right)^{-1}, \tag{20}$$

$$K_k = P_k^- H^T \left(HP_K^- H^T + R\right)^{-1}, \tag{21}$$

$$P_k = (I - K_k H)P_k^-, \tag{22}$$

where x is the state vector to be estimated by the Kalman filter, A denotes the state transition model, u is the optional control input to the state x, B is the control input matrix, and P is the error covariance estimate [45]. Z denotes the measurement of the state vector, H is the observation matrix, and the process and measurement noise covariance matrices are 2 and R, respectively. K is the Kalman filter gain.

Here is an example, so we can understand the Kalman filter more intuitively.

As shown in Figure 2, the state vector of an object at the period t obeys the magenta normal distribution, according to which the position of the object at period $t+1$ can be predicted, and the predicted value is the blue normal

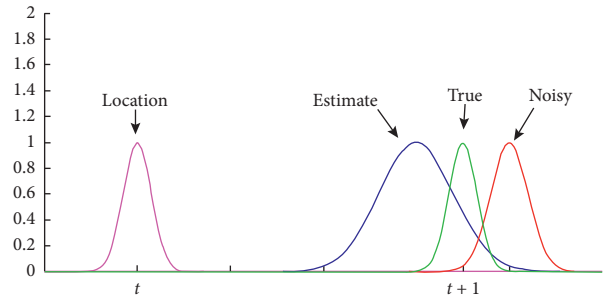


Figure 2: Kalman filter principle explanation.

distribution. This blue normal distribution is getting fatter, because a layer of noise to the recursion is added, so the uncertainty is getting bigger. In order to avoid the deviation caused by pure estimation, we made a measurement on the position of the object at the period of $t+1$, and the measurement results are subject to the red normal distribution. Through the five equations of Kalman filter, the real state vector in the car at the time of $t+1$ can be obtained. The state vector follows the green normal distribution, which means that the prediction of Kalman filter can be carried out iteratively.

Kalman filter has been widely applied to the dynamic multiobjective algorithm to ensure that the dynamic multiobjective algorithm can converge to the Pareto-optimal front end in time when the problem changes. It can be said that, at present, Kalman filter is one of the most effective methods to make the population converge to the Pareto-optimal set and the solution set converge to the Pareto-optimal front end. Therefore, this paper reversely applies it to the static multiobjective algorithm to promote the static multiobjective algorithm to converge to the Pareto-optimal front end faster.

In order to overcome the defects of MOGWO described above, this paper proposes a multiobjective Grey Wolf Algorithm based on Kalman filter transformation, which is hereinafter referred to as MKGWO.

After each iteration, a new grey wolf population is generated by the newly generated grey wolf population and the previous-generation grey wolf population using Kalman filter through update probability Pu, which promotes the convergence of solution set to the Pareto-optimal front end. This strategy is called Kalman filter operator, and the formula is described as follows:

$$\left\{x_{t+1}^i \,\middle|\, i \in pu * n\right\}, \tag{23}$$

$$x_{t+1} = \text{Kalman filter}\left(x_{t+1}^-, x_t\right). \tag{24}$$

3.2. Reinforcement Learning. In the field of big data and machine learning, data mining and learning technology can be divided into supervised learning, unsupervised learning, and reinforcement learning. Reinforcement learning grew out from animal learning and parameter perturbation adaptive control theory, referring to the mapping from

environmental state to the action. It is a machine learning method that can adapt to and interact with the environment. This method is different from supervised learning through positive cases and counterexamples to advise the agent of what action to take, but by trial and error to find the optimal behavioral strategy.

As is shown in Figure 3, the basic principle of reinforcement learning is test evaluation process, and the agent chooses an action for the environment, after which the action state change is accepted at the same time producing a reinforcement signal. Then, the agent selects next action according to the reinforcement signal and environmental current state, and the selection principle is increased by the probability of positive reinforcement. The selected action not only affects the immediate reinforcement value, but also the state of the environment at the next moment and the final reinforcement value.

In MOGWO improved based on Kalman filter operator, we found that Pu (update probability) could not be fixed. For different problems, the optimal vaccination probability was different. For the same problem, the optimal vaccination probability is also different at different periods of iteration, so we choose the reinforcement learning method here to dynamically determine the vaccination probability. In this paper, this improved method is collectively called reinforcement learning operator, as shown in Figure 4.

The reinforcement learning method used here is based on snap–drift neural network. It switches between snap mode and drift mode. In this operator, agent (MOGWO) accepts the state (snap or drift) and reward value (Pm) at time t then takes an action (increase or decrease Pu) to convert to a new state:

$$Pm = \frac{Se}{N}, \tag{25}$$

$$Pu = \begin{cases} \max(0, Pu - \omega), & \mu = \text{snap}, \\ \min(1, Pu + \omega), & \mu = \text{drif}. \end{cases} \tag{26}$$

Se is the number of nondominated solutions generated in this iteration; Pm is the conversion probability, that is determined by the proportion of update individuals in population this iteration. Snap mode is used for less than 50%, and drift mode is used for more than or equal to 50%. ω is the step size of Pu with each change.

3.3. MKGWO Flow. MKGWO's algorithm flow framework is as follows (Algorithm 1).

4. Simulation Experiments

To test the performance of MKGWO, MKGWO and MOGWO, MOPSO, NSGA2, MOEA/D, and PESA2 simulation experiment, the benchmark functions and correlation index are analyzed in this section.

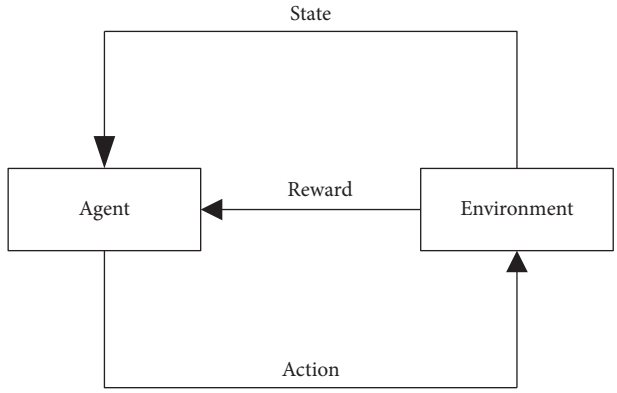


Figure 3: Principle of reinforcement learning.

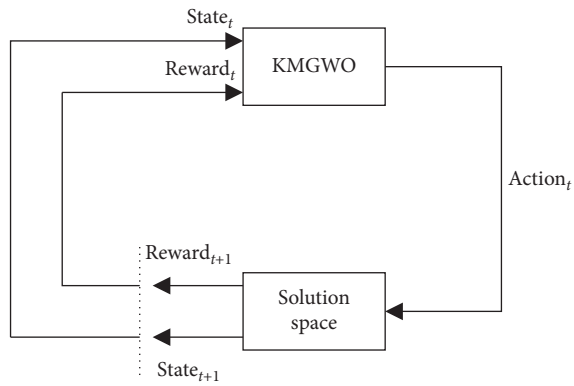


Figure 4: Reinforcement learning operator operation process.

4.1. Experimental Environment and Benchmark Function

4.1.1. Experimental Environment. The operating environment of the simulation experiment is as follows: the machine is Dawning 5000A supercomputer. Xeon X5620 CPU (4 cores) ∗2, 24 GB memory, 300 GB SAS hard disk. Equipped with Rhel5.6 operating system. The programming tool is MATLAB 2012a (for Linux).

4.1.2. Benchmark Function. In this paper, six benchmark functions are selected to evaluate the performance of the algorithm. This group of benchmark functions is widely used in the test of multiobjective optimization algorithm. The function names, dimensions, ranges, and expressions are shown in Table 1. These six benchmark functions can be divided into two categories: Kursawe, Schaffer, ZDT1, and ZDT6 are two-dimensional test functions used to investigate the search ability of the algorithm at low latitude; Viennet2 and Viennet3 are 3D test functions, adding more Pareto points and increasing the difficulty of searching, for further detecting the overall performance of the algorithm. These test problems are considered as the most challenging test problems in the literature that provide different multiobjective search spaces with different Pareto-optimal fronts: convex, nonconvex, discontinuous, and multimodal.

Initialize the grey wolf population X_i $(i = 1, 2, \ldots, n)$
Initialize a, A, and C
Calculate the objective values for each search agent
Find the nondominated solutions and initialize the archive with them
X_α = Select Leader (archive)
Exclude alpha from the archive temporarily to avoid selecting the same leader
X_β = Select Leader (archive)
Exclude beta from the archive temporarily to avoid selecting the same leader
X_δ = Select Leader (archive)
Add back alpha and beta to the archive
$t = 1$
while ($t <$ Max number of iterations).
for each search agent
  Update the position of the current search agent by equations (6)–(16)
end for
Update a, A, and C
Invoke Kalman filter by equations (18)–(24)
Invoke reinforcement learning operator by equations (25) and (26)
Calculate the objective values of all search agents
Find the nondominated solutions
Update the archive with respect to the obtained nondomination solutions
If the archive is full
  Run the grid mechanism to omit one of the current archive solutions
  Add the new solution to the archives
End if
X_α = Select Leader (archive)
Exclude alpha from the archive temporarily to avoid selecting the same leader
X_β = Select Leader (archive)
Exclude beta from the archive temporarily to avoid selecting the same leader
Add back alpha and beta to the archive

Algorithm 1: MKGWO flow framework.

Table 1: Benchmark function equation.

Function name	Equation	Search domain	Search boundary
Kursawe	$\text{Minimize} = \begin{cases} f_1(x) = \sum_{i=1}^{2}[-10\exp(-0.2\sqrt{x_i^2 + x_{i+1}^2})], \\ f_2(x) = \sum_{i=1}^{3}[\lvert x_i\rvert^{0.8} + 5\sin(x_i^3)] \end{cases}$	3	$-5 \le x_i \le 5$
Schaffer	$\text{Minimize} = \begin{cases} f_1(x) = x^2, \\ f_2(x) = (x-2)^2 \end{cases}$	1	$-5 \le x_i \le 5$
Viennet2	$\text{Minimize} = \begin{cases} f_1(x) = 1/2(x_1 - 2)^2 + 1/13(x_2 + 1)^2 + 3, \\ f_2(x) = 1/36(x_1 + x_2)^2 + 1/8(x_2 - x_1 + 2)^2 - 17, \\ f_3(x) = 1/175(x_1 + 2x_2 - 1)^2 + 1/17(2x_2 - x_1)^2 - 13 \end{cases}$	2	$-4 \le x_i \le 4$
Viennet3	$\text{Minimize} = \begin{cases} f_1(x) = 1/2(x_1^2 + x_2^2) + \sin(x_1^2 + x_2^2), \\ f_2(x) = 1/8(3x_1 - 2x_2 + 4)^2 + 1/27(x_1 - x_2 + 1)^2 + 15, \\ f_3(x) = 1/x_1^2 + x_2^2 + 1 - 1.1\exp(-(x_1^2 + x_2^2)) \end{cases}$	2	$-3 \le x_1 \le 10, -10 \le x_2 \le 3$
ZDT1	$\text{Minimize} = \begin{cases} f_1(x) = x_1, \\ f_2(x) = g(x)h(f_1(x), g(x)), \\ g(x) = 1 + 9/29\sum_{i=2}^{30} x_i, \\ h(f_1(x), g(x)) = 1 - \sqrt{f_1(x)/g(x)} \end{cases}$	30	$0 \le x_i \le 1$
ZDT6	$\text{Minimize} = \begin{cases} f_1(x) = 1 - \exp(-4x_1)\sin^6(6\pi x_1), \\ f_2(x) = g(x)h(f_1(x), g(x)), \\ g(x) = 1 + 9[\sum_{i=2}^{10} x_i/9]^{0.25}, \\ h(f_1(x), g(x)) = 1 - (f_1(x)/g(x))^2 \end{cases}$	10	$0 \le x_i \le 1$

4.2. Contrast Indicators and Algorithm Parameters. For the performance metric, we have used Inverted Generational Distance (IGD) for measuring convergence. The Spacing (SP) is employed to quantify and measure the coverage. The mathematical formulation of IGD is similar to that of Generational Distance (GD). This modified measure formula is as follows:

$$\mathrm{IGD} = \frac{\sqrt{\sum_{i=1}^{n} d_i^2}}{n}, \tag{27}$$

where n is the number of true Pareto-optimal solutions and d indicates the Euclidean distance between the ith true Pareto-optimal solution and the closest obtained Pareto-optimal solutions in the reference set. The Euclidean distance between obtained solutions and reference set is different here. In IGD, the Euclidean distance is calculated for every true solution with respect to its nearest obtained Pareto-optimal solutions in the objective space.

The mathematical formulation of the SP and MS measures is as follows:

$$\mathrm{SP} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\overline{d} - d_i\right)^2}. \tag{28}$$

In the simulation experiment, the population number of each algorithm is 200, the number of archives is 200, and the number of iterations is 500. Each algorithm ran independently for 30 generations, and its minimum value, maximum value, average value, and variance were taken as the results. The remaining parameters are shown in Table 2.

4.3. Comparative Analysis of Experimental Results. The Pareto diagram in Figures 5–10 shows that the Pareto nondominant solution generated by KMGWO is basically consistent with the real Pareto front end and the solution set distribution is relatively uniform. Then, the simulation experiment results are analyzed by data comparison. In the Kursawe function test results, the indicators of KMGWO and MOEA/D are the best, around 0.12, slightly larger than PESA2. The worst is NSGA2 algorithm. It shows that in addition to KMGWO, MOEA/D is also suitable for solving low-dimensional multiobjective problems. The test results of Schaffer's function show that KMGWO has the best effect. MOGWO and MOEA/D are slightly worse than KMGWO, and the standard deviation of MOGWO is smaller. In both Viennet2 and Viennet3, the results of the three-target benchmark test functions, PESA2 and KMGWO, have the best effect of two algorithms, respectively. In Viennet2, KMGWO is only slightly less than PESA2 algorithm. Taken together, KMGWO is very suitable for solving three-target optimization problem; this may be associated with the algorithm search ability being stronger. The next step research direction can test the effect of this algorithm in the target problem; it is very exciting, and in the two test functions, KMGWO's standard deviation is also very good. It shows that the stability of the algorithm in this kind of test function is excellent. In this paper, two of the functions of ZDT system

Table 2: Algorithms' parameters.

Algorithms	Parameters
KMGWO	Alpha = 0.1; beta = 4; gamma = 2; $Pu = 0.2$; $w = 0.1$
MOGWO	Alpha = 0.1; beta = 4; gamma = 2
MOPSO	$w = 0.4$; $c1 = 2$; $c2 = 2$
NSGA2	pc = 0.9; pm = 0.5
MOEA/D	Gamma = 0.5
PESA2	$p = 0.5$; $h = 0.3$; gamma = 0.15

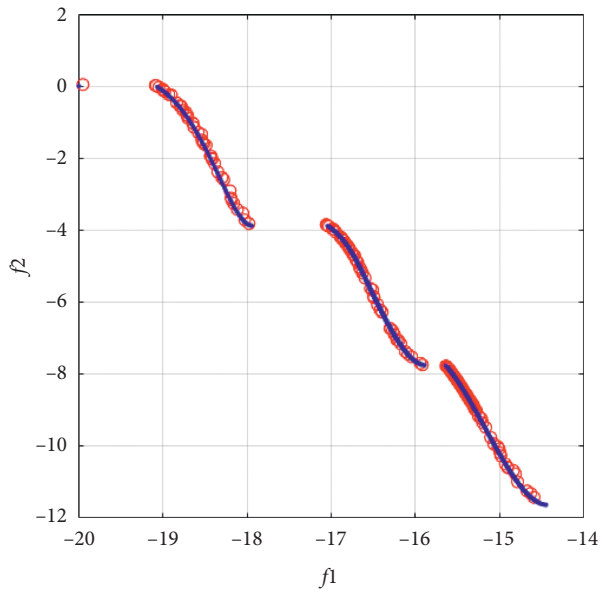


Figure 5: Kursawe's test result.

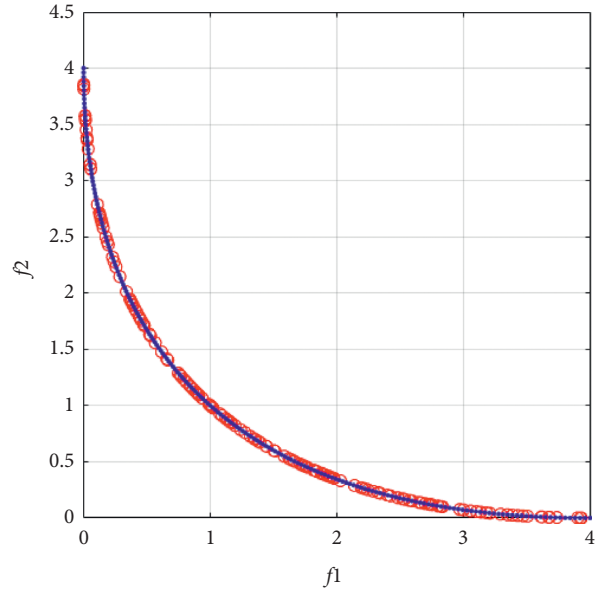


Figure 6: Schaffer's test result.

proposed by Deb are selected: ZDT1 and ZDT6. The test results of ZDT1 function showed that the best algorithm was MOGWO, MOPSO was slightly inferior, and KMGWO was excellent and ranked third. The test results of ZDT6 function

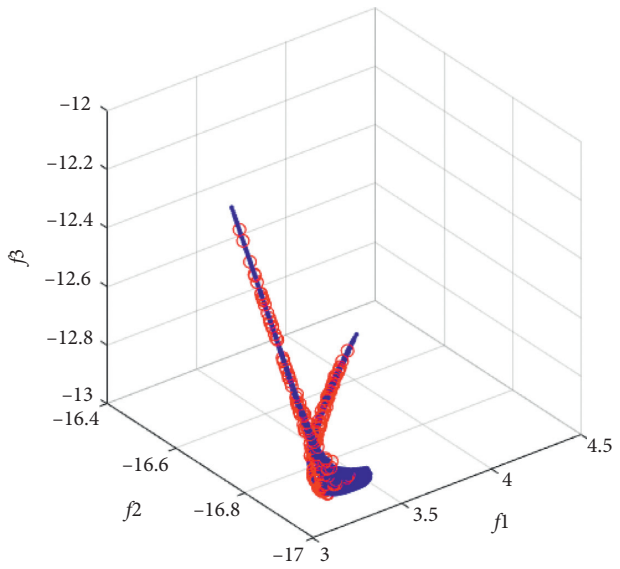


Figure 7: Viennet2 test result.

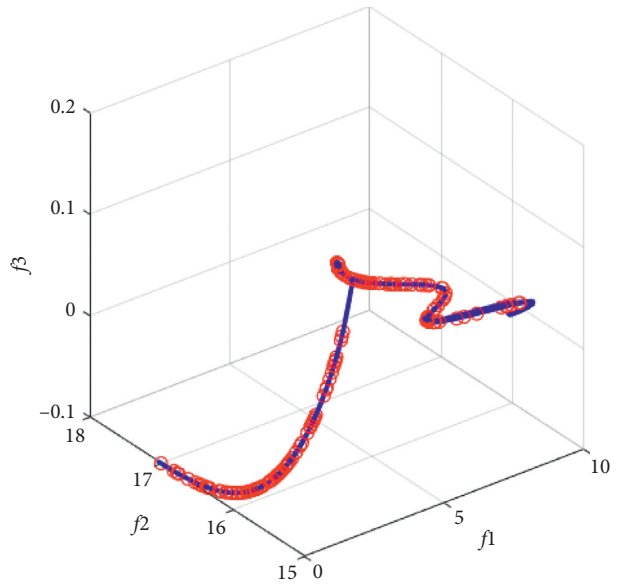


Figure 8: Viennet3 test result.

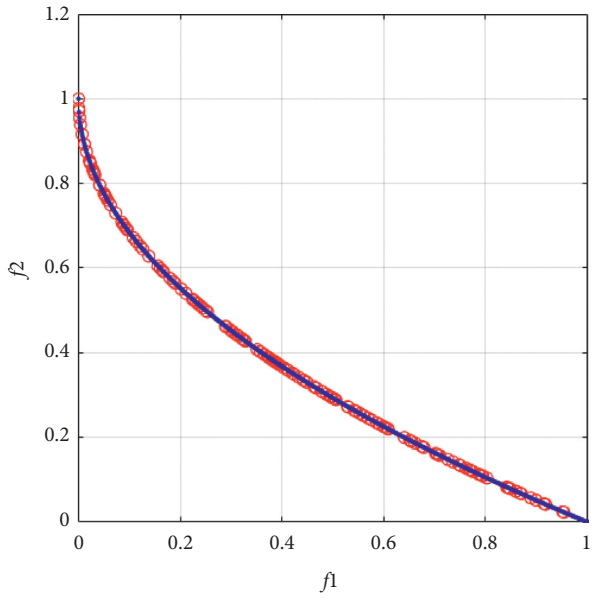


Figure 9: ZDT1 test result.

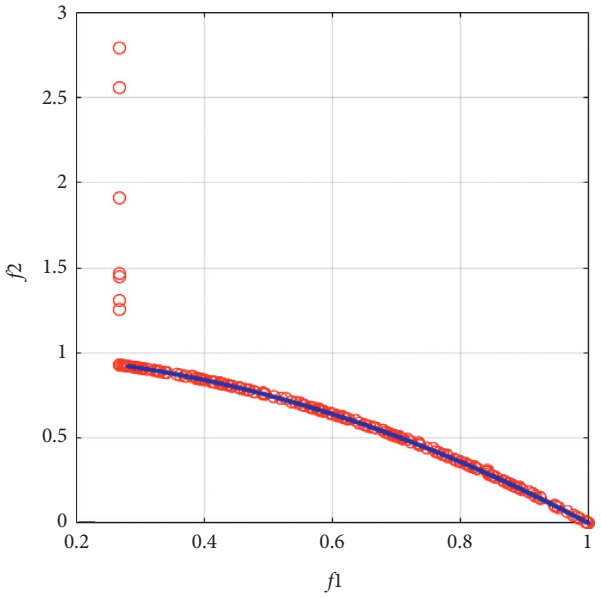


Figure 10: ZDT6 test result.

show that KMGWO is the best algorithm with excellent stability. In general, KMGWO's ability to approach the real Pareto front end is very strong, especially when the number of targets is high, so the algorithm will be widely used in production.

Concerning IGD metric, the merit is clear that the KMGWO significantly dominates over the KMGWO on almost the problems. As shown in Table 3, the KMGWO is the best performing method in our comparison experiment. This strongly demonstrates that the reinforcement learning operator can effectively improve the overall performance of the algorithm. The reason for this superiority of MKGWO lies in reinforcement learning operator as follows: in reinforcement learning operator, the optimal vaccination probability is also different at different periods of iteration; population is divided into many subpopulations and each solution has its own neighbors.

Table 4 is the statistics of SP. SP value represents the degree of uniformity between Pareto solutions. The smaller this value is, the more homogeneous the Pareto solutions obtained by the algorithm will be, and the smaller the distance difference between them will be. Kursawe's function test results show that NSGA2 SP value is the smallest of these six algorithms, 1.4, and MOGWO, MOEA/D, and PESA2 with poor results, suggesting that NSGA2 simple double objective test functions can generate the most homogeneous solution set. It is worth noting that KMGWO effect is better than MOGWO; Kalman filtering and reinforcement

Table 3: Comparison of algorithm running results' IGD.

		Kursawe	Schaffer	Viennet2	Viennet3	ZDT1	ZDT6
KMGWO	Min	0.001268	0.000548	0.000195	0.00013	$4.15E-05$	0.000336
	Max	0.001467	0.000606	0.000221	0.000663	$5.97E-05$	0.000473
	Mean	0.001367	0.000577	0.000208	0.000397	$5.06E-05$	0.000404
	Std	0.00014	$4.1E-05$	$1.84E-05$	0.000377	$1.28E-05$	$9.7E-05$
MOGWO	Min	0.001429	0.000626	0.000207	0.000222	$2.96E-05$	0.002448
	Max	0.00208	0.000689	0.000218	0.000265	$3.66E-05$	0.01636
	Mean	0.001755	0.000657	0.000213	0.000244	$3.31E-05$	0.009404
	Std	0.000461	$4.47E-05$	$7.86E-06$	$2.99E-05$	$4.93E-06$	0.009838
MOPSO	Min	0.002178	0.000669	0.00036	0.0002	0.000267	0.026323
	Max	0.002426	0.000705	0.00039	0.000221	0.000296	0.030315
	Mean	0.002302	0.000687	0.000375	0.00021	0.000281	0.028319
	Std	0.000175	$2.61E-05$	$2.08E-05$	$1.44E-05$	$2.11E-05$	0.002823
NSGA2	Min	0.402172	0.09739	1.538041	0.85775	0.092674	0.134332
	Max	0.410001	0.097398	1.592757	0.870869	0.109724	0.185229
	Mean	0.406086	0.097394	1.565399	0.864309	0.101199	0.15978
	Std	0.005536	$5.83E-06$	0.03869	0.009277	0.012056	0.035989
MOEA/D	Min	0.001268	0.000605	$2.33E-05$	0.000105	$3.44E-05$	0.000864
	Max	0.004135	0.000613	$4.94E-05$	0.000208	$4.66E-05$	0.018473
	Mean	0.002701	0.000609	$3.64E-05$	0.000156	$4.05E-05$	0.009668
	Std	0.002027	$5.98E-06$	$1.85E-05$	$7.28E-05$	$8.6E-06$	0.012451
PESA2	Min	0.001274	0.000628	0.000183	0.000195	0.002272	0.021336
	Max	0.001549	0.000644	0.000347	0.000323	0.002847	0.052005
	Mean	0.001412	0.000636	0.000265	0.000259	0.00256	0.036671
	Std	0.000194	$1.15E-05$	0.000116	$9.05E-05$	0.000407	0.021687

Table 4: Comparison of algorithm running results' SP.

SP		Kursawe	Schaffer	Viennet2	Viennet3	ZDT1	ZDT6
KMGWO	Min	1.869705	0.464979	0.229326	1.681636	0.057698	0.170041
	Max	2.140332	0.628268	0.354561	2.394785	0.067846	0.255117
	Mean	2.005019	0.546624	0.291944	2.038211	0.062772	0.212579
	Std	0.191362	0.115463	0.088554	0.504272	0.007176	0.060158
MOGWO	Min	2.071943	0.544736	0.338835	1.691569	0.058589	0.082575
	Max	2.255649	0.588034	0.395661	2.20073	0.061941	0.180822
	Mean	2.163796	0.566385	0.367248	1.946149	0.060265	0.131699
	Std	0.129899	0.030616	0.040182	0.360031	0.00237	0.069471
MOPSO	Min	1.812654	0.597622	0.406334	2.205442	0.074537	0.309447
	Max	1.948818	0.601261	0.41442	2.221322	0.076896	0.435343
	Mean	1.880736	0.599441	0.410377	2.213382	0.075716	0.372395
	Std	0.096282	0.002573	0.005718	0.011228	0.001668	0.089021
NSGA2	Min	1.442097	1.198419	0.225471	2.55751	0.342942	0.245284
	Max	1.490551	1.228526	0.2299	2.645984	0.343657	0.372529
	Mean	1.466324	1.213472	0.227685	2.601747	0.343299	0.308907
	Std	0.034262	0.021289	0.003132	0.06256	0.000505	0.089976
MOEA/D	Min	1.645038	0.234028	0.238373	0.097528	0.054964	0.072255
	Max	1.655435	0.295418	0.312842	0.145355	0.063264	0.15472
	Mean	1.650237	0.264723	0.275607	0.121441	0.059114	0.113488
	Std	0.007352	0.04341	0.052658	0.033819	0.005869	0.058311
PESA2	Min	2.091793	0.599484	0.276225	2.362664	0.069884	0.22178
	Max	2.160544	0.638922	0.34284	2.398225	0.076021	0.814949
	Mean	2.126169	0.619203	0.309533	2.380444	0.072953	0.518364
	Std	0.048614	0.027887	0.047104	0.025145	0.00434	0.419434

learning operator under the condition of synergy, multi-objective algorithm, can improve the static solution set of uniformity. The test results of Schaffer's function show that the values of KMGWO and MOEA/D are optimal, both around 0.2, but the standard deviation of KMGWO is better. In this test function, the performance of KMGWO is the

best. The test results of the Viennet2 function show that the SP values of KMGWO, MOEA/D, PESA2, and NSGA2 are roughly the same, all around 0.2. From the image, we can know that the test function is a three-objective test function, and its image is a three-dimensional image, but not complicated, indicating that most algorithms can find a relatively uniform solution set on it. The test results of Viennet3 show that the SP value of the MOEA/D algorithm is optimal, reaching 0.9. The results of KMGWO and MOGWO come next, reaching 1.6. The other test algorithms are far inferior to the three algorithms. In the test results of ZDT1, the opposite was found. Except for NSGA2, all the algorithms reached the optimal value of around 0.05, and the uniformity of NSGA2 was the worst, which was only 0.3. This situation may be caused by the cross search ability of NSGA2. In the test results of ZDT6 function, the SP value of MOEA/D, PESA2, and MOGWO was the smallest, around 0.8, followed by KMGWO, with SP value of 0.17, and then MOPSO, with 0.3. To sum up, KMGWO is quite a competitive algorithm in terms of the uniformity of the generated Pareto solution set, but other algorithms, such as MOEA/D, cannot be ignored, and their performance in this index is relatively excellent.

5. Application to Energy Saving considering Flow-Shop Scheduling

5.1. Energy Consumption considering Flow-Shop Scheduling. The low-carbon scheduling problem in the flow shop studied in this section can be described as follows: n jobs need to go through m stages in the same flow direction, and each job has only one working procedure in each stage. Considering the preparation time of the machine, the preparation time is related to the ordering of the two adjacent jobs. Therefore, the preparation time is based on the sequence correlation of the artifacts, and the machine startup is linked to the processing time of the artifacts.

At different stages, the machine has different speed gears for production and can be adjusted. From the point of view of energy consumption, the machine has four different states: (machine) on the machining processing status, start state in preparation for a new jobs (machine), standby (machine is in idle), and turned off (machine is turned off). Under normal circumstances, when the machine is working at a higher rate, the processing time will be shortened, but the corresponding energy consumption will increase. Therefore, this problem aims to maximize the completion time and energy consumption index. Due to the characteristics of the problem studied in this chapter, this problem is much more complicated than the traditional flow-shop scheduling problem. In this problem, other settings are as follows.

The job is processed continuously in the workshop. In other words, the process cannot be interrupted. Machines are allowed free time and have unlimited buffers between phases. When there is a first job processing, the machine boots. When all the jobs are finished, the machine is shut down. The machine speed cannot be adjusted in the course of a job processing.

In order to present the mathematical model of the problem, we first defined the following related mathematical symbols according to the above description of the problem.

Symbol definition are below:

I: Set of jobs, $I = \{1, \ldots, i, \ldots, n\}$.
J: Set of processes, $J = \{1, \ldots, j, \ldots, m\}$.
V_j: The speed set of the process j machine, $V_j = \{1, \ldots, v, \ldots, |V_j|\}$.
$p_{i,j,v}$: The processing time in process j speed v job i.
$s_{i',i,j}$: The start time in process j job i' transform to i. $i' = i, s_{i,i,j}$ indicates the job i is the first job allotted this machine.
$pp_{j,v}$: Process j the power consumed by the machine running at speed v.
sp_j: Process j power consumed by machine start-up.
ip_j: The power consumed by the machine in idle running in process j.

Decision variables are as below:

$b_{i,j}$: The initial processing time of job I at stage j.
$e_{i,j}$: The end processing time of job I at stage j.
$x_{i,j,v}$: The two-dimensional decision variable, when the job from I to j is processed at a rate of k, is 1; otherwise the value is 0.
$z_{i,i',j}$: Two-dimensional decision variable, when the job I' is processed onstage j, the value is 1; otherwise the value is 0.
T_j^{on}: Auxiliary decision variable, indicating the start-up time of stage j machine , decision by $b_{i,j}$ and $x_{i,j,k,v}$.
T_j^{off}: Auxiliary decision variable, indicating the halt time of stage j machine , decision by $b_{i,j}$ and $x_{i,j,k,v}$.

Based on these mathematical symbols, the mixed-integer programming model of the flow-shop low-carbon scheduling problem is presented as follows:

Objective function:

$$\text{Min} \cdot C_{\max} = \max_{i \in I} e_{i,m}, \tag{29}$$

$$\text{Min TEC} = E_1 + E_2 + E_3, \tag{30}$$

$$E_1 = \sum_{j \in J} \sum_{i \in I} \sum_{v \in V_j} x_{i,j,v} \cdot p_{i,j,v} \cdot pp_{j,v}, \tag{31}$$

$$E_2 = \sum_{j\in J}\sum_{i\in I}\sum_{i'\in I} x_{i,j,v}\cdot z_{i',i,j}\cdot s_{i',i}\cdot sp_j, \tag{32}$$

$$E_3 = \sum_{j\in J}\left(T_j^{\text{off}} - T_j^{\text{on}} - \sum_{i\in I}\sum_{v\in V_j} x_{i,j,k,v}\cdot\left(p_{i,j,v} + \sum_{i'\in I} z_{i',i,j}\cdot s_{i',i}\right)\right)\cdot ip_j. \tag{33}$$

Constraint condition:

$$\sum_{v\in V_j} x_{i,j,v} = 1, \quad \forall i\in I,\ j\in J, \tag{34}$$

$$e_{i,j} - b_{i,j} = \sum_{v\in V_j} p_{i,j,v}\cdot x_{i,j,v}, \quad \forall i\in I,\ j\in J, \tag{35}$$

$$b_{i,j} - e_{i,j-1} \geq 0, \quad \forall i\in I, j\in\{2,\ldots,m\}, \tag{36}$$

$$z_{i,i',j} + z_{i',i,j} \leq 1, \quad \forall i,i'\in I,\ j\in J, \tag{37}$$

$$b_{i,j} - \sum_{i'\in I} e_{i',j}\cdot z_{i',i,j} - \sum_{i'\in I} s_{i',i,j}\cdot z_{i',i,j} \geq 0, \quad \forall i\in I,\ j\in J, \tag{38}$$

$$e_{i,j} - \sum_{v\in V_j} p_{i,j,v}\cdot x_{i,j,v} - \sum_{i'\in I} s_{i',i,j}\cdot z_{i',i,j} = 0, \quad \forall i\in I,\ j\in J, \tag{39}$$

$$T_j^{\text{on}} = \min\left\{b_{i,j}\cdot x_{i,j,v} - \sum_{i'\in I} s_{i',i,j}\cdot z_{i',i,j}\right\}, \quad \forall i\in I,\ j\in J,\ v\in V_j, \tag{40}$$

$$T_j^{\text{off}} = \max\left\{e_{i,j}\cdot x_{i,j,v}\right\}, \quad \forall i\in I,\ j\in J,\ v\in V_j, \tag{41}$$

$$x_{i,j,v}\in\{0,1\}, \quad \forall i\in I,\ j\in J,\ v\in V_j, \tag{42}$$

$$z_{i,i',j}\in\{0,1\}, \quad \forall i,i'\in I,\ j\in J, \tag{43}$$

$$e_{i,j} > 0, b_{i,j}\geq 0, \quad \forall i\in I, j\in J. \tag{44}$$

Formula (29) stands for the minimum completion time. Formula (30) represents the minimum energy consumption. Formula (31) represents the total energy consumption when the machine is in the processing state, while formula (32) represents the total energy consumption when the machine is in the starting state, and formula (33) represents the total energy consumption when the machine is in the standby state. Formula (34) means that each job traverses all stages, and in a specific stage, each job is assigned to a machine and processed at a speed level. Formula (35) means that interruption is not allowed during processing. Formula (36) ensures that the job operation can only be started after the operation of the previous stage is completely processed. Formulas (37) and (38) guarantee the machine capacity limit, which means the job can be processed only after the previous work is completed. Formula (39) means that the machine starting to process immediately after installation is complete. Formula (40) represents the calculation of auxiliary variables, which is equal to the minimum value between the start time and the set time of the job assigned to the corresponding machine. Formula (41) represents the calculation of auxiliary variables, which is equal to the maximum value of the end time of the job assigned to the corresponding machine. Formulas (42)–(44), respectively, define the feasible range of decision variables.

This section gives a simple example of three jobs and three stages, each with three different processing speeds. Table 5 shows the processing time and corresponding power consumption of the machine at each stage; that is, the element in the table represents $(p_{i,j}, pp_{i,j,v})$.

Table 6 shows the sequence-dependent start times; the element in the table represents $s_{i',i,j}$. Set $sp_j = 3$; $ip_j = 1$.

Table 5: Processing time and corresponding power.

Process	Job $i=1$			Job $i=2$			Job $i=3$		
	$v=1$	$v=2$	$v=3$	$v=1$	$v=2$	$v=3$	$v=1$	$v=2$	$v=3$
$j=1$	(28, 4)	(22, 7)	(15, 11)	(22, 4)	(20, 7)	(18, 11)	(20, 4)	(18, 7)	(17, 11)
$j=2$	(20, 5)	(16, 8)	(12, 12)	(21, 5)	(18, 8)	(15, 12)	(23, 5)	(20, 8)	(17, 12)
$j=3$	(19, 5)	(15, 7)	(11, 10)	(20, 5)	(16, 7)	(12, 10)	(23, 5)	(20, 7)	(16, 10)

Table 6: Sequence-dependent start-up time.

Job	Job $i=1$			Job $i=2$			Job $i=3$		
	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$	$j=1$	$j=2$	$j=3$
$i'=1$	5	8	4	4	6	8	6	4	4
$i'=2$	6	4	6	2	3	2	4	6	2
$i'=3$	2	4	8	2	4	2	6	4	4

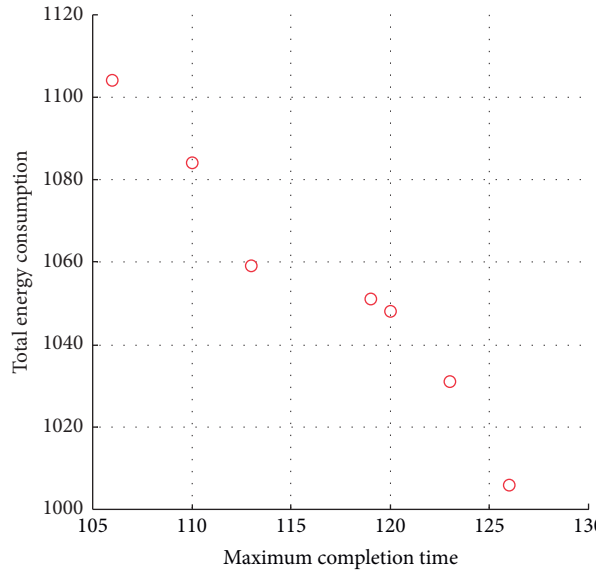


Figure 11: Experimental results of Pareto chart.

Table 7: Flow-shop scheduling.

	Job 2				Job 3				Job 1			
	Begin	Start	Work	End	Begin	Start	Work	End	Begin	Start	Work	End
Process 1	0	2	15	17	17	2	22	41	41	6	20	67
Process 2	14	3	20	37	37	4	21	62	63	4	23	90
Process 3	35	2	19	56	60	2	20	82	86	4	23	113

5.2. Experimental Result. To solve the above pipeline scheduling problem, this paper sets the parameters of KMGWO as follows: population number 20, warehouse number 20, 100 iterations. Figure 11 shows the Pareto front end obtained by KMGWO algorithm. It can be clearly observed that this problem is not complicated for KMGWO, and the Pareto-optimal solution set is easily obtained. Because there was not enough prior knowledge, we chose the scheme calculated at the Pareto point with the longest time of 113 and energy consumption of 1059.

Table 7 is the Pareto point calculated plan, showing the start time of the job in each process, warm-up time, working time, and end time. We, through these data, drew a Gantt chart, as shown in Figure 12; this is a time and energy balance options assembly line scheduling Gantt chart. In

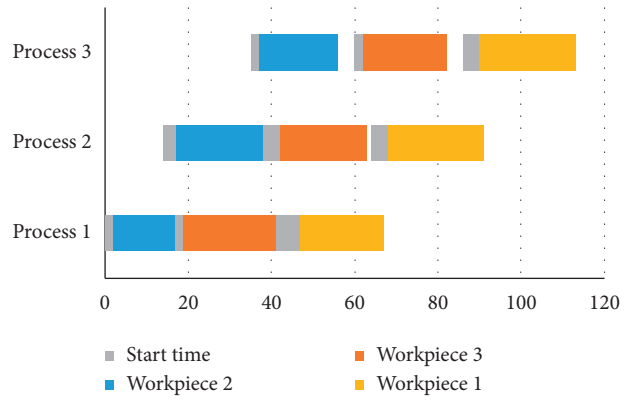


Figure 12: Experimental results Gantt chart.

practice, we can choose according to the actual demand of enterprise decision point countless times and can deal with complex situation; this is the power of KMGWO algorithm.

6. Conclusions

In this paper, an improved multiobjective operational model based on Grey Wolf Optimization Algorithm related to Kalman filtering and reinforcement learning (KMGWO) is proposed, which is the combination of data mining technology and mathematical logic. With Kalman filter, the algorithm promotes the understanding set to the real Pareto front end. The reinforcement learning operator is applied to enhance the utilization of the dominant position of the group, and adaptive parameters are used instead of human intervention. The results of six benchmark functions show that the algorithm performs better than the comparison algorithm in terms of approximating the real Pareto optimal solution set and keeping the solution set uniform. Considering the energy saving of the assembly line scheduling solution, KMGWO performance is excellent and accordingly suitable for solving the practical optimization problems. This operational model has the formidable superiority in the field of mathematical optimization, which can be applied to machine learning, engineering optimization design, and other important areas, thus enhancing the performance of energy saving in production management.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Social Science Foundation of China (NSSFC) under Grant no. 17BGL238.

References

[1] Y. Yao, Y. Chun-Ming, and Y. Feng, "Solving bi-objective reentrant hybrid flow shop scheduling problems by a hybrid discrete grey wolf optimizer," *Operations Research and Management Science*, vol. 28, no. 8, pp. 190–199, 2019.

[2] X. Zhang and X. Wang, "Comprehensive review of grey wolf optimization algorithm," *Computer Science*, vol. 46, no. 3, pp. 30–38, 2019.

[3] H. Meng, X. zhang, and L. San-Yang, "Two-group method for constrained multi-objective optimization problems," *Chinese Journal of Computers*, vol. 31, no. 2, pp. 229–235, 2008.

[4] Y. G. Woldesenbet, G. G. Yen, and B. G. Tessema, "Constraint handling in multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 514–525, 2009.

[5] Y. Zhang, "Concise multi-objective particle swarm optimization algorithm for constrained optimization," *Electronic Journals*, vol. 39, no. 6, pp. 1437–1440, 2011.

[6] H. Jain and K. Deb, "An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: handling constraints and extending to an adaptive approach," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 602–622, 2014.

[7] W. F. Gao, G. G. Yen, and S. Y. Liu, "A dual-population differential evolution with coevolution for constrained optimization," *IEEE Transactions on Cybenetics*, vol. 45, no. 5, pp. 1094–1107, 2015.

[8] B. Xiaojun and L. Zhang, "Constrained multi-objective optimization algorithm based on adaptive e-truncation strategy," *Journal of Electronics and Information Technology*, vol. 38, no. 8, pp. 2047–2053, 2016.

[9] J. Branke and K. Deb, *Integrating User Preferences into Evolutionary Multi-Objective Optimization Knowledge Incorporation in Evolutionary Computation*, pp. 461–477, Springer, Berlin, Germany, 2005.

[10] J. Branke, T. Kaußler, and H. Schmeck, "Guidance in evolutionary multi-objective optimization," *Advances in Engineering Software*, vol. 32, no. 6, pp. 499–507, 2001.

[11] C. A. C. Coello, "Evolutionary multi-objective optimization: some current research trends and topics that remain to be explored," *Frontiers of Computer Science in China*, vol. 3, no. 1, pp. 18–30, 2009.

[12] D. Lei, M. Li, and L. Wang, "A two-phase meta-heuristic for multiobjective flexible job shop scheduling problem with total energy consumption threshold," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 1097–1109, 2019.

[13] C. A. C. Coello, G. B. Lamont, and D. A. Van Veldhuisen, *Evolutionary Algorithms for Solving Multi-Objective Problems*, Springer, Berlin, Germany, 2007.

[14] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, 2004.

[15] X. Wu, X. Shen, and Q. Cui, "Multi-objective flexible flow shop scheduling problem considering variable processing time due to renewable energy," *Sustainability*, vol. 10, no. 3, p. 841, 2018.

[16] C. A. Coello Coello and M. S. Lechuga, "MOPSO: a proposal for multiple objective particle swarm optimization," in *Proceedings of the 2002 Congress on Paper presented at the Evolutionary Computation*, Honolulu, HI, USA, May 2002.

[17] I. Das and J. E. Dennis, "Normal-boundary intersection: a new method for generating the pareto surface in nonlinear

multicriteria optimization problems," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.

[18] K. Deb, *Advances in Evolutionary Multi-Objective Optimization Search Based Software Engineering*, Springer, Berlin, Germany, 2012.

[19] L. Yin, X. Li, L. Gao, C. Lu, and Z. Zhang, "A novel mathematical model and multi-objective method for the low-carbon flexible job shop scheduling problem," *Sustainable Computing: Informatics and Systems*, vol. 13, no. 3, pp. 15–30, 2017.

[20] F. Y. Edgeworth, *Mathematical Physics*, Routledge, London, UK, 1881.

[21] L. Yin, X. Li, C. Lu, and L. Gao, "Energy-efficient scheduling problem using an effective hybrid multi-objective evolutionary algorithm," *Sustainability*, vol. 8, no. 12, pp. 1268–1301, 2016.

[22] P. Mahat, Z. Chen, and B. Bak-Jensen, "Review of islanding detection methods for distributed generation," in *Proceedings of the 2008 Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, Nanjing, China, April 2008.

[23] Z. Zhang, L. Wu, T. Peng, and S. Jia, "An improved scheduling approach for minimizing total energy consumption and makespan in a flexible job shop environment," *Sustainability*, vol. 11, p. 179, 2019.

[24] A. R. Malekpour, A. R. Seifi, M. R. Hesamzadeh, and N. Hosseinzadeh, "An optimal load shedding approach for distribution networks with DGs considering capacity deficiency modelling of bulked power supply," in *Proceedings of the Australasian Universities Power Engineering Conference (AUPEC'08)*, Sydney, Australia, December 2008.

[25] W. Liao and T. Wang, "Promoting green and sustainability: a multi-objective optimization method for the job-shop scheduling problem," *Sustainability*, vol. 10, no. 11, p. 4205, 2018.

[26] M. H. Haque, "A linear static voltage stability margin for radial distribution system," in *Proceedings of the 2006 IEEE Power Engineering Society General Meeting*, pp. 1–6, Montreal, Canada, June 2006.

[27] S. Banerjee, D. Maity, and C. Chanda, "Teaching learning based optimization for economic load dispatch problem considering valve point leading effect," *Electrical Power and Energy Systems*, vol. 73, 2015.

[28] A. Khamis, H. Shareef, A. Mohamed, and Z. Dong, "A load shedding scheme for DG integrated islanded power system utilizing backtracking search algorithm," *Ain Shams Engineering Journal*, vol. 9, no. 1, 2015.

[29] A. Khamis, H. Shareef, A. Mohamed, and E. Bizkevelci, "An optimal load shedding methodology for radial power distribution systems to improve static voltage stability margin using gravity search," *Jurnal Teknolog*, vol. 68, no. 3, pp. 71–76, 2014.

[30] Z. Sun and X. Gu, "Hybrid algorithm based on an estimation of distribution algorithm and cuckoo search for the no idle permutation flow shop scheduling problem with the total tardiness criterion minimization," *Sustainability*, vol. 9, p. 953, 2017.

[31] X. Li, C. Lu, L. Gao, S. Xiao, and L. Wen, "An effective multiobjective algorithm for energy-efficient scheduling in a real-life welding shop," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 2, pp. 5400–5409, 2018.

[32] J. H. Holland, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, MI, USA, 1975.

[33] C. Lu, S. Xiao, X. Li, and L. Gao, "An effective multi-objective discrete grey wolf optimizer for a real-world scheduling problem in welding production," *Advances in Engineering Software*, vol. 99, no. 99, pp. 161–176, 2016.

[34] X. Zhu and N. Wang, "Cuckoo search algorithm with membrane communication mechanism for modeling overhead crane systems using RBF neural networks," *Applied Soft Computing*, vol. 56, no. 56, pp. 458–471, 2017.

[35] C. Lu, L. Gao, X. Li, and S. Xiao, "A hybrid multi-objective grey wolf optimizer for dynamic scheduling in a real-world welding industry," *Engineering Applications of Artificial Intelligence*, vol. 57, no. 57, pp. 61–79, 2017.

[36] L. Zhang and N. Wang, "Application of coRNA-GA based RBF-NN to model proton exchange membrane fuel cells," *International Journal of Hydrogen Energy*, vol. 43, no. 1, pp. 329–340, 2018.

[37] K. Wang and N. Wang, "A novel RNA genetic algorithm for parameter estimation of dynamic systems," *Chemical Engineering Research and Design*, vol. 88, no. 11, pp. 1485–1493, 2010.

[38] L. Zhang and N. Wang, "An adaptive RNA genetic algorithm for modeling of proton exchange membrane fuel cells," *International Journal of Hydrogen Energy*, vol. 38, no. 1, pp. 219–228, 2013.

[39] Q. Zhu, N. Wang, and L. Zhang, "Circular genetic operators based RNA genetic algorithm for modeling proton exchange membrane fuel cells," *International Journal of Hydrogen Energy*, vol. 39, no. 31, pp. 17779–17790, 2014.

[40] K. Wang and N. Wang, "A protein inspired RNA genetic algorithm for parameter estimation in hydrocracking of heavy oil," *Chemical Engineering Journal*, vol. 167, no. 1, pp. 228–239, 2011.

[41] C. Neuhauser and S. M Krone, "The genealogy of samples in models with selection," *Genetics*, vol. 145, no. 2, pp. 519–534, 1997.

[42] F. Wang, Y. Rao, C. Zhang, Q. Tang, and L. Zhang, "Estimation of distribution algorithm for energy-efficient scheduling in turning processes," *Sustainability*, vol. 8, no. 8, p. 762, 2016.

[43] S. M. Werner and A. J. Ruthenburg, "Nuclear fractionation reveals thousands of chromatin-tethered noncoding RNAs adjacent to active," *Cell Reports*, vol. 12, pp. 1089–1098, 2015.

[44] J. M. Engreitz, "Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression," *Nature Reviews Molecular Cell Biology*, vol. 17, pp. 756–770, 2016.

[45] M. D. Simon, C. I. Wang, P. V. Kharchenko et al., "The genomic binding sites of a noncoding RNA," *Proceedings of the National Academy of Sciences*, vol. 108, no. 51, pp. 20497–20502, 2011.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 8036154, 7 pages
https://doi.org/10.1155/2020/8036154

Research Article

Establishment of a Financial Crisis Early Warning System for Domestic Listed Companies Based on Three Decision Tree Models

Gang Wang,[1] Keming Wang,[2] Yingying Zhou,[1] and Xiaoyan Mo[3]

[1]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510006, China*
[2]*Center for Accounting, Finance and Institutions, Sun Yat-Sen Business School, Sun Yat-Sen University, Guangzhou 510275, China*
[3]*School of English and Education, Guangdong University of Foreign Studies, Guangzhou 510006, China*

Correspondence should be addressed to Keming Wang; wangkm@mail.sysu.edu.cn

Received 25 May 2020; Accepted 2 June 2020; Published 5 October 2020

Guest Editor: Weilin Xiao

The financial crisis is a realistic problem that the general enterprise must encounter in the process of financial management. Due to the impact of the COVID-19 and the Sino-US trade war, domestic companies with unsound financial conditions are at risk of shutdowns and bankruptcies. Therefore, it is urgently needed to study the financial warning of enterprises. In this study, three decision tree models are used to establish the financial crisis early warning system. These three decision tree models include C50, CART, and random forest decision trees. In addition, the ROC curve was used for comprehensive evaluation of the accuracy analysis of the model to confirm the predictive ability of each model. This result can provide reference for domestic financial departments and provide financial management basis for the investing public.

1. Introduction

With the continuous improvement of enterprise development capabilities, diversified investment has become an important mode for many enterprises to carry out investment activities, which plays an important supporting role in promoting the continuous development and growth of enterprises. Despite the increasing number and scale of diversified investments by enterprises, there are still some companies that lack awareness of financial risks, resulting in a series of financial risks and the failure of diversified investments. In the current situation of the COVID-19 epidemic and the Sino-US trade war environment, it is particularly important for investors and relevant government departments to grasp the viability of enterprises. This requires enterprises to deeply understand the series of financial risks they may face and take practical and effective measures to vigorously strengthen the early warning and prevention of financial risks and promote the diversified investment of enterprises to achieve tangible results.

Financial crisis early warning models are commonly based on the listed company's financial data, business methods, and published financial information. Most of them use mathematics or statistics to establish models for prediction and analysis, so as to discover the financial crisis in the process of enterprise management, give early warning to the operation and management of listed companies before the crisis, and then, take measures to prevent the crisis from happening. These prevent the crisis from causing harm to enterprises and effectively safeguard the interests of relevant financial departments, enterprises, and the investing public.

The result of financial crisis early warning is not simply whether a financial crisis has occurred. There will be some fraudulent accounting and financial matters in the financial data disclosed by listed companies. The packaging and whitewashing of the company's financial statements will seriously distort the information and make the predicted results inconsistent with the actual results. Like this type of enterprise, although the early warning results do not have the risk of financial crisis, it deserves our special attention.

While most of the current domestic literature use mathematical models to build financial crisis early warning systems, this study uses three decision tree models to establish financial crisis early warning systems. These three decision tree models include C50, CART, and random forest decision trees. In addition, in this study, the ROC curve was used for comprehensive evaluation on the accuracy analysis of the model to confirm the predictive ability of each model. This result can provide reference for domestic financial departments and provide financial management basis for the investing public.

The structure of the following part of the article is as follows: The second part is about the domestic and foreign financial crisis warning-related literature, providing background information for this part. In the third part, this article will discuss three decision tree models, including C50, CART, and random forest decision tree, analyze the financial crisis prediction of 168 listed enterprises, and use modeling and cross-validation methods to test the accuracy of the model. In the last part, conclusions and recommendations will be made for the overall study.

2. Literature Review

Since the 1960s, many scholars have conducted in-depth exploration of the prediction model of financial crisis in enterprises, the most famous of which is Altman's *Z*-score model (Z-SeoreModel). Our country's research started in 1996. Zhou et al. [1] established an *F*-score model with cash flow. Coincidentally, Yang and Xu et al. [2] also established a *Y*-score model based on the *Z*-score model, using principal component analysis Empirical Research. Zhang [3] used discriminant analysis methods to establish functions to analyze financial crisis enterprises and nonfinancial crisis enterprises. In 2005, Yang and Huang [4] used the BP artificial neural network to establish the early warning model, which increased the prediction accuracy. Xu and Chen [5] summarized the previous results and created four new financial early warning models, including the ideal distance discriminant model, the closest distance discriminant model, the principal component discriminant model of minimum deviation, and the fuzzy comprehensive evaluation model and used empirical evidence. The analysis verifies the abovementioned four models and puts forward the conclusion that the current financial early warning model needs to be combined with qualitative indicators.

In the research of the past three years, in 2018, Wang [6] used principal component analysis and innovative thinking to select 11 characteristic financial ratios with strong E-correlation of financial variables to establish a linear model and obtained remarkable results. There are significant differences between the main characteristic variables of financial risk companies and financial health companies. Lin [7] first proposed to take ST companies and non-ST companies with a 1 : 2 ratio as the research object, built a number of financial crisis early warning indicator systems, used nonparametric tests to eliminate nonsignificant indicators, then extracted the principal components as explanatory variables through principal component analysis, so as to establish a comprehensive evaluation function of the financial situation to divide the financial situation twice, and finally, build an Ologit financial crisis early warning model. Lin [8] found a new way to study the construction and significance of the financial crisis early warning index system for listed companies. Based on the principles of scientificity, effectiveness, and availability, 23 indicators were selected to build a financial crisis early warning system for listed companies, which provided more basis for studying the selection of financial early warning indicators. In the same year, Yang [9] also conducted research on indicators. Based on the existing domestic research results, he selected financial indicators with a high discrimination rate to construct a new financial risk assessment system for listed companies. In terms of model innovation, Ou [10] innovatively used the factor analysis method to build an early warning model of financial risk for real estate companies and, then, conducted financial risk research and analysis on real estate companies. Hu [11] used the F-point model to conduct an empirical analysis of listed companies in the construction machinery industry, proving that the F-point model is still a more effective early warning method in this industry.

In 2019, in terms of model innovation, Song [12] innovatively used deep learning to build a neural network model for prediction, and the prediction accuracy rate reached more than 72%. Zhang [13] innovatively established the Aalen additive model to predict the financial distress of listed enterprises, especially the time-varying characteristics of influencing factors, based on the negative net profit of enterprises for two consecutive years and the survival analysis method, so as to find the dynamic situation of financial risks of enterprises. In terms of indicator innovation, Xiong [14] added four types of nonfinancial indicators, including earnings management degree indicators, market price indicators, governance structure indicators, and auditor-related indicators on the basis of traditional financial indicators. He reconstructed the financial risk warning model and built a financial risk early warning model based on the Logit model. The research results show that adding nonfinancial indicators to the traditional early warning model can improve the early warning accuracy of the model.

In 2020, in terms of model innovation, Wu [15] took the listed companies on China's GEM as the research object, and used the Twin-SVM to construct the financial crisis early warning model for the nonequilibrium sample characteristics composed of different financial conditions of the enterprise. The model is not only superior in prediction accuracy to other models but also significantly superior in the robustness of prediction. The generalization performance of the two subindustries of manufacturing and information transmission and software and information technology service industry is also significantly superior to the rest several models. Wang [16] developed an early warning model of financial risk of punishment constraint based on cluster analysis. In the empirical analysis of listed enterprises in the a-share market, the new model showed good classification effect and robustness. According to the

current development of science and technology, Li [17] proposed that enterprises need to establish a financial risk early warning system that conforms to the development trend of big data. Taking big data and finance as the fundamental entry point, this paper actively finds out the deficiencies of the early warning system of enterprise financial risk under the background of big data and aims to improve it. Lian [18] focused on the analysis of the existing problems of financial risk warning and prevention under the diversified investment of enterprises and drew the conclusion that the focus should be on the innovation of the financial risk warning concept and the improvement of the financial risk prevention system, management mode, and monitoring mechanism.

From the review presented above, it can be seen that most of the current domestic and foreign financial crisis early warning literature use mathematical equations or statistical methods to establish financial crisis early warning systems and compare various predictive capabilities using mathematical equations or statistical methods for analysis. According to the previous literature results, many scholars pointed out that the accuracy of the mathematical model is not high and can only make linear prediction of numerical values, and the prediction results of nonlinear numerical trends are relatively inaccurate. This study adopts the data mining method to establish a financial early warning system with three decision tree models, which is innovative.

3. Research Methods and Indicators

Decision tree is a common data mining method. Its basic idea is to continuously make decisions based on certain characteristics according to human thinking and, finally, derive the most suitable classification. Each node represents a sample set with certain characteristics, and the nodes are directly generated according to some characteristics of the samples. However, among the classification prediction methods, decision tree is one of the most typical data mining methods. It is often used in combination with other data mining methods. For example, fuzzy decision tree is a mixed fuzzy theory and decision tree and rough set decision tree is a mixed rough set theory and decision tree. Mixing different methods can improve the classification and prediction ability of the decision tree. The decision tree will establish a tree-like structure, which is composed of root nodes, child nodes, and category leaf nodes. If the decision tree stops growing, each piece of data representing the sample data will be processed, and there will be no unprocessed data.

3.1. C50 Decision Tree. C50 decision tree has a wide range of applications and is suitable for discrete and continuous type of numerical or category prediction. On the branch criterion, the category type is based on the information gain ratio, while the continuous type is branched on the basis of variance reduction. In the branch method, categorical variables adopt multivariate branch and continuous variables adopt binary branch.

3.2. CART Decision Tree. The application of CART decision tree is also very common. Similar to C50, it is also suitable for discrete and continuous numerical or category prediction. The binary recursive segmentation technique is adopted in the branching criterion. The segmentation of each sample set by CART algorithm is the GINI coefficient calculation, and the smaller the value is, the more reasonable the segmentation is. CART decision tree divides the current sample set into two subsample sets, so that each nonleaf node of the growing decision tree has only binary branches. Therefore, the decision tree generated by CART algorithm is a binary tree with simple structure.

3.3. Random Forest. This study mentions a very special decision tree called random forest, which is a classifier that contains multiple decision trees, and its output category is determined by the mode of the category of the individual tree output. The algorithm of the random forest is as follows: First, N is the number of training samples, and M is the number of feature variables. Then, the number of feature variables m is used to determine the decision result of a node on the decision tree, which should be much smaller than M. Then, we sample N times (Bootstrap) from N batches of training data in the way of put back sampling to form a data set and use the undrawn samples to make prediction and evaluate the error. For each node, m feature variables are randomly selected, and the decision of each node in the decision tree is based on these feature variables. According to the m feature variables, we calculate the best segmentation method. Note that each tree will grow completely without pruning.

3.4. Index Selection and Data Sources. This study uses data from domestic companies listed on the Shanghai Stock Exchange and Shenzhen Stock Exchange, deducting companies with incomplete data and using a two-to-one approach, which is two-thirds of normal enterprise and a third of the crisis, a total of 168 data. These explanatory variables (X) data are based on the turnover rate(receivables turnover rate ($X1$) and total assets turnover rate ($X2$)), growth rate (main business income growth rate ($X3$) and total asset growth rate ($X4$)), and rate of return (return on equity ($X5$) and rate of return on assets ($X6$)), and other indicators are collected. The narrative statistics is shown in Table 1. Among them, the crisis enterprise (ST) included by the explained variable (Y) refers to the enterprise that has suffered losses for two consecutive years and is specially treated, which is represented by the value of 1 and the normal enterprise by the value of 0. In addition, in data processing, we divided all 168 data into 6 groups of data, and each group of data has 28 data. In the decision tree modeling, 5 groups of 140 data are used to establish a decision tree model, and a group of 28 data is substituted into the decision tree model to test the accuracy of the model for cross validation. Therefore, there were 6 sets of data, that is, the first group was modeled by 1–5 groups and the sixth group was tested. The second group adopted 2–6 group modeling and the first group, testing. The third group adopted 3–1 group modeling and the second

Table 1: Sample data description statistics.

	$X1$	$X2$	$X3$	$X4$	$X5$	$X6$	Y
Maximum	5147.7006	9.0348	8748.3663	3702.8692	988.2194	115.8828	1.0000
Minimum	0.2053	0.0223	−73.3733	−45.2444	−72.7279	−42.4491	0.0000
Average	126.6895	0.6820	167.5339	93.5607	145.3563	15.0707	0.3362
Standard deviation	595.8717	1.0842	978.7711	436.6452	194.7686	22.2904	0.4694

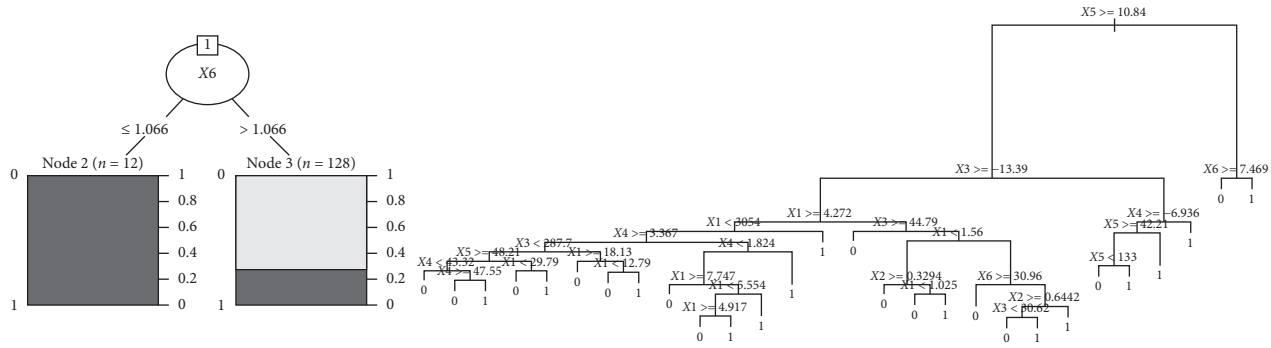


Figure 1: The first set of data modeling tree diagram.

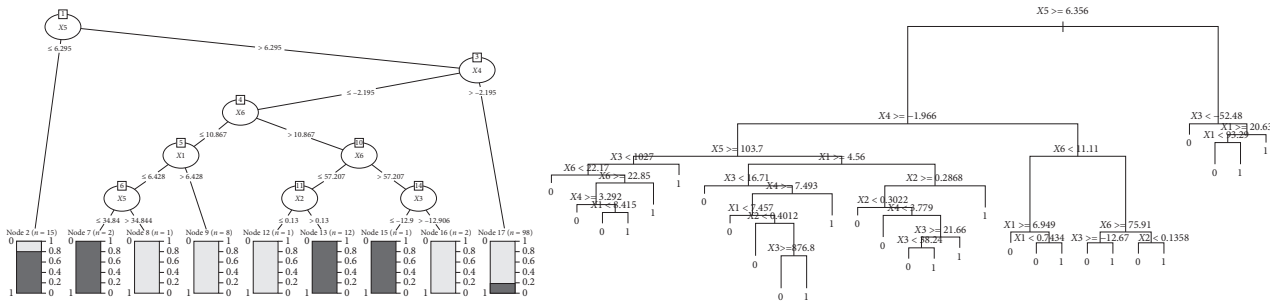


Figure 2: The second set of data modeling tree diagram.

group, testing. The fourth group adopted 4–2 group modeling and the third group, testing. The fifth group adopted group 5–3 modeling and group 4, testing. The sixth group adopted the 6–4 group modeling and the fifth group, testing. In this way, 168 pieces of data would be collected under the prediction of the three decision tree models, which could provide us with the accuracy of further analysis of the financial warning model of the three decision trees.

4. Empirical Analysis

4.1. Establishment of the Financial Early Warning Model of Three Decision Trees. In this study, three decision tree models were used to establish the domestic enterprise financial crisis early warning system, and it was divided into 6 groups and 5 groups of data to establish the decision tree model. First of all, we carried out the establishment of three decision tree models, and these models include C50 decision tree, CART decision tree, and random forest. The modeling tree diagram can be seen in Figures 1–6.

Through the tree diagram, we can find that the classification of the decision tree in each group is not the same, but the classification key factors of each model in each group are generally consistent. $X1$ appeared as a key factor to create the category 1 round, $X5$ appeared as a key factor to create the category 2 rounds, and $X6$ appeared as a key factor to create the category 3 rounds. They are, respectively, corresponding to the accounts receivable turnover ratio ($X1$), rate of return (return on equity ($X5$), and return on assets ($X6$)) and fully show that the three indicators are main factors to measure whether the listed companies are of the financial crisis. The error rate of overall modeling and testing of decision tree is shown in Table 2.

From Table 2, it is found that the classification accuracy of the C50 decision tree is not as good as CART and random forest in all training stages, and the classification error of C50 in the test stage is mostly lower than that of CART and random forest. In other words, during the testing phase, the C50 decision tree performed relatively well. If we ignore the type of model and only look at the sum of the training stage and the test stage, it is found from the table that the classification error of the training stage is lower than that of the test stage. If we observe the sum of the three classification errors, we will find that the classification error of C50 is higher than that of CART and random forest, but the difference is not obvious. Therefore, the ROC curve is needed to further explore the accuracy of the three models.

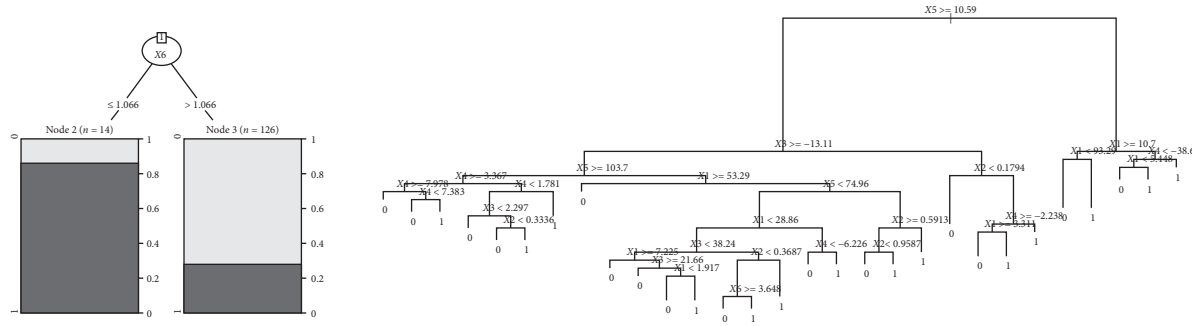


Figure 3: The third set of data modeling tree diagram.

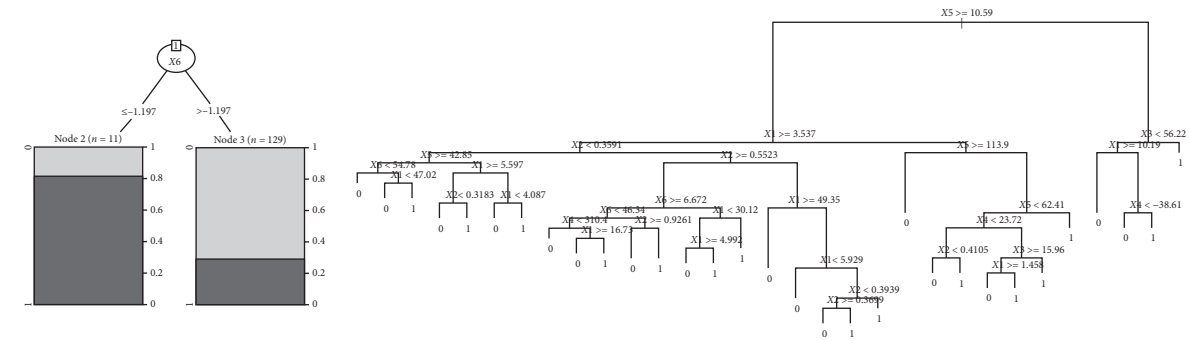


Figure 4: The fourth set of data modeling tree diagram.

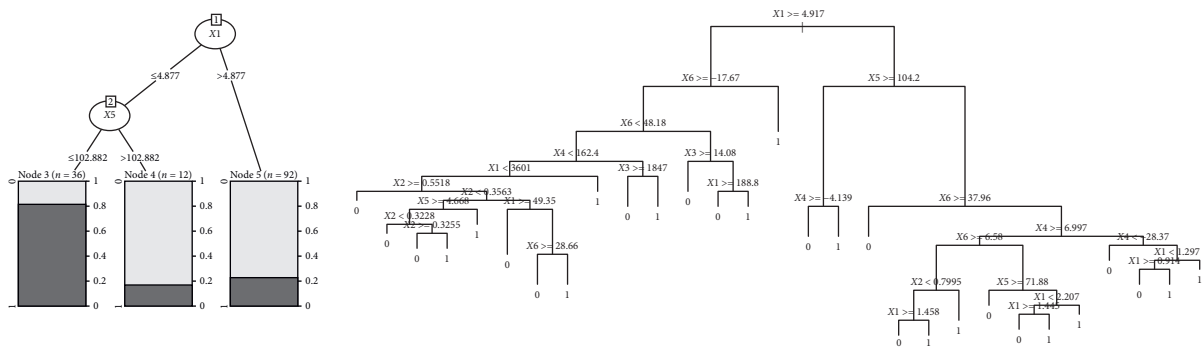


Figure 5: The fifth set of data modeling tree diagram.

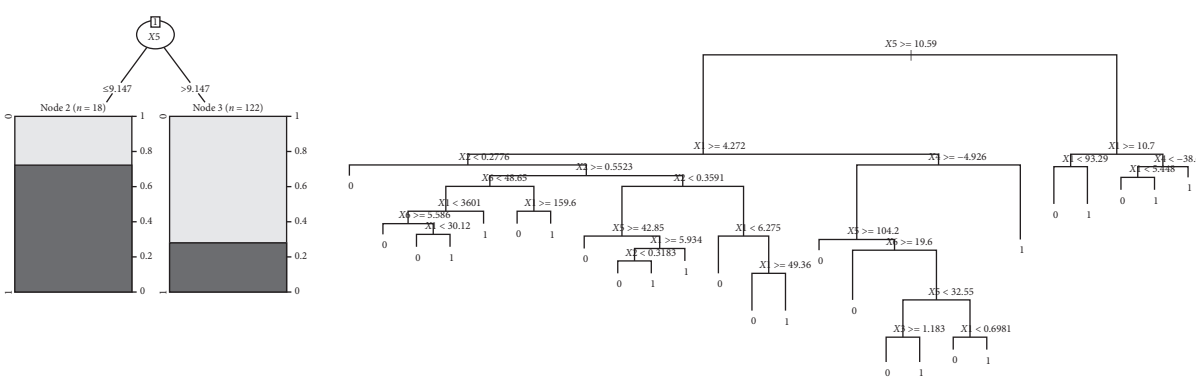


Figure 6: The sixth set of data modeling tree diagram.

Table 2: Misclassification rate of each group during training and testing.

Group	Model	C50	CART	Random forest	Total
Group 1 to group 5, training and group 6, testing	Training	0.242	0.014	0.021	0.277
	Testing	0.429	0.571	0.464	1.464
Group 2 to group 6, training and group 1, testing	Training	0.164	0	0	0.164
	Testing	0.464	0.571	0.571	1.606
Group 3 to group 1, training and group 2, testing	Training	0.264	0.021	0.014	0.299
	Testing	0.321	0.464	0.321	1.106
Group 4 to group 2, training and group 3, testing	Training	0.279	0.014	0.014	0.307
	Testing	0.285	0.25	0.25	0.785
Group 5 to group 3, training and group, 4 testing	Training	0.25	0.014	0.014	0.278
	Testing	0.321	0.286	0.321	0.928
Group 6 to group 4, training and group 5, testing	Training	0.279	0	0	0.279
	Testing	0.214	0.25	0.286	0.75
Total		3.512	2.455	2.276	8.243

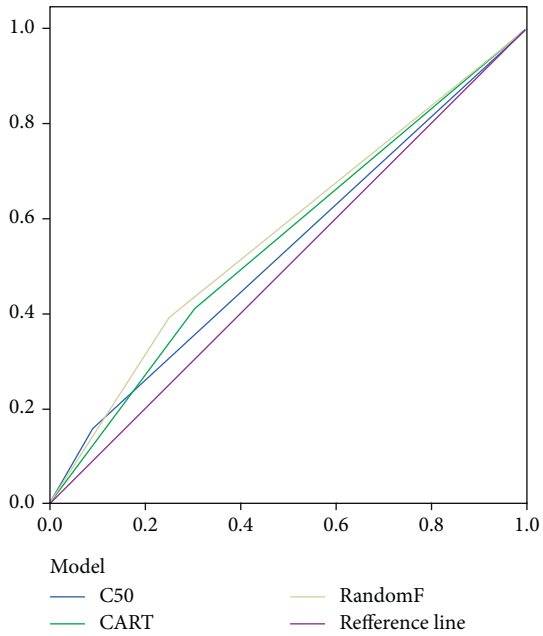


Figure 7: ROC curves of classified prediction results of three decision tree early warning models. Note: the x-axis represents specificity and the y-axis represents sensitivity.

4.2. ROC Curve of Classified Prediction Results. Figure 7 is the ROC curve of the financial warning model of 168 domestic listed enterprises, which is cross verified by the data of group 6. Bradley (1997) pointed out that the larger the area of the reference line and the area under the curve, the more accurate the classification ability of the model. It is obvious from the figure that the blue line represents the C50 model, the green line represents the CART, the yellow line represents the random forest, and the purple line represents the reference line. The area under the curve of random forest is larger than that of C50 and CART decision tree, which means that the classification ability of random forest is more accurate.

According to the output results of ROC curve analysis in Table 3, the area under the Random F curve (AUC) is 571, which is higher than that of the C50 and CART model, so it has a good ability of early warning and detection.

Table 3: Area under curve of three models.

Test result variable	Area	Standard error[a]	Asymp Sig.[b]	95% confidence interval	
				Lower limit	Upper limit
C50	0.536	0.048	0.451	0.441	0.630
CART	0.554	0.048	0.258	0.460	0.647
RandomF	0.571	0.048	0.132	0.478	0.665

Test result variables: C50, CART, and RandomF have at least one knot between the positive and negative actual state groups. Statistics can be skewed.

5. Conclusions

The main aim of this research is to discuss how to use decision trees to establish a domestic enterprise financial

early warning model. Especially in the current situation of the COVID-19 epidemic and the Sino-US trade war environment, it is very important for investors and relevant government departments to grasp the viability of enterprises. This study uses the financial data of 168 domestic enterprises in 2016 and divides the data into 6 groups, with 5 groups for modeling, 1 group for testing, and cross-validation to verify the classification accuracy of the 3 decision models. First, it can be concluded from the research results that the random forest model highlights its good classification and prediction ability in comparison. Secondly, after research, it is concluded that the three indicators of accounts receivable turnover rate, return on equity and return on assets are important factors to measure whether a listed company is experiencing financial crisis. In the future, enterprises and investors can attach importance to these indexes when referring to official data.

The disadvantage of this study is that this study only uses three decision tree models, and there are many methods in data mining that are very suitable for building financial early warning models, such as the neural network, Bayesian classification, and especially, the artificial neural network that is with a number of different species such as the fall risk transmission neural network, neural network and support vector machine(SVM), and gray neural network. Therefore, this study suggests that a similar neural network can be used to establish an enterprise financial crisis early warning system in the future.

Data Availability

All enterprise index data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Social Science Fund of China Key Research Project (Project No. 17ZDA086): Research on Reforms and Innovations of Monitoring System in State-Owned Enterprises.

References

[1] S. Zhou, J. Yang, and P. Wang, "On the early warning analysis of financial crisis—F-score model," *Accounting Research*, vol. 8, no. 8, pp. 8–11, 1996.

[2] S. Yang and W. Xu, "Financial early warning model of listed companies—an empirical study of Y score model," *China Soft Science*, vol. 1, no. 1, pp. 56–60, 2003.

[3] L. Zhang, "Discriminant model for early warning analysis of financial crisis," *Quantitative Economic Technical Economic Research*, vol. 3, no. 3, pp. 49–51, 2000.

[4] S. Yang and L. Huang, "Listed company financial early warning model based on BP neural network," *System Engineering Theory and Practice*, vol. 1, no. 1, pp. 12–18, 2005.

[5] W. Xu and D. Chen, "Financial risk early warning modeling principles and several new early warning models," *Statistics and Decision*, vol. 8, no. 8, pp. 150–153, 2016.

[6] J. Wang, "The establishment of financial early warning mathematical model of listed companies," *Commercial Accounting*, vol. 12, pp. 135-136, 2018.

[7] Q. Lin, "Research on early warning of financial crisis of listed companies based on ologit model," *Management*, vol. 12, no. 12, pp. 95–97, 2018.

[8] D. Lin, Construction and Significance of Early Warning Index System for Financial Crisis of Listed Companies.

[9] Y. Yang, "Research on information disclosure of listed companies and construction of financial risk early warning system," *Modern Business*, vol. 1, pp. 151-152, 2018.

[10] G. Ou, "Research on early warning of financial risk of real estate enterprises based on factor analysis method," *Social Scientist*, vol. 9, no. 56, pp. 56–63, 2018.

[11] S. Hu, "Financial early warning analysis based on F-score model_taking listed companies in the construction machinery industry as an example," *Financial Management and Capital Operation*, vol. 22, no. 33, pp. 33–36, 2018.

[12] G. Song, "Research on the financial risk warning model of listed companies based on deep learning," *Value Engineering*, vol. 1, no. 1, pp. 53–56, 2019.

[13] M. Zhang, "ST forecast of Chinese listed companies based on aalen additive model," *Journal of Systems Management*, vol. 28, no. 1, p. 10, 2019.

[14] Y. Xiong, "Research on financial risk early warning of listed companies based on F-score," *Risk Management*, vol. 1, no. 27, pp. 111–115, 2019.

[15] Q. Wu, "Identification and early warning of financial crisis of listed companies on the growth enterprise market," *Finance and Accounting Monthly*, vol. 2, no. 56, pp. 56–64, 2020.

[16] X. Wang, "Penalty-constrained financial risk early warning model based on cluster analysis," *Finance and Economics*, vol. 2, no. 35, pp. 153–156, 2020.

[17] D. Li, "Research on early warning of corporate financial risks in the era of big data," *Operation and Management*, vol. 1, pp. 116-117, 2020.

[18] Y. Lian, "Research on financial risk early warning and prevention under enterprise diversified investment," *Economic Management Space*, vol. 1, pp. 131-132, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 7064897, 10 pages
https://doi.org/10.1155/2020/7064897

Research Article

Research on the Impact of Logistics Technology Progress on Employment Structure Based on DEA-Malmquist Method

Xiao-qing Lei,[1] Jia-jia Yang ,[2] Jian-bo Zou,[1] and Mei-er Zhuang[3]

[1]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510006, China*
[2]*Department of Public Policy, King's College London, London WC2R 2LS, UK*
[3]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510000, China*

Correspondence should be addressed to Jia-jia Yang; 1298256371@qq.com

Received 20 August 2020; Revised 3 September 2020; Accepted 4 September 2020; Published 15 September 2020

Academic Editor: Weilin Xiao

In recent years, with the vigorous development of e-commerce, the logistics industry has also developed rapidly and has gradually become one of China's important industries. This paper takes 49 listed companies in the logistics industry in China as the research object and uses the DEA-Malmquist index method to measure the technology progress index of the logistics industry based on their panel data for a total of 10 years from 2008 to 2017. In this way, this article explores the impact of technological progress in the logistics industry and provides a reference for solving the contradiction of employment structure in China

1. Introduction

After more than 30 years of development, the logistics industry not only has become a pillar industry of China's national economy but also serves as an important modern service. In 2013, the size of China's logistics market surpassed that of the United States for the first time. Since then, China has become a veritable logistics power country. However, the development quality of China's logistics industry is generally not high, and its dependence on factor inputs is strong. Because its development model is extensive, the logistics industry has always been regarded as a labor-intensive industry. In recent years, modern logistics equipment technologies such as automated three-dimensional warehouses, automatic sorting equipment, and smart logistics equipment have developed rapidly. Information technologies such as the Internet of Things, big data, and cloud computing have been widely used in the logistics industry. The automation level of the logistics industry has been continuously improved, showing that capital is obviously deepening and technological progress is accelerating. This indicates that China's logistics industry will undergo a profound technological change, which will not only changes the development model of the logistics industry but also have an significant impact on employment in the logistics industry.

Technological progress has a dual effect on employment, namely, creation effect and shock effect. From the previous literature, it can be seen that most scholars mainly study the relationship between technological progress and employment and seldom analyze the degree of matching between technological progress and employment structure. Therefore, exploring the mechanism of technological progress in the logistics industry on employment and finding the best matching relationship between them has important practical and theoretical significance for the long-term development of the logistics industry. This article will explore the impact of technological progress in the logistics industry on the employment structure and find a balance between the two, providing a reference for scientifically formulating technological progress and employment policies in China's logistics industry.

2. Literature Review

For the research on the impact of technological progress in the logistics industry on the employment structure, this section will summarize from three perspectives, namely, the

perspective of industrial structure, the perspective of biased technological progress, and the perspective of rising unemployment.

Technological progress can indirectly drive changes in the employment structure by causing adjustments to the industrial structure. Therefore, some scholars have tried to systematically analyze the impact of technological progress and employment structure from the perspective of industrial structure. For example, Xiong and Zhu [1] started from the aspect of industrial structure changes and discussed the relationship among technological progress, changes in the three industrial structure layouts, and employment, and they claimed that changes in the industrial structure layout are negative for China's employment, which will increase the number of the unemployed. Additionally, Jiang [2] found that technological progress will first affect the change of industrial structure to a certain extent and then indirectly affect the change of employment status. Bo and Wen [3] used the data envelopment analysis method to divide the industry into resource-intensive, labor-intensive, capital-intensive, and technology-intensive sectors based on the total factor productivity of the industrial sector and analyzed the impact of technological progress on the employment of different industrial sectors in China. Furthermore, some scholars also conducted empirical analysis on the relationship between technological progress and employment structure based on the perspective of technological progress. For example, Srour et al. [4] and Buera et al. [5] used the number of patents and the internal R&D expenditure of enterprises to replace the indicators of technological progress and found that technological innovation promoted changes in the employment structure of the labor market. Moreover, Berman et al. [6] studied the application of new technologies such as computers in the United States in the 1980s, and Spitz-Oener [7] conducted empirical research based on sample data from the former Federal Republic of Germany from 1979 to 1999. The results showed that technological progress would significantly increase the relative demand for high-level personnel. Similarly, Kaplan [8] also believes that the faster the development of advanced information technology such as AI, the more types of low-skilled jobs to be replaced by information and automation equipment. In addition, some scholars have studied the relationship between technological progress and employment structure from the perspective of unemployment rate. For instance, Zhu and Li [9] used artificial intelligence as the research object, and found that some employees with low academic qualifications and first-line production type will be replaced by artificial intelligence technology. Besides, Mincer and Danninger [10] empirically studied the relationship between technological progress and the structure of employment skills from the perspective of the unemployment rate based on the sample statistics of the United States from 1970 to 1995. The results showed that the increase in technological progress would lead to the increasing unemployment of low-skilled personnel. To be more specific, through research Acemoglu [11] argued that the technological progress contributes to the simultaneous rises of unemployment in both low-skilled workers and high-skilled staff. Apart from that, Katz and Murphy [12] argued that high-skilled skewed technological progress will lead to an increasingly relative demand for high-skilled personnel, but at the same time, it will reduce the relative demand for low-skilled labor.

In summary, related research on this topic mainly presents the following characteristics: first, it is mainly based on the entire national economic sector or industry. In other words, manufacturing is the study background, and there is a lack of relevant research from the microenterprises or industry level; second, empirical analysis is mostly based on provincial panel data, and many studies use sample data before 2010, while technological progress has changed rapidly, especially in recent years, so the current level of technological progress will be largely ignored. Third, the employment structure should reflect changes in the relative numbers of different types of labor, including education level, gender ratio, professional structure, skill training, and age ratio. The most significant feature is the employment changes of highly skilled and low-skilled personnel in terms of their quantity and structure. In order to make up for the limitations of previous research, this article starts from a microperspective. Based on the panel data of 49 listed companies in China's logistics industry, this study on the relationship between technological progress and employment structure can fill up some of the gaps in the field of research, which has strong practical significance.

3. DEA-Malmquist Method

Rolf Fare proposed the Malmquist index method in 1994 on the basis of the nonparametric linear programming method. This new method is to measure the total factor productivity (TFP) index and then to study the technological progress indicators. Meanwhile, the total factor productivity index was effectively decomposed into a technical efficiency index and a technology change index based on a distance function for the first time:

$$M_0^{t+1} = \left[\frac{D^t\left(x_0^{t+1}, y_0^{t+1}\right)}{D^t\left(x_0^t, y_0^t\right)} \times \frac{D^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)}{D^{t+1}\left(x_0^t, y_0^t\right)}\right]^{1/2} = \frac{D^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)}{D^t\left(x_0^t, y_0^t\right)} \left[\frac{D^t\left(x_0^{t+1}, y_0^{t+1}\right)}{D^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)} \times \frac{D^t\left(x_0^t, y_0^t\right)}{D^{t+1}\left(x_0^t, y_0^t\right)}\right]^{1/2}. \tag{1}$$

Under the condition of constant return to scale, TFP can be decomposed into the technical progress index (tech) and technical efficiency index (effch):

$$M_0\left(x^{t+1}, y^{t+1}, x^t, y^t\right) = \left[\frac{D_0^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_0^{t+1}\left(x^t, y^t\right)} \times \frac{D_0^t\left(x^{t+1}, y^{t+1}\right)}{D_0^t\left(x^t, y^t\right)}\right]^{1/2} \times \frac{D_0^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_0^{t+1}\left(x^t, y^t\right)} = \text{effch}^* \text{tech}. \tag{2}$$

If the technical efficiency index effch > 1, it indicates that the technical efficiency is improved, and vice versa. This

decomposition can clearly understand the changes and trends of total factor productivity, further analyzes the factors that lead to changes, explores the influence coefficient of each factor, and then explores the focus and improvement space for its progress and growth. Through further exploration of technical efficiency, the changes can be further decomposed into pure technical efficiency and scale efficiency:

$$M_0\left(x^{t+1}, y^{t+1}, x^t, y^t\right) = \left[\frac{D_0^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_0^{t+1}\left(x^t, y^t\right)} \times \frac{D_0^t\left(x^{t+1}, y^{t+1}\right)}{D_0^t\left(x^t, y^t\right)}\right]^{1/2} \times \frac{D_0^{t+1}\left(x^{t+1}, y^{t+1}\right)}{D_0^{t+1}\left(x^t, y^t\right)} = \text{pech}^*\text{sech}^*\text{tech}. \tag{3}$$

This article uses the DEA-Malmquist method for analysis, mainly for the following reasons: first, the production model function corresponding to the DEA-Malmquist index method does not need to be uniquely determined, which can effectively avoid the problem of inaccurate measurement results caused by improper model processing. Secondly, the Malmquist index method does not require various research hypotheses, it can better reflect the true situation of the current market, and the calculated results are more convincing. Finally, the results calculated by this method can not only help explain the differences in economic development levels but also clearly know how to effectively increase productivity levels through the decomposition of total factor productivity. Compared with traditional estimation methods, various factors affecting productivity can be considered in many aspects, including management, technical level, resource allocation, and technological mastery. Traditional estimation methods can only estimate a specific value but cannot reflect these influencing factors play a role in productivity improvement.

However, due to differences in research objects or perspectives, different scholars have different opinions and choices on input variables and output variables, which leads to the various selections of input and output elements when they use the DEA-Malmquist index method to measure the total factor productivity index and its decomposition. Based on previous results and combining the characteristics of China's logistics industry, this paper selects input elements from the perspectives of equipment input, human capital, and business operations. They are net fixed assets, total wages, and operating costs as input indicators, while output indicators are main business income, using DEAP2.1 software to measure the technological progress index of the logistics industry.

4. Empirical Analysis of Technological Progress in China's Logistics Industry on Employment Structure

This chapter uses the statistical data of 49 listed companies in the logistics industry of Shanghai and Shenzhen A shares from 2009 to 2017 to conduct an empirical analysis of the relationship between technological progress and employment structure in China's logistics industry through research methods such as unit root test, cointegration analysis, and regression analysis.

4.1. Construction of Regression Model. Based on the theory of biased technological progress and combined with the status quo of China's logistics industry, relevant indicators are selected to construct an econometric regression model:

$$y1 = C + \alpha 1x1 + \beta 1x2 + \gamma 1x3 + \delta 1x4, \tag{4}$$

$$y2 = C + \alpha 2x1 + \beta 2x2 + \gamma 2x3 + \delta 2x4. \tag{5}$$

Among them, the explained variables $y1$ and $y2$, respectively, represent the proportion of personnel with a college degree or above and the proportion of nonproduction personnel. The explanatory variable $x1$ is the technological progress of the logistics industry replaced by the technological progress index; at the same time, three control variables ($x2$, $x3$, and $x4$) are added to improve the interpretation of the regression model. They are the size of the enterprise measured by the main business income and the employee training structure replaced by per capita training investment and salary structure represented by per capita salary of employees.

4.2. Variable Description and Data Source and Processing

4.2.1. Variable Description

(1) Selection of Explained Variables. For the definition of the explanatory variable indicators, there are currently two main methods of classification: the first is to regard the proportion of persons with a college degree or above as high-skilled labor, and those below college are low-skilled labor; that is, the observable education premium is used to reflect the skill gains; The second is to regard nonproductive labor as high-skilled labor, and those engaged in productive labor are low-skilled labor. The second classification method is widely used in the study of wage differences in industrial enterprises above designated size or large- and medium-sized industrial enterprises. For example, the scientific and technological personnel of large- and medium-sized manufacturing enterprises used this way as a substitute index for skilled labor in manufacturing, and the labor service fees represent the wage level of skilled labor. In previous studies, only one of these methods was often selected for index measurement, which lacked comprehensiveness to a certain extent. Therefore, this article uses both college degree and above and nonproductive labor as substitute variables for high-skilled labor, which not only fully reflects the changes in the structure of high- and low-skilled labor in the industry but also enriches the related information in the research through comparative analysis.

(2) Selection of Explanatory Variables. In the previous literature, explanatory variables usually used indicators such as total factor productivity index, technological progress index,

R&D, number of patents granted, and value of machinery and equipment per capita. Since this article is based on listed companies in the logistics industry as the research object on the basis of combining industry conditions and considering the availability of data and the objectivity of the research, this paper use the technology progress index measured by the DEA-Malmquist index method to represent the technology progress, and it can fully reflect the current situation of China's logistics industry through selecting reasonable input-output indicators.

(3) Selection of Control Variables. In order to accurately reflect the impact of technological progress in the logistics industry and control other factors that may affect the structure of employment skills, this paper takes enterprise scale, training structure, and salary structure as the control variables of the model. Generally speaking, the larger the scale of an enterprise, the more manpower and material resources are needed in order to improve production efficiency, so the demand for highly skilled talents will increase relatively. In addition, training investment is also an indispensable factor in the development of a company. Good training can improve the overall quality and abilities of the company's employees and increase the company's production efficiency, which in turn will attract some external talents. Finally, this article uses per capita wage to replace the wage structure. Except for a good platform or a complete training system, high wages will also attract new blood. At the same time, companies with higher wage levels tend to have more powerful company's strength, promising development prospects, and higher management level, and the greater demand for high-level talents In order to eliminate the dimensional relationship between variables and make the data comparable, this article standardizes the analysis data of all variables before the empirical test, as shown in Table 1.

4.2.2. Data Source and Processing. Based on the definition and classification of the logistics industry, there are currently 115 listed logistics companies in China. In order to ensure the validity and accuracy of the data and research results, this article needs to eliminate some noncompliant companies. This article exclude companies that do not meet the conditions in accordance with the following principles: first, companies with insufficient data years; second, companies whose annual report data classification rules differ significantly from those of other companies; third, companies that disclose data between 2008 and 2017 years; fourth, companies that have too few staff and do not have required indicator data. Finally, there are 49 listed companies in the logistics industry that meet the needs of this study, including 8 in the air transport industry, 21 in the road transport industry, 3 in the railway transport industry, and 15 in the water transport industry, There is one firm in each industry, namely, loading and unloading industry transportation agency, warehousing, and business service industry as shown in Table 2.

4.3. Empirical Analysis

4.3.1. Unit Root Test. Before performing regression analysis on panel sequence data, it is necessary to test the stationarity of the research data first to determine whether it is a stationary sequence of the same order. Among the methods for testing the stationarity of panel data, the commonly used method is unit root testing. This paper adopts the unit root test and uses the three methods of LLC, ADF-Fisher, and PP-Fisher to analyze the stability of panel data (see Table 3).

It can be found from Tables 2-4 that the results of $X2$ and $X4$ in the ADF-Fisher test are not significant, and the results of $X4$ in the PP-Fisher test are not significant, which indicates that the test values of these variables cannot reject the null hypothesis with existence of panel unit roots. The first-order difference of each variable is significant at the 1% significance level, indicating that there is no panel unit root.

4.3.2. Cointegration Test. In addition, in order to ensure the long-term stability between the variables in this model and effectively prevent spurious regression, it is necessary to conduct a cointegration test on each variable. Based on the research of previous scholars, this paper uses the Pedroni test and Kao test at the same time and uses Eviews9.0 analysis software to do the cointegration analysis test for each research variable. The specific inspection results are shown in Table 4.

From the results of the cointegration test in Table 4, we can find that the Pedroni test's test statistic value divided by the panel ADF test value in model (1) is significant at the 5% significance level, and the other test values are all at 1% significant. The test statistics of the Kao test are significant at the 5% and 1% significance levels, respectively. Both test methods are said to obviously reject the null hypothesis. This shows that there is a cointegration relationship between the explanatory variable and the explained variable, which meets the stability requirements of time or panel data, and there is a long-term stable relationship between the variables, which is suitable for regression analysis.

4.4. Regression Analysis. The estimation of panel data model usually adopts the individual fixed effect model or individual random effect model. Therefore, before performing regression analysis, performing the Hausman test to determine which regression model should be based on its test results. The test results are shown in Table 5; that is, the P values in model (1) and model (2) are both less than 0.05, significantly rejecting the null hypothesis and indicating that the fixed-effects model should be used for regression analysis.

In addition, considering that there may be heteroscedasticity problems in the model, the regression results of the model seem not robust. Therefore, the "fixed effect model + robust standard error" is adopted; that is, the panel calibration standard error (PCSE) estimates model (1) and the model (2). The regression results of the model are shown in Table 6.

Table 1: Index selection and data sources.

Research variables	First-level indicator	Metrics	Data sources
Explanatory variable ($X1$)	Technological progress in the logistics industry	Technology progress index ($X1$)	Cathay pacific database
Explained variables ($Y1$, $Y2$)	Employment skill structure	Proportion of employees with college degree and above ($Y1$)	Listed company annual report
Proportion of nonproduction personnel ($Y2$)			
Control variable ($X2$, $X3$, $X4$)	Enterprise size ($X2$)	Main business income	Cathay pacific database
Training structure ($X3$)	Training investment per employee	Listed company annual report	
Salary structure ($X4$)	Salary per employee	Listed company annual report	

Table 2: Sample enterprises.

Enterprise code	Company abbreviation	Industry
000088	Yantian Port	Road transport industry
000089	Shenzhen Airport	Air transport industry
000099	CITIC Haizhi	Air transport industry
000429	Guangdong Expressway A	Road transport industry
000507	Zhuhai Port	Water transport industry
000582	Beibu Gulf Port	Water transport industry
000828	Dongguan Holdings	Road transport industry
000885	City Development Environment	Road transport industry
000900	Modern Investment	Road transport industry
000905	Xiamen Port	Water transport industry
002010	Chuanhua Zhilian	Business service industry
002040	Nanjing Port	Water transport industry
600009	Shanghai Airport	Air transport industry
600012	Anhui Expressway	Road transport industry
600017	Rizhao Port	Water transport industry
600020	Zhongyuan Expressway	Road transport industry
600026	COSCO Shipping Energy	Water transport industry
600029	Southern Airline	Air transport industry
600033	Fujian Express	Road transport industry
600035	Chutian Expressway	Road transport industry
600106	Chongqing Road and Bridge	Road transport industry
600115	Eastern Airlines	Air transport industry
600119	Yangtze River Investment	Road transport industry
600125	Iron Dragon Logistics	Railway transportation industry
600190	Port of Jinzhou	Water transport industry
600221	HNA Holdings	Air transport industry
600269	Ganyue Expressway	Road transport industry
600270	Sinotrans Development	Handling
600279	Chongqing Port and Kowloon	Water transport industry
600317	Yingkou Port	Water transport industry
600350	Shandong Expressway	Road transport industry
600368	Wuzhou Transportation	Road transport industry
600377	Nanjing-Shanghai Expressway	Road transport industry
600428	COSCO Haite	Water transport industry
600548	Shenzhen Expressway	Road transport industry
600561	Jiangxi Changyun	Road transport industry
600575	Wanjiang Logistics	Water transport industry
600611	Public Transportation	Road transport industry
600650	Jinjiang Investment	Xiamen airport
600662	Johnson & Johnson Holdings	Daqin Railway
600676	Delivery Shares	Lianyungang
600717	Tianjin Harbor	Air China
600787	China Reserve	Guangzhou-Shenzhen Railway
600798	Ningbo Shipping	
600897	Xiamen Airport	Air transport industry
601006	Daqin Railway	Railway transportation industry
601008	Lianyungang	Water transport industry
601111	Air China	Air transport industry
601333	Guangzhou-Shenzhen Railway	Railway transportation industry

Table 3: Unit root test results.

Variables	LLC	ADF-Fisher	PP-Fisher
$Y1$	−4.29407***	132.633**	159.769***
$\Delta Y1$	−16.7506***	260.739***	381.058***
$Y2$	−17.5232***	156.507***	128.048**
$\Delta Y2$	−24.5752***	282.135***	348.022***
$X1$	−14.7307***	227.555**	280.123***
$\Delta X1$	−24.5421***	341.662***	529.502***
$X2$	−3.38952***	101.830	152.253***
$\Delta X2$	−5.38774***	177.549***	220.121***
$X3$	−8.32027***	153.784***	205.803***
$\Delta X3$	−30.8146***	272.010***	306.273***
$X4$	−2.64603***	92.6991	110.325
$\Delta X4$	−20.4947***	268.751***	379.493***

Note. *, **, and ***, respectively, indicate significant at the 10% level, significant at the 5% level, and significant at the 1% level.

Table 4: Cointegration test results.

	Statistics	Model (1)	Model (2)
Pedroni test	Panel v	−2.841026 (0.9978)	−3.668634 (0.9999)
	Panel rho	6.335616 (1.0000)	5.741516 (1.0000)
	Panel PP	−2.994059 (0.0014)	−3.768672 (0.0001)
	Panel ADF	−2.187057 (0.0144)	−3.269702 (0.0005)
	Group rho	8.957185 (1.0000)	9.126544 (1.0000)
	Group PP	−11.45461 (0.0000)	−11.36620 (0.0000)
	Group ADF	−5.795894 (0.0000)	−4.448520 (0.0000)
Kao test	ADF	−1.931086 (0.0267)	−4.780048 (0.0000)

Table 5: Hausman test.

Detection value	Model (1)	Model (2)
Chi-sq.	7.225	10.442
Prob.	0.0245	0.0336
Result	FE	FE

Note. Chi-sq. means chi-square value, FE means individual fixed effect, and RE means individual random effect.

(1) Technological progress is positively correlated with the proportion of personnel with a college degree or above and is significant at a significance level of 5%; technological advancement positively and significantly affects the proportion of nonproductive personnel at a significance level of 10% ratio. This shows that with the continuous investment in technology in China's logistics industry, the relative demand for high-skilled labor is increasing, and there is a certain degree of skill-biased technological progress in China's logistics industry.

(2) There is a positive correlation between the scale of the enterprise and the proportion of employees with a college degree or above, and it is significant at the 1% significance level, which indicates that the expansion of the scale of logistics enterprises will lead to a significant increase in the proportion of employees with a high degree of education. Generally speaking, the higher the management level of the company, the more the resources it has. The more attractive it is to the new generation of young people, the faster costs can be obtained. At the same time, the scale of the enterprise and the proportion of nonproductive employees show a negative correlation. It is significant at the 1% significance level, which shows that the expansion of the logistics enterprise scale will reduce the growth of the proportion of nonproductive employees.

(3) The relationship between the training structure and the proportion of personnel with college degree or above and the proportion of nonproductive employees is not significant, indicating that the training investment in Chinese logistics companies does not actually affect the number of highly skilled personnel and nonproductive personnel. To what extent the increase in quantity has contributed to the increase, it can also be seen from the other hand that China's logistics industry still has major problems in the construction of training system and training input and output. The training mechanism needs to be improved as soon as possible to improve the efficiency of enterprise input and output.

(4) The 10% significance level of the wage structure has a positive and significant impact on the proportion of nonproductive personnel, and the coefficient of the wage structure is 0.463, but there is no obvious significant relationship with the proportion of personnel with a college degree or above. This shows that the increase in employee wages will significantly increase the proportion of high-educated and management personnel in the company.

4.5. Treatment of Endogenous Problems. From the results of Hausman's test, we can know that the regression of panel data in this paper uses an individual fixed-effects model, and the consistency of the model estimates requires that the explanatory variables are independent of the random error term.

This article will use the method of lagging all the explanatory variables in the econometric model to deal with the endogenous problem and still use the individual fixed-effects model to estimate. However, since the current data of listed companies in the logistics industry do not have statistics for 2018, the data of all the explained variables are preceded by a model regression. This approach can not only effectively reduce the estimation bias caused by the possible endogenous problems of explanatory variables but also effectively reflect the time lag effect of technological progress and other explanatory variables on the employment skill structure. The specific regression results are shown in Tables 4–6.

The results show that compared with the regression results of individual fixed effects without considering endogeneity, the relationship between technological

Table 6: Model regression analysis results.

Model	Model (1)	Model (2)	Endogenous treatment (explanatory variables lag one period)	
Research variables	Proportion of employees with college degree and above ($Y1$)	Percentage of nonproductive personnel ($Y2$)	Proportion of employees with college degree and above ($Y1$)	Percentage of nonproductive personnel ($Y2$)
Constant (C)	0.440194*** (0.016699)	0.333155*** (0.013669)	0.418368*** (0.014612)	0.331739*** (0.00925)
Technical surgery progress ($X1$)	0.022331** (0.081772)	0.117531* (0.067472)	0.092411*** (0.031450)	0.092193* (0.066930)
Enterprise size ($X2$)	0.458099*** (0.144343)	−0.298354*** (0.095644)	0.441151*** (0.141801)	−0.181941** (0.087253)
Training structure ($X3$)	0.058520 (0.155537)	−0.294763 (0.221910)	0.109641 (0.148890)	0.005133 (0.095176)
Salary structure ($X4$)	0.235782 (0.159337)	0.462559* (0.247141)	0.061670 (0.148312)	−0.131753 (0.096665)
Adjusted R^2	0.821132	0.726680	0.821009	0.712306

Note. * Represents the 10% significance level, ** represents the 5% significance level, and *** represents the 1% significance level.

progress and enterprise size and employment structure all presents a significant positive correlation. However, the number and significance level of the estimated coefficients of the above variables have decreased to varying degrees due to the decrease in sample observations, which indicates that the significantly positive estimated coefficients of technological progress and firm size in the econometric regression equation are not affected by the endogenous problems caused by the correlation between the current explanatory variables and the current residuals influences.

4.6. Empirical Analysis from the Perspective of Industry Segments. Based on China's specific conditions, different subsectors of the logistics industry have their own characteristics in terms of input and output of technological progress and employment structure. Therefore, in order to better promote the development of the logistics industry, a clear understanding of the differences and characteristics of the technological progress and employment structure of each logistics subindustry will help the progress and development of the logistics industry, and also help clarify government's policy direction. The most important thing is that the classification of logistics companies enables companies to position themselves more clearly and enable them to give full play to their advantages. Therefore, this section analyzes the relationship between technological progress and employment structure from the perspective of subindustry. However, because the index and related data that do not meet the research requirements of this article are eliminated at the beginning of the data processing stage, this article only selects the three major subdivisions, namely, the air transport industry, road transport industry, and water transport industry to analyze and compare after considering the sample data volume.

First, the Hausman test is performed on different measurement models of each subindustry to determine which model to choose for regression analysis. The specific analysis results are shown in Table 7.

As shown in Tables 4–7, for the air transport industry and the road transport industry, the Hausman test results indicate that the individual fixed-effects model should be selected for model regression analysis. For the water transport industry, the P value of the model is greater than 0.05. The effect model is subjected to regression analysis, and the specific model regression results are shown in Table 8.

Table 7: Hausman test results of the measurement model of each subindustry.

Industry	Detection value	Model (3)	Model (4)
Air transport industry	Chi-sq.	22.038	14.216
	Prob.	0.0002	0.0066
	Result	FE	FE
Road transport industry	Chi-sq.	8.075	19.757
	Prob.	0.0089	0.0006
	Result	FE	FE
Water transport industry	Chi-sq.	1.966	7.231
	Prob.	0.7420	0.1242
	Result	RE	RE

Note. Chi-sq. means the chi-square value, FE means the fixed-effects model, and RE means the random effects model.

5. Discussion

From the analysis of the results in Tables 4–7, the following conclusions can be drawn:

(1) First, from the perspective of technological progress, the relationship between technological progress in the air transport industry and the proportion of personnel with a college degree or above is positively significant at the 5% significance level, and the coefficient of the technological progress index is 0.394, which is much larger than the other two industries; the road transport industry technological progress and the proportion of employees with a college degree and above are positively significant at a significance level of 10%, and the proportion of nonproduction employees is positively significant at a significance level of 5%; In the water transportation industry, technological progress has not had a significant impact on the employment skill structure. It can be seen from the nature of the industry that the air transportation industry is an industry that places great emphasis on safety and stability, so the requirements for employee skills are also extremely strict. At the same time, the air transportation

Table 8: Regression analysis results of each logistics segment industry model.

Research study variables	Subdivided industry	Model (3) Proportion of employees with college degree and above ($Y1$)	Model (4) Percentage of nonproductive personnel ($Y2$)
Constant Ci	Air transport industry	0.592990***	0.470530***
	Road transport industry	0.343011***	0.337696***
	Water transport industry	0.391089***	0.265077***
Skill improved $X1i$	Air transport industry	0.394020**	−0.132142
	Road transport industry	0.096592*	0.284305**
	Water transport industry	0.044584	0.011655
Enterprise size $X2i$	Air transport industry	0.323837*	−0.318394***
	Road transport industry	1.499344*	0.488669
	Water transport industry	0.487271**	−0.553805
Training structure $X3i$	Air transport industry	0.036073*	−0.103128
	Road transport industry	0.389798	0.408890
	Water transport industry	0.242129	−0.787076***
Salary structure $X4i$	Air transport industry	0.130568	0.020490
	Road transport industry	3.757804***	−2.709841**
	Water transport industry	0.865081***	2.224235***

Note. $X1i$, $X2i$, $X3i$, and $X4i$ in the table represent three different subindustries corresponding to technological progress, enterprise scale, training structure, and work structure variables, respectively. For example, $X11$ represents the technological progress of the air transport industry, where $i = 1, 2, 3$. In addition, *** and *** indicate the significance level of 10%, 5%, and 1%, respectively.

industry is a large state-owned monopoly industry with a high level of personnel management. In addition, the professional classification of the air transportation industry includes a large number of engineering and information technology personnel. It can be seen that the air transportation industry has a greater demand for highly skilled personnel and management requirements. For the road transportation industry and water transportation industry, it is mainly based on skill research and job title evaluation, and the requirements for academic qualifications are not very high. There is even no classification of technical, engineering, or information personnel in many companies, especially in the water transportation industry.

(2) From the perspective of enterprise scale, the enterprise scale of each subindustry has a positive correlation with the proportion of employees with college degree and above, and they are significant at the significance levels of 10%, 10% and 5%, respectively. In addition, the technological progress of the air transport industry is negatively correlated with the proportion of nonproductive personnel at the 1% significance level. To sum up, the expansion of the scale of air transportation, road transportation, and water transportation companies is conducive to increasing the proportion of highly educated employees in the enterprise, and the role of road transportation is the most obvious. The expansion of the scale of the air transport industry will reduce the proportion of nonproductive employees in the company. This may be because with the continuous development of the company, the number of nonproductive personnel has increased to a certain extent, but the growth rate of personnel is still relatively not very obvious.

Although China's logistics industry has developed rapidly in recent years, the scale of the company has continued to expand, and the equipment, R&D investment, human capital investment, resource allocation, and management level have all been greatly improved, but overall there is still a big gap compared with the logistics level of developed countries such as the United States in lots of aspects, such as insufficient output efficiency, unbalanced overall quality of employees, older employment age, and more productive front-line employees and so forth. Therefore, it can be considered that although the scale of enterprises continues to increase, there is still a large shortage of supply and demand for high-skilled management talents in the logistics industry.

(3) From the perspective of training structure, the training structure of the air transport industry and the proportion of personnel with a college degree or above are positively significant at the significance level of 10%, indicating that the increase in the investment in personnel training in the air transport industry will have positive impacts on the increase in the proportion of personnel with education qualifications and above. For the road transportation industry, the regression results of the model are not significant, indicating that the road transportation industry's investment in skills training is not significantly related to the increase in the proportion of highly educated or nonproductive personnel. In addition, there is no significant correlation between the training structure of the water transportation industry and the proportion of personnel with a college degree or above, but the water transportation industry is negatively correlated with the proportion of nonproductive personnel at the 1% significance level. Investing in training will significantly increase

the proportion of production personnel. From the perspective of the professional structure of personnel in the water transport industry, it mainly includes engineering and technical personnel, management personnel, transport crews, political workers, and other categories, and the proportion of crew members is much larger than that of management and technical personnel, and even some companies do not have technology personnel classification. It can be seen that the water transportation industry pays more attention to the operation training of production labor, which is conducive to the improvement of input and output efficiency on the one hand and can reduce the company's employment costs on the other hand.

(4) From the perspective of wage structure, with the exception of the air transportation industry, there is no significant correlation between the variables, and the relationship between the variables in the other two industries shows a significant correlation. Among them, the wage structure of the road transportation industry not only has a positive correlation with the proportion of employees with a college degree or above at a significant level of 1% but also has a negative correlation with the proportion of nonproductive personnel at a significant level of 5%. Besides, the wage structure of the water transportation industry has a significant positive correlation with the proportion of employees with a college degree or above and the proportion of nonproductive personnel at a significant level of 1%. This shows that the increase in employee wages in the road and water transportation industry will significantly increase the proportion of high-educated employees in the company.

6. Conclusion

This article mainly analyzes the relationship between technological progress and employment structure in China's logistics industry and analyzes the industry characteristics reflected in the level of technological progress and employment structure in various subsectors of the logistics industry, and then it explores the adaptability between employment structures and level of technological progress. On the whole, technological progress in China's logistics industry will have a significant positive impact on the structure of employment skills. With the continuous improvement of the level of information technology, both the proportion of personnel with a college degree or above and the proportion of nonproduction personnel are showing an increasing trend. Secondly, the expansion of the scale of Chinese logistics companies will significantly increase the relative demand for highly educated personnel, but it will also inhibit the growth of the proportion of nonproductive personnel. That is, the larger the scale of the company, the greater the proportion of personnel with a college degree or above and the more decreases of the relative proportion of nonproduction personnel. Finally, the increase in the wage level of employees in China's logistics industry will significantly increase the proportion of nonproductive personnel in the company, but it will not have a significant impact on the increase in personnel with a college degree or above.

The contributions of this article are as follows: first, this article is an industry-specific microresearch, which bridge the research gap. At present, most of the literature on the relationship between technological progress and employment is at the macrolevel, and the research objects are mainly industry or manufacturing. There is very little research in discussing the relationship between technological progress and the structure of employment skills in the context of technological progress, let alone on the logistics industry. Therefore, starting from a microperspective, this article has conducted in-depth research on the panel data of 49 listed companies in Chinese logistics industry to make up for this research gap. Second, this essay has innovatively enriched the measurement system of technological progress indicators. It can be found that most scholars directly choose a single indicator or the arithmetic average of two indicators to replace technological progress for analysis, but because there are many factors that reflect technological progress in the logistics industry, such choice will ignore some important details and characteristics of subindustry, which inevitably causes large errors in the research results. This article not only selects net fixed assets, total employee wages, and operating costs as input indicators from the perspective of machinery and equipment, human capital investment, and management and operation but also takes the main business income as the output indicators to measure the technological progress index of the logistics industry and use it to represent technological progress. In this way, it greatly reflects the characteristics of the industry and increases the accuracy of the research results.

However, there are also some limitations in the research. To be more specific, this article uses data from listed companies in the logistics industry. However, there are certain differences in the data classification rules for the employment structure of listed companies in the logistics industry. Some companies do not have the classification of production and technical employees, and they have different subindustries, which will inevitably affect the results of the research to a certain extent and cause certain errors. Therefore, future research can consider making more improvements in this area to make up for the limitations of this research.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This article was supported by theGuangdong Province Philosophy and Social Science Planning Project, "Research

on the Optimization of Labor Skills Structure of China's Service Industry under the New Normal Based on the Perspective of Technological Progress" (GD15CYJ04); Guangdong Province General University Characteristic Innovation Project, "Technical Progress Direction and Labor Skills of China's Service Industry under the New Normal Research on Structural Optimization" (15T12); 2017 Guangdong Undergraduate Higher Education Teaching Reform Project, "Research on Logistics Professional Course System Based on Service Science"; and the Key Scientific Research Projects of Universities in Guangdong Province (2018WTSCX034).

References

[1] S. Xiong and Y. Zhu, "The experience of technological progress and employment growth in the eastern region," *Contemporary Economy*, vol. 36, no. 12, pp. 100-101, 2008.

[2] M. Jiang, "Analysis of the impact of technological progress on employment," *Science and Technology Management Research*, vol. 40, no. 1, pp. 70–73, 2007.

[3] L. Bo and J. Wen, "The employment effect of technological progress in China's industrial sector," *Economic Developments*, vol. 61, no. 10, pp. 34–37, 2010.

[4] I. Srour, E. Taymaz, and M. Vivarelli, *Skill-Biased Technological Change and Skill-Enhancing Trade in Turkey: Evidence from Longitudinal Microdata*, IZA—Institute of Labor Economics, Bonn, Germany, 2013.

[5] R. Rogerson, J. P. Kaboski, and F. Buera, "Skill-biased structural change and the skill premium," in *Proceedings of the 2015 Meeting Papers, Society for Economic Dynamics*, Warsaw, Poland, 2015, https://ideas.repec.org/p/red/sed015/895.html#download.

[6] E. Berman, J. Bound, and S. Machin, "Implications of skill-biased technological change: international evidence," *The Quarterly Journal of Economics*, vol. 113, no. 4, pp. 1245–1279, 1998.

[7] A. Spitz-Oener, "Technical change, job tasks, and rising educational demands: looking outside the wage structure," *Journal of Labor Economics*, vol. 24, no. 2, pp. 235–270, 2006.

[8] J. Kaplan, *Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence*, Yale University Press, London, UK, 2015.

[9] Q. Zhu and M. Li, "Research on countermeasures for artificial intelligence, technological progress and labor structure optimization," *Science and Technology Progress and Countermeasures*, vol. 35, no. 6, pp. 36–41, 2018.

[10] J. Mincer and S. Danninger, "Technology, unemployment and inflation," *NBER Working Paper*, National Bureau of Economic Research, Cambridge, MA, USA, 2000.

[11] D. Acemoglu, "Changes in unemployment and wage inequality: an alternative theory and some evidence," *American Economic Review*, vol. 89, no. 5, pp. 1259–1278, 1999.

[12] L. F. Katz and K. M. Murphy, "Changes in relative wages, 1963–1987: supply and demand factors," *The Quarterly Journal of Economics*, vol. 107, no. 1, pp. 35–78, 1992.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 5045207, 9 pages
https://doi.org/10.1155/2020/5045207

Research Article

The Establishment of a Financial Crisis Early Warning System for Domestic Listed Companies Based on Two Neural Network Models in the Context of COVID-19

Feixiong-Ma,[1] Yingying-Zhou ,[1] Xiaoyan-Mo,[2] and Yiwei-Xia[1]

[1]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510006, China*
[2]*School of English and Education, Guangdong University of Foreign Studies, Guangzhou 510006, China*

Correspondence should be addressed to Yingying-Zhou; 20170401730@gdufs.edu.cn

Received 15 May 2020; Accepted 15 July 2020; Published 24 August 2020

Academic Editor: Kalyana C. Veluvolu

In the context of COVID-19, many companies have been affected by the financial crisis. In order to carry out a comparative study on the accuracy of the company's financial crisis early warning method, this study used RPROP artificial neural network and support vector machine, with 162 listed companies' two-year panel financial indicator data as a model sample, and the test sample established a financial crisis early warning model. The theory of comprehensive evaluation combining two kinds of neural network methods is put forward innovatively. The predicted results can strengthen the supervision of the listed companies with risks by themselves and others and have important economic and social significance to ensure the stable operation of the listed companies, the securities market, and the national economy.

1. Introduction

Due to the impact of COVID-19 epidemic, many companies are facing financial crisis. The number of companies that run into difficulties or even go bankrupt due to financial risks has greatly increased. The financial crisis of company operation is usually caused by the accumulation of various problems in the company's long-term business activities, and these problems can be discovered and solved in advance. In general, companies can observe potential risks by regularly analyzing and observing financial statement data. In order to help companies to avoid and disperse financial risks in a timely and effective manner, the financial crisis warning system is particularly important in company's risk management. Many scholars have established financial crisis warning models for different industries, but the accuracy of the models still needs to be improved. So, this study adopts the method of data mining, using RPROP neural network and support vector machine (SVM), two years in 162 listed companies panel financial index data as the modeling sample, test sample, and a financial crisis warning model, in order to manage the financial risk of listed companies and guarantee smooth running of listed companies, securities market, and the national economy.

At present, most of the domestic relevant literatures use mathematical models to build financial crisis early warning systems or use various methods for comparison. This study also uses a comparative method and uses two neural network models to establish financial crisis early warning systems. These two neural network models include RPROP neural network and support vector machine. This result can provide a reference for the domestic financial sector and provide a basis for financial management for the investment public.

The structure of the following part of the article is as follows: the second part is the domestic and foreign financial crisis warning related literature, providing the background information of this part. In the third part, this paper will discuss two kinds of neural network models, including RPROP neural network and support vector machine, use these two methods to conduct financial crisis prediction analysis on two years' data of 162 listed companies, and use

modeling and cross-validation to verify the accuracy of the models. In the last part, conclusions and suggestions will be made for the overall study.

2. Literature Review

Scholars at home and abroad have made many explorations on the construction of financial risk early warning system design model. Beaver [1] first used the traditional financial indicators to raise the rough judgment problem of a company's financial crisis and formally constructed a single-factor analysis model through statistical theory. However, the single-factor analysis model does not fully reflect the financial risk of a company. Therefore, the single-factor quantitative model has developed into a multifactor quantitative model. Edward Altman [2] proposed a multivariate linear multifactor model "Z-Score" to judge the severity of financial risks of companies by scoring results. However, because the Z-score model was not fully considered when establishing cash flow changes, it has a relatively large limitation. Therefore, Zhou [3] and other scholars transformed the Z-score model and established a new model for financial crisis prediction, the F-score model (Failure Score Model). With the establishment of the concept of information flow, Aziz et al. [4] proposed a model of forecasting financial distress with cash flow information in 1988. In 1993, Coats and Fant [5] built a financial early warning model based on the neural network theory. The model imitates the learning process of biological brain neural network without considering whether it conforms to the normal hypothesis, and can deal with nonquantifiable variables, which is widely applied. Also in 1999, univariate analysis and multivariate linear decision analysis were performed using financial statement data from 27 ST companies and 27 non-ST companies [6].

With the deepening of scientific research and the development of computer science and technology, other disciplines such as artificial intelligence and neural networks have also been applied to the theory of company's financial early warning systems. Continuously improving forecast accuracy has gradually become an important development direction in this field. Wilson and Sharda [7] used the method of artificial neural network to explore the prediction of company bankruptcy. Zhou and Wang [8] proposed a company financial crisis warning method based on system fuzzy optimization and neural network model and integrated the functions of company's financial crisis measurement, neural network dynamic learning and reasoning of financial crisis warning reasoning knowledge, and prediction of financial crisis indicators. From the perspective of cash flow, Liang [9] established the canonical rule judgment model and Fisher linear judgment model and made the prediction and the comparison of the results. Lin [10] took ST company and non-ST company with a ratio of 1 : 2 as the research object and constructed the Ologit financial crisis early warning model. Song [11] used relevant financial data of listed companies as a research sample and built a neural network model with deep learning as a research tool to predict whether a company would fall into financial crisis. The prediction accuracy could reach more than 72%. Wu [12] took the listed companies on China's GEM as the research object and used Twin-SVM to construct the financial crisis early warning model for the nonequilibrium sample characteristics of the company's different financial conditions. Wang [13] presented a financial crisis early warning detection algorithm based on fuzzy cognitive map. In addition, some scholars have innovated to construct the financial risk warning model by fuzzy analytic hierarchy process [14], PCA-SVM [15], improved MRMR algorithm [16], and other methods. Although the research on financial early warning models in China has been deepened in recent years, it is still in the exploration stage, and there is still room for improvement in terms of prediction accuracy. Therefore, this study applies the R programming language, using RPROP neural network and support vector machine (SVM), with 162 listed companies panel two years' financial index data as the modeling sample, the test sample, and a financial crisis warning model, in order for managers of listed companies and other stakeholders to predict the company's financial situation to provide scientific decision means and reliable basis.

3. Description of Research Methods and Indicators

3.1. Resilient Back Propagation Neural Network (RPROPN). Back Propagation Network (BPN) is the most widely used artificial neural network. It is a multilayer feedforward network trained by error back propagation (referred to as error back propagation). Its algorithm is called BPN algorithm, and its basic idea is to use the gradient descent method search technology, in order to minimize the mean square error of the predicted output value and expected output value of the network. In the backward transfer phase, the link weights between nodes will be corrected in reverse according to the prediction error, which will make the predicted output of BPN network close to the target value after repeated corrections. However, the algorithm still has some shortcomings. The first is that its learning efficiency η needs to be determined by us in advance; in addition, the weight change is based on the rate of change of the error gradient. Although at first glance this seems to be no problem, we dare not guarantee that it will always be correct and effective. To this end, Riedmiller et al. proposed the RPROP algorithm (resilient backpropagation) to improve the Backprop algorithm. In this study, we also adopted the improved BP algorithm, that is, the Resilient Back Propagation (RPROP), to construct a neural network prediction model.

3.2. Support Vector Machine (SVM). Support vector machine (SVM), proposed by Professor Vapnik (1995), is a supervised machine learning algorithm. The basic model is to find the optimal separation hyperplane in the feature space to maximize the positive and negative sample interval on the training set. The core idea is to separate the data by

constructing a segmentation plane. It is most commonly used for categorical prediction problems.

3.3. Sample Collection and Index Selection. The data in this study uses data from companies listed on the Shanghai Stock Exchange and Shenzhen Stock Exchange. In a two-to-one way, 162 listed companies were selected, including 55 ST companies and 107 non-ST companies, and each group of three companies was in the same or similar industry in order to have better comparability. In this study, the panel data is formed with the data of each explanatory variable from 2015 to 2016. These sample companies are mostly in the real estate sector and the comprehensive sector. It should be noted that, for each ST company, we randomly select two non-ST companies in the same sector, and the overall number of samples is 1:2. In terms of scope, it should be objective. The sample composition of this study is shown in Table 1.

At present, there is a lack of more specific economic theory guidance for the selection of financial indicators, and the essential reasons for the company's special treatment are not necessarily the same. It is difficult to fully describe with simple financial ratio indicators. Therefore, we mainly take quantitative analysis and provide six indicators: turnover rate, total asset turnover rate, main business revenue growth rate, total asset growth rate, equity return rate, and asset return rate, which are based on the five financial strengths. And in order to better test the accuracy of the model, the 2015 and 2016 data of the indicators are selected for comparison, and the defined variables are shown in Table 2.

And according to the difference between 2015 and 2016, these explanatory variables (X) data are based on the turnover rate (receivables turnover rate (X1), total asset turnover rate (X2)), growth rate (main business revenue growth rate (X3), total asset growth rate (X4)), and return rate (return on equity (X5), return on assets (X6)) and other indicators are collected. The narrative statistics are shown in Table 3.

Among them, the crisis company (ST) included in the explanatory variable (Y) refers to the company that has suffered losses for two consecutive years and is specially treated. It is represented by the value 1 and the normal company by the value 0. In addition, in data processing, we divided all 162 pieces of data into 6 groups, and each group had 27 pieces of data. In the decision tree modeling, 5 groups of 135 data pieces were used to establish a decision tree model, and a group of 27 data pieces was substituted into the decision tree model to test the accuracy of the model for cross-validation. Therefore, there are 6 sets of data in the same year. In other words, the first group uses group 1–5 modeling and group 6 testing. The second group adopts group 2–6 modeling and group 1 testing. The third group adopts group 3–1 modeling and group 2 testing. The fourth group adopts group 4–2 modeling and the group 3 testing. The fifth adopts group 5–3 modeling and group 4 testing. The sixth group adopts group 6–4 modeling and group 5 testing. The same was done for two years, but the second year is marked with 7–12 groups. This method can help us to further analyze the accuracy of two kinds of neural network financial warning models.

Table 1: Sample statistics table.

	ST	Non-ST	Total
Total	55	107	162

Table 2: Variable definition.

Symbol	Meaning
Y	Whether the company is special or not (ST = 1, non-ST = 0)
X_{1j}	Turnover rate of company j in 2015
X_{2j}	Total asset turnover rate of company j in 2015
X_{3j}	Main business revenue growth rate of company j in 2015
X_{4j}	Total asset growth rate of company j in 2015
X_{5j}	Equity return rate of company j in 2015
X_{6j}	Asset return rate of company j in 2015
X_{7j}	Turnover rate of company j in 2016
X_{8j}	Total asset turnover rate of company j in 2016
X_{9j}	Main business revenue growth rate of company j in 2016
X_{10j}	Total asset growth rate of company j in 2016
X_{11j}	Equity return rate of company j in 2016
X_{12j}	Asset return rate of company j in 2016

4. Empirical Analysis

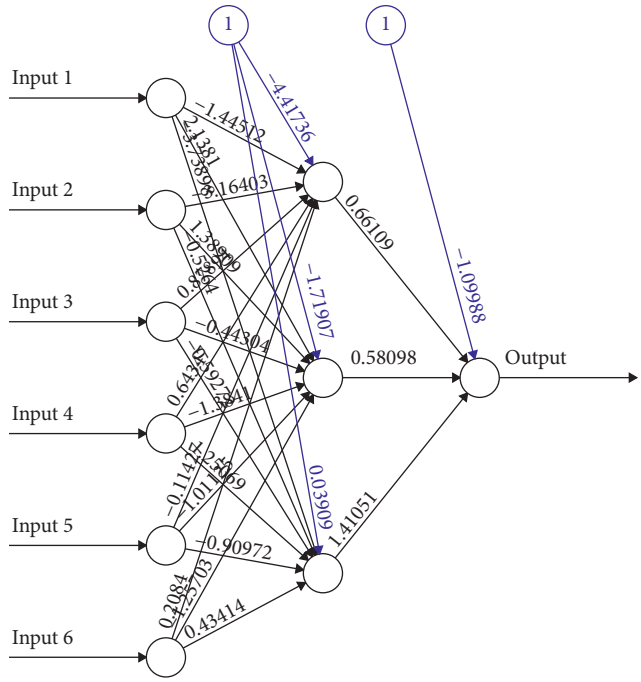
4.1. Resilient Back Propagation Neural Network (RPROPN). In this study, the Resilient Back Propagation Neural Network is mainly implemented with the help of *R* tool. For the purpose of comparative study, we confirmed the input nodes of the RPROP model by using the six same original financial indicators retained as the benchmark. The number of output nodes is 1. If the network output result is less than the threshold value of the output layer, we would consider it to be an ST company. Otherwise, it is a non-ST company, and the threshold of the output layer is determined by the network training itself. The number of intermediate layers is set as 3 neurons, and the number of repetitions is 5. A decision tree model is established based on a total of 12 sets of 135 data pieces in two years. The results are shown in Figures 1–12.

Through the established model, we, respectively, tested the 12 groups through 27 pieces of data in each group. Among them, we set the result as $Y = 0$ below 0.5 and $Y = 1$ above 0.5, which is called correct judgment of the model. The test results are as follows.

4.2. Support Vector Machine (SVM). The support vector machine uses the Gaussian kernel function (RBF). The output of 0 represents financial normality and the output of 1 represents financial crisis. This experiment is based on the "linear space inseparable problem" and "nonlinear problem" to establish the SVM model. Under this model, two parameters need to be selected: "RBF kernel function density" and "penalty factor C." We know that gamma is a parameter that comes with the RBF function as the kernel. It implicitly determines the distribution of the data after mapping into a new feature space. The larger the gamma, the fewer the support vectors, and the smaller the gamma value, the more the support vectors. The number of support vectors affects

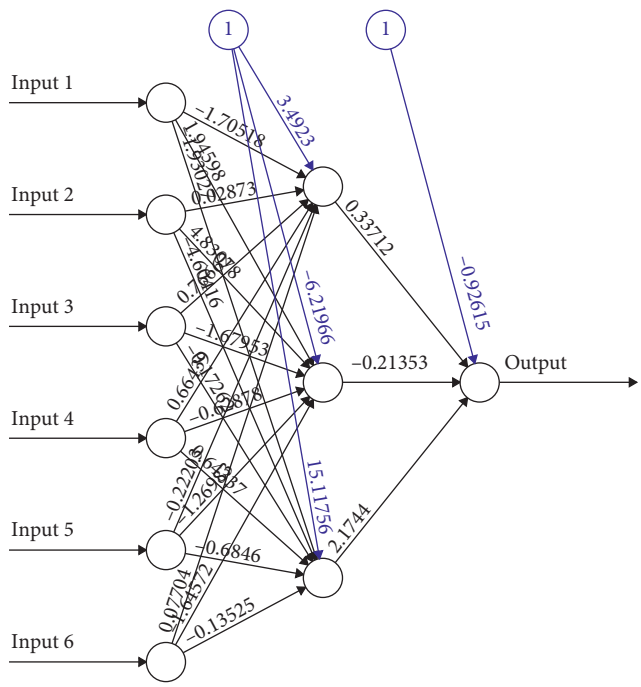
Table 3: Descriptive statistics of each index.

Descriptive statistics	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
Maximum	7172.23	10.08	4335.66	924.57	1106.21	161.65	7172.23	10.08	4335.66	924.57	1106.21	161.65
Minimum	0.24	0.02	−90.05	−82.97	−65.62	−29.98	0.24	0.02	−90.05	−82.97	−65.62	−29.98
Average	91.11	0.65	46.31	23.44	135.51	16.14	91.11	0.65	46.31	23.44	135.51	16.14
Standard deviation	581.29	0.96	351.33	85.30	170.64	19.53	581.29	0.96	351.33	85.30	170.64	19.53



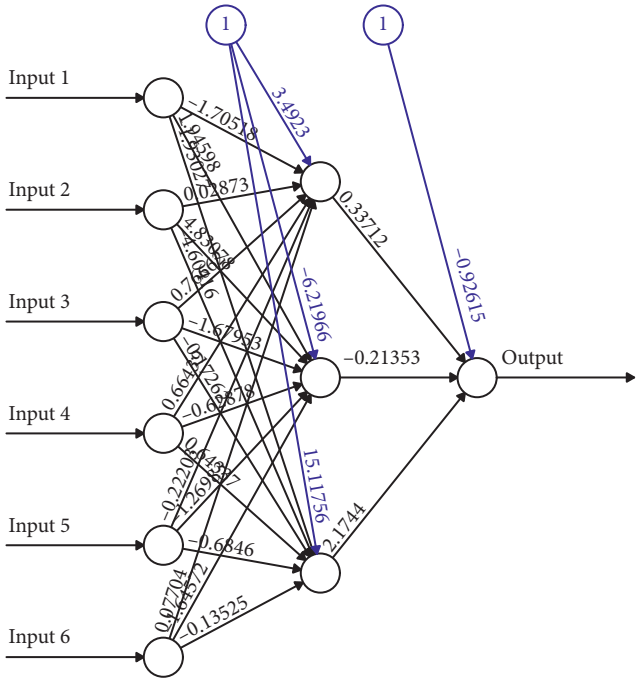
Error: 14.152619 Steps: 92

Figure 1: Group 1 modeling in 2015.



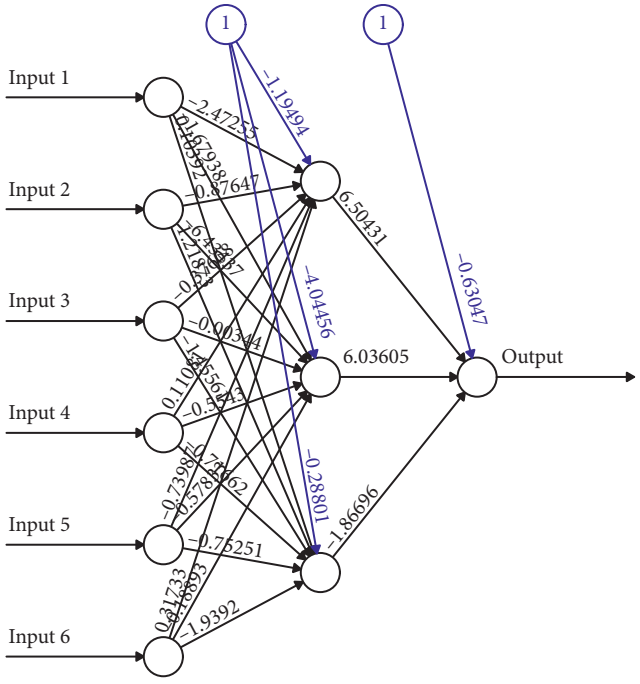
Error: 13.528471 Steps: 220

Figure 2: Group 2 modeling in 2015.



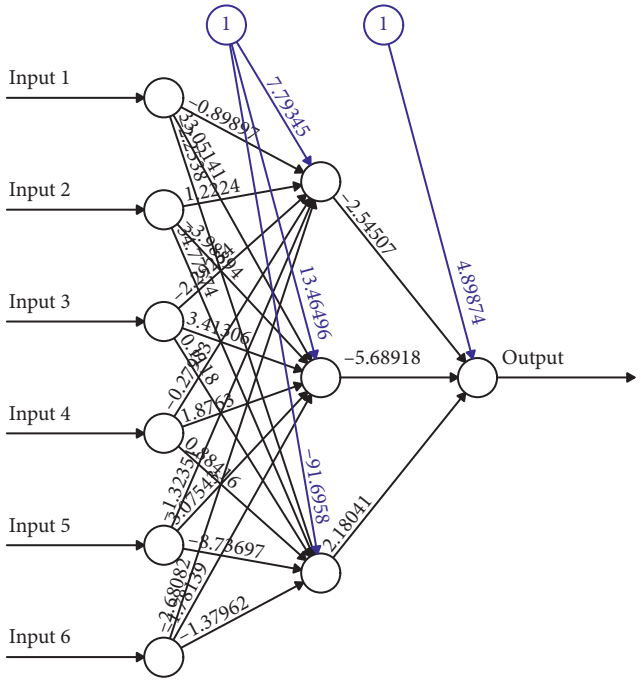
Error: 13.528471 Steps: 220

Figure 3: Group 3 modeling in 2015.



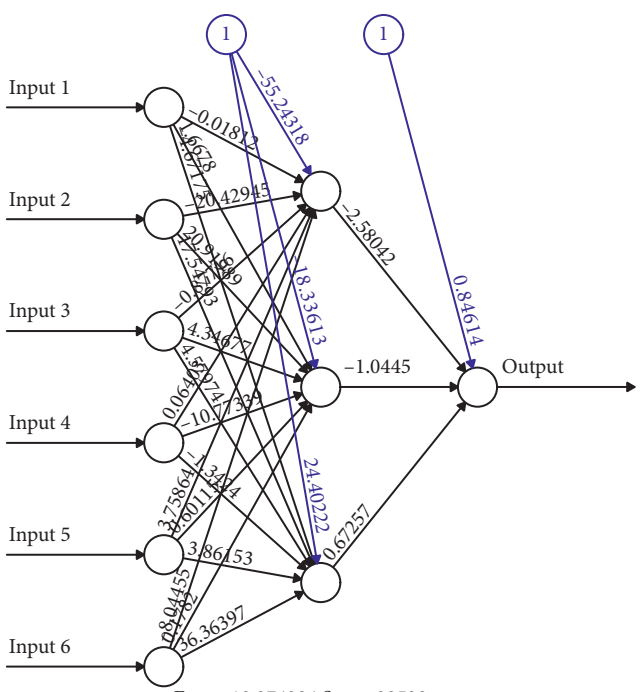
Error: 13.84064 Steps: 63

Figure 4: Group 4 modeling in 2015.



Error: 12.452405 Steps: 1827

Figure 5: Group 5 modeling in 2015.



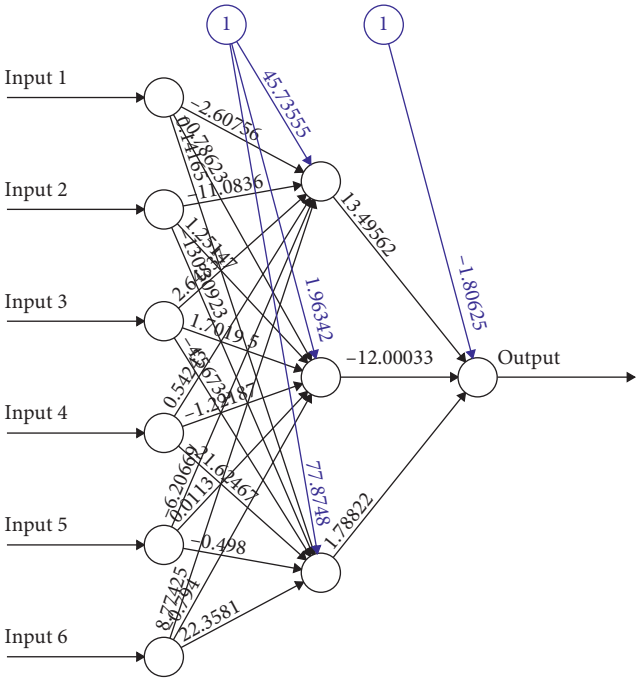
Error: 10.974294 Steps: 22528

Figure 6: Group 6 modeling in 2015.

the speed of training and prediction. Through calculation research and case matching, here we set gamma = 0.01, cost = 1.

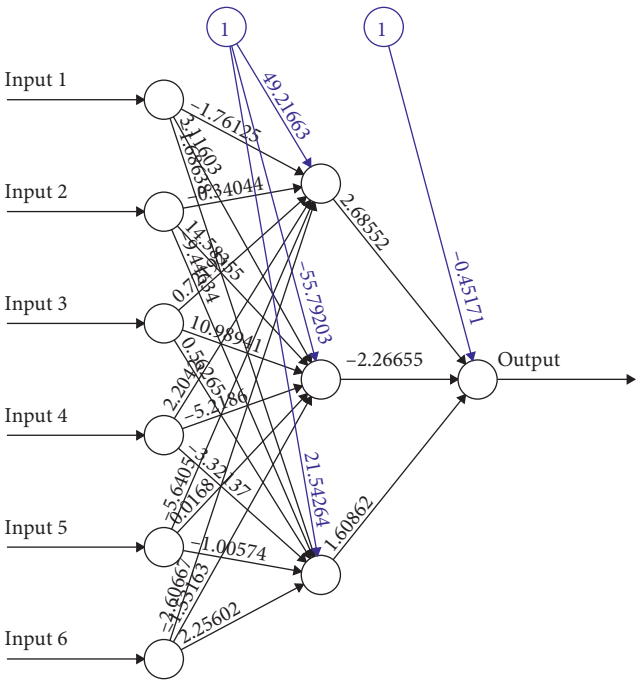
Finally, use the selected parameters and regularized data to train the SVM model, and the results are as follows.

By comparing and analyzing the judgment results of Tables 4 and 5, the judgment results of applying RPROP artificial neural network model and support vector machine to predict the financial crisis of China's listed companies show the following.



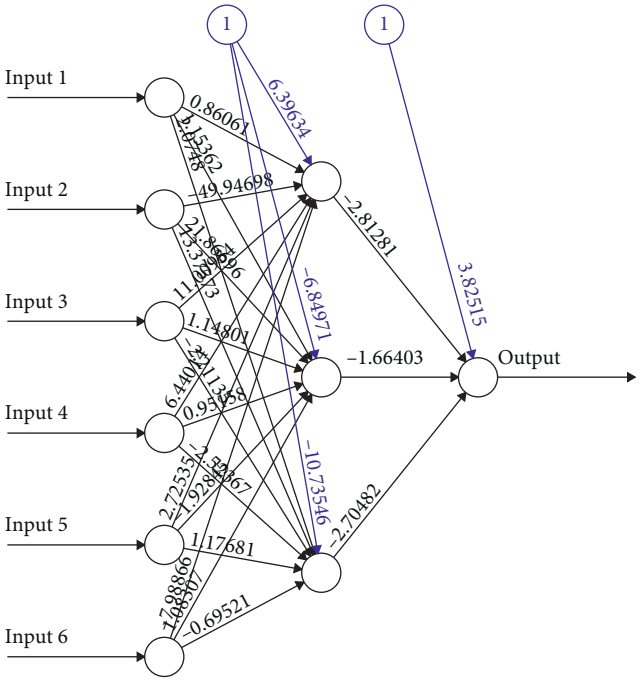
Error: 10.979337 Steps: 10674

Figure 7: Group 1 modeling in 2016.



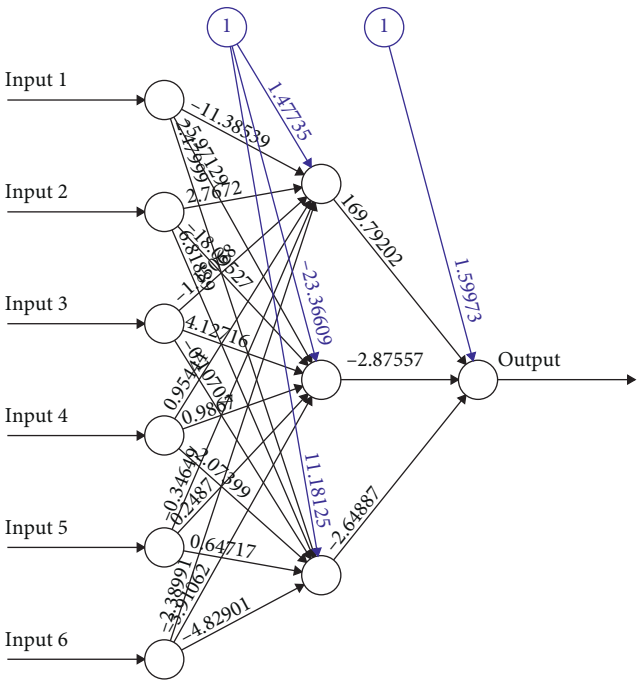
Error: 9.899326 Steps: 2033

Figure 8: Group 2 modeling in 2016.



Error: 10.571473 Steps: 2979

Figure 9: Group 3 modeling in 2016.



Error: 9.850746 Steps: 4739

Figure 10: Group 4 modeling in 2016.

First, the judgment accuracy of both models is acceptable, but in contrast, the judgment accuracy of the RPROP neural network model for ST companies is overall higher, indicating that the RPROP neural network model can fully learn the data information contained in the samples, which is conducive to improving the ability to judge the unidentified samples.

Second, the discriminant accuracy rate of the two models for ST companies is lower than that of non-ST companies, which is particularly obvious in the support vector machine

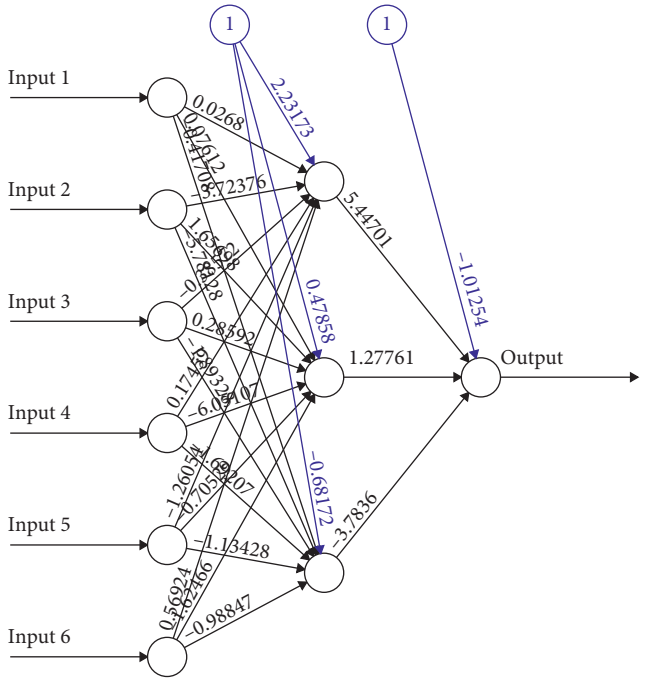


Figure 11: Group 5 modeling in 2016.

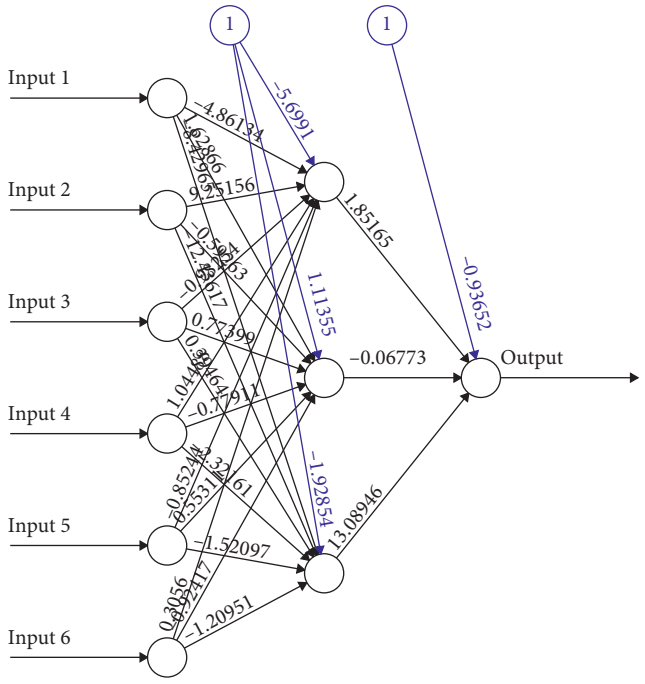


Figure 12: Group 6 modeling in 2016.

Table 4: RPROP's judgment on modeling samples.

	2015		2016	
	Actual number	The average of the correctly judged number	Actual number	The average of the correctly judged number
ST	9	0.17	9	3.83
Non-ST	18	16.83	18	15.33
Total number of companies	28	19.17	28	20.17
Accurate judgment rate		0.69		0.72

Table 5: SVM's judgment on test data.

	2015		2016	
	Actual number	The average of the correctly judged number	Actual number	The average of the correctly judged number
ST	9	0.67	9	1.0
Non-ST	18	18	18	18
Total number of companies	28	18.67	28	19
Accurate judgment rate		0.67		0.68

model. Through the analysis of a single sample of misjudgment, it can be found that the wrong samples are in a critical state of the ST and the ST listed company. But compared with the BP artificial neural network model, support vector machine (SVM) model for ST's discriminant accuracy is higher than the BP artificial neural network model. This is a conclusion worthy of in-depth discussion.

5. Conclusions

This study compares RPROP neural network and support vector machine algorithm to analyze the financial crisis of listed companies. The research results show that the RPROP artificial neural network method based on financial indicator information combined with support vector machine is a more effective method to predict whether financial crisis will occur in a company's finance. First, RPROP artificial neural network is used for the first round of troubleshooting, and then support vector machine is used for verification, which will make the whole result more reliable.

The research results show that the occurrence of financial crisis can be predicted in advance, and the experimental results show that the occurrence of financial crisis is the result of multiple factors. Using the R programming language, with the help of computer tools, the algorithm design and data calculation of the RPROP network model can be easily completed, and an early warning model of the financial crisis of listed companies can be established to provide a scientific basis for investors, creditors, and other stakeholders to predict the financial situation of the company, as well as to provide a reference basis for listed companies to strengthen their internal management and get rid of the financial crisis in a timely manner. It is suggested that listed companies can combine nonfinancial quantitative indicators to deeply explore financial statement information of listed companies and give early warning to companies with financial risks. The predicted results can enable the managers of listed companies to timely adjust their strategies, change their business model, and reduce or avoid losses. Investors have insight into the development prospects of listed companies and make rational investment. Creditors can avoid loan risks. In addition, it also provides a basis for the securities regulatory authorities to strengthen supervision over the listed companies with risks, which is of great economic and social significance to ensure the smooth operation of the listed companies, the securities market, and the national economy. Compared with other methods, the reason why the prediction accuracy is improved is that the RPROP neural network model is a nonlinear separable mapping, and the data requirements for model variables are not particularly demanding. The lack of individual data has little effect on the prediction results of the model. The RPROP neural network learns more fully and contains more information on the original data. Therefore, we believe that the RPROP neural network model has broad application prospects and application value in analyzing and studying the financial situation of China's listed companies.

However, it should be noted that, first, in the research process, we have placed too much emphasis on the elements of the financial five-force model, and nonfinancial indicators and some nonquantitative indicators should also remain crucial in listed companies facing an increasingly fierce market. These are not noticed or ignored in our research and can be referenced in future research. Second, this study is conducted in two perspectives: cross-section and longitudinal tracking. This study only uses the cross-section financial data of the first two years of the listed company to establish a prediction model and then uses the listed companies in the overall sample of the same period to make a judgment test, indicating that the prediction model is effective for the prediction of the first two years of financial crisis. If a longer time chain is considered, whether the same principle or even the same model can be used for prediction requires a continuous follow-up investigation and research on the longitudinal data samples of the same company.

Data Availability

The original report data of all enterprise index data were obtained from the public data of listed companies on the official websites of Shanghai Stock Exchange and Shenzhen Stock Exchange. The data include indicators such as accounts receivable turnover rate, total asset turnover rate, main business income growth rate, total asset growth rate and return on equity, return on assets, and other indicators. All enterprise index data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was funded by a Key Research Project of Universities in Guangdong Province (2018WTSCX034).

References

[1] W. H. Beaver, "Financial ratios as predictors of failure," *Journal of Accounting Research*, vol. 4, pp. 71–111, 1966.

[2] E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *The Journal of Finance*, vol. 23, no. 4, pp. 589–609, 1968.

[3] S. Zhou, J. Yang, and P. Wang, "On the early warning analysis of financial crisis——F score model," *Accounting Research*, vol. 8, p. 4, 1996.

[4] A. Aziz, D. C. Emanuel, and G. H. Lawson, "Bankruptcy prediction—an investigation of cash flow based models," *Journal of Management Studies*, vol. 25, no. 5, pp. 419–437, 1988.

[5] P. K. Coats and L. F. Fant, "Recognizing financial distress patterns using a neural network tool," *Financial Management*, vol. 22, no. 3, pp. 142–155, 1993.

[6] J. Chen, "Empirical analysis of financial deterioration forecast of listed companies," *Accounting Research*, vol. 4, pp. 31–36, 1999.

[7] M. D. Odom and R. Sharda, "A neural network model for bankruptcy prediction," in *Proceedings of the 1990 IJCNN International Joint Conference on Neural Networks*, pp. 163–168, IEEE, San Diego, CA, USA, June 1990.

[8] M. Zhou and X. Wang, "Enterprise financial crisis early warning based on fuzzy optimization and neural network," *Journal of Management Science*, vol. 3, pp. 86–90, 2002.

[9] F. Liang, "Empirical research on early warning system of financial crisis based on cash flow—a case study of machinery and equipment industry," *Finance and Trade Economics*, vol. 2, pp. 23–27, 2005.

[10] Q. Lin, "Research on early warning of financial crisis of listed companies based on Ologit model," *Management*, vol. 12, pp. 95–97, 2018.

[11] Ge Song, "Research on early warning model of financial risk for listed companies based on deep learning," *Value Engineering*, vol. 1, pp. 53–56, 2019.

[12] Q. Wu, "Identification and early warning of financial crisis of listed companies on GEM," *Finance and Accounting Monthly*, vol. 2, no. 0056, pp. 56–64, 2020.

[13] Q. Wang, F. Hui, X. Wang et al., "Research on early warning and monitoring algorithm of financial crisis based on fuzzy cognitive map," *Cluster Computing*, vol. 22, no. 2, pp. 1–9, 2019.

[14] S. Gao, "Application research on financial risk early warning of small and medium-sized enterprises under fuzzy analytic hierarchy process," *China Business*, vol. 11, pp. 46-47, 2020.

[15] X. Shi, "Construction of early-warning model of corporate financial crisis based on PCA-SVM," *Journal of Finance and Accounting*, vol. 10, pp. 131–134, 2020.

[16] K. Luo and G. Wang, "Research on financial early warning based on improved MRMR algorithm and cost sensitive classification," *Statistics and Information Forum*, vol. 35, no. 03, pp. 77–85, 2020.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 9746153, 7 pages
https://doi.org/10.1155/2020/9746153

Research Article

Investment Payback Period Calculating Model for Airport Bridge Facility

Jian Wan,[1] Zheng-hong Xia ,[2] and Xin-ping Zhu[2]

[1]*China Academy of Civil Aviation Science and Technology, Beijing 100028, China*
[2]*School of Air Traffic Control, Civil Aviation Flight University of China, Guanghan 618307, China*

Correspondence should be addressed to Zheng-hong Xia; 66949827@qq.com

Received 4 March 2020; Revised 4 June 2020; Accepted 27 July 2020; Published 12 August 2020

Guest Editor: Weilin Xiao

Due to the uncertainty and difficulty of estimating the investment payback period of the airport bridge facility, a model for calculating the investment payback period of bridge facility is proposed in this paper from the perspective of airport routine operation. Based on the actual operational data of Kunming Changshui Airport, Wuhan Tianhe Airport, and Lijiang Sanyi Airport in 2018, the factors influencing the payback period of bridge facility are the number of bridge facilities and service time, which have been discussed in this paper. According to the simulation, it is concluded that the number of bridge facilities and service time are the key points to the length of the investment payback period, and the annual operating cost of the airline can be saved quite a lot. The research results can be used to assist the leaders' decision-making of airports and airlines for the promotion of the service of bridge facilities.

1. Introduction

Auxiliary power unit (APU) can supply the aircraft with power and air source. Since its fuel consumption rate is lower than that of main engines, APU can often be used as an alternative to the main engines for providing lighting, air condition, and air source when the aircraft is in the process of maintenance and transition. However, APU will generate emissions such as carbon dioxide and nitrogen oxides as the main engines do, as well as airport noise of about 95 dB. With the "energy-saving emission reduction" work promoted by CAAC recently, Guangzhou Baiyun International Airport, Chengdu Shuangliu International Airport, and Xi'an Xianyang Airport have installed the bridge facility on the gallery bridge (including the 400 Hz/115 V power supply and air-conditioning equipment) to provide the lighting, air condition, and air source for the aircraft instead of using APU, which can reduce the exhausted gas emission and noise pollution to a certain extent. However, due to the cost of installation and maintenance of the bridge facility, the lack of strong promotion policy and standard operation process and other factors such as airline company and pilot's incompatibility, there also appears uncertainty of the investment recovery period and the annual operating profit of these facilities. Therefore, the popularization of bridge facilities in domestic trunk and hub airports will meet with many difficulties, and the research on the operation cost, income, and payback period of investment of the bridge facilities will be imperative. It is of great significance to promote the use of bridge facilities in civil aviation airports and airlines, to promote the implementation of energy-saving and emission reduction policies of the civil aviation bureau, and to achieve the goal of "green airport, green civil aviation."

Xu and Li established the cost analysis of flight delays and simulation in the ground-holding model in 2006 [1]. Li and Zhang studied the necessity and feasibility of reducing the usage of APU and proposed the saving operation mode of the APU in 2013 [2]. Xia et al. constructed the ground-holding cost analysis models for delayed flights, discussed the optimization of ground waiting cost with the bridge facility supplying electrical and pneumatic power, and also pointed out the advantages of the bridge facility over APU in ground waiting during flight delays [3, 4]. Pan et al.

improved the flight delay cost model by considering different power supply modes, put forward the concept of waiting tolerance time limit in the cabin, and the trend of total ground delay cost of the flight with time was also analyzed when using the bridge facility and APU power supply, respectively [5]. Olsen et al. presented an invention related to an aircraft auxiliary power unit and a method of isolating an auxiliary power unit in an aircraft in 2013 [6]. Ebrahimi et al. proposed a comprehensive model and characterization of the advanced aircraft electric power systems with a hybrid fuel cell battery auxiliary power unit in 2015 [7]. Chen et al. designed an automation management system which can achieve the fine management for energy, safety, and using the effect of bridge-borne equipment and then can reflect the implementation situation and enforcement of replacing APU with bridge-borne equipment at the airport in 2014 [8]. Li et al. proposed a mathematic model of the system by using the method based on the combination of mechanism modeling and parameter identification in 2017 [9].

According to the foregoing, these studies have pointed the advantages of bridge facilities and the disadvantages of using APU. However, there have been no related research studies such as the operating cost of the bridge facilities and the payback period from the view of the airport. Therefore, the model of the payback period of investment into bridge facilities was proposed in this paper, taking into account of these factors such as the number of bridge facilities, docking rate, and the period of using the bridge equipment. This paper analyzed the operational cost of the bridge facility and the annual operation cost savings of airlines in detail. And the number of bridge facilities of the hub, trunk, and feeder airports was also predicted based on the shortest payback period and the maximum annual profits. The research results can be used to assist the leaders' decision-making of the airports and airlines for the promotion of the service of the bridge facilities.

2. Payback Period of Investment Models for Bridge Facilities

2.1. Definitions

$\text{Cost}_{\text{operation}}$: from the view of the airport, $\text{Cost}_{\text{operation}}$ is the annual average operational cost for bridge facilities.

Income is the annual income that the airport can obtain from these bridge facilities.

Profit is the average annual profits for operating the bridge facility.

$\text{Cost}_{\text{device}}$ is the cost for installing these bridge facilities including 400 Hz/115 V power supply, air-conditioning equipment, and related cables.

$\text{Cost}_{\text{elceccharge}}$ is the annual cost of electricity for the operation of the bridge facilities.

$\text{Cost}_{\text{maintenance}}$ is the cost for maintenance of bridge facilities including external maintenance and the costs for spare parts.

$\text{Cost}_{\text{staff}}$ is the annual cost for the workers who guarantee the normal operation of the airport bridge facilities.

Cost_{tax} is the taxes that the airport must pay for operating the bridge facility.

$\text{Cost}_{\text{depreciation}}$ is the depreciation cost of these bridge facilities.

AirlineCostOpt is the total annual operating cost savings for airlines after using the bridge-mounted equipment.

$\text{CostOptimize}_{\text{airline}}$ is the operational cost savings for each aircraft using the bridge-mounted equipment instead of using APU to provide electrical and pneumatic power.

$\text{Cost}_{\text{APU}}^{\text{FuelConsumption}}$ is the cost of fuel consumption per aircraft when using APU to supply electrical and pneumatic power.

$\text{Cost}_{\text{APU}}^{\text{maintain}}$ is the APU maintenance cost per aircraft.

$\text{Cost}_{\text{power}}$ is the payment from the airport for electrical and pneumatic power supply from the bridge facility for each aircraft.

power_1 is the rated output power of the power supply, and its unit is Kilowatt.

power_2 is the rated output power of the air-conditioning equipment, and its unit is Kilowatt.

η_1 is the operating efficiency of the power supply.

η_2 is the operating efficiency of the air-conditioning equipment.

p is the requested number of operators for running these bridge facilities.

year is the payback period of investment of the bridge facilities.

n is the number of boarding bridges or bridge facilities. We assume the boarding bridge was assembled as the same as the bridge facility.

k is the annual departing and landing number of aircrafts.

α is the annual number of aircrafts that will be docked to the bridge.

β is the annual docking rate of the aircraft.

λ is the proportion of these docked aircrafts which only use the power supply.

ξ is the proportion of these docked aircrafts which use the power supply and air-conditioning equipment simultaneously.

t is the average service time of the bridge facility.

q is the annual profit growth rate of the operational bridge facility.

$c_{\max}$ is the maximum number of departure and landing aircrafts per hour.

fuel is the aircrafts' APU fuel consumption rate per hour in kg/h.

m_1 is the hourly electricity price of the airport.

m_2 is the annual cost for each worker who operates the bridge facility.

m_3 is the annual cost for external maintenance of each bridge facility.

m_4 is the annual spare part costs for the maintenance of every bridge.

m_5 is the fuel price per ton.

2.2. Models

$$\text{Income} = k * \beta * \lambda * \left[\frac{t}{15}\right] * \frac{(450 * \xi + 170 * (1 - \xi))}{4}. \tag{1}$$

According to the charge standard for the usage of bridge facility, the service of 400 Hz/115 V power supply per hour will be charged at 170 RMB, and 280 RMB will be charged per hour for the service of air-conditioning equipment [10]. Every 15 minutes will be considered as one charge unit. If the service time is shorter than 15 minutes, it will be considered as one charge unit as 15 minutes as well. Therefore, the annual income for bridge facility utilization that the airport can obtain is from formula (1) as follows: k is the annual departing and landing aircraft, β is the annual docking rate of the aircraft, λ is the proportion of these docked aircrafts that only use the power supply, ξ is the proportion of these docked aircrafts that use both power supply and air-conditioning equipment.

$$\begin{aligned}\text{Cost}_{\text{operation}} = {} & \text{Cost}_{\text{eleccharge}} + \text{Cost}_{\text{maintenance}} \\ & + \text{Cost}_{\text{depreciation}} + \text{Cost}_{\text{staff}} + \text{Cost}_{\text{tax}}.\end{aligned} \tag{2}$$

The annual cost for the operation of bridge facility $\text{Cost}_{\text{operation}}$ is equal to the summation of electricity cost, maintenance cost, depreciation cost, operators' wages, and taxes, which can be stated as formula (2).

$$\text{Cost}_{\text{eleccharge}} = \left(\text{power}_1 * \eta_1 + \text{power}_2 * \eta_2\right) * m_1 * n. \tag{3}$$

The annual electricity cost mainly consists of the rated power of power supply and air-conditioning equipment power_1 and power_2, the electricity charge m_1, and the number of bridge facilities n.

$$\text{Cost}_{\text{staff}} = p * m_2. \tag{4}$$

The annual cost for each bridge facility operator m_2 must include wages, social security, and provident funds.

$$\text{Cost}_{\text{tax}} = \text{Income} * 3\%. \tag{5}$$

The annual airport tax costs mainly consist of business tax. According to the State Price Bureau, the standard for air transport business tax is about 3% of the average annual income of the airport for the operation of bridge facilities.

$$\text{Cost}_{\text{depreciation}} = \text{Cost}_{\text{device}} * \beta = \text{Cost}_{\text{device}} * \frac{(1 - \delta)}{k} * 100\%. \tag{6}$$

The annual depreciation costs of these bridge facilities can be calculated by formula (6), where β is the average depreciation rate of the fixed assets that is equal to the depreciation amount each year and is expected to be equal to the depreciation fixed assets, k is the rated maximum service life of the bridge facility, and δ is the estimated net residual value rate.

$$\text{Cost}_{\text{maintenance}} = n * \left(m_3 + m_4\right). \tag{7}$$

The annual cost for the maintenance of the bridge facility can be obtained from formula (7), where m_3 and m_4 are the annual external maintenance and spare part cost separately.

$$\text{Profit} = \text{Income} - \text{Cost}_{\text{operation}}, \tag{8}$$

$$\text{year} = \log_{\bar{q}}^{\left[1+(q-1)*\left(\text{Cost}_{\text{device}}/\left(\text{Income}-\text{Cost}_{\text{operation}}\right)\right)\right]}, \tag{9}$$

$$\begin{aligned}\text{AirlineCostOpt} = {} & \text{CostOptimize}_{\text{airline}} \\ & * 365 * k * \beta.\end{aligned} \tag{10}$$

The total annual operating cost savings of an airline using the bridge facility are equal to the cost savings of a single aircraft using the bridge facility multiplied by the airline's annual takeoff and landing sorties k at the airport and the average annual bridge facility utilization rate of the airline β.

$$\begin{aligned}\text{CostOptimize}_{\text{airline}} = {} & \text{Cost}_{\text{APU}}^{\text{FuelConsumption}} + \text{Cost}_{\text{APU}}^{\text{maintain}} \\ & - \text{Cost}_{\text{power}}, \\ \text{Cost}_{\text{power}} = {} & \left[\frac{t}{15}\right] * \frac{(450 * \xi + 170 * (1 - \xi))}{4}, \\ \text{Cost}_{\text{APU}}^{\text{FuelConsumption}} = {} & \frac{\text{fuel} * m_5 * t}{60 * 1000}, \\ \text{Cost}_{\text{APU}}^{\text{maintain}} = {} & \frac{m_3 * t}{60 * 365 * 24}, \\ \omega = {} & \frac{\text{CostOptimize}_{\text{airline}}}{\text{Cost}_{\text{APU}}^{\text{FuelConsumption}}}.\end{aligned} \tag{11}$$

Since the use of bridge facility can effectively reduce the operation cost for airlines and prolong the lifespan of the aircraft APU, the airline companies fully cooperate with the promotion of the use of the bridge, which will greatly improve the aircraft docking rate and increase the airport operational income and cost. Finally, the payback period of investment at these bridge facilities can be calculated by formula (9) based on the equation of annual profit and installation cost of these bridge facilities:

$$p = \begin{cases} 12, & n \leq 15, \\ 12 + \left[\dfrac{[n - 15]}{5}\right] * 3, & n > 15. \end{cases} \tag{12}$$

To maintain a higher service level, we assumed at least four teams of which each team must have three workers who are responsible to keep the normal operation of these bridge facilities. If the number of these facilities is less than 15, 12 operators are required, and one team will be added if the number of bridges is increased by 5. Then, the number of operators for these bridge facilities can be calculated by formula (10).

$$\alpha = k * \beta. \tag{13}$$

The annual docking aircraft can be obtained from annual departing and landing aircraft and docking rate.

$$\beta = \begin{cases} 1 - 0.1\left[\dfrac{(n-10)}{10}\right], & n > 10, \\ 1, & n \leq 10. \end{cases} \tag{14}$$

The average annual bridge facility utilization rate of an aircraft at the airport is assumed to be 1 when the number of the company's aircraft takeoffs and landings is less than 10 sorties. When the number of takeoff and landing sorties increases, it may lead to inadequate gallery bridges. At this time, it can be assumed that the bridge facility utilization rate of the aircraft decreases stepwise with the increase of takeoff and landing sorties.

3. Example Analysis

Based on the investment payback period calculating model for bridge facilities in this paper, we will discuss the airport income and profit as well as the payback period of these bridge facilities with the examples of Kunming, Wuhan, and Lijiang airports. Meanwhile, the paper will also take into account of the factors such as the number of bridge facilities, the annual number of aircrafts which will be docked on the bridge, the proportion of these docked aircrafts which use power supply and air-condition equipment separately, and the bridge facility service time of these docked aircrafts. Firstly, the installation cost of each bridge facility is about 1.5 million (RMB is assumed in this paper), which is used for the most popular brands such as Weihai Guangtai power supply and Hua Shengqiang air-conditioning equipment with rated power 72 kilowatt and 87 kilowatt and operating efficiency 80% and 90% separately. The annual cost for external maintenance for each bridge facility and maintenance cost for each bridge facility are about 17,500 RMB and 10,000 RMB, the rated service life span of these facilities is 10 years, the estimated net residual value rate is 3%, and the electricity charge fee is 0.7 RMB per degree. The business tax is about 3% of the airport annual income for the bridge facility operation, and the annual cost for the bridge facility operator is about 85,000 RMB. The annual departing and landing number of aircrafts of these three airports is 360,785, 187,699, and 56,932 in 2018, and the maximum number of departure and landing aircrafts per hour is 58, 39, and 16 separately.

3.1. Influence of the Number of Bridges on the Investment Payback Period. The docking rate of aircrafts at Kunming, Wuhan, and Lijiang airports is assumed to be 70%, considering these factors such as how many transition aircrafts will use the bridge facility, its service period is about 45 minutes, 40% of these aircrafts will use power supply and air-conditioning equipment simultaneously, and the number of bridges is the same as the amount of the bridge facility. Based on the proposed models in the first paragraph, the influence of the number of bridges on the payback period of bridge facility investment can be obtained from Figure 1.

It can be seen from Figure 1 that the payback period of bridge facility investment will decrease first and then increase with the increase of the number of bridges. If $n \in (5, 16)$, the docking rate of aircrafts and the aircrafts which use the bridge facility will increase with the increase of the number of bridges, and the annual airport profits will increase linearly, which will shorten the payback period of bridge facility investment; and the trend of Lijiang Sanyi Airport dropped most steeply. If $n = 16$, the maximum number of departure and landing aircrafts per hour is equal to the bridge facility number, then the bridge facility utilization rates reach the maximum value, and the payback period of bridge facility investment in Lijiang Sanyi Airport reaches the lowest value, 5.13 years. If $n \in (17, 22)$, the bridge facilities will be left unused, leading to the increase of operational cost and installation cost without any income, which will result in the increase of the payback period of bridge investment drastically. And if $n = 22$, the bridge facility running at Lijiang Sanyi Airport will be deficit. If $n \in (5, 39)$, the payback period of bridge investment will decrease slowly. If $n = 39$, the payback period of the bridge facility reaches the shortest value, 2.47 years. If $n > 39$, the airport bridge facility income will be the same, but the increase of installation and running cost will lead to the increase of the payback period dramatically. If $n = 58$, the payback period of Kunming Changshui Airport will be the shortest value, 2.08 years. If $n > 58$, the payback period will increase and result in a deficit finally. Therefore, the key point to decide the length of the payback period for bridge facility investment is the rational allocation of the number of bridges, which is closely related to the maximum number of aircrafts at the peak hour. Therefore, it is suggested that the numbers of bridge facilities in these three airports are 16, 39, and 58, and the payback period of bridge facility investment will be the shortest value.

3.2. Influence of Bridge Facility Service Time on the Payback Period of the Investment. Based on the interim regulations of bridge facility operation, every 15 minutes will be considered as one charge unit, and the transition time is about 30 to 60 minutes. The influence of service time of the bridge facility on the payback period can be seen from Figure 2, where the proportion $\lambda = 0.7$ and the percentage of docked aircrafts that use air-conditioning equipment is 40% on the basis of the departure and landing aircraft and the docking rate of Kunming, Wuhan, and Lijiang airports in 2018.

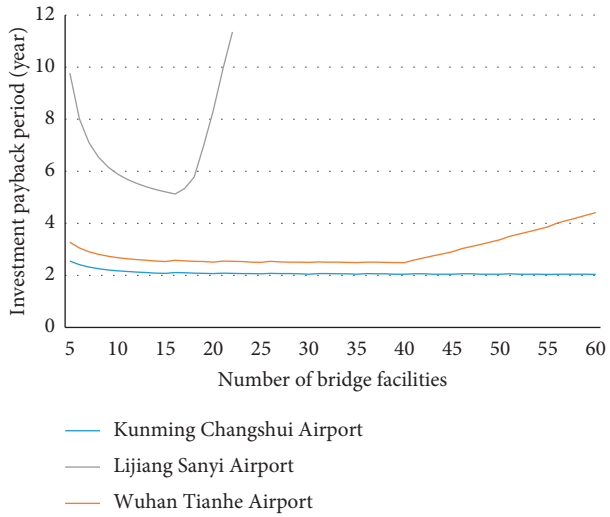


Figure 1: Influence of the number of bridges on the investment payback period of the bridge facility.

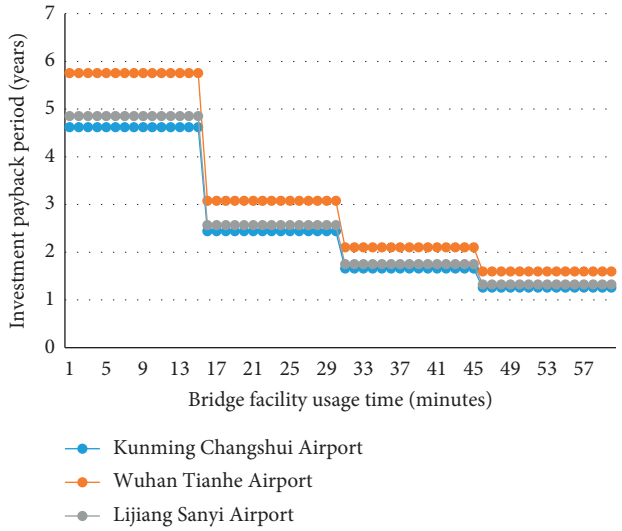


Figure 2: Influence of bridge facility service time on the investment payback period.

Apparently, as the bridge facility service time of the transition aircraft increased, the payback period of the bridge facility will decrease sharply with obvious latter shape characteristics, the magnitude of the change is becoming smaller and smaller with the increase of service time t, and the change in the feeder airport is most clear. If $t \in (1, 15]$, the operation cost of the bridge facility is more than its income in Wuhan Tianhe Airport, while the bridge facility operation in Kunming and Lijiang airports is profitable, and the payback period of bridge facility investment of these airports is 5.75, 4.62, and 4.86 years, respectively. If $t \in (15, 30]$, the bridge operation at the three airports will be profitable, and the payback period is decreased sharply to 3.08, 2.57, and 2.44 years, respectively. If $t \in (30, 45]$, the payback period is about 2.11, 1.66, and 1.75 years. And if $t \in (45, 60]$, the payback period will be decreased to 1.60, 1.25, and 1.34 years, respectively.

3.3. Influence of the Amount of Bridge Facility and Service Time on the Investment Payback Period. In order to discuss which is more important for the recycling cycle, this paper has taken a large-sized airport as an example based on this model. It is assumed that 700 aircrafts will land or takeoff every day from this airport, and 40 sorties will takeoff and land at peak hours. The average rate of relying on bridges is 70%, and the installation of bridge facilities will cost about 1.5 million yuan per set. Each set of spare parts is maintained by the commission on preparation and foreign affairs at about 17.5 million yuan per year, and the cost of fitting parts is about 10,000 yuan per year. The rated output powers of power supply and air conditioning are 72 kW and 87 kW, respectively. The operating efficiency is 80% and 90%, the rated service life of the equipment is 15 years, the estimated net residual value is 3%, and the business tax paid by the airport is 3%. Assume that 70% of the aircrafts near the bridge at the airport will use the bridge facility equipment about 95% of which are medium-sized aircrafts (e.g., B737 and A320 only need to access one set of bridge facility equipment) and 5% are heavy aircrafts (e.g., B747 and A340 need to access two sets of bridge facility equipment). The frequency for the usage of the bridge facility equipment is about 3-4 times a day and half an hour to 2 hours in average each time. The charge for airport electricity is 0.6 yuan per hour, and the salary and welfare benefits of staff who maintain the bridge facilities are about 75,000 yuan per year. At the same time, the annual average ratio of bridge facility air conditioning and power supply is 0.4, and the investment payback period of the bridge facility equipment varies with the number of the bridge facility equipment sets and service time as shown in Figure 3.

It can be seen that the investment payback period of the bridge facility is inversely correlated with the number of bridge facility sets. It decreases with the increase of bridge facility sets, and when the number of bridge facility sets increases gradually, the investment payback period of the bridge facility changes more insignificantly with its service time. At the same time, the investment payback period of the bridge facility shows a distinct multistep decline with the increase of service time, and when the service time increases gradually, the investment payback period becomes increasingly insignificant with the increase of the number of bridge facility sets. $n \in (6, 30]$ is the "rapid reduction stage" of the static investment payback period of the bridge facility. At this stage, the number of the bridge equipment is small. Although the bridge leaning rate is high, there are fewer aircrafts that can actually use bridge equipment. The income of bridge equipment obtained from the airport is low. At the same time, the labor cost of bridge equipment operation accounts for a high proportion of the total operation cost each year. Therefore, at this stage, the payback period of investment is longer. When the payback period of static investment in bridge-mounted equipment reaches 16.5 years, the investment payback period decreases rapidly to 7 years and tends to be flat. At that time, the number of aircraft sorties docked to bridges and using bridge-borne equipment grows rapidly, resulting in a rapid increase in revenue. At this time, the proportion of labor costs for

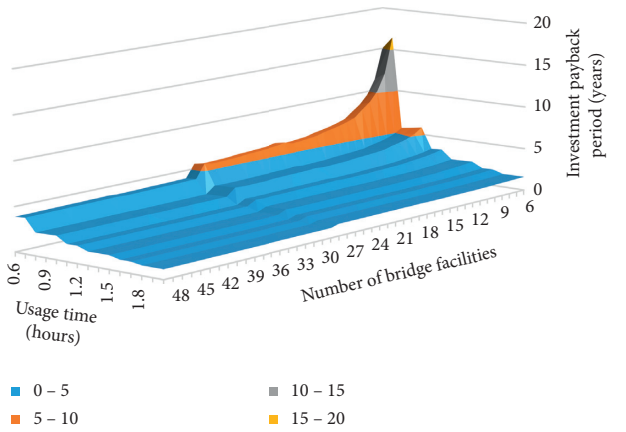


Figure 3: Influence of the number of bridge facilities and service time on the investment payback period.

annual operating costs is getting lower and lower. Therefore, the recovery period of static investment for bridge-borne equipment decreases rapidly and becomes stable. At the same time, since the bridge equipment charges for 15 minutes, the investment payback period of the bridge equipment presents a multistep decline and gradually reduces with the service time.

3.4. Analysis of Annual Operation Cost Savings for Airlines. The cost for the aircraft using the bridge facility while waiting at the station will be much less than APU hourly fuel consumption. At the same time, it can save APU time, reduce maintenance costs, and improve flight safety. In this section, from the point of airline operation, assuming that all the takeoff and landing aircrafts are B737–800, the influence of daily sorties and average time of the bridge facilities on the annual operating cost of airlines are discussed in detail as shown in Figure 4.

It can be seen that the annual operating cost saved for airlines increases linearly with the average service time of the bridge facility, and when the average service time of the bridge facility increases gradually, the annual operating cost for airlines changes more significantly with the increase of takeoff and landing sorties. The annual operating cost saved for airlines increases stepwise with the increase of takeoff and landing aircrafts at the airport. When the number of takeoff and landing sorties increases gradually, the annual operating cost saved for airlines becomes increasingly significant with the increase of time spent at the bridge facilities. This is because the utilization rate of the aircraft on the bridge decreases stepwise with the increase of takeoff and landing sorties. Therefore, the annual cost savings will have obvious inflection points at the demarcation point of the utilization rate of the aircraft on the bridge. According to the rules for the operation of the civil aircraft [6], the waiting time of B737 is about 45 minutes. Assuming that there are 50 B737 taking off and landing at the airport every day, the bridge utilization rate is 0.6. Therefore, the annual operating cost for the airline can be saved by 6,274,350 yuan, let alone the airline can save millions of operations per year after using the bridge-mounted equipment.

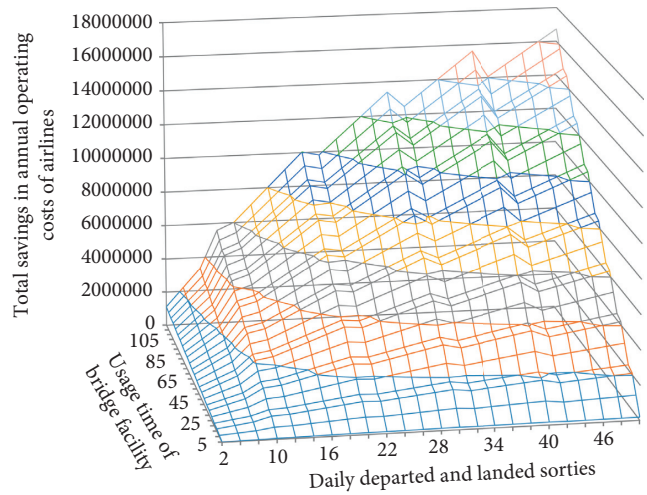


Figure 4: Consequence of daily departure/landing number of aircrafts and average service time on airlines' operational cost.

4. Conclusions

The investment payback period calculating model for bridge facility from the perspective of airports and airlines was proposed in this paper. Based on the actual operational data of Kunming Changshui Airport, Wuhan Tianhe Airport, and Lijiang Sanyi Airport in 2018, the influence of the annual bridge operational airport profit and the payback period of the bridge facility were studied with the factors such as the number of bridge facilities and the service time of bridge facility of the transition aircraft. The conclusion is that the number of bridge facilities and service time are the key points to the length of the payback period, and it changes more insignificantly with its service time. And the number of bridge facilities on the condition of the shortest payback period was pointed out in this paper. Practice shows that the bridge facilities can replace APU to provide the electrical and pneumatic power needed by the aircraft to achieve a win-win situation for airports and airlines. The research results can enhance the confidence and determination of airports and airlines in using the bridge facility and effectively promote the implementation of civil aviation energy-saving and emission reduction policies at airports.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key Research and Development Program of China Sichuan Science and Technology under Grant no. 2019FYG0308.

References

[1] X.-h. Xu and X. Li, "Cost analysis of flight delays and simulation in ground-holding model," *Journal of Nanjing University of Aeronautics & Astronautics*, vol. 38, no. 1, pp. 115–120, 2006.

[2] L.-h. Li and J.-h. Zhang, "Pollution emission and energy-saving operation for airborne APU," *Environment Science & Technology*, vol. 36, no. 10, pp. 34–38, 2013.

[3] Z.-h. Xia, R. Kang, G.-j. Zhou et al., "Ground-holding cost analysis of the delayed flights," *Journal of Sichuan University*, vol. 52, no. 4, pp. 793–799, 2015.

[4] Z.-h. Xia, "Estimation model for investment payback period for bridge facilities in civil airport," *China Science Paper*, vol. 11, no. 19, pp. 2193–2197, 2016.

[5] W.-j. Pan, Z.-j. Wang, Z.-h. Xia, and G. Zhou, "Research on flight delay loss difference of cost optimization," *Science Technology and Industry*, vol. 15, no. 12, pp. 38–42, 2015.

[6] K. W. Olsen, G. P. Whiteford, P. Sheridan, and S. K. Thompson, *Aircraft Auxiliary Power Unit Suspension System for Isolating an Aircraft Auxiliary Power Unit*, European Patent Office, Munich, Germany, 2013.

[7] H. Ebrahimi, J. R. Gatabi, and H. El-Kishky, "An auxiliary power unit for advanced aircraft electric power systems," *Electric Power Systems Research*, vol. 119, pp. 393–406, 2015.

[8] W.-X. Chen, J.-Q. Lin, and J.-H. Zhang, "Design of network-based energy management system for airport bridge-borne equipment," *Energy and Energy Conservation*, vol. 5, 2014.

[9] Z. Li, J. Lin, and X. Wang, *Modeling and Control of Aircraft Cabin Temperature System Based on Airport Bridge-Borne Air Conditioner*, Modern Electronics Technique, Riyadh, Saudi Arabia, 2017.

[10] China Civil Aviation Administration, *Interim Measures for the Operation of Bridge Facility in Civil Airports as a Substitute for Aircraft Auxiliary Power Devices*, China Civil Aviation Administration, Beijing, China, 2013.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 6217872, 9 pages
https://doi.org/10.1155/2020/6217872

Research Article

Construction of Risk Evaluation Index System for Power Grid Engineering Cost by Applying WBS-RBS and Membership Degree Methods

Jiao Wang

School of Business Economic Management, Northeast Electric Power University, Jilin 132012, China

Correspondence should be addressed to Jiao Wang; dbdl2037134@126.com

Received 3 March 2020; Accepted 29 June 2020; Published 4 August 2020

Guest Editor: Weilin Xiao

Based on the perspective of cost risk management of the entire process of power grid construction project, the WBS-RBS risk identification method is applied. The risk factors of power grid engineering cost are comprehensively identified from the five major engineering cost formation stages: investment decision estimate stage, design estimate stage, bidding price stage, construction budget stage, and final account stage. Combined with the basic data obtained from the questionnaire, the membership degree method is adopted to screen the key indicators of the preliminary factors of cost risk and then eliminate the redundant indicators. Furthermore, a set of simple and efficient risk evaluation index system for power grid engineering cost is constructed in order to improve the risk evaluation and management level of power grid construction project in China.

1. Introduction

An important content of power grid construction project management is cost risk evaluation. Scientific project risk evaluation can ensure that the cost of the entire process of power grid engineering can be controlled effectively. Through the rational use of funds, the goal of optimal return on project investment can be achieved.

The research on the risk evaluation of power grid engineering cost mainly began at the beginning of this century so that the related literature is still relatively limited. Among these literature studies, the study in [1] adopts data envelopment analysis (DEA) to investigate the cost risk existed in the process of power grid construction. The study in [2] adopts Monte Carlo method to simulate and predict the risks encountered in power grid engineering projects. The study in [3], in order to solve the problem of uncertainty in the power grid engineering cost caused by the adoption of single prediction model, adopts artificial neural network (ANN) and genetic algorithm (GA). It has optimized the prediction model of power grid engineering cost and improved the accuracy. However, none of the above literature has established a relatively complete cost risk evaluation index system for power grid engineering, which is not conducive to the project cost prevention and control personnel to evaluate and deal with risks in a timely and effective manner. The study in [4] analyzes risk sources of power grid engineering through the application of the fault tree analysis (FTA). Although engineering cost risk evaluation index system is constructed, the subjective factors are too heavy, and it lacks objective credibility when used for evaluation. The study in [5] analyzes the risks that may be involved in the power grid engineering cost qualitatively and quantitatively and adopts fuzzy evaluation method to calculate risk degree of engineering projects. Nevertheless, it fails to identify key risk indicators so that the practical application efficiency of the risk evaluation index system is reduced. The study in [6] systematically classifies power grid cost risk factors, draws the fishbone diagram, and analyzes the sensitivity of individual cost factors. Although it identifies the risk factors that affect cost, it does not evaluate the extent to which they may cause cost losses.

However, in other large-scale engineering construction fields, some scholars had adopted WBS-RBS method to

evaluate the project cost risk and obtained relatively ideal results. The study in [7] had utilized WBS-RBS method combined with literature reading to identify the risk factors of civil engineering projects and constructed a risk evaluation index system of engineering cost including 28 secondary risk factors, which can provide a basis for risk managers to make risk decisions. The study in [8] had adopted work breakdown structure-risk breakdown structure (i.e., WBS-RBS) method to identify the risk of water conservancy project in the whole process management according to the complicated characteristics of water conservancy project, so as to effectively control the cost risk of the whole process of water conservancy projects. The study in [9] had applied WBS-RBS method to identify and analyze the risks of subway construction, which improved the risk management and control level of subway construction cost. However, the WBS-RBS risk assessment method used in the above mainly focuses on the evaluation of the importance of risk but neglects the importance of the project itself, which also affects the risk. The study in [10] proposed a risk assessment method based on WBS-RBS-DSM, which is on the basis of the traditional WBS-RBS matrix, to calculate the comprehensive risk value faced by the project. But the application conditions of DSM model are relatively strict, which is difficult to meet in practical application. The study in [11], in the process of risk identification, had applied WBS-RBS risk identification method to decompose and analyze the project systematically. He used Delphi method to identify each preliminary risk factor of the identification matrix systematically and screened out the key factors affecting the project cost. However, due to the existence of many subjective factors in the discrimination process, the evaluation results are not objective.

In conclusion, on the basis of previous research on cost risk assessment of nongrid projects by WBS-RBS method, this paper improves WBS-RBS method through the analysis of membership degree. WBS-RBS method based on the analysis of membership degree is also applied to the risk evaluation of power grid engineering cost, which reflects the time, stage, and management object of project risk in detail, so as to enable the relevant organizations of the project to defend against the risk in advance. In addition, compared with the traditional risk identification method, the WBS-RBS membership analysis method could identify risks in each stage of the power grid engineering life cycle in detail, which makes up for the omission of traditional risk identification methods such as brainstorming, Delphi, or expert investigation in practical application due to the randomness and subjectivity.

Therefore, based on the research results on current risk evaluation of power grid engineering cost, this paper starts with the entire process of power grid construction and adopts WBS-RBS method to identify cost risk factors for the entire process of power grid construction project in China. Then, membership degree method is applied to select key factors that affect power grid engineering cost from the preliminary identification indicators of power grid construction project risks, so as to establish a relatively scientific, streamlined, and appropriate system for risk evaluation of power grid engineering cost. This enables China's power grid engineering cost analysts and project risk-related managers to fully and clearly recognize the risk factors of cost in each stage of the entire project process, thereby achieving the goal of effective control of power grid engineering cost in advance.

2. Influence of Each Stage of Power Grid Construction on Cost

Risk evaluation of power grid engineering cost is integral to power grid construction project management. Because the project cost management runs through the entire process of power grid project construction, the risk evaluation of power grid engineering cost is staged and complex. The extent to which each stage of the current power grid engineering affects the cost is illustrated in Figure 1.

According to Figure 1, the impact of the investment decision-making stage on project cost is 75%–95%, the design stage is 35%–75%, the bidding stage is 20%–30%, the construction stage is 35%–50%, and the final account stage is 0–5%. It is clear from Figure 1 that the investment decision-making stage and the design stage before the implementation of the project are the key points of cost control. Consequently, relevant work such as identifying risk factors that affect engineering cost, screening evaluation indicators, and constructing index system should be carried out in the investment decision-making stage, which is the early stage of power grid engineering.

3. Features of Power Grid Engineering Cost Risk Management and Design Principles of the Index System

3.1. Features of Power Grid Engineering Cost Risk Management. The risk of power grid engineering cost widely exists in the entire process of power grid construction mainly due to the following three aspects: Firstly, the value of construction products is the basis for the formation of engineering cost. Secondly, the power grid construction itself is unique and complex. Thirdly, there are many differences between the composition and calculation of power grid engineering cost and other large- and medium-sized construction projects. The study in [12] analyzes the factors that affect power grid engineering cost from both macroperspectives and microperspectives. From the macroperspective, government behavior, market supply and demand, construction science and technology, and socioeconomic development may all have a direct or indirect impact on the power grid engineering cost. From the microperspective, for a specific power grid construction project, its design level, construction level, natural environment conditions, owners' value orientation, and management behavior could directly or indirectly affect its engineering cost and run through the entire process of power grid construction.

In addition, due to the influence of natural conditions such as geology and geomorphology, as well as the unbalanced economic development level of various provinces in

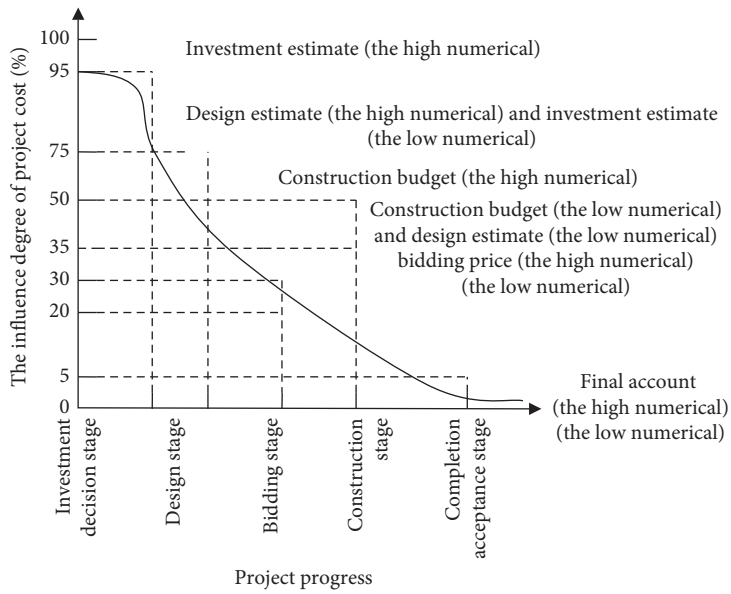


Figure 1: Project cost process.

China, the risk of power grid engineering cost varies in different regions. In view of similar natural and geological conditions and economic levels in Northeast China, it is more appropriate to analyze the risk factors of power grid engineering cost. Therefore, this paper focuses on the risk factors that affect the power grid engineering cost in Northeast China. The study in [13] considers that no matter it is the cost management of power transmission and transformation project or the cost management of other types of engineering projects, it is necessary to identify and analyze the cost risk. This work should be carried out in the early stage of risk management and run through the entire process of engineering cost. Identifying and analyzing the risk factors of power grid engineering cost is the premise and basis of cost risk estimation, cost risk evaluation, and cost risk response. The WBS-RBS risk identification method has a good effect on the risk identification process of other large- and medium-sized construction projects. This method is comprehensive and clear in analyzing the risk factors, and the conclusions obtained are more accurate. Based on the above factors, this paper introduces the WBS-RBS method for the first time to systematically decompose and initially identify the risk factors of power grid engineering cost risks and then combines expert questionnaire survey opinions and adopts the membership degree method to further screen the preliminary risk indicators, so as to acquire the key risk factors that need to be focused on prevention and control in power grid engineering cost.

3.2. Design Principles of the Cost Risk Evaluation Index System. Risk evaluation index system for power grid engineering cost is a collection of potential risk factors at all stages in the entire process of power grid construction. This paper starts from the perspective of risk management in the entire process of power grid engineering cost, combines the actual conditions of power grid engineering, and identifies and screens out key risk factors that affect power grid engineering cost in China based on the design principles of evaluation indicators to establish a clear and simple cost risk evaluation index system. The specific design principles are as follows.

(1) Principle of scientificity: in the selection of indicators, the calculation of the selected indicators is required to be clear and accurate. The content reflected should be objective, comprehensive, and standardized. Those fuzzy factors cannot be selected as evaluation indicators.

(2) Principle of hierarchy: the cost risk evaluation indicators cover a wide range and the evaluation system constructed is relatively complex. Therefore, the identification and selection of indicators should not only meet the integrity of the evaluation system, but also reflect the hierarchical relationship between cost risk factors.

(3) Principle of feasibility: when constructing cost risk evaluation index system, it is necessary to fully consider whether the information expressed by the indicators is accurate and available or not, which determine whether the evaluation index system can be effective or not in practical applications.

(4) Principle of combining qualitative and quantitative analysis: in order to avoid too many subjective factors or improper processing of objective data involved in evaluation process that lead to inaccurate evaluation results, qualitative and quantitative analysis methods should be combined to design a balanced and scientific evaluation index system.

(5) Principle of relative independence: in order to acquire objective and accurate evaluation results, it is required to maintain great relative independence between the indicators in the evaluation system and

to avoid the high overlap of the risk information attribute reflected between the indicators to cause problems with indicator membership attribute.

4. WBS-RBS Model Introduction

WBS (work breakdown structure) refers to forming a WBS tree, which is a relatively complete work breakdown structure in project management, by decomposing the tasks of an engineering project layer by layer into a minimum work package (WP). RBS (risk breakdown structure) is derived from the construction principle of the work breakdown structure (WBS), and the RBS tree is a more comprehensive risk breakdown structure in project risk management. Since the breakdown structures of WBS and RBS are both guided by project goals, the two can be combined to form a WBS-RBS matrix to determine the risks existing at different stages of the project and analyze the extent of the loss caused by the risk factors to the project and probability of occurrence [11]. The WBS-RBS model is illustrated in Figure 2.

The WBS tree is shown in the upper part of Figure 2. The American Project Management Institute (PMI) defines WBS as a hierarchical decomposition of the work that the project team needs to perform to create the required deliverables and achieve the goal of completing the project. The planned work included in the lowest level of its structure is called work package (WP) [10]. The RBS tree is shown in the left half of Figure 2. RBS refers to the risk breakdown structure formed by degrading the risks faced by the project to the most basic risk events through hierarchical decomposition technology.

The lowest level work package (WP) of the work breakdown structure tree is regarded as the column of the WBS-RBS matrix, and the lowest level basic risk event of the risk breakdown structure tree is regarded as the row of the WBS-RBS matrix to construct a complete WBS-RBS risk identification matrix. Then, whether the work package (WP) faces corresponding risks can be judged one by one according to the matrix elements [14].

5. Identification of Influencing Factors of Power Grid Engineering Cost Risk Based on the WBS-RBS Method

The risk evaluation of power grid engineering cost should run through the entire process of the implementation of power grid construction project and is continuous and systematic. It mainly includes five stages: investment decision estimate, design estimate, bidding price, construction budget, and final account. This paper adopts WBS-RBS method to decompose the entire process of power grid engineering cost and analyze the potential risk factors at the corresponding stage.

5.1. WBS Tree in the Entire Process of Power Grid Engineering Cost. The entire process of power grid engineering cost management mainly includes the following five stages:

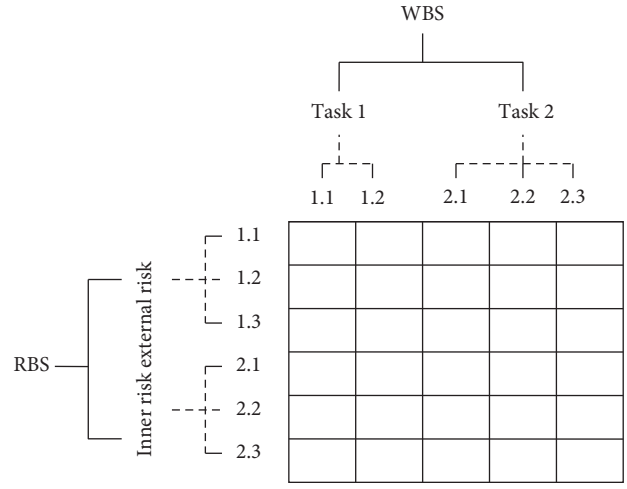


Figure 2: WBS-RBS matrix diagram.

estimation management in investment decision-making stage, budget estimation management in design stage, price management in bidding stage, budget management in construction stage, and final account management (including settlement) in completion stage, which form a complete cost management system for the entire process of power grid engineering. The cost management work of each stage continues to be broken down to the specific activity work package (WP), and finally the WBS tree can be constructed as illustrated in Figure 3.

5.2. RBS Tree of Cost Risk in Each Stage of Power Grid Engineering Cost. Due to the large amount of complex power grid construction project and the long construction cycle, there are many risk factors that affect the project cost. In order to accurately identify the cost risk factors, based on the management theory of the entire process of power grid construction, this paper divides the risk sources into five categories according to the five stages of power grid construction: risks in investment decision-making stage, risks in design stage, risks in bidding stage, risks in construction stage, and risks in final account stage [12]. Then, the risk sources of each stage are broken down layer by layer, and finally the RBS tree is constructed, as illustrated in Figure 4.

5.3. Construction of WBS-RBS Risk Identification Matrix for Power Grid Engineering Cost. Since the key risk factors of cost will be identified and the evaluation index will be optimized later, the probability of risk events and the degree of possible loss to the project are not considered here. The risk factors that affect power grid engineering cost are listed, as shown in Table 1.

6. Screening of Key Indicators of Cost Risk by Using the Membership Degree Method

The cost risk factors for the entire process of power grid construction project listed in Table 1 are the preliminary identification index system. The number of indicators in this

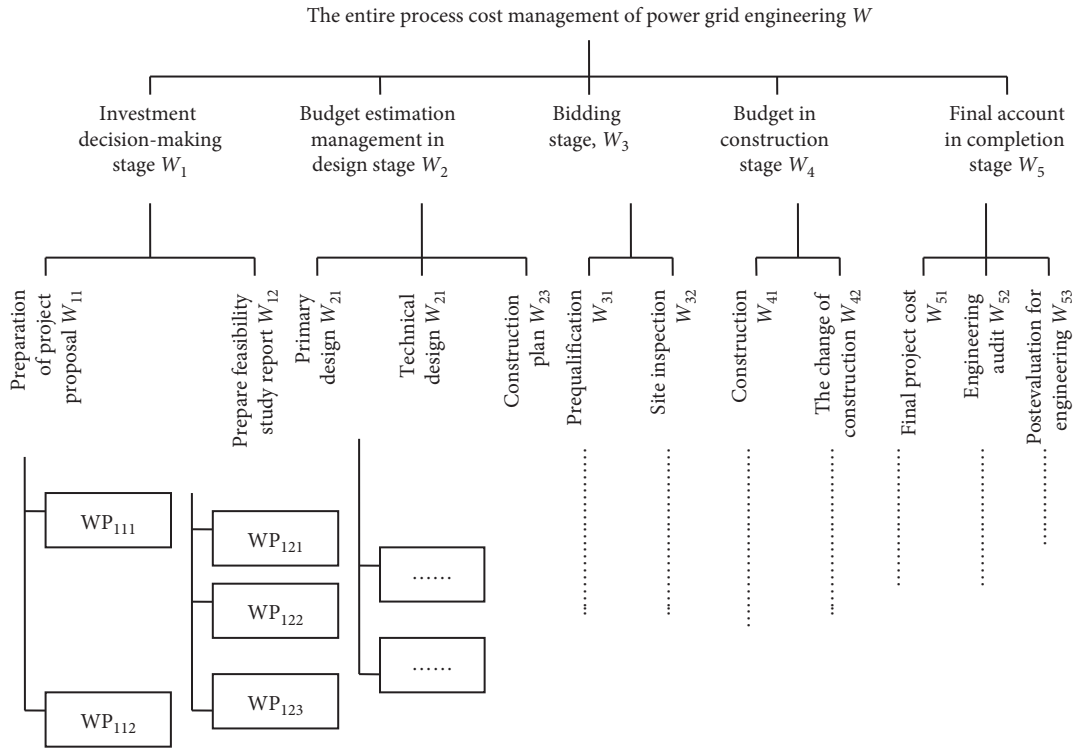


Figure 3: WBS diagram for power grid construction project.

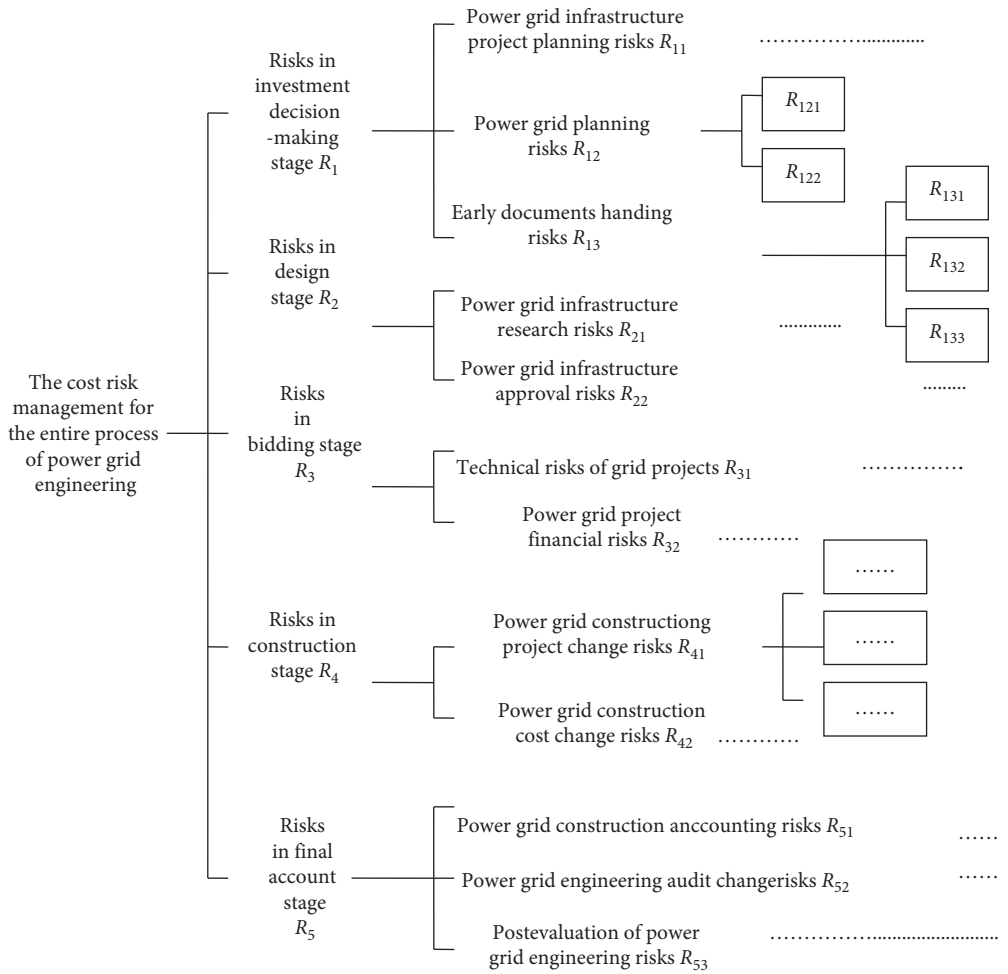


Figure 4: RBS diagram for power grid engineering project.

Table 1: Preliminary identification of risk factors in the entire process of power grid engineering cost.

The first-level indicator	The second-level indicator	The third-level indicator
Risk of investment decision (R_1)	Risk of power grid planning (R_{11})	Risk of power load forecasting (R_{111}), risk of economic situation evaluation (R_{112})
	Risk of planning of power grid capital construction project (R_{12})	Risk of scarcity of capital construction land resource (R_{121}), risk of planning connection for power grid engineering cities (R_{122}), risk of planning and coordination in cross-sectoral engineering (R_{123}), risk of planning of supporting projects for power grid construction (R_{124})
	Risk of handling of preliminary documents (R_{13})	Risk of law and policy (R_{131})
Risk of design (R_2)	Risk of project establishment and research in power grid capital construction (R_{21})	Risk of purchase price of capital construction equipment (R_{211}), risk of participation in power grid engineering audit (R_{212}), risk of young crop compensation for power grid capital construction project (R_{213}), risk of selection of design company (R_{214})
	Risk of project establishment and approval in power grid capital construction (R_{22})	Risk of project approval cycle (R_{221}), risk of allocation of power grid engineering approval personnel (R_{222})
Risk of bidding (R_3)	Risk of power grid engineering technology (R_{31})	Risk of experience of tendering and bidding companies (R_{311}), risk of technical level of construction company (R_{312})
	Risk of power grid engineering finance (R_{32})	Risk of inner ability to pay wages in advance of construction companies (R_{321}), risk of target profit margin (R_{322})
Risk of construction (R_4)	Risk of project change in power grid capital construction (R_{41})	Risk of project change notice (R_{411}), risk of tracking of capital construction project planning (R_{412})
	Risk of cost change in power grid capital construction (X_{42})	Risk of land acquisition fund (R_{421}), risk of machinery, material and labor cost (R_{422})
Risk of final accounts (R_5)	Risk of final accounts in power grid engineering (R_{51})	Risk of final accounts (R_{511}), risk of engineering payment (R_{512})
	Risk of audit in power grid engineering (R_{52})	Risk of audit quality of final accounts (R_{521}), risk of audit of final accounts (R_{522}), risk of separate audit of "quantity and price" in final accounts (R_{523})
	Risk of postevaluation for power grid engineering (R_{53})	Risk of completion rate of power grid engineering (R_{531}), risk of evaluation after project completion (R_{532})

system is complex. In the actual project risk analysis, the relevant risk prevention and control personnel are required to collect massive data and information, so as to further analyze the complex data and draw conclusions. This is undoubtedly a time-consuming and labor-consuming huge project, which greatly reduces the evaluation efficiency of the final risk indicators. Moreover, there are subjective evaluation indicators in the preliminary identification index system, which may reduce the objectivity and accuracy of the evaluation results. Therefore, this paper takes the form of a questionnaire. Firstly, it collects the opinions of experts in the related fields of engineering cost extensively and adopts Likert scale method to ask the corresponding experts to score the preliminary indicators and further adjust the preliminary indicators according to the score results. Secondly, it combines the membership degree method to calculate the membership degree value of the risk index, selects the key indicators that affect the cost risk of power grid engineering based on the membership threshold, and establishes the risk evaluation index system so that it can be applied to the entire process of risk evaluation of power grid engineering cost in practice.

6.1. Distribution and Recovery of Questionnaires. The questionnaire design is based on Likert scale method, and it adjusts the measurement method appropriately. Specified score is as follows: 1–5 points, of which 1-not important, 2-generally important, 3-relatively important, 4-important, and 5-very important. The lower the score of an indicator, the less important it is. It can be eliminated within a certain standard and range, and the remaining indicators can establish the risk evaluation index system. This paper adopts e-mail, on-site distribution, and other methods to distribute the questionnaire. The questionnaire will be distributed to experts in the industries such as construction companies, survey and design companies, and supervision companies, as well as relevant researchers in universities and research institutes.

A total of 120 questionnaires were distributed to relevant parties involved in the construction project, and 116 questionnaires were recovered as required. The validity of the questionnaire is divided according to the degree of convergence of the selection results. If the degree of convergence is too high, it is regarded as an invalid questionnaire, and the invalid questionnaire is eliminated. Based on this criterion, 7 questionnaires returned were invalid and 109 questionnaires were effectively recovered, with a recovery rate of 90.8%. The statistical results of the entire process of the risk evaluation system of power grid engineering cost are shown in Table 2.

Table 2: Score statistics.

Index sign	Score distribution				
	5	4	3	2	1
R_{111}	9	25	37	48	0
R_{112}	6	35	30	38	0
R_{121}	10	40	33	20	6
R_{122}	0	1	28	59	21
R_{123}	24	37	33	15	0
R_{124}	15	36	21	37	0
R_{131}	0	3	34	56	16
R_{211}	0	18	39	42	10
R_{212}	2	30	35	31	11
R_{213}	0	14	56	26	13
R_{214}	10	35	45	10	9
R_{221}	0	0	37	58	14
R_{222}	1	13	24	32	39
R_{311}	0	38	37	24	10
R_{312}	14	46	32	17	0
R_{321}	11	29	41	17	11
R_{322}	0	19	30	35	25
R_{411}	2	22	60	25	0
R_{412}	11	23	36	31	8
R_{421}	9	23	37	49	0
R_{422}	0	14	35	36	24
R_{511}	0	14	36	40	19
R_{512}	0	24	47	38	0
R_{521}	0	0	25	53	31
R_{522}	0	21	47	37	4
R_{523}	33	40	23	13	0
R_{531}	21	53	32	3	0
R_{532}	0	15	59	33	2

6.2. Reliability and Validity Tests of Questionnaire Results. In order to ensure the objectivity and accuracy of the questionnaire results, the corresponding data will be tested for reliability and validity. SPSS 17.0 mathematical statistical analysis software was adopted to perform Cronbach test on the data collected from a valid questionnaire. The results show (as shown in Table 3) that Cronbach's Alpha coefficient is 0.811. When this value is greater than 0.8, it indicates that the preliminary evaluation index has passed the reliability test and has a relatively ideal reliability. The validity test results show (as shown in Table 4) that the KMO sample measurement value is 0.712, which is greater than 0.5. Bartlett sphericity test value is extremely small and much less than 0.001, which indicates that the preliminary risk index is suitable for factor analysis, so it passes the validity test.

6.3. Selection of Key Risk Indicators and Establishment of the Indicator System by Using the Membership Degree Method. In view of the process of social science research, membership degree is a concept adopted to analyze the evaluation ability of a specific indicator on the target to be evaluated. As the number of preliminary risk indicators constructed by WBS-RBS matrix is relatively complex, which affects the application effect of risk evaluation index system in actual projects, the membership degree method is adopted to optimize the preliminary index system and screen the key indicators that affect the risk of power grid engineering cost, so as to establish a flexible, simple, and efficient risk evaluation index system suitable for the actual operation in China [8].

Let $\{R\}$ represent a fuzzy set, which represents the risk index system of power grid engineering cost. Each element in $\{R\}$ (the risk index) is represented by R_t; then the analysis of membership degree can be performed on R_t. And ask the n experts selected in the previous questionnaire to score the ability to evaluate many indicators of the preliminary cost risks of power grid engineering. The score result obtained for a certain R_t indicator is represented by the membership degree value Z_t. Z_t is

Table 3: Reliability analysis results of the questionnaire.

Cronbach's alpha	Standardized item-based Cronbach's alpha	The number of terms
0.811	0.920	28

Table 4: Validity analysis results of the questionnaire.

Adequate sampling Kaiser-Meyer-Olkin measurement	0.712
Bartlett spherical degree of inspection	
The approximate chi-square	1619.023
Df	286
Sig.	0.000

Table 5: Cost risk indicator score.

Index code	The number of specialists	The membership degree (Z_t)
R_{111}	71	0.651376
R_{112}	71	0.651376
R_{121}	83	0.761468
R_{122}	29	0.266055
R_{123}	94	0.862385
R_{124}	72	0.660550
R_{131}	37	0.339450
R_{211}	57	0.523936
R_{212}	67	0.614679
R_{213}	63	0.642202
R_{214}	90	0.825688
R_{221}	37	0.339450
R_{222}	38	0.348624
R_{311}	75	0.688073
R_{312}	92	0.844037
R_{321}	81	0.743119
R_{322}	49	0.449541
R_{411}	64	0.770642
R_{412}	70	0.642202
R_{421}	69	0.633028
R_{422}	49	0.449541
R_{511}	50	0.458716
R_{512}	66	0.651376
R_{521}	25	0.229358
R_{522}	18	0.623853
R_{523}	96	0.880734
R_{531}	106	0.972477
R_{532}	74	0.678899

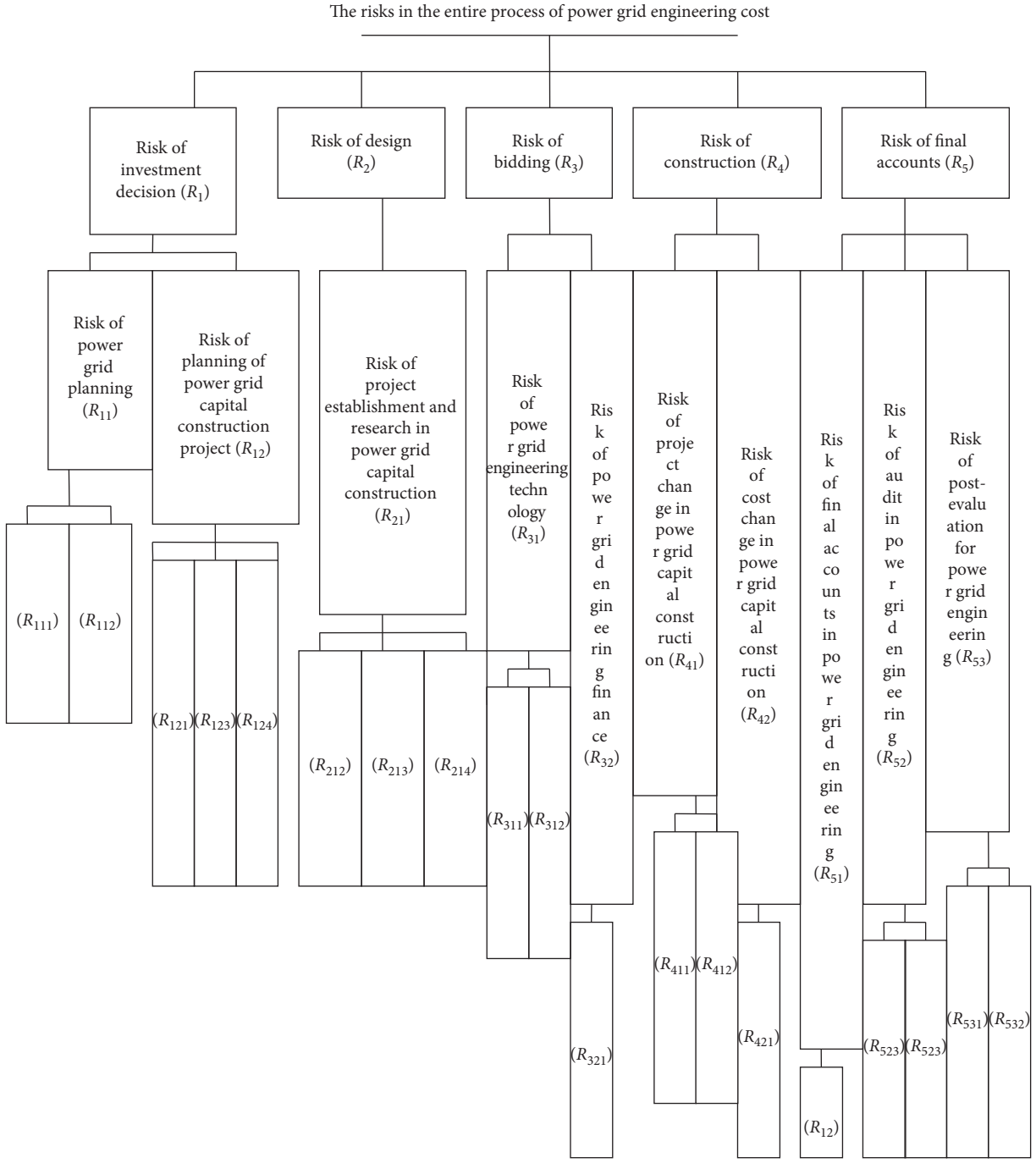


Figure 5: Risk evaluation index system of power grid engineering cost.

calculated by the formula $Z_t = m_p/n$, where m_p represents the number of experts who choose very important, important, and relatively important and n represents the total number of experts who participate in scoring. The larger the value of Z_t, the greater the possibility that the indicator R_t belongs to the fuzzy set $\{R\}$. That is, it can be judged that this indicator is critical in the evaluation system and can be retained as a formal evaluation indicator; otherwise, it is necessary to eliminate it to achieve the purpose of optimizing the evaluation indicator system. This paper analyzes the 109 valid questionnaires collected and calculates the membership degree of the risk evaluation index system for the entire process of power grid engineering through the formula $Z_t = m_p/n$, as shown in Table 5.

At present, there is no authoritative conclusion on the determination of the threshold value in the analysis of membership degree. Therefore, this paper calculates the sum of the membership values of all the preliminary cost risk indicators shown in Table 3 and then takes the average value as the threshold value of the selected risk evaluation index system. That is, if the membership degree value Z_t of the risk indicator R_t is less than the average membership degree 0.614387, the

evaluation system is excluded. The risk indicators that have not been excluded in the end are the key indicators for the risk evaluation of power grid engineering cost, and the key indicators are adopted to establish the risk evaluation index system of power grid engineering cost, as illustrated in Figure 5 (note: due to space limitations, the third-level indicators are only represented by letters as seen in Table 1).

7. Conclusion

Due to the influence of various internal and external uncertainties, there are great risks which often lead to the overrun of project budget in the power grid projects. This paper combines the characteristics of power grid engineering and based on the characteristics of project cost management at each stage of power grid construction, the connotation of cost risk management, and the effect of risk evaluation, this paper analyzes from the perspective of cost risk management of the entire process of power grid engineering and adopts the WBS-RBS method for the first time to identify risks in the entire process of power grid engineering by referring to the application of WBS-RBS method in the cost management of hydraulic project engineering. This is more comprehensive than traditional methods such as decision tree to identify power grid cost risk factors. At the same time, this paper studies the efficient application of the membership degree method in government performance evaluation and highway engineering management performance evaluation and identifies and adopts the membership degree method to identify and screen the complicated preliminary indicators of power grid cost risk, so as to acquire key indicators that affect project cost risk.

The establishment of WBS-RBS membership analysis method not only enables the power grid construction company to prevent risks well in the construction process, but also enables the supervisor to conduct a key investigation on the behaviors of all parties involved in the project construction process.

Finally, the WBS-RBS membership analysis method was applied to study the risk of power grid engineering cost in China. A simple scientific risk evaluation index system suitable for the power grid engineering cost in China is constructed in order to provide beneficial research ideas for China's risk evaluation and management of power grid engineering cost.

Data Availability

The data in this article can be shared with relevant institutions or individuals for research.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the doctoral research fund of North–East China Electric Power University that started the project BSJXM-2019106.

References

[1] W. Zhang and Y. Duan, "Social stability risk relationship model of major engineering projects based on fuzzy-ISM," *Journal of Civil Engineering and Management*, vol. 36, no. 5, pp. 102–108, 2019.

[2] X. Wu, L. Lin, and J. Gu, "Analysis on comprehensive benefit of life-cycle engineering consulting project in power transmission and transformation engineering: taking seven pilot projects of state grid as examples," *Construction Economy*, vol. 41, no. 1, pp. 15–21, 2020.

[3] X. Ding and F. Xu, "Bayesian network modeling for project risk management: a case study of Hong Kong-zhuhai-Macao bridge," *Journal of Systems and Management*, vol. 27, no. 1, pp. 176–192, 2018.

[4] K. Nan, S. Gui, J. Wang et al., "Cost of the distributed power distribution network risk assessment simulation research," *Computer Simulation*, vol. 34, no. 3, pp. 96–100, 2017.

[5] Z. Zhao and Q. Man, "Evolutionary game analysis of risk management behavior of major infrastructure projects based on prospect theory," *Journal of Systems and Management*, vol. 27, no. 1, pp. 99–117, 2018.

[6] X. Huang, "Study on monte-carlo method based risk assessment model of power transmission project cost," *Modern Electronics Technique*, vol. 40, no. 20, pp. 178–180, 2017.

[7] L. Z. Kong, "Risk evaluation of engineering project cost based on set pair analysis," *Journal of Civil Engineering and Management*, vol. 33, no. 1, pp. 90–96, 2016.

[8] Y. Xue, "Risk identification of water conservancy projects in the whole process of WBS-RBS method," *Water Conservancy and Hydropower in Rural China*, vol. 23, no. 2, pp. 71–78, 2018.

[9] J. Chen, "Risk Identification of WBS-RBS method in subway construction," *The North Project*, vol. 11, no. 1, pp. 109–111, 2017.

[10] N. RenB. Han et al., "Project risk identification and assessment based on WBS-RBS-DSM," *Systems Engineering*, vol. 32, no. 11, pp. 97–100, 2014.

[11] X. Fang W. Zhang et al., "Optimization of evaluation index cost risk based system from power transmission project on rough seta theory," *Journal of Civil Engineering and Management*, vol. 32, no. 4, pp. 40–47, 2015.

[12] J. Wang and Y. C. Li, "Prediction model for construction cost based on grey relational analysis PSO-SVR," *Journal of Huaqiao University (Natural Science Edition)*, vol. 37, no. 6, pp. 708–713, 2016.

[13] X. Cui Y. Zhang et al., "Power transmission project cost risk assessment based on matter element and extension theory," *Journal of Civil Engineering and Management*, vol. 12, no. 4, pp. 90–94, 2015.

[14] J. Wang and L. Ding, "Analysising the main impact factors on the build price of the 500 kV power line engineering," *Journal of Northeast Dianli University (Natural Science Edition)*, vol. 32, no. 5, pp. 9–11, 2012.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 3568470, 16 pages
https://doi.org/10.1155/2020/3568470

Research Article

Research on Utility Evaluation of Grid Investment considering Risk Preference of Decision-Makers

Hongliang Wu, Daoxin Peng, and Ling Wang

Energy Development Research Institute, CSG, 510623, Guangzhou, China

Correspondence should be addressed to Daoxin Peng; pengdaoxin92@163.com

Received 20 March 2020; Revised 16 June 2020; Accepted 22 June 2020; Published 29 July 2020

Guest Editor: Shianghau Wu

Effectiveness evaluations are one of the important ways to guide grid investment and to improve investment efficiency. Improving the effectiveness of grid investment evaluations is studied based on the optimization of the investment evaluation index system and the utility evaluation model. The index system is optimized by establishing an evaluation index system of grid investment effectiveness, considering the redundancy between the indices, and constructing an ISM-DEA model. The utility function model was introduced to fully consider the different risk appetites of decision-makers, and a utility evaluation model that takes risk appetite into account was established. An improved weight integration model based on multiobjective optimization was established by considering the minimum deviation and the trend-optimal objective function when setting the index weights. The calculation results show that the feasibility of the index system optimization model and utility evaluation model constructed in this study is verified under the premise of satisfying the assumptions. By adjusting the risk preference coefficient of decision-makers, the dynamic optimization of the grid investment utility evaluation results can be realized.

1. Introduction

The Central Committee of the Communist Party of China and the State Council formally issued "Several Opinions on Further Deepening the Reform of the Electric Power System" on March 15, 2015. With the continuous deepening of this round of the power system reform, the role of power grid companies in the power industry chain is required to gradually change to that of middlemen, who collect electricity transaction fees in the market environment. Subsequent documents, such as the supervision and review of transmission and distribution pricing costs, have caused grid companies to face greater investment risks—especially for researching and determining the effectiveness of grid investment—and have established higher investment evaluation work requirements for grid companies [1, 2].

The traditional assessment of power grid investment primarily focuses on technical indicators, and insufficient attention is given to economic and environmental benefit indicators. The many studies that address power grid investment assessment have limitations, and research on grid investment evaluation can be divided into the construction of evaluation index systems and the construction of evaluation models.

1.1. Evaluation Indicator System. Some scholars focus on the evaluation of a single indicator to establish an indicator system. For example, Psarros and Georgios N use the levelized cost of energy as the only primary evaluation index; a comprehensive evaluation index is developed to evaluate the potential of the energy system, which integrates all the technical, economic, and environmental performance indicators [3, 4]. A single indicator focuses on one aspect of the problem, and results are one-sided. More scholars now conduct research on the indicator system. For example, Yu [5, 6] constructed a multi-dimensional evaluation index system when studying the investment benefits of power grids and distribution network investments and applied them to power grid assessments. Jiang [7] used a neural network model to construct a grid investment evaluation index system that considered value at risk. Han [8] considered the

effectiveness of micro-grid investment for photovoltaic and energy storage access and constructed a multi-dimensional evaluation index system for the effectiveness of micro-grid investment.

1.2. Evaluation Model. Scholars have completed abundant research in fields related to evaluation models. For example, Guo [9] used the correlation feature fusion analysis, correlation tests, and statistical analysis for the fuzzy attribute feature decomposition of power grid investment benefit evaluation. A regression analysis of the samples helps achieve accurate prediction and quantitative evaluation of power grid investment benefit evaluation. Xiaobao Yu [10, 11] performed evaluation research on the value of power customers by constructing an index weighting model under multi-objective function optimization. Xing [12] used the set pair analysis method to construct a grid investment evaluation model and a single factor sensitivity analysis method and model to calculate the grid project investment benefit. He [13] established a sound evaluation index system for power grid investment projects and adopted a DEA analysis model to perform evaluation research. Sun [14] realized evaluation research on grid investment by constructing an evaluation model based on an improved DEA and Monte Carlo simulation; Yu [15, 16] used the value-at-risk theory introduced into a comprehensive energy system assessment and constructed an assessment based on the CVaR theory and multiobjective optimization index weighting models.

In addition to the evaluation index system and the evaluation model, research on power grid investment evaluation also involves risk assessment issues. Negnevitsky [17] presented a risk assessment approach to analyze power grid system security for operation planning with high wind power generation penetration. Liang [18] proposed a novel risk-based uncertainty set optimization method for the energy management of typical hybrid AC/DC microgrids, where RPG outputs are considered to be the major uncertainties. Karki [19] presented a generalized and approximate risk-based method that is relatively simple to apply. Cao [20] presented a methodology based on cyber-power joint analysis to quantitatively evaluate the operational risk of ADNs caused by cyber contingencies.

Analysis of previous research reveals many studies on power grid investment evaluation, but deficiencies exist. First, in the construction of a power grid investment evaluation index system, more consideration is given to technical indexes than to social and environmental benefits, and the coupling between indexes is not addressed. Second, conventional single evaluation methods are typically selected; some scholars use a variety of evaluation theories, producing a lack of research results that consider the risk preference of evaluation subjects. Therefore, this study optimizes two aspects of the research on power grid investment effectiveness evaluation. A comprehensive dimension and oriented and optimized index system is important for the construction of a multidimensional evaluation index system. The utility function theory is used during the construction of the evaluation model to embed the risk preference of the decision-maker. The specific novelties of this study are the following:

① Optimization of the evaluation index system of conventional power grid investment effectiveness, construction of an improved ISM-DEA model, division of index types by using the relationship between indexes, and measurement of the influence of each index on the evaluation results by using the control variable method to eliminate the indexes with the least influence, solving the problem of index coupling and optimizing the index system.

② Combination of two mutually informative methods considering a variety of evaluation weight results, consideration of the advantages and disadvantages—any subjective or objective weighting method has disadvantages when used alone—of different methods, maximization of the characteristics of various weight results, and a proposal for the construction of a comprehensive weight calculation model based on the moment estimation theory.

③ Construction of the utility evaluation power grid investment model based on risk preference, configuration of risk preference parameters by embedding the degree of risk preference of decision-makers in the utility function to influence the evaluation model results by adjusting the risk preference coefficient to realize the dynamic evaluation adjustment results according to different people.

This article is divided into five parts. The first is an introduction that outlines the research background and current development status of power grid investment assessment. The second and third parts are the main bodies of this dissertation. The second part primarily screens power grid investment effectiveness indicators and establishes a conventional grid investment index system and an index optimization model. The third part introduces the utility function theory, designs the risk preference coefficient of the decision-maker, and inserts said coefficient into the utility function theory to construct the utility evaluation model of power grid investment based on risk preference. The fourth part performs the simulation analysis of the model and the routine and comparison analyses of the simulation result. The fifth part is the conclusion and summarizes the results and future research prospects.

2. Construction and Optimization of the Investment Effectiveness Evaluation Index System

2.1. Establishment of the Evaluation Indicator System. The evaluation of the effectiveness of power grid investment primarily considers the multiple economic, environmental, social, and technical benefits. Literature searches and investigation reveal many indicators that evaluate the effectiveness of power grid investment and highlight its importance in the construction of an indicator system. The

Table 1: Power grid investment effectiveness evaluation index system.

First-level indicator	Second-level indicator	Third-level indicators	Index number
Grid investment effectiveness evaluation	Economic benefits	New investment included in fixed assets ratio	A_1
		Unit new asset pricing cost	A_2
		Asset utilization efficiency	A_3
		Weighted average return on capital	A_4
		Unit asset pricing depreciation fee	A_5
		Life-cycle return on investment	A_6
	Social benefits	Investment power growth ratio	B_1
		User reliability	B_2
		System reliability	B_3
	Environmental benefits	CO_2 emission reduction rate	C_1
		Proportion of clean energy power generation capacity	C_2
		Pollutant emission reduction rate	C_3
	Technical benefits	Qualified rate of voltage	D_1
		Node voltage ratio	D_2
		Average transfer ratio of transmission line	D_3
		Network coordination	D_4

following basic principles of indicator system construction must be met:

(1) Systematic: there must be a certain logical relationship between the indicators, which must reflect the principal characteristics and status of the distribution network from different aspects and the characteristics of the internal relationship of safety-reliability-economics. The construction of the indicator system must have a hierarchical structure, from top to bottom and from the macro- to micro-level layers to form an organic unity.

(2) Scientific: the design and selection of each index must have a scientific basis, which can objectively, representatively, and truly reflect the engineering characteristics and social economic benefits of the distribution network. The design should not be excessively small or simple to avoid the omission of object information and inaccurate evaluation results.

(3) Comparability: the calculation measures and calculation methods for index selection must be consistent and unified. Each index is simple and clear and has a strong microcosm.

(4) Measurability: the selection should consider the convenience of the indicator data for collection and measurement; the ability to be quantitatively processed is important to facilitate mathematical calculation and analysis.

(5) Independence: the physical and social attributes of the objects reflected by the indicators should be independent of each other and should not overlap.

On this basis, a set of index systems applicable to the evaluation of grid investment effectiveness is constructed and divided into three levels. The first level is the target level, which is the grid investment effectiveness evaluation. The second level is the criterion level, including the four major technical, economic, social, and environmental index attributes. The third-level (index layer) contains 16 indicators (Table 1).

2.2. Index Optimization Model. Effective research is necessary for the grid investment evaluation index system focused on the construction of an effectiveness evaluation model. Before constructing the model, the index system is optimized and screened to improve scientificity and rationality. Therefore, the effectiveness evaluation model of power grid investment is divided into two parts. The first part and second part are the model index optimization and the effectiveness evaluation, respectively. Index optimization modeling primarily uses the improved ISM-DEA model and the control variable method to optimize the index. The effectiveness analysis of the optimized index system was performed based on the utility function to measure the investment effectiveness results of each evaluation object.

2.2.1. Standardization of Indicators. Considering the complexity of the grid investment effectiveness evaluation indicator data, there are many different types of indicators; the original data of the indicators must be processed first and can generally be classified into three categories.

(1) Positive Index. Positive indicators are converted into standard values by setting indicator thresholds, including upper ($x^{\max}$) and lower ($x^{\min}$) limits.

The standardization formula of the indicator is as follows:

$$y = \frac{x - x^{\min}}{x^{\max} - x^{\min}} \times 0.3 + 0.7. \quad (1)$$

(2) Reverse Indicator. The reverse indicator sets the indicator threshold, including the upper ($x^{\max}$) and lower ($x^{\min}$) limits. The lower limit of the forward indicator generally refers to the optimal value, and that of the reverse indicator is converted into the standard value.

The standardization formula of the indicator is as follows:

$$y = \frac{x^{\max} - x}{x^{\max} - x^{\min}} \times 0.3 + 0.7. \quad (2)$$

(3) Moderate Index. The normalization of the moderate index is based on the inverse index processing function; the first step is to set an optimal moderate valuex^{mid} and to convert the moderate index into the reverse index according to the following formula:

$$x' = \left| x - x^{\text{mid}} \right|. \quad (3)$$

The index is then processed as an inverse indicator and converted to a standard value.

2.2.2. ISM Model. The interpreted structure model method is a specialized research method used to analyze the related structure of complex elements in educational technology research. The function uses the known random relationship between system elements to reveal the internal structure of the system. This study improves the structure interpretation model. The basic idea is based on the relational link diagram generated by the reachability matrix, selecting output and input indicators from the top and bottom, respectively. The workflow for explaining the structural model can be divided into the decision-making and calculation processing phases [14].

Step 1: enter the original data to form an $n \times m$ rank data matrix A_0:

$$A_0 = \begin{bmatrix} x_{11} & x_{21} & x_{31} & \cdots & x_{n1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{n2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{1m} & x_{2m} & \cdots & \cdots & x_{nm} \end{bmatrix}. \quad (4)$$

Step 2: process A_0 and calculate the correlation coefficient between n indicators. We take the x_1 and x_2 indicators as examples and calculate the correlation coefficient formula as follows:

$$r_{12} = \frac{\sum x_{1j}x_{2j} - \left(\sum x_{1j} \sum x_{2j}/m\right)}{\sqrt{\left(\sum x_2^{1j} - \left(\sum x_{1j}/m\right)^2\right)\left(\sum x_2^{2j} - \left(\sum x_{2j}/m\right)^2\right)}}. \quad (5)$$

Obtain $n \times n$ coefficient matrix A_1.

$$A_1 = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & \cdots & \cdots & r_{nm} \end{bmatrix}. \quad (6)$$

Step 3: process matrix A_1 in the following manner: if $r_{ij} > \alpha$, let $r_{ij} = 1$; otherwise, let $r_{ij} = 0$; the value of α can be defined (generally 0.85) to obtain the $n \times n$ order adjacency matrix A_2:

$$A_2 = \begin{bmatrix} 1 & 1 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (7)$$

Step 4: process matrix A_2; the $n \times n$ reachable matrix A_3 can be obtained by processing. A_2 and A_3 satisfy the following rules:

$$(A_2 + I)^{k-1} \neq (A_2 + I)^k = (A_2 + I)^{k+1} = A_3. \quad (8)$$

Step 5: division and level division. First, the matrix A_3 is processed to find the reachable set P, antecedent set Q, and intersection S:

Solve P: find the column corresponding to element 1 in each row
Solve Q: find the row with element 1 in each column
Solve S: apply the function that comes with MATLAB

The reachable set P, antecedent set Q, and intersection S are obtained through processing.

Step 6: perform processing based on the PQS set obtained in step 5, and perform level division. Through processing, several levels and indicators contained in each level class can be obtained. According to this classification, an indicator level map can be drawn.

Step 7: select the index from the highest level and select the previous index as the output-based index; select the index from the lowest level and select the previous index as the input-based index.

Step 8: output the result and driving indicator groups.

Figure 1 shows the specific process.

2.2.3. Improving the Data Envelope Analysis. Data envelopment analysis (DEA) is a method based on the concept of relative efficiency that evaluates the effectiveness of multi-input and multi-output decision units with the same type [12]. The basic idea is to treat an economic system or the production process (one unit) as an entity within a certain range of possibilities by assigning a number of production factors and output of a certain number of "product" activities as entities (units) called decision-making units (DMU); the group of numerous DMU is evaluated by analysis of the input or output ratios, with the weight of each DMU input or output indicator to evaluate for the variable operation and determine an efficient production frontier. The effectiveness of each DMU is determined according to the distance between each DMU and the effective production front. The projection method is used to signal the

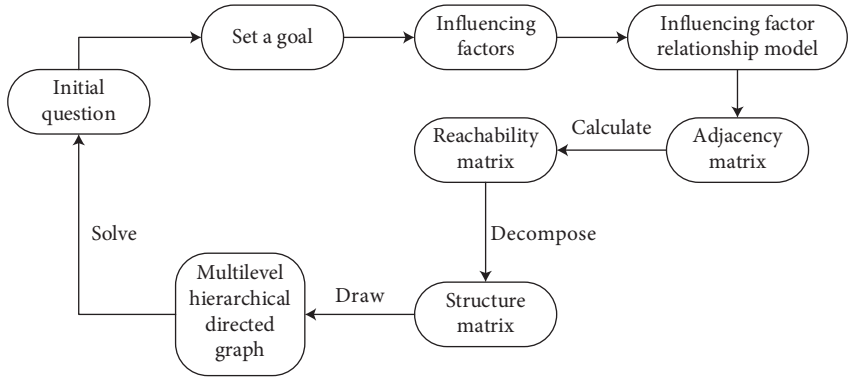


Figure 1: Structural model flow diagram.

reasons for non-DEA effective or weak DEA effective DMU and the direction and degree of improvement.

The data envelope evaluation model evaluates the investment efficiency of power grid companies and has been improved to consider the continuity of the development and construction of power grid companies; thus, the power grid construction effect at the end of the previous time period should translate to the next time period. The initial conditions for the start and the management efficiency evaluation model established in this study are dynamic and defined as the D-dynamic data envelope model (D-DEA). The benefits provided by the construction investment of one attribute can become the construction investment of another attribute. The objective function of the data envelope evaluation model constructed in this study is as follows:

$$\min \frac{\sum_{t=1}^{T} W^t \left(\sum_{k=1}^{K} \omega^k \left(1 - (1/m_k) \sum_{i=1}^{m_k} \left(s_{\text{iok}}^{t-} / x_{\text{iok}}^t\right)\right)\right)}{\sum_{t=1}^{T} W^t \left(\sum_{k=1}^{K} \omega^k \left(1 + (1/r_k) \sum_{i=1}^{r_k} \left(s_{\text{rok}}^{t+} / y_{\text{rok}}^t\right)\right)\right)}. \quad (9)$$

Here, x_{ijk}^t represents the value of the input index i under the attribute k of the decision unit j in t years. y_{rjk}^t represents the value of the output index i under the attribute k of the decision unit j in t years. s_{ijk}^{t-} represents the relaxation variable of the input index i under the attribute k of the decision unit j in t years. s_{rjk}^{t+} represents the relaxation variable of the output index i under the attribute k of the decision unit j in t years.

The constraints are as follows:

$$s.t. \begin{cases} \sum_{j=1}^{n} \lambda_{jk}^t x_{ijk}^t + s_{ik}^{t-} = x_{iok}^t, o = 1, \ldots, n; k = 1, \ldots, K; i = 1, \ldots, m; t = 1, \ldots, T, \\ \sum_{j=1}^{n} \lambda_{jk}^t y_{rjk}^t - s_{rk}^{t-} = y_{rok}^t, o = 1, \ldots, n; k = 1, \ldots, K; r = 1, \ldots, s; t = 1, \ldots, T, \\ \sum_{j=1}^{n} \lambda_{jk}^t z_{jk_i}^{t,t+1} = \sum_{j=1}^{n} \lambda_{jk}^{t+1} z_{jk_i}^{t,t+1}, \\ \sum_{j=1}^{n} \lambda_{jk}^t z_{j(k,h)_i}^t = \sum_{j=1}^{n} \lambda_{jh}^{t+1} z_{j(k,h)}^t, \\ \sum_{j=1}^{n} \lambda_j^k = 1, \ k = 1, \ldots, K, \\ \lambda_{jk}^t \geq 0, s_i^{k-} \geq 0, s_r^{k+} \geq 0, i = 1, \ldots, m; r = 1, \ldots, s; k = 1, \ldots, K. \end{cases} \quad (10)$$

Here, $z_{j(k,h)_i}^t$ represents the correlation value of the decision unit j from the k attribute to the h attribute in the t period. $z_{jk_i}^{t,t+1}$ represents the influence value of the attribute k of the decision-making unit j from year t to year $t+1$.

Through optimization and adjustment of the model, the setting parameters change as follows:

$$\sum_{j=1}^{n} \lambda_{jk}^t z_{j(k,h)_i}^t = z_{o(k,h)_i}^t, \quad (11)$$

$$\sum_{j=1}^{n} \lambda_{jh}^t z_{j(k,h)_i}^t = z_{o(k,h)_i}^t. \quad (12)$$

From this, the efficiency value of each decision unit for each year can be obtained as follows:

$$\theta_o^t = \frac{\sum_{k=1}^{K} \omega^k \left(1 - (1/m_k) \sum_{i=1}^{m_k} \left(s_{iok}^{t-} / x_{iok}^t\right)\right)}{\sum_{k=1}^{K} \omega^k \left(1 + (1/r_k) \sum_{i=1}^{r_k} \left(s_{rok}^{t+} / y_{rok}^t\right)\right)}. \quad (13)$$

The efficiency value of each attribute of the decision unit DMU is calculated as follows:

$$\theta_{ok} = \frac{\sum_{t=1}^{T} W^t \left(1 - (1/m_k) \sum_{i=1}^{m_k} \left(s_{iok}^{t-} / x_{iok}^t\right)\right)}{\sum_{t=1}^{T} W^t \left(1 + (1/r_k) \sum_{i=1}^{r_k} \left(s_{rok}^{t+} / y_{rok}^t\right)\right)}. \quad (14)$$

The numerous indicators imply extensive evaluation when constructing the grid investment effectiveness evaluation index system. In this study, an improved DEA model is used to optimize the index system, thereby improving the efficiency and feasibility of the subsequent evaluation. Based on the principle of the control variable method, DEA is used to calculate the efficiency value of each evaluation object under the premise that different indexes are missing separately. The indexes are screened according to the variation in the efficiency value of each evaluation object. The first N indexes with the greatest impact are reserved according to actual needs that serve as the basis for the follow-up utility evaluation.

Taking the index data of each evaluation object in the year t as an example, x_{ij} indicates the input index value of i in region j, and y_{ij} indicates the output index value of i in region j. Under normal circumstances, the evaluation efficiency value of each evaluation object is θ_{j0}. Assuming that one of the input indicators x_l is eliminated, and the other indicators are unchanged, the evaluation efficiency value of

each evaluation object is θ_{jl}, and the change in the efficiency value of each evaluation object can be obtained as follows:

$$\Delta\theta_{jl} = \frac{\left|\theta_{jl} - \theta_{j0}\right|}{\theta_{j0}} \times 100\%, \tag{15}$$

$$\delta_{x_l} = \frac{1}{j}\sum_{j=1}^{M} \Delta\theta_{jl}. \tag{16}$$

The average efficiency change value δ_{x_l} of the input indicator x_l is sorted in ascending order, the first n ($n < N$) input indicators are eliminated as needed, and the remaining indicators are used as key input indicators for subsequent evaluation.

Assuming that an output index y_k is eliminated and other indicators remain unchanged, the evaluation efficiency value of each evaluation object is θ_{jk}, and the overall deviation rate of the evaluation object efficiency value is calculated as follows:

$$S_{y_k} = \sqrt{\sum_{j=1}^{M}\left(\theta_{jk} - \frac{1}{j}\sum_{j=1}^{M}\theta_{jk}\right)^2}. \tag{17}$$

The overall deviation rate S_{y_k} of the output indicator y_k is sorted in ascending order, the first n ($n < N$) output indicators are eliminated as needed, and the remaining indicators are used as key output indicators for subsequent evaluation.

3. Utility Evaluation Model considering Risk Preference

3.1. Utility Theory. Utility theory is an economic theory that describes the ability of an article to satisfy human desires or the subjective psychological evaluation of the utility of an article to explain the value and its formation process [21]. Preference analysis, a principal component of utility theory, can reflect the risk preference of decision-makers. In this study, utility theory is applied to the evaluation of grid project investment effectiveness, and the corresponding effect function is obtained according to the risk appetite (conservative, aggressive, or neutral) of the investor:

$$U = \alpha + \beta \ln\left[\sigma + \frac{\varepsilon^2}{1 - 2\varepsilon}\right], \quad \varepsilon < 0.5, \tag{18}$$

$$\begin{aligned} U &= \sigma, \\ \varepsilon &= 0.5, \end{aligned} \tag{19}$$

$$U = 1 - \left|\alpha + \beta \ln\left\{1 - \sigma + \frac{1 - \varepsilon^2}{2\varepsilon - 1}\right\}\right|, \quad \varepsilon > 0.5. \tag{20}$$

Here, ε is the conversion coefficient; σ is the normalized value of each index; and U is the utility value. The conversion coefficient is determined according to the risk appetite of the decision-maker. The α and β coefficients are obtained by using linear regression to obtain special utility points and their corresponding σ values. Figure 2 shows the utility function curve. The utility value and index value are standard values.

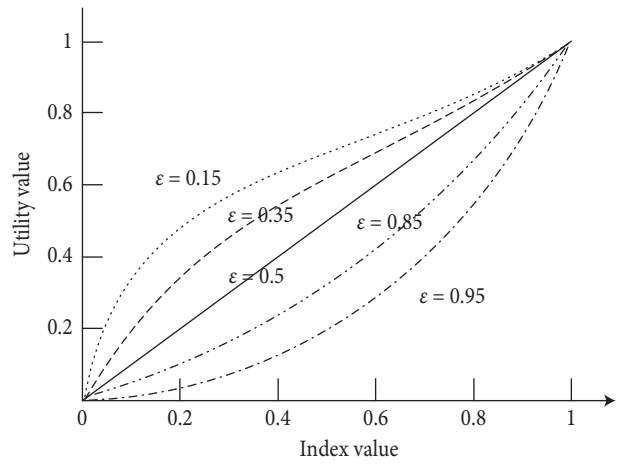


Figure 2: Utility function curve.

(1) When $\varepsilon < 0.5$, the utility function curve is an upward convex curve, indicating that decision-makers are averse to risk, sensitive to loss responses, and conservative

(2) When $\varepsilon = 0.5$, the utility function curve is a straight line, and decision-makers are risk-neutral

(3) When $\varepsilon > 0.5$, the utility function curve is an upward concave curve. Decision-makers like risk, are sensitive to the benefit response, and are aggressive

When making decisions based on the utility function method, the risk attitudes of decision-makers under different ε are different. Smaller and larger ε values indicate more conservative and aggressive risk attitudes, respectively.

Because the goals of grid planning in different regions are different, the same grid investment criteria cannot be used to quantify the evaluation indicators. In the quantification process, both the characteristics of the indicators and the risk tolerance of the planning area must be considered. For example, in economically developed regions with high power grid quality requirements, excessive power outages are not allowed. From a risk perspective, reducing risks can be classified as risk aversion, and an evaluation plan for conservative decision-makers should be adopted. Less-developed areas can withstand the risk of power outages, pay more attention to economic development, and are classified as risky decision-makers. In this case, the evaluation scheme of aggressive decision-makers should be adopted.

3.2. Calculation of the Weight of Each Indicator. This study proposes an improved weight integration method based on multi-objective optimization to resolve the shortcomings of the single evaluation method and the disadvantages of combination weighting. This method takes into account the subjective consciousness of decision-makers and the objective characteristics of the data, minimizes deviations, and optimizes trends as the objective function of the integration

weights. This study selects the reference object comparison judgment method and analytic hierarchy process as the subjective weighting method and selects the coefficient of variation method and the entropy weight method to perform objective observation of the index weights. The specific calculation model [22] and specific steps are as follows:

Step 1: define the objective function.

Setting the type of objective weighting method as q, the type of subjective weighting method as p, and the number of evaluation indicators as n, define the subjective weighting result as W_s and the objective weighting result as W_t:

$$W_s = \left\{w_{sj} | 1 \le s \le p, 1 \le j \le n\right\}, \sum_{j=1}^{n} w_{sj} = 1, w_{sj} \ge 0, \tag{21}$$

$$W_t = \left\{w_{tj} | 1 \le t \le q, 1 \le j \le n\right\}, \sum_{j=1}^{n} w_{tj} = 1, w_{tj} \ge 0. \tag{22}$$

Here, w_{sj} indicates the weight of the index j of the s-th subjective weighting method and w_{tj} indicates the weight of the index j of the t-th subjective weighting method.

The integration weight $W_j = [w_1, w_2, ..., w_n]$ may first be assumed to minimize the deviation of the integration weight from the results of each weighting method; the objective function is

$$\min H\left(w_j\right) = \zeta \sum_{s=1}^{p} \left(w_j - w_{sj}\right)^2 + \psi \sum_{t=1}^{q} \left(w_j - w_{tj}\right)^2. \tag{23}$$

Here, ζ represents the degree of emphasis on subjective weighting results, and ψ represents the degree of emphasis on objective weighting results.

Step 2: calculate the bias coefficient ζ, ψ.

According to the basic idea of the moment estimation theory for each evaluation index, the expected values of the subjective weight and the expected value of the objective weight are

$$E\left(w_{sj}\right) = \frac{\left(\sum_{s=1}^{p} w_{sj}\right)}{p}, 1 \le j \le n, \tag{24}$$

$$E\left(w_{tj}\right) = \frac{\left(\sum_{t=1}^{q} w_{tj}\right)}{q}, 1 \le j \le n. \tag{25}$$

Then, the weighting coefficient ζ_j, ψ_j of each indicator is

$$\zeta_j = \frac{E\left(w_{sj}\right)}{E\left(w_{sj}\right) + E\left(w_{tj}\right)}, \tag{26}$$

$$\psi_j = \frac{E\left(w_{tj}\right)}{E\left(w_{sj}\right) + E\left(w_{tj}\right)}. \tag{27}$$

From this, the bias coefficient ζ, ψ can be calculated as

$$\zeta = \sum_{j=1}^{n} \frac{\zeta_j}{n}, \tag{28}$$

$$\psi = \sum_{j=1}^{n} \frac{\psi_j}{n}. \tag{29}$$

Step 3: solve the minimum combined weight set.

Because of the difference in the index deviation and taking into account the calculation rationality, the degree of emphasis of all indicators is assumed to be the same, and the objective function is obtained as follows:

$$\min H = \zeta \sum_{j=1}^{n} \sum_{s=1}^{p} \left(w_j - w_{sj}\right)^2 + \psi \sum_{j=1}^{n} \sum_{t=1}^{q} \left(w_j - w_{tj}\right)^2. \tag{30}$$

The constraint function is

$$s.t. \begin{cases} \sum_{j=1}^{n} w_j = 1, 0 \le w_j \le 1, \ 1 \le j \le n. \end{cases} \tag{31}$$

Step 4: solve the optimal combination weight set.

The fourth step is to optimize the comparison degree relationship between the indicators. Because the results (especially of subjective weighting) of different weighting theories on the importance of indicators differ, the following trend-optimal objective function is constructed to integrate the weight results and better reflect the importance degree relationship between indicators:

$$\min G = \zeta \sum_{j=1,k=1,k \ne j}^{n} \sum_{s=1}^{p} \left(\frac{w_j}{w_k} - \frac{w_{sj}}{w_k}\right)^2 + \psi \sum_{j=1,k=1,k \ne j}^{n} \sum_{t=1}^{q} \left(\frac{w_j}{w_k} - \frac{w_{tj}}{w_k}\right)^2. \tag{32}$$

The constraint function is

$$s.t. \begin{cases} \sum_{j=1}^{n} w_j = 1, \sum_{k=1}^{n} w_k = 1, 0 \le w_j \le 1, \ 1 \le j \le n, \\ 0 \le w_k \le 1, \ 1 \le k \le n. \end{cases} \tag{33}$$

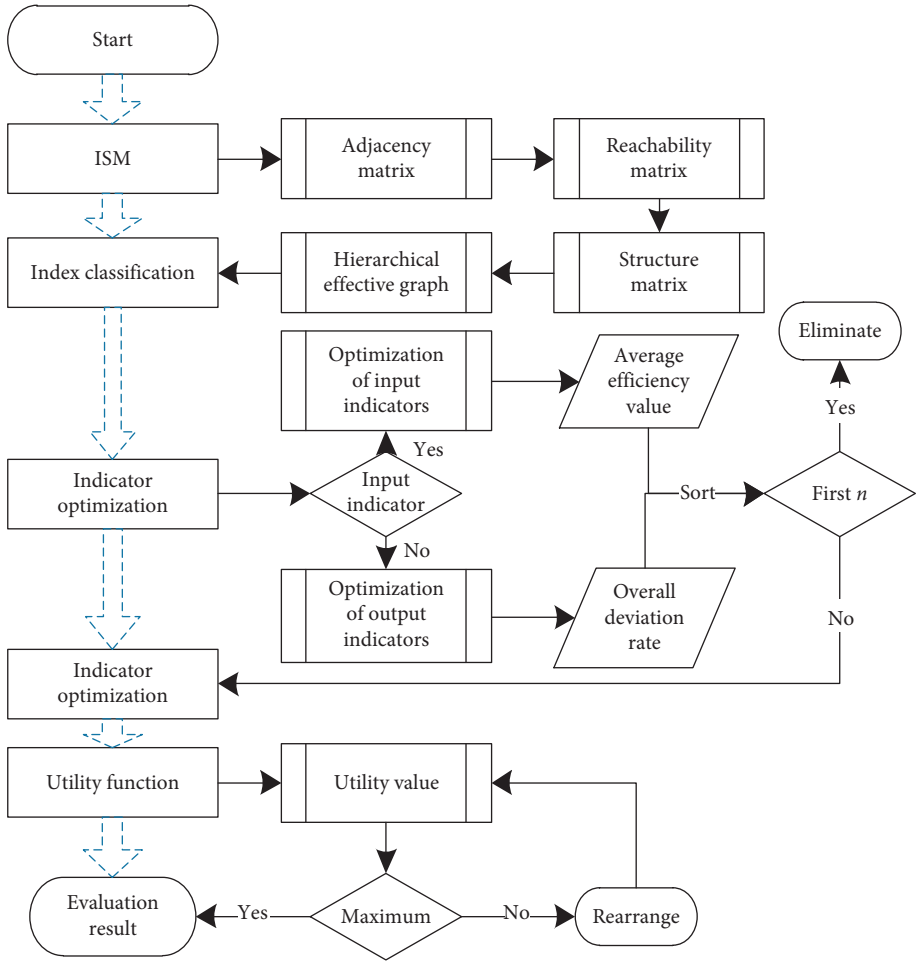


Figure 3: Algorithm flow.

Here, w_k represents the integration weight of the indicator, and w_j represents the weight of different indicators.

Step 5: solve the integrated weight set.

The last step is to combine the two objective functions, consider the minimum deviation and the optimal trend, and assume the same degree of importance to obtain the multi-objective function, as follows:

$$\min Z = \frac{1}{2}\min H + \frac{1}{2}\min G. \tag{34}$$

The weight combination under the multi-objective function is solved, and the weight result is defined as the final comprehensive evaluation weight.

3.3. Evaluation Model of Investment Effectiveness. The total utility value U_T of the investment project i is calculated based on the utility function. The existing method often uses a linear weighting method to obtain it, as shown in

$$U_T(i) = \sum_{j=1}^{n} w_{ij} U_{ij}. \tag{35}$$

The linear weighting method must be independent; however, the indicators in the system interact with each other. Using the linear weighting method creates evaluation result bias. Therefore, the total utility value of the target layer is obtained based on the distance vector merging rule, and the investment ranking of the construction projects is performed accordingly.

The utility value can be reduced to a real-valued function (consisting of the θ variable) to obtain the total utility from θ. The function value of each point in the θ-dimensional space constitutes a utility value surface. On a surface formed in b-dimensional space, there must be a single optimal point E (1, 1, ..., 1) and a single least ideal point D (0, 0, . . ., 0). The utility value corresponding to the ideal point is 1. The utility value corresponding to the least ideal point is 0. The other points are inferior to point E and superior to point D to varying degrees. The utility function value is between 0 and 1. Assume that the utility value surface is smooth and continuous. According to the distance vector merging rule, points on the plane closer to the optimal point have larger utility values, and vice versa. The distance from the point on the utility surface to the optimal point is

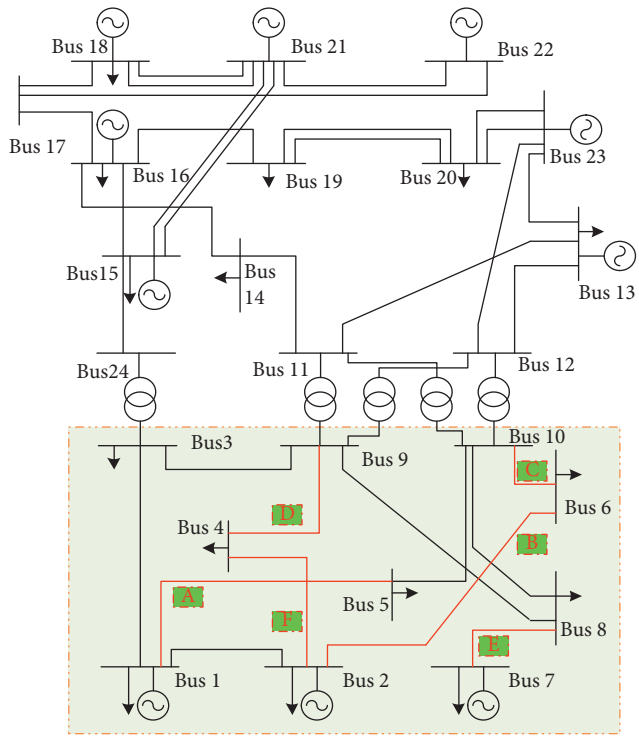


Figure 4: IEEE-RTS79 system diagram.

Table 2: Project operation parameter setting.

	Add point	Add load	Line length	New transformer capacity	Annual investment cost	Annual operating cost
—	—	MW	km	MV·A	Ten thousand	Ten thousand
Project A	Bus 1–bus 5	220	36	200	6520	520
Project B	Bus 2–bus 6	180	75	150	7215	560
Project C	Bus 6–bus 10	136	25	80	3443	500
Project D	Bus 4–bus 9	210	44	60	4040	590
Project E	Bus 7-bus 8	100	26	100	3820	500
Project F	Bus 4–bus 2	150	50	80	4120	550

$$d_1(i) = \sqrt{\sum_{j=1}^{n} \left[w_{ij}\left(1 - U_{ij}\right)\right]^2}. \tag{36}$$

The distance from point E to point D is $d_2(i) = \sqrt{\sum_{j=1}^{n} w_{ij}^2}$. Then, the total utility value U_T of the project i in the utility space can be expressed as

$$U_T(i) = 1 - \frac{d_1(i)}{d_2(i)}. \tag{37}$$

3.4. Algorithm Flow Design. By integrating the above algorithm models, the final power grid investment effectiveness evaluation model is obtained. Figure 3 shows the specific algorithm flow.

4. Example Analysis

4.1. Basic Assumptions. Regional power grid IEEE-RTS79 system (Figure 4) data was collected, and a number of project simulation runs were performed to sort out the simulation operation results data and to verify the feasibility of the power grid investment effectiveness evaluation index and evaluation model constructed in this study. The investment effectiveness of the project is evaluated according to the proposed evaluation model.

Suppose the transformer and transmission line are overloaded because of regional load growth, and the existing system needs to be planned. The marketization rate was 0.32. Considering the complex terrain in this area, the cost of the transmission line project is 1.6 million yuan/km, the cost of the transformer project is 400,000 yuan/(MV·A), the operating life is 16 y, and the construction period of new projects is 2 y. Table 2 shows the basic data of the 6 projects.

The index value was calculated according to the index type. The qualitative index reference [20] uses the trapezoidal fuzzy number method for quantification. The part of the quantitative index that involves future risk estimation and profit estimation is analyzed by a Monte Carlo simulation. Finally, we perform dimensionless processing (Table 3).

Table 3: Dimensionless results of indicator data.

	Project A	Project B	Project C	Project D	Project E	Project F
A_1	0.500	0.000	1.000	0.773	0.636	0.409
A_2	0.333	0.667	1.000	0.333	0.667	0.000
A_3	0.700	0.000	0.800	0.400	1.000	0.400
A_4	0.000	0.889	0.444	0.667	1.000	0.444
A_5	0.600	1.000	0.400	0.200	0.000	0.800
A_6	1.000	0.676	0.265	0.000	0.794	0.588
B_1	1.000	0.143	0.000	0.571	0.429	0.071
B_2	0.000	0.750	1.000	0.875	0.500	0.125
B_3	1.000	0.500	0.000	0.500	0.000	1.000
C_1	0.000	0.222	0.667	1.000	0.444	0.333
C_2	0.167	1.000	0.500	0.667	0.667	0.000
C_3	0.400	0.400	1.000	0.600	0.200	0.000
D_1	0.000	0.833	0.417	0.917	1.000	0.833
D_2	0.778	0.250	0.000	1.000	0.278	0.639
D_3	1.000	0.000	0.529	0.706	0.412	0.353
D_4	0.667	0.364	1.000	0.303	0.424	0.000

4.2. Index Optimization. The system indicators are classified into input-type and output-type. According to the ISM model, an element relationship table is first constructed according to the influence relationship between the indicators; an adjacency matrix A_1 can be established according to the element relationship table. Matrix operations are performed on the adjacency matrix A_1 to find the reachable matrix M_1(Figure 5).

The reachability matrix is decomposed to find the intersection $R(S_i) \cap Q(S_i)$ of the reachability set $R(S_i)$ and the antecedent set, $Q(S_i)$. The reachability matrix is then decomposed according to the $R(S_i) \cap Q(S_i) = R(S_i)$ condition, and the qualified index elements S_i are extracted and defined as the top-level index. Next, the remaining index elements are re-decomposed: the top-level index is removed, and the qualified index elements are extracted and defined as second-level indicators. The same is true until all indicators are extracted. A schematic diagram of the relationship between the index elements can be constructed according to the extraction process and results (Figure 6).

According to the schematic diagram of the relationship between indicator elements, 16 indicators are distributed at both ends of the hierarchy. According to the different driving effects, there are two key indicator categories: input and output indicators. According to the improved D-DEA algorithm, control variable analysis was performed on 16 indicators; calculation results are shown in Table 4 (gray bottom indicators are input indicators; others are output indicators).

The average efficiency change value of the three input-type indicators and the overall deviation rate of the three output-type indicators are the smallest (Table 4). The impact of these six indicators on the subsequent evaluation is very small. The index is screened, and the optimized index system is shown in Table 5.

4.3. Utility Evaluation

4.3.1. Calculation of Indicator Weight. For the calculation of the index weights, a comprehensive weighting method is used and combined with the design of the weighting model. Tables 6 and 7 show the specific weight comparisons and final results.

4.3.2. Comprehensive Utility Evaluation. Based on the conversion coefficient setting differences, the overall utility of the project is evaluated and ranked from risk preference and risk avoidance perspectives. The risk preference conversion and risk avoidance conversion coefficients were set to 0.8 and 0.4, respectively.

(1) Risk Appetite. For risk appraisers, the conversion factor and parameters α and β are set to 0.8 and 0.5, respectively. The dimensionless values of each indicator are substituted into the utility function to obtain the utility value of the corresponding indicator for each item (Table 8).

The distance vector merging rule is used to obtain the utility values of various project category layers (Figure 7). The utility advantages of each project differ in different dimensions. For example, in the economic benefit dimension, the evaluation values of projects B and D are greater than 0.8, better than those of the other projects. Thus, these projects have superior economic utility. In the social benefit dimension, only projects E and F had evaluation values greater than 0.8, whereas other projects had unsatisfactory evaluation results. Only projects A and F have environmental benefit evaluation results greater than 0.8, whereas other projects are expected to have poor results. The evaluation results of the technical benefits of each project are approximately 0.7.

Similarly, the distance vector merging rule is used to obtain the comprehensive utility value of various projects (Figure 8). The comprehensive utility evaluation results are ranked in the following order: Project B > Project F > Project A > Project D> Project E> Project C.

(2) Risk Aversion. For risk-avoidance evaluators, the conversion factor and the parameters α and β are set to 0.4 and 0.5, respectively. The non-dimensionalized values of each

	A_1	A_2	A_3	A_4	A_5	A_6	B_1	B_2	B_3	C_1	C_2	C_3	D_1	D_2	D_3	D_4
A_1	1	1	1	0	0	0	1	0	0	0	1	0	0	0	0	1
A_2	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1
A_3	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
A_4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
A_5	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
A_6	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
B_1	0	0	0	0	0	0	1	0	0	1	1	1	1	1	0	1
B_2	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	1
B_3	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	1
C_1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
C_2	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
C_3	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
D_1	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1
D_2	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0
D_3	1	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0
D_4	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1

Figure 5: Reachable matrix.

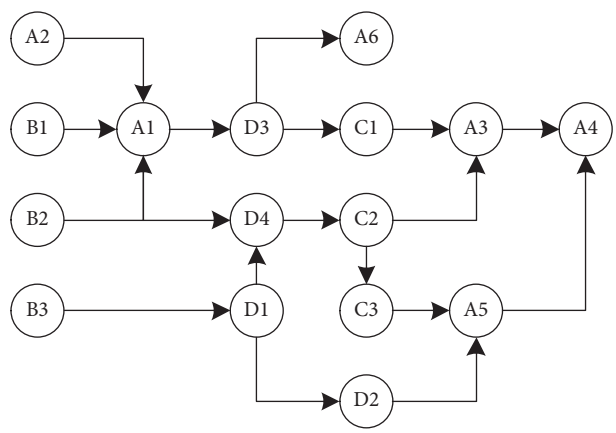


Figure 6: Indicator relationship diagram.

Table 4: Evaluation results of the improved D-DEA algorithm.

Index/project	A	B	C	D	E	F	Average efficiency change (%)	Overall deviation rate (%)
A_1	0.917	1	0.944	1	0.911	0.964	3.90	8.73
A_2	1	0.937	0.92	0.927	0.93	0.963	2.49	6.77
A_3	1	0.94	0.974	0.955	0.976	0.905	4.88	7.41
A_4	1	1	0.972	0.936	0.976	1	4.65	5.67
A_5	0.935	0.97	1	0.934	0.943	0.91	6.64	7.08
A_6	0.967	0.923	0.956	1	0.961	0.959	3.19	5.50
B_1	1	0.942	0.948	0.907	0.914	0.953	3.30	7.42
B_2	0.908	0.942	0.932	1	0.961	0.903	4.68	8.06
B_3	1	1	0.921	0.946	0.931	1	2.72	8.44
C_1	1	0.935	0.906	0.965	0.953	0.943	2.35	7.03
C_2	0.93	0.908	0.92	0.904	1	0.934	6.14	7.83
C_3	0.95	0.95	0.972	0.965	1	0.953	5.32	4.32
D_1	1	0.929	0.923	0.945	1	0.992	3.01	8.15
D_2	0.957	1	0.927	0.968	0.939	0.914	4.76	6.94
D_3	1	0.952	0.913	1	0.929	0.902	2.41	9.55
D_4	0.953	1	0.927	0.924	0.904	0.928	4.98	7.50
Original value	1	0.932	0.907	1	0.912	1	—	—

Table 5: Index system optimization results.

Second layer	Third layer	A	B	C	D	E	F
Economic benefits	New investment included in fixed assets ratio	0.500	0.000	1.000	0.773	0.636	0.409
	Asset utilization efficiency	0.700	0.000	0.800	0.400	1.000	0.400
	Unit asset pricing depreciation fee	0.600	1.000	0.400	0.200	0.000	0.800
Social benefit	Investment power growth ratio	1.000	0.143	0.000	0.571	0.429	0.071
	User reliability	0.000	0.750	1.000	0.875	0.500	0.125
Environmental benefits	CO_2 emission reduction rate	0.000	0.222	0.667	1.000	0.444	0.333
	Proportion of clean energy power generation capacity	0.167	1.000	0.500	0.667	0.667	0.000
Technical benefits	Qualified rate of voltage	0.000	0.833	0.417	0.917	1.000	0.833
	Node voltage ratio	0.778	0.250	0.000	1.000	0.278	0.639
	Network coordination	0.667	0.364	1.000	0.303	0.424	0.000

Table 6: Comparison result of three-level index weights.

	Only reference comparison method	AHP	Coefficient of variation method	Entropy weight method	Comprehensive weight
A_1	0.412	0.512	0.287	0.254	0.321
A_3	0.425	0.321	0.612	0.595	0.538
A_5	0.163	0.167	0.101	0.151	0.141
B_1	0.652	0.576	0.333	0.376	0.451
B_2	0.348	0.424	0.667	0.624	0.549
C_1	0.432	0.641	0.772	0.661	0.672
C_2	0.568	0.359	0.228	0.339	0.328
D_1	0.428	0.418	0.287	0.332	0.336
D_2	0.331	0.411	0.256	0.279	0.298
D_4	0.241	0.171	0.457	0.389	0.366

Table 7: Indicator weight results.

Secondary	Weight of this layer	Third	Weight of this layer	Comprehensive weight
Economic benefits	0.46	New investment included in fixed assets ratio	0.321	0.148
		Asset utilization efficiency	0.538	0.247
		Unit asset pricing depreciation fee	0.141	0.065
Social benefit	0.201	Investment power growth ratio	0.451	0.091
		User reliability	0.549	0.110
Environmental benefits	0.201	CO_2 emission reduction rate	0.672	0.135
		Proportion of clean energy power generation capacity	0.328	0.066
Technical benefits	0.139	Qualified rate of voltage	0.336	0.047
		Node voltage ratio	0.298	0.041
		Network coordination	0.366	0.051

Table 8: Utility value of the corresponding indicator of the investment project.

Index\project	Project A	Project B	Project C	Project D	Project E	Project F
A_1	0.548	0.735	0.245	0.405	0.481	0.587
A_3	0.447	0.735	0.388	0.591	0.245	0.591
A_5	0.500	0.245	0.591	0.668	0.735	0.388
B_1	0.245	0.688	0.735	0.514	0.579	0.712
B_2	0.735	0.419	0.245	0.339	0.548	0.694
C_1	0.735	0.660	0.466	0.245	0.572	0.618
C_2	0.680	0.245	0.548	0.466	0.466	0.735
D_1	0.735	0.367	0.584	0.310	0.245	0.367
D_2	0.402	0.650	0.735	0.245	0.640	0.480
D_4	0.466	0.606	0.245	0.630	0.581	0.735

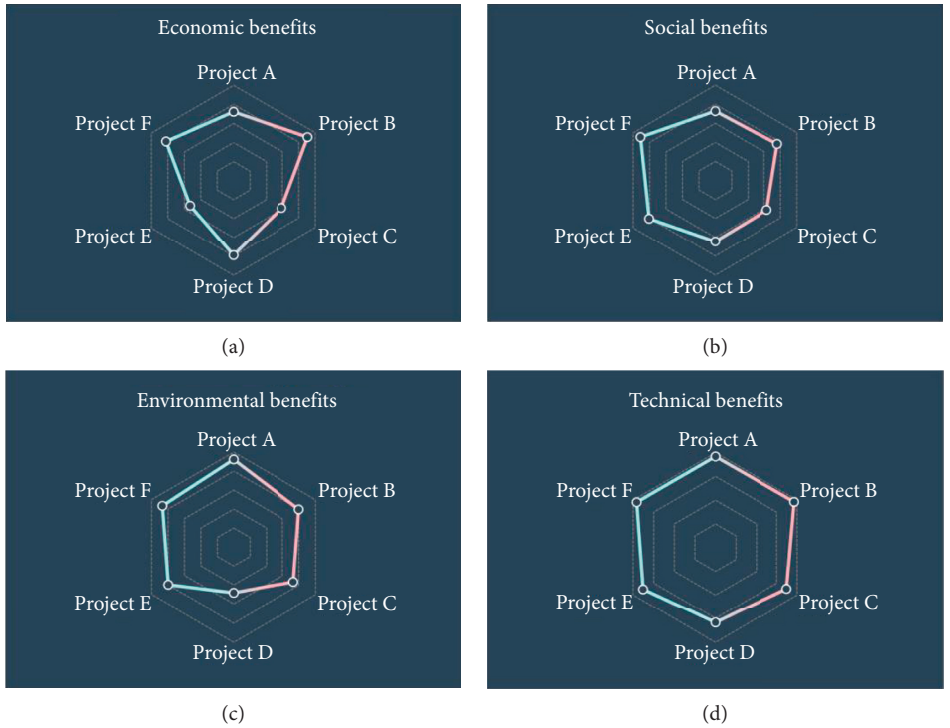


Figure 7: Results of project utility evaluation in each dimension.

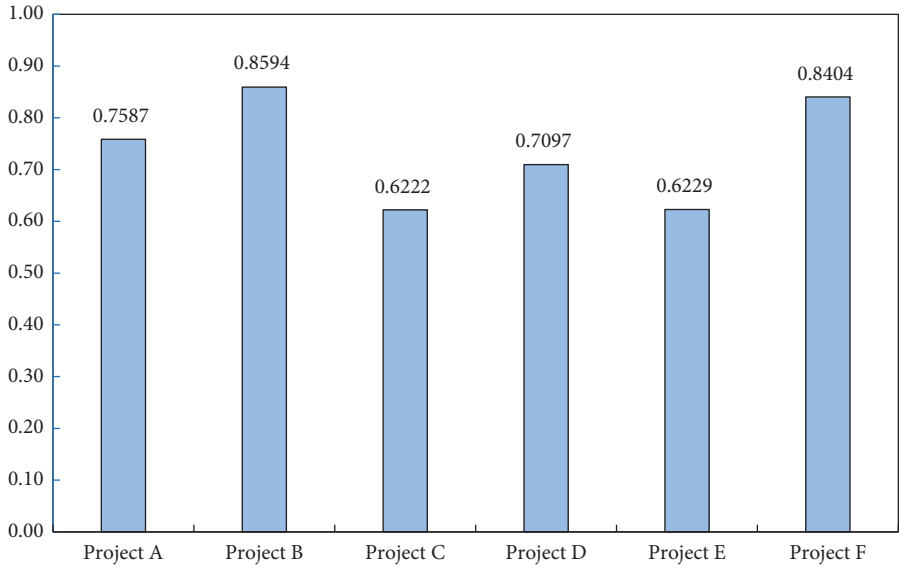


Figure 8: Comprehensive utility value of each project.

indicator are substituted into the utility function to obtain the utility value of the corresponding indicator for each project (Table 9).

The distance vector merging rule is used to obtain the utility value of various project category layers (Figure 9). The utility advantages of each project differ in different dimensions. For example, in the economic benefit dimension, the evaluation value of projects C and E is greater than 0.9, which is significantly better than that of the other projects. Thus, these projects have superior economic utility. In the social benefit dimension, only project D had an evaluation value greater than 0.9, whereas other projects had unsatisfactory evaluation results. Only project D had environmental benefit evaluation results greater than 0.8, whereas other projects are expected to have poor results. The evaluation results of the technical benefits of each project are approximately 0.85.

The distance vector merging rule is used to obtain the comprehensive utility value of various projects (Figure 10). The results of the comprehensive utility evaluation are

Table 9: Utility value of the corresponding indicator of the investment project.

Index\project	Project A	Project B	Project C	Project D	Project E	Project F
A_1	0.631	0.388	0.794	0.726	0.681	0.595
A_3	0.703	0.388	0.735	0.591	0.794	0.591
A_5	0.668	0.794	0.591	0.500	0.388	0.735
B_1	0.794	0.471	0.388	0.658	0.603	0.431
B_2	0.388	0.719	0.794	0.758	0.631	0.461
C_1	0.388	0.511	0.691	0.794	0.609	0.563
C_2	0.483	0.794	0.631	0.691	0.691	0.388
D_1	0.388	0.745	0.598	0.770	0.794	0.745
D_2	0.728	0.524	0.388	0.794	0.537	0.682
D_4	0.691	0.576	0.794	0.549	0.601	0.388

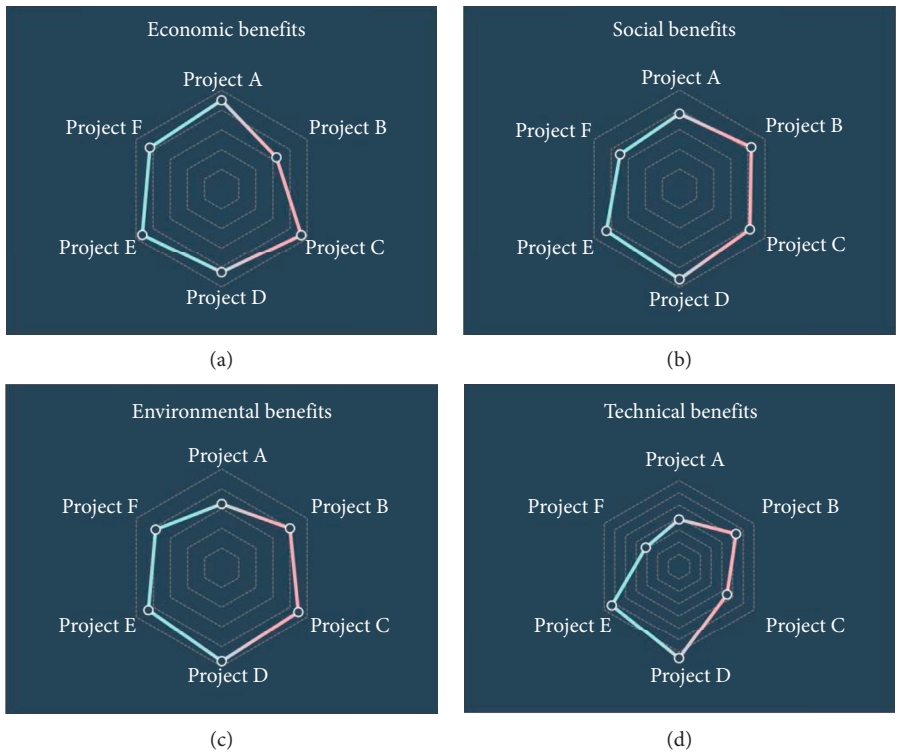


Figure 9: Results of project utility evaluation in each dimension.

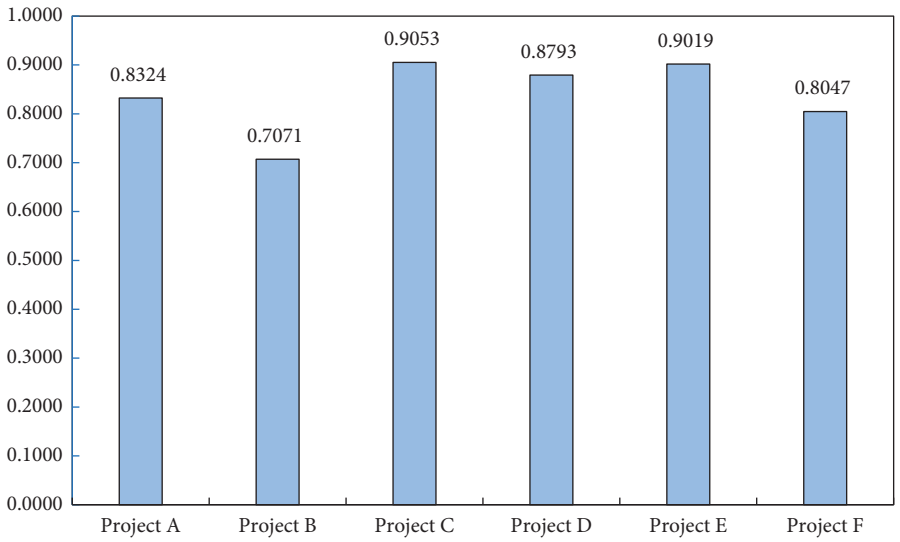


Figure 10: Comprehensive utility value of each project.

ranked in the following order: Project C > Project E > Project D > Project A > Project F > Project B.

The resulting rankings of the comprehensive utility evaluation of risk appetite and risk avoidance are opposite, partially verifying that the constructed utility evaluation model can provide different evaluation results for different risk appraisers. Therefore, the risk appetite of the decision-maker can be considered when making a decision, and the decision-maker can be guided towards a rational decision.

5. Conclusion

This study optimized the evaluation system and resolved deficiencies with the previous power grid investment evaluation index by constructing the grid investment effectiveness evaluation index system and the utility evaluation model based on risk preference. Index redundancy is high but changes the past grid investment evaluation results of a single problem. Taking a regional power grid IEEE-RTS79 system as an example, the data results of six power grid investment projects are selected for analysis, and the validity of the constructed model is verified. The conclusions of this study are as follows:

(1) When constructing the grid investment effectiveness evaluation index system, there is a high correlation between certain indicators that makes the evaluation results unreadable. ISM-DEA modeling reduces this risk and improves the evaluation effectiveness.

(2) Compared with the results of traditional indicator weights, the multi-objective comprehensive weighting method proposed in this study not only considers the characteristics of subjective and objective weights but also considers the limitations of various methods, with minimum deviation and the optimal trend as the optimization goals. The specified index weights better agree with the requirements of evaluation rationality.

(3) The example results show that the utility function model of the risk preferences of the decision-makers is adopted by adjusting the risk preference coefficient; two more extreme coefficients are set in this study to achieve two different evaluation results, indicating that the evaluation result of the effectiveness of power grid investment is not fixed but rather affected by future risk prediction.

The effectiveness of grid investment is not limited to the research scope of this article but is affected by factors such as future changes in transmission, distribution price policies, and fluctuations in power market supply and demand. More in-depth exploration of the effectiveness of grid investment is the key to improve the scientific rationality of grid investment and to guide future grid investment work. Additionally, the popularity of electric vehicle charging piles will lead to distributed power access and will become an important factor affecting the optimization of grid investment. The evaluation effectiveness of the grid investment constructed in this study is also a very important research direction in the future for analyzing the charge and discharge uncertainty of electric vehicles.

Data Availability

The initial data of the dissertation mainly come from the project research. Some data have confidentiality agreements. Except for the data mentioned in the dissertation that can be disclosed, other data cannot be disclosed due to confidentiality issues.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was supported by China Southern Power Grid Technology Project (Grant no. ZBKJXM20180369).

References

[1] State Council of the CPC Central Committee: Several opinions on further deepening the reform of the power system. 2015.

[2] National Development and Reform Commission: Provincial Grid Transmission and Distribution Pricing Measures Development and Reform Prices. 2020.

[3] G. N. Psarros and S. A. Papathanassiou, "Evaluation of battery-renewable hybrid stations in small-isolated systems," *IET Renewable Power Generation*, vol. 14, no. 1, pp. 39–51, 2020.

[4] W. Ma, J. Fan, S. Fang, and G. Liu, "Techno-economic potential evaluation of small-scale grid-onnected renewable power systems in China," *Energy Conversion and Management*, vol. 196, pp. 430–442, 2019.

[5] X. Yu and Z. Tan, "Research on the differential evaluation of smart grid in the improved hall 3D perspective," *Modern Electric Power*, vol. 32, no. 4, pp. 42–48, 2015.

[6] X. Yu, W. Zhang, and Z. Tan, "Evaluation of technical efficiency of distribution network considering environmental differences," *Electric Power Construction*, vol. 37, no. 7, pp. 112–118, 2016.

[7] Q. Jiang, R. Huang, Y. Huang, S. Chen et al., "Application of BP neural network based on genetic algorithm optimization in evaluation of power grid investment risk," *IEEE Access*, vol. 7, pp. 154827–154835, 2019.

[8] X. Han, H. Zhang, and X. Yu, L. Wang, "Economic evaluation of grid-connected micro-grid system with photovoltaic and energy storage under different investment and financing models," *Applied Energy*, vol. 184, pp. 103–118, 2016.

[9] W. Guo and H. Jin, "Research on evaluation of investment benefits of power grid based on decision model analysis," *Automation and Instrumentation*, vol. 5, pp. 175–178, 2019.

[10] X. Yu, Z. Tan, and G. Qu, "Research on differentiated electricity price package based on electricity customer evaluation," *China Electric Power*, vol. 53, no. 2, pp. 9–19, 2020.

[11] X. Yu, *Research on Optimization Model of Electricity Purchase and Sale Transactions of Electricity Sale Companies under the Electricity Market Environment*, North China of Electric Power University, Beijing, China, 2018.

[12] D. Xing, *Research on Grid Investment Evaluation Based on Set Pair Analysis Method under New Power Reform*, North China Electric Power University, Beijing, China, 2019.

[13] Y. He, W. Liu, J. Jiao, and J. Guan, "Evaluation method of benefits and efficiency of grid investment in China: a case study," *Engineering Economist*, vol. 63, no. 1, pp. 66–86, 2018.
[14] Y. Sun, X. Yu, Z. Tan, X. Xu, and Q. Yan, "Efficiency evaluation of operation analysis systems based on dynamic data envelope analysis models from a big data perspective," *Applied Sciences*, vol. 7, no. 6, p. 624, 2017.
[15] X. Yu and Y. Sun, "Trading risk control model of electricity retailers in multi-level power market of China," *Energy Science & Engineering*, vol. 7, no. 6, pp. 2756–2767, 2019.
[16] X. Yu and Y. Geng, "Complementary configuration research of new combined cooling, heating, and power system driven by renewable energy under energy management modes," *Energy Technology*, vol. 7, no. 10, Article ID 1900409, 2019.
[17] M. Negnevitsky, D. H. Nguyen, and M. Piekutowski, "Risk assessment for power system operation planning with high wind power penetration," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1359–1368, 2015.
[18] Z. Liang, H. Chen, X. Wang, S. Chen, and C. Zhang, "Risk-based uncertainty set optimization method for energy management of hybrid AC/DC microgrids with uncertain renewable generation," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1526–1542, 2020.
[19] R. Karki, S. Thapa, and R. B. Billinton, "A simplified risk-based method for short-term wind power commitment," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 3, pp. 498–505, 2012.
[20] G. Cao, W. Gu, P. Li et al., "Operational risk evaluation of active distribution networks considering cyber contingencies," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3849–3861, 2020.
[21] Ma Qian, Z. Wang, X. Pan et al., "Evaluation method of grid investment decision-making based on utility function in the environment of new power reform," *Electric Power Automation Equipment*, vol. 39, no. 12, pp. 198–204, 2019.
[22] Y. He, *Research on Multi-User Assisted Wind Power Consumption Optimization Model Based on Demand Response*, North China Electric Power University, Beijing, China, 2016.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 5702519, 9 pages
https://doi.org/10.1155/2020/5702519

Research Article

Enhanced Unsupervised Graph Embedding via Hierarchical Graph Convolution Network

H. Zhang,[1] **J. J. Zhou**,[1] **and R. Li**[1,2]

[1]*School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China*
[2]*Key Laboratory of Computation Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan 030006, China*

Correspondence should be addressed to H. Zhang; zhanghu@sxu.edu.cn

Received 14 May 2020; Accepted 19 June 2020; Published 26 July 2020

Guest Editor: Shianghau Wu

Graph embedding aims to learn the low-dimensional representation of nodes in the network, which has been paid more and more attention in many graph-based tasks recently. Graph Convolution Network (GCN) is a typical deep semisupervised graph embedding model, which can acquire node representation from the complex network. However, GCN usually needs to use a lot of labeled data and additional expressive features in the graph embedding learning process, so the model cannot be effectively applied to undirected graphs with only network structure information. In this paper, we propose a novel unsupervised graph embedding method via hierarchical graph convolution network (HGCN). Firstly, HGCN builds the initial node embedding and pseudo-labels for the undirected graphs, and then further uses GCNs to learn the node embedding and update labels, finally combines HGCN output representation with the initial embedding to get the graph embedding. Furthermore, we improve the model to match the different undirected networks according to the number of network node label types. Comprehensive experiments demonstrate that our proposed HGCN and HGCN* can significantly enhance the performance of the node classification task.

1. Introduction

Graph embedding aims to map high-dimensional sparse network data to low-dimensional dense real-value vector space, which can adaptively extract features and be applied to the network analysis tasks, such as node classification, link prediction, and dimensional visualization [1, 2]. Generally, graph embedding models mainly include the shallow embedding model and the deep embedding model. The shallow embedding model represented by Deepwalk is directly learned by defining the loss function for network representation, which is usually an unsupervised learning method [3]. This model is suitable for the undirected graph, but it also has the disadvantage of ignoring the feature information in graphs. On the contrary, the deep embedding model represented by the GCNs uses the additional feature and label information of the graph at the same time, which is generally a semi-supervised or supervised learning method. However, the deep embedding model is not more suitable for undirected graphs with only structure information, at the same time many real networks do not include the node features and labels due to privacy protection and annotation difficulty, so the study of deep unsupervised graph embedding method for a real undirected graph is urgent.

To help address these issues, we propose a novel unsupervised graph embedding method with a hierarchical graph convolution network, which does not need the extra features and labeled data. The general process is shown in Figure 1, which includes 3 layers with the initial, update, and output layer. Among them, the input of the model is a graph, the output is the low-dimensional vector representation of nodes that relate to the number of label types, and the detailed structure of the model is partly omitted.

Our contributions are the following: (1) we propose an enhanced hierarchical graph convolution network for the undirected graph with only structure information. (2) According to the number of label types in the different experimental data sets, we introduce max-pooling to

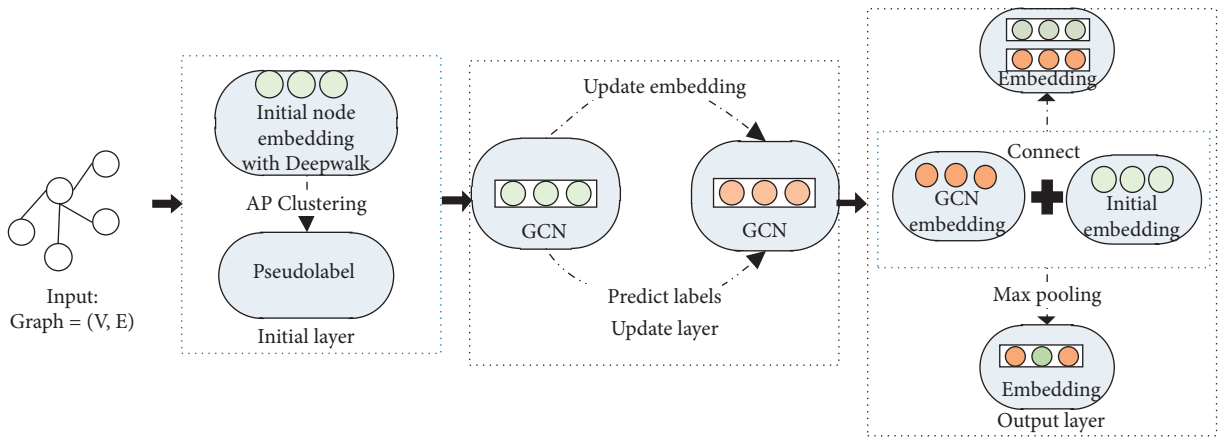


Figure 1: Model workflow. The model includes 3 layers with the initial, update, and output layer. In the output layer, the proposed model is improved according to the number of label types.

improve the proposed model. (3) We conduct the node classification experiments on Wikipedia, American air-traffic, Cora, and Citeseer data sets separately, and our proposed method achieves significant improvements than other state-of-the-art baselines.

2. Related Work

Graph embedding methods have different classification systems. In this section, firstly, we introduce the traditional graph embedding methods. Then, we review the different graph embedding methods in the field of machine learning. In the end, we classify the graph embedding models according to the structure of the embedding methods, and a profound analysis of the disadvantages of the existing graph embedding methods is given. Additionally, we further introduce existing GCN variants for graphs.

2.1. Traditional Graph Embedding Methods. Traditional network representation learning methods are realized by dimension reduction technology [4]. Classic dimension reduction methods include Principal Component Analysis (PCA) [5] and Multi-dimensional Scale (MDS) [6]. Both methods can capture linear structure information, but they cannot acquire a nonlinear structure in input data.

From a linear algebraic perspective, all unsupervised graph embedding methods are generally represented by the various graph matrices, especially the Laplace operator and the adjacent matrix [7]. In terms of computational efficiency, the feature decomposition of the data matrix is expensive.

2.2. Graph Embedding Methods for Machine Learning. In the representation learning for machine learning, the semi-supervised graph embedding method requires a set of features that can distinguish nodes [8]. Typical model mainly uses manual feature extraction to learn node representation for specific professional fields, which has the disadvantages of the inaccuracy and poor robustness. Another method mainly learns node representation such as MMDW [9] by solving optimization problems, which improves the accuracy of the extracted feature. However, the number of estimation parameters is large and the time complexity is high [10].

As for the unsupervised graph embedding method, to balance accuracy and computational efficiency, the feature representation needs to define an objective function independent of downstream tasks, such as Deepwalk and LINE. The above unsupervised graph embedding method only uses network structure information to obtain low-dimensional embedding. Moreover, the graph embedding method also learns from the rich content of node attribute and edge attribute. TADW [11] incorporates text features of vertices into network representation learning under the framework of matrix factorization. CENE [12] integrates text modelling and structure modelling in a general framework by treating the content information as a special kind of node. HSCA [13] is a network embedding method of attribution graph, which simulates homophone, network topology, and node characteristics at the same time. Inspired by pre-training methods, pre-training models based on GCN are proposed. DeepGraphInfoMax [14] relies on maximizing mutual information between patch representations and corresponding high-level summaries of graphs, while both of them derived using established graph convolutional network architectures. The pre-trained GCN model proposed by Hu et al. [15] can capture generic graph structural information that is transferable across tasks.

Compared with the methods of manually extracting features, the representation method by optimizing objective function can obtain more comprehensive features, which is closely related to the downstream predicting task [16]. Therefore, the unsupervised graph embedding method improves the disadvantages of semi-supervised machine learning method, such as difficult extensibility and high training complexity.

2.3. Graph Embedding Methods with Different Structures. According to the structure of the representation methods [17], graph embedding methods are divided into shallow embedding method and deep embedding method. In the

shallow embedding model, inspired by the successful application of skip-gram [18] in natural language processing [19], a series of skip-gram based models are proposed to encode network structures into continuous spatial vector representations in recent years, such as Deepwalk, LINE, and Node2vec.

Deepwalk [20] simulates word sequences to generate node sequences by random walk from each node, which forms a "corpus" based on those node sequences. Furthermore, Deepwalk sets the size of the background window and then imports the "corpus" into the skip-gram model to get the node representation. The first-order similarity of direct connection and the second-order similarity of shared neighbor nodes are optimized respectively, and the two similarities are combined at the output of the nodes in LINE [21]. Node2vec [22] uses two additional parameters to control the direction of the random walk on the generation step of the "corpus" of Deepwalk. Struc2vec [23] defines the nodes that are not structurally adjacent but have the same structural roles. The shallow embedding models usually use unsupervised machine learning methods, so they have the disadvantage of ignoring the node features.

In the deep embedding model, the mainstream methods use the deep learning model to capture the nonlinear relationship between nodes [24]. Typical deep learning models are NE-FLGC, SDNE, GCN, and a series of GCN variants.

NE-FLGC [25] studies the problem of representation learning for network with textual information, which aims to learn low-dimensional vectors for nodes by leveraging network structure and textual information. SDNE [26] automatically captures the local relationship of the nodes by using the unsupervised learning method and takes the second-order neighbor of the nodes as the input to learn the low-dimensional representation of graphs, but the model still does not consider the node features.

GCN [27] is a deep semi-supervised graph embedding model with incorporating the extra feature and labeled data into the graph embedding learning process, which cannot be easily applied to the undirected graph with only structure information. In terms of applicability, the existing GCN variants are mainly studied in two aspects. One variant assumes the graph is attributed, such as GraphSAGE [28], GAT [29], N-GCN [30], and Fast-GCN [31]. GraphSAGE [28] can be used to generate node embeddings for previously unseen nodes or entirely new input graphs, as long as these graphs have the same attribute schema as the training data. GAT [29] uses attention mechanism to address the shortcomings of prior methods based on graph convolutions or their approximations. N-GCN [30] improves the scalability of GCN on the whole graph by setting the size of the convolution kernel. Fast-GCN [31] is a batch training algorithm combining importance sampling, and it can not only make GCN training more efficient but also generalize well for inference. Generally, they take the structure information or the node degree information as the feature and then use the correct label to train GCN to learn the graph embedding. However, the model still needs to incorporate the correct labeled data into the graph embedding processing. Another variant is proposed to solve the label problem, such as M3S and GMNN. M3S [32] uses the correct extra feature and enlarges the labeled data by self-training to learn the undirected graph embedding. GMNN [33] similarly applies the correct extra feature and neighbor nodes to generate pseudo-labels for unsupervised learning. As a consequence, the previous method still cannot consider the feature and the label of the node at the same time.

3. Model

In this section, we describe the proposed method and give a detailed framework for the model. Firstly, we generate the initial node embedding using Deepwalk. Simultaneously, the pseudo-labels are built by the AP clustering algorithm [34]. Then, we build the hierarchical GCN to update the node embedding and labels. Finally, we introduce max-pooling to enhance our model according to the number of label types in the different experimental datasets.

3.1. Initial Node Embedding. The first step of the model is to generate the initial node embedding using Deepwalk. Deepwalk is the shallow embedding model with random walking, and it can better learn the global structure and the context of the node in the undirected graph. Therefore, Deepwalk is used as a pre-training method to generate the initial node embedding in this paper.

Let $G = (V, E)$ be a given undirected graph with vertex set V and edge set E, where $n = |V|$ denotes the number of nodes in the graphs. Our formal definition is general and can apply to any undirected and unweighted network. Firstly, we acquire a specific length random walk sequence $w(v_i)$ for the node v_i, which is shown in the following formula (1):

$$w(v_i) = \left(w(v_i)^1, w(v_i)^2, \ldots, w(v_i)^k\right), \tag{1}$$

where $w(v_i)^n$ is the n^{th} node. And then, we use a language model to learn from these random walk sequences and denote the initial embedding as follows:

$$E_M(w(v_i)) = \text{skip} - \text{gram}(w_i^u), \tag{2}$$

where w_i^u is the word sequence in the language model, and skip – gram is the language model. $E_M(.)$ is denoted as the output function. We use $w_i^u = (w_1, w_2, \ldots w_n)$ to denote the word sequence, where w_i is a specific word in the language model. The detailed processing flow is shown in Figure 2.

3.2. Pseudolabels. The second step of the model is to construct the pseudo-labels and predict labels with GCN. After obtaining the initial node embedding of the undirected graph, we use the clustering method to acquire pseudo-label y_i for the unlabeled node i, where $y_i \in Y$ and Y is a pseudo-label set. The specific processing flow is shown in Figure 3.

And now, we update G to denote the graph, as shown in the following formula (3):

$$G = (V, E, E_M(w(v_i)), Y). \tag{3}$$

In this paper, AP clustering model called near-neighbor propagation clustering algorithm is adopted to generate

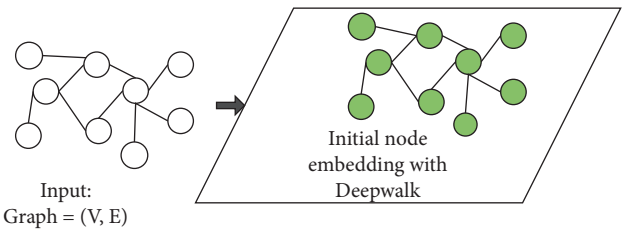


Figure 2: The process of generating the initial node embedding. The input of the model is an undirected graph.

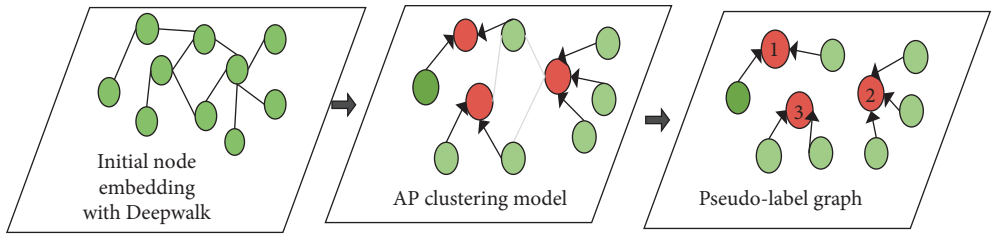


Figure 3: The processes of generating pseudo-labels with the AP clustering method.

pseudo-labels, which is mainly because the method contains the following three advantages. (1) Because the number of label categories is unknown, the model needs to choose a clustering method that does not need to specify the number of clusters. Unlike k-means and k-center algorithms, AP clustering model is not necessary to set the final number of clusters when clustering. (2) The cluster centers of AP algorithm are the actual nodes in the data set, which are the representative of each class, so AP clustering model is not sensitive to the initial negative value of the node embedding. (3) If the sum of squares of errors is used to measure the performance of the algorithms, the sum of squares of errors of AP clustering is lower than that of other methods. This evaluation index shows that AP clustering method is effective.

Based on this, we further use GCN to predict the label for the unlabeled node and define the predicted maximum value as the label y_i', where $y_i' \in Y'$ and Y' is the predication label set. And now, formula (3) is updated as following in formula (4):

$$G = \left(V, E, E_M\left(w\left(v_i\right)\right), Y'\right). \tag{4}$$

3.3. Hierarchical GCN. In this section, we focus on the hierarchical GCN model, which includes two GCN [27] models. At the start, we introduce GCN. Next, we set up hierarchical graph convolution network to propose the HGCN model and the HGCN* model by analyzing the updated undirected graphs. Additionally, we further explain the reason for the improved model.

3.3.1. Graph Convolution Network. Different from the conventional convolutional neural networks (CNN), the convolution operation for a graph is defined as the weighted average of neighbors of one particular node. Mathematically, the graph convolution layer is defined as

$$H^{(l+1)} = \sigma\left(PH^{(l)}W^{(l)}\right), \tag{5}$$

where $P = D^{-1/2}(A + I)D^{-1/2}$ is the normalized Laplacian matrix of the graph G. $H^{(l)}$ is the input of the l^{th} hidden layer of GCN, i.e., the output of the $(l-1)^{\text{th}}$ hidden layer. $W^{(l)}$ is the weight matrix in the l^{th} hidden layer that would be trained. A is the adjacency matrix of the undirected graph. And $\sigma(\cdot)$ is the activation function. For any given node, a GCN layer aggregates the previous layer's embedding of its neighbor with A, followed by linear transformation W and nonlinear activation σ, so as to obtain a contextualized node representation. We denote $F_w(\cdot)$ as L-layer GCNs, parametrized by $\{w_i\}_{i=1}^{L}$. For each given graph G = (V, A) with input features $H^{(0)}$, $F(\cdot)$ can get node representations by:

$$F_w(G) = \sigma\left(P\left(..\sigma\left(PH^{(1)}W^{(1)}\right)..\right)W^{(L)}\right). \tag{6}$$

3.3.2. HGCN and HGCN.* After the initial node embedding section and the pseudo-labels section, we have learned the graph information from formula (4). After predicting labels, we also update the graph embedding with GCN, which is as shown in formula (7):

$$G = \left(V, E, E_M'\left(w_{v_i}\right), Y'\right), \tag{7}$$

where $E_M'(w_{v_i})$ is the updated node embedding by the first GCN.

According to formula (6), $E_M'(w_{v_i})$ is defined as formula (8):

$$E_M'\left(w_{v_i}\right) = \sigma\left(P\left(..\sigma\left(PE_M\left(w\left(v_i\right)\right)W^{(0)}\right)..\right)W^{(2)}\right), \tag{8}$$

where $H^{(0)} = E_M(w(v_i))$ is the input feature, which is the initial node embedding using Deepwalk. In this paper, we

use 3-layer GCNs with 128-dimensional hidden vectors to accomplish the pre-training and adaptation.

And then, we further use the second GCN to learn the undirected graph again. The final embedding $E''_M(w_{v_i})$ is denoted as formula (9):

$$E''_M\left(w_{v_i}\right) = \sigma\left(P\left(..\sigma\left(PE'_M\left(w_{v_i}\right)W^{(0)}\right)..\right)W^{(2)}\right). \tag{9}$$

Moreover, after this second learning process, GCN will lead to too smooth for the embedding, which can affect node classification. Therefore, to relieve smoothing, we combine the secondary learning representation $E''_M(w_{v_i})$ with the initial node embedding $E_M(w_{v_i})$. Finally, we get the ultimate embedding $O(G)$ for the undirected graph as shown in formula (10):

$$O(G) = E''_M\left(w_{v_i}\right) \oplus E_M\left(w_{v_i}\right), \tag{10}$$

where $\oplus$ is a vector splicing operation.

The previous HGCN method can be applied to learn a better embedding, when the graph is an undirected graph with a few numbers of node label types. However, when the undirected graph has a large number of node label types, the proposed model will also decompose labels into multi-dimensional features and add them to the learning processing, which can lead to the feature redundancy. Therefore, motivated by this analysis, we introduce max-pooling to improve the embedding model for the graph with a large number of label types, namely HGCN*. The exact processing flow is shown in Figure 4.

4. Experiments

4.1. Datasets. To verify the effectiveness of the proposed model, we conduct several comparative experiments on four open available graph datasets:

(i) Cora dataset [35] consists of 2,708 machine learning papers classified into one of seven classes and 5,429 links between them.

(ii) Citeseer dataset [35] is a link dataset built with permission from the Citeseer web database. It contains 3,327 publications from six classes and 4,732 links among them.

(iii) Wikipedia dataset [36] consists of 2,405 Wikipedia pages from 17 categories and 17,981 links between them. It is much denser than Citeseer and Cora datasets.

(iv) American air-traffic dataset [23] is from the United States Air Traffic Network. It contains 1190 airports and 13599 relationships between airports and airports, and the airport labels are divided into four groups.

All specific nodes and edges in our experimental datasets are shown in Table 1.

Note that all datasets are used as undirected graphs with only network structure information and the correct labels in the datasets are only used in the node classification verification task in this paper.

4.2. Baselines. For comparison, we adopt six kinds of baselines as follows:

(i) Deepwalk [20] is a typical unsupervised graph embedding method which adopts the skip-gram language model.

(ii) LINE [21] is also a popular unsupervised method which considers the first-order and second-order proximity information.

(iii) Node2vec [22] learns low-dimensional representations for nodes in a graph by optimizing a neighborhood preserving objective, which explores the structure of the network by controlling the p and q parameters, and learns the d dimension feature representation by simulating biased random walks.

(iv) GraRep [37] captures the long-distance nodes and uses the matrix decomposition method to learn the d dimension feature representation of the node.

(v) Hope [38] introduces higher-order similarity and uses the matrix decomposition method to learn the d dimension feature representation of the node.

(vi) SDNE [26] learns the d dimension node representation by capturing the network nonlinear structure using multi-layer nonlinear functions, which is a semi-supervised deep graph embedding model.

Several existing methods include shallow embedding models, matrix decomposition models, and deep embedding models. Note that since some existing work ([27–29]) utilizes extra node features, we do not compare with them directly.

4.3. Node Classification Experiments. To verify the effectiveness, we follow the same settings used in Hamilton et al. (2016) [27], including the following benchmark tasks: (1) using HGCN to classify nodes on American air-traffic, Cora, and Citeseer datasets; (2) using HGCN* to classify nodes on Wikipedia dataset.

4.3.1. Experimental Settings. In this experiment, Micro-F1, Macro-F1, and Accuracy are used as measurements. The model uses the same proportion of training and test data as the existing published papers, i.e., 50% of the training data and 50% of the test data.

To facilitate the comparison, on the Cora and Citeseer datasets, we report the results of one experiment for comparison with the baseline models. However, on the American air-traffic network dataset, we repeat the experiment 5 times using random samples for training and demonstrate the average performance. The experimental results are given in Table 2. "*" denotes the published results of the existing papers [29], and the roughened numbers represent the best results.

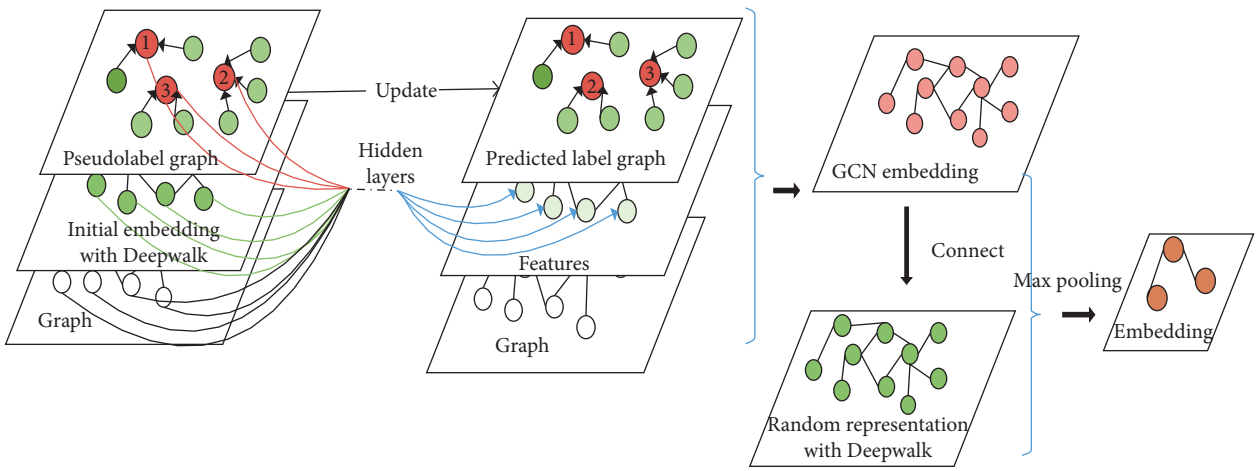


Figure 4: HGCN*. When the undirected graph node has a large number of label types, the model selects features with max pooling in the output layer.

Table 1: Datasets.

Datasets	Cora	Citeseer	Wikipedia	American air-traffic
Nodes	2708	3327	2405	1190
Edges	5429	4732	17981	13599
Classes	7	6	17	4

Table 2: Node classification results with a few label types.

Datasets	American air-traffic			Cora			Citeseer		
Evaluating indicator	Micro-F1	Macro-F1	Accuracy	Micro-F1	Macro-F1	Accuracy	Micro-F1	Macro-F1	Accuracy
Deepwalk	0.523	0.518	0.523	—	—	0.672*	—	—	0.432*
LINE(2nd)	0.504	0.499	0.504	0.704	0.684	0.704	0.444	0.399	0.444
Node2vec	0.493	0.459	0.493	0.771	0.765	0.771	0.567	0.486	0.567
SDNE	**0.580**	0.560	**0.580**	0.660	0.648	0.660	0.400	0.351	0.400
HGCN	0.571	**0.568**	0.571	**0.797**	**0.780**	**0.797**	**0.568**	**0.518**	**0.568**

On the Wikipedia dataset, we use the same measurements to compare with the results of open-source framework OpenNE [39]. The best experimental results are presented in Table 3.

4.3.2. Results and Discussion. As shown in Table 2, it can be seen that our proposed model achieves the best results compared to all shallow baseline models, and the outcomes are very close to the deep baseline model SDNE. On the Cora and Citeseer datasets, the experimental results of our model are higher than all baseline models in all measurements, which show that the model can effectively learn the undirected graphs.

Obviously, on the American air-traffic dataset, the performances of HGCN are close to the results of SDNE. The main reason is that the nodes on American air-traffic dataset are labeled by the network global structure. In other words, the node label of the dataset is more related to their structure identity than the labels of their neighbors. HGCN uses the label information of neighbor nodes to update and learn the labels for the unlabeled data, so the node classification performances of HGCN are significantly affected. To visually display the performances of HGCN, the accuracy is presented in Figure 5.

As illustrated in Figure 5, we compared the accuracy of all baselines including the shallow and the deep embedding model on the three undirected graph datasets. The average performances of the shallow embedding model are higher than SDNE model on the Cora dataset and the Citeseer dataset. And the result of the SDNE is better than the other shallow baselines on the American air-traffic dataset. This suggests that SDNE can learn well the special structure network, which node labels are related according to the network global structure. While the proposed model HGCN has a significant increase on most experimental datasets, which further verifies the applicability of our model on undirected graphs.

As for the improved model according to the number of label types, we verify its effectiveness on the Wikipedia dataset and compare it with the results of the open-source framework OpenNE [39]. The experimental results are shown in Table 3.

According to the results of Table 3, when the HGCN model uses only vector splicing, the performance degrades in Micro-F1. In Macro-F1, the results are better than the other all baseline models except for Deepwalk. That suggests that

Table 3: Node classification results on Wikipedia.

	Baselines	Micro-F1	Macro-F1
Random walk	Deepwalk	0.669*	0.560*
	LINE(2nd)	0.576*	0.387*
	Node2vec	0.651*	0.541*
Matrix decomposition	GraRep	0.633*	0.476*
	Hope	0.601*	0.438*
Deep model	SDNE	0.643*	0.498*
Our model	HGCN	**0.599**	**0.549**
	HGCN*	**0.717**	**0.590**

The result with * is from the open-source OpenNE [39].

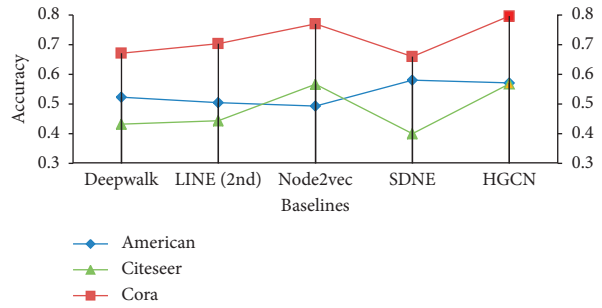


Figure 5: Node classification performances (accuracy) with fewer label types.

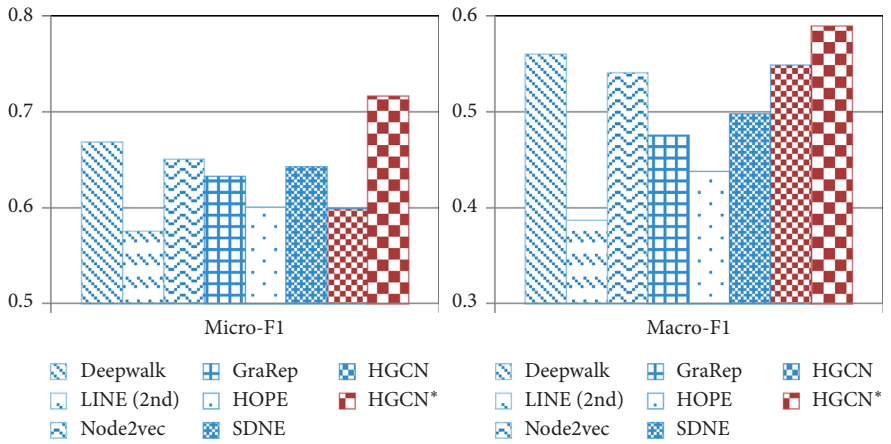


Figure 6: Node classification performances (Micro-F1; Macro-F1) with baseline models on the Wikipedia dataset.

the HGCN model breaks up pseudo-labels into the features, while the Wikipedia dataset leads to features redundant because of more label types.

To help address this problem, the max-pooling layer is introduced to the HGCN model, which can select the features to improve the effect of node classification on the Wikipedia dataset. As can be seen from the experimental results in Table 3, compared to Deepwalk, LINE (2nd), and Node2vec, the Micro-F1 of HGCN* is improved by 4.8%, 14.1%, and 6.6%. And in Macro-F1, the results are enhanced by 3%, 20.3%, and 4.9%, separately. Compared to GraRep, HOPE, and SDNE, the Micro-F1 of HGCN* is improved by 8.4%, 11.6%, and 7.4%, and in the Macro-F1, the results are improved by 11.4%, 15.2%, and 9.2%, separately.

To further illustrate the interesting trends, we plot the Micro-F1 and Macro-F1 to investigate the changes of the different baseline methods in Figure 6.

As can be seen from the results in Figure 6, compared to shallow embedding models, matrix decomposition models, and deep embedding models, the Micro-F1 and Macro-F1 of HGCN* are all improved on the Wikipedia dataset. Figure 6 also shows that the Micro-F1 and Macro-F1 of HGCN have been significantly improved by using the max-pooling. The experimental results further indicate that the improved

model for the dataset with a large number of label types can get better graph embedding effectively.

5. Conclusion

In this study, we explore an unsupervised graph embedding method for the undirected graph. Comparing with conventional graph embedding models, we introduce hierarchical graph convolution network to propose the HGCN and HGCN* methods respectively. Besides, we come to the following conclusions:

(1) In this paper, a hierarchical GCN for the undirected graph with only structure information is proposed. The model does not need the extra features and labeled data, which improves the applicability of the GCN in the real network.

(2) Moreover, we introduce max-pooling to improve the proposed model according to the number of label types. Experimental results show that our model achieves considerable improvement than all the baselines.

However, due to the limitations of the pre-training and the clustering methods, the proposed model in this paper still has the disadvantage of obtaining the initial node embedding and the poor label accuracy. In the future, we will further focus on dealing with the problems and plan to explore the impact of initialing models and clustering methods on the proposed model, which ultimately can provide a better method of unsupervised graph embedding.

Data Availability

All experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Thanks are due to the teachers and classmates of the project team for their guidance and help. This research was supported by the National Natural Science Fund of China (no. 61936012 and 61806117), the National Social Science Fund of China (no. 18BYY074), the Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (No. 201802012), and the Innovation Project for College Graduates of Shanxi Province (no. 2020SY015).

References

[1] Y. Cheng, D. Wang, P. Zhou, and T. Zhang, "A survey of model compression and acceleration for deep neural networks," 2017, https://arxiv.org/abs/1710.09282.

[2] G. Palash and F. Emilio, "Graph embedding techniques, applications, and performance: a survey," *Knowledge-Based Systems*, vol. 151, pp. 78–94, 2018.

[3] K. S. Mehran, G. Rishab, J. Kshitij et al., "Relational representation learning for dynamic (knowledge) graphs: a survey," 2019, https://arxiv.org/abs/1905.11485.

[4] I. K. Fodor, *A Survey of Dimension Reduction Techniques*, Neoplasia, Amsterdam, Netherlands, 2002.

[5] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1–3, pp. 37–52, 1987.

[6] J. B. Kruskal, M. Wish, and E. M. Uslaner, *Multidimensional Scaling, Methods*, Vol. 116, SAGE, New York, NY, USA, 1978.

[7] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," *Advances in Neural Information Processing Systems*, vol. 14, no. 6, pp. 585–591, 2002.

[8] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: methods and applications," 2017, https://arxiv.org/abs/1709.05584.

[9] C. C. Tu, W. C. Zhang, Z. Y. Liu, and S. M. Sun, "Max-margin deepwalk: discriminative learning of network representation," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 3889–3895, New York, NY, USA, July 2016.

[10] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[11] C. Yang, Z. Y. Liu, D. L. Zhao, M. S. Sun, and E. Y. Chang, "Network representation learning with rich text information," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 2111–2117, Buenos Aires, Argentina, July-August 2015.

[12] X. F. Sun, J. Guo, X. Ding, and T. Liu, "A general framework for content-enhanced network representation learning," 2016, https://arxiv.org/abs/1610.02906.

[13] D. K. Zhang, J. Yin, X. Q. Zhu, and C. Q. Zhang, "Homophily, structure, and content augmented network representation learning," in *Proceedings of the International Conference on Data Mining (ICDM)*, pp. 609–618, Barcelona, Spain, December 2016.

[14] P. Velickovic, W. Fedus, W. L. Hamilton, P. Lio, Y. Bengio, and R. D. Hjelm, "Deep graph infomax," in *Proceedings of the International Conference on Learning Representations (ICLR)*, New Orleans, LA, USA, May 2019.

[15] Z. N. Hu, C. J. Fan, T. Chen, K. W. Chang, and Y. Z. Sun, "Pre-training graph neural networks for generic structural feature extraction," 2019, https://arxiv.org/abs/1905.13728.

[16] J. Pennington, R. Sorcher, and C. D. Manning, "Glove: global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, October 2014.

[17] J. Zhou, G. Q. Cui, Z. Y. Zhang et al., "Graph neural networks: a review of methods and applications," 2018, https://arxiv.org/abs/1812.08434.

[18] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, pp. 3111–3119, Lake Tahoe, NV, USA, December 2013.

[19] T. Tang and H. Liu, "Leveraging social media networks for classification," *Data Mining and Knowledge Discovery*, vol. 23, no. 3, pp. 447–478, 2011.

[20] B. Perozzi, R. Alrfou, and S. Skiena, "Deepwalk: online learning of social representations," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge*

Discovery and Data Mining (KDD), pp. 701–710, New York, NY, USA, August 2014.

[21] J. Tang, M. Qu, M. Z. Wang, M. Zhang, J. Yan, and Q. Z. Mei, "Line: large-scale information network embedding," in *Proceedings of the 24th International Conference on World Wide Web (WWW)*, pp. 1067–1077, Florence, Italy, May 2015.

[22] A. Grover and J. Leskovec, "Node2vec: scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 855–864, San Francisco, CA, USA, August 2016.

[23] L. F. R. Ribeiro, P. H. P. Saverese, and D. R. Figueiredo, "Struc2vec: learning node representations from structural identity," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 385–394, Halifax, Canada, August 2017.

[24] J. Y. Xie, R. Girshick, and A. Farhadi, "Unsupervised deep embedding for clustering analysis," in *Proceedings of International Conference on Machine Learning (ICML)*, pp. 478–487, New York, NY, USA, June 2016.

[25] H. Y. Xu, H. T. Liu, W. J. Wang, Y. H. Sun, and P. F. Jiao, "NEFLGC: network embedding based on fusing local (First-Order) and global (Second-Order) network structure with node content," *Pacific-Asia Conference on Knowledge discovery and data mining (PAKDD)*, pp. 260–271, 2018.

[26] D. X. Wang, P. Cui, and W. W. Zhu, "Structural deep network embedding," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1225–1234, San Francisco, CA, USA, August 2016.

[27] T. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, https://arxiv.org/abs/1609.02907.

[28] W. L. Hamilton, Z. T. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proceedings of Annual Conference Neural Information Processing Systems (NIPS)*, pp. 1024–1034, Long Beach, CA, USA, December 2017.

[29] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," 2017, https://arxiv.org/abs/1710.10903.

[30] S. Abuelhaija, A. Kapoor, B. Perozzi, and J. Lee, "N-GCN: multi-scale graph convolution for semi-supervised node classification," 2018, https://arxiv.org/abs/1802.08888.

[31] J. Chen, T. F. Ma, and C. Xiao, "FastGCN: fast learning with graph convolutional networks via importance sampling," 2018, https://arxiv.org/abs/1801.10247.

[32] K. Sun, Z. X. Zhu, and Z. C. Lin, "Multi-stage self-supervised learning for graph convolutional networks," 2019, https://arxiv.org/abs/1902.11038.

[33] M. Qu, Y. Bengio, and J. Tang, "GMNN: graph markov neural networks," 2019, https://arxiv.org/abs/1905.06214.

[34] U. Bodenhofer, A. Kothmeier, and S. Hochreiter, "AP cluster: an R package for affinity propagation clustering," *Bioinformatics*, vol. 27, no. 17, pp. 2463-2464, 2011.

[35] A. K. McCallum, K. Nigam, J. Rennie, and K. Seymore, "Automating the construction of internet portals with machine learning," *Information Retrieval*, vol. 3, no. 2, pp. 127–163, 2000.

[36] P. Sen, G. Namata, M. Bilgic, L. Getoor, B. Galligher, and T. Eliassi-Rad, "Collective classification in network data," *AI Magazine*, vol. 29, no. 3, pp. 93–106, 2008.

[37] S. S. Cao, W. Lu, and Q. K. Xu, "Grarep: learning graph representations with global structural information," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM)*, pp. 891–900, Melbourne Australia, October 2015.

[38] M. D. Ou, P. Cui, J. Pei, Z. W. Zhang, and W. W. Zhu, "Asymmetric transitivity preserving graph embedding," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1105–1114, San Francisco, CA, USA, August 2016.

[39] Qinghua University, OpenNE, https://github.com/thunlp/OpenNE.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 2891572, 14 pages
https://doi.org/10.1155/2020/2891572

Research Article

A Novel Neural Network-Based SINS/DVL Integrated Navigation Approach to Deal with DVL Malfunction for Underwater Vehicles

Wanli Li, Mingjian Chen, Chao Zhang, Lundong Zhang, and Rui Chen

Institute of Geospatial Information, Information Engineering University, Zhengzhou 450000, China

Correspondence should be addressed to Wanli Li; liwanli1201@hotmail.com

Received 18 February 2020; Revised 5 April 2020; Accepted 7 July 2020; Published 26 July 2020

Guest Editor: Shianghau Wu

A navigation grade Strapdown Inertial Navigation System (SINS) combined with a Doppler Velocity Log (DVL) is widely used for autonomous navigation of underwater vehicles. Whether the DVL is able to provide continuous velocity measurements is of crucial importance to the integrated navigation precision. Considering that the DVL may fail during the missions, a novel neural network-based SINS/DVL integrated navigation approach is proposed. The nonlinear autoregressive exogenous (NARX) neural network, which is able to provide reliable predictions, is employed. While the DVL is available, the neural network is trained by the body frame velocity and its increment from the SINS and the DVL measurements. Once the DVL fails, the well trained network is able to forecast the velocity which can be used for the subsequent navigation. From the experimental results, it is clearly shown that the neural network is able to provide reliable velocity predictions for about 200 s–300 s during DVL malfunction and hence maintain the short-term accuracy of the integrated navigation.

1. Introduction

The ocean covers 71% of the surface of the Earth. It is rich in biological resource and minerals resource. Underwater vehicles, including remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs), are widely used in the exploration of the ocean. However, the lack of reliable navigation techniques is still a key fact which limits the application of the underwater vehicles. Although Global Navigation Satellite System (GNSS) is widely used in surface and air navigation [1], its signals rapidly attenuate in water. The acoustic navigation systems [2], such as Long Baseline (LBL) and Ultra Short Baseline (USBL), have a limited range. For example, the 12 kHz LBL typically operates at up to 10 km ranges [3]. Strapdown Inertial Navigation System (SINS), which is able to provide complete navigation parameters, such as position, velocity, and attitude, is usually employed in autonomous underwater navigation. However, as the positioning errors of the SINS increase with time, an external aiding sensor is necessary.

Based on Doppler Effect, the Doppler Velocity Log (DVL) is able to provide velocity measurements relative to the seafloor. It is regarded as one of the most potential aiding sensors which are able to limit the error growth of SINS [4, 5]. However, the DVL provide velocities in the Doppler instrumental frame, but not the geodetic frame. There are still challenges in DVL aided integrated navigation. For example,, misalignments between the SINS body frame and Doppler instrumental frame will occur during manufacture [6]. Therefore, a calibration process is usually required before conducting a mission [7, 8]. Even so, the position error of SINS/DVL will accumulate without any external position observations. Furthermore, the azimuth error of the integrated navigation shows slow convergence in the case of DVL aiding. Fast SINS in-motion initial alignment is also a challenge issue which has aroused great attention [9].

In addition, DVL has a limited maximum range. Table 1 shows the specifications of the DVLs [10] produced by Teledyne RD Instruments, Inc. As can be seen from the table, DVL working at a higher operating rate can provide more accurate velocity measurements, whereas its maximum range is smaller and vice versa. In the case that the acoustic wave emitted by the DVL cannot reach the seafloor, the DVL will work in the mode of water-tracked, which provides water reference velocities, or fail to provide velocities.

SINS/DVL integrated is typically accomplished by using a Kalman filter (KF), which fuses the data from the SINS and

Table 1: Specifications of the RDI workhorse navigator Doppler Velocity Log.

Parameter item	WHN 300	WHN 600	WHN 1200
Long-term accuracy	±0.4%V ± 0.2 cm/s	±0.2%V ± 0.1 cm/s	±0.2%V ± 0.1 cm/s
Operating rate	300 kHz	600 kHz	1200 kHz
Maximum range	200 m	90 m	25 m
Velocity range	±10 m/s	±10 m/s	±10 m/s

the DVL. Therefore, the accuracy of the DVL measurements is of vital importance in optimally estimating the subsequent navigation solutions. In [11], the authors proposed a pretreatment method for the velocity of DVL based on the motion constraint of underwater vehicles. It is able to restrain random noise in the DVL. Adaptive Kalman filter (AKF) is also regarded as an effective approach to deal with the accuracy decrease of the DVL [12, 13]. SINS/DVL integrated navigation, in case that only partial DVL measurements are available, is also investigated. In [14, 15], tightly coupled navigation structures are proposed to deal with the situation that DVL has fewer than three beam measurements.

Under certain conditions, such as sailing across sea creatures or the DVL exceeds its maximum measuring range, the DVL will fail to maintain bottom-lock and provide velocity updates [16, 17]. To circumvent this problem, Tal et al. [18] derived an extended loosely coupled (ELC) approach which is able to provide virtual vehicle velocity by using partial raw data of the DVL and additional information. In [19], the authors proposed a hybrid approach to forecast the measurements of the DVL while it malfunctions. The current and past velocities obtained from SINS are taken as the predictor's inputs. In this paper, we proposed a neural network-based approach to predict the body frame velocity when DVL is unavailable. Related methodologies in similar situations can be found in the literature on SINS/GNSS integrated navigation [20–22]. The main idea of these approaches is to build the map between SINS measurements (angular rate, specific force, velocity, position, and so forth) and GNSS measurements [21]. If the GNSS signals are available, the neural network is trained by the SINS and GNSS measurements. Once the outages of the GNSS happen, the virtual GNSS measurements can be obtained by the well trained neural network. Therefore, it is able to maintain the short-term stability of the SINS/GNSS integration. Inspired by these works, a nonlinear autoregressive exogenous (NARX) neural network model, which is able to forecast the velocity during the DVL malfunction, is constructed. The NARX neural network has been widely used in time series forecasting [23]. The prediction of the velocity can also be regarded as a time series forecasting problem. When DVL is available, the body frame velocity and its increment obtained from the SINS and DVL are collected to train the neural network. Once DVL is unavailable, the velocity forecasted by the well trained neural network is utilized to assist the SINS and hence maintain the accuracy of the integrated navigation.

The rest of this paper is organized as follows. Section 2 is devoted to the presentation of the SINS/DVL integrated navigation scheme. Both the SINS error dynamics model and the DVL error model are derived. Section 3 gives the algorithmic description of the neural network-based SINS/DVL integrated navigation scheme. In Section 4, the performance of the proposed algorithm is evaluated with ship-mounted experimental data collected in the Yangtze River. Conclusions are drawn in Section 5.

2. SINS/DVL Integrated Navigation

Figure 1 shows the structure of the SINS/DVL integration. Normally, the SINS is consisted of three orthogonal gyroscopes and three orthogonal accelerometers, which are able to provide angular rate and specific force in the SINS body frame (denoted as b). With these measurements, the navigation solutions including position, attitude, and velocity can be obtained. The DVL provides velocity in the Doppler instrumental frame (denoted as d). To match the velocity of SINS, firstly, it should be transformed into the body frame with alignment calibration parameters. Then, the DVL velocity can be transformed into the navigation frame (denoted as n, local-level frame, and its orientation is north-east-sown (NED)) with the attitude of the SINS. As a velocity matching integration scheme, the velocities of the DVL and SINS in the navigation frame are chosen as the input of the Kalman Filter. With the estimation of the Kalman Filter, the navigation solutions can be reset and the sensor errors such as gyro bias and accelerometer bias can be corrected. As can be seen from Figure 1, with a feedback scheme, the navigation errors of the SINS, such as attitude error and velocity error, are compensating each Kalman filter cycle.

2.1. SINS Error Dynamics Model. Due to the inherent biases (gyro bias and accelerometer bias) of the SINS, the position error, attitude error, and the velocity error accumulate with time. It is known to us that, in the case of velocity matching integrated navigation, the position error is unobservable. Therefore, the position error is abandoned to be chosen as a state of the integration. The velocity and attitude error propagation process of the SINS can be described as follows [7]:

$$\delta\dot{v}^n = [f^n\times]\phi - (2\omega_{ie}^n + \omega_{en}^n)\times\delta v^n + C_b^n\nabla, \tag{1}$$

$$\dot{\phi} = -(\omega_{ie}^n + \omega_{en}^n)\times\phi + (\delta\omega_{ie}^n + \delta\omega_{en}^n) - C_b^n\varepsilon, \tag{2}$$

where the subscript n denotes the components described in the navigation frame. δv^n is the velocity error. ϕ is the attitude error. f^n is the specific force in the navigation frame. C_b^n is the direction cosine matrix of the body frame to the navigation frame. ω_{ie}^n denotes the angular rate of the Earth frame e relative to the inertial frame i. ω_{en}^n is the angular rate of the navigation frame relative to the Earth frame. ∇ and ε are the biases of the gyro and the accelerometer in the body

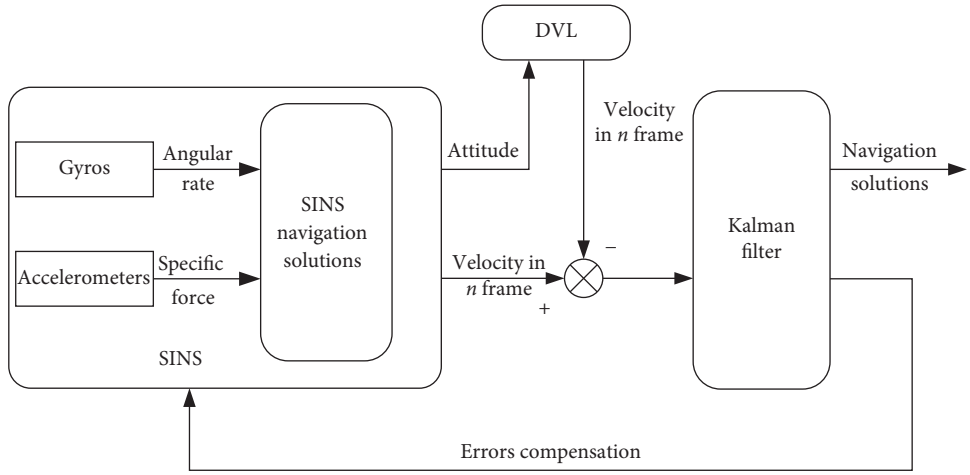


Figure 1: SINS/DVL integrated navigation structure.

frame, respectively. In SINS/DVL integration, ∇ and ε are usually modeled as random constants:

$$\dot{\nabla} = 0, \tag{3}$$

$$\dot{\varepsilon} = 0. \tag{4}$$

The error states of the SINS are chosen as follows:

$$x = \left[\delta v_N \;\; \delta v_E \;\; \delta v_D \;\; \phi_N \;\; \phi_E \;\; \phi_D \;\; \nabla_x \;\; \nabla_y \;\; \nabla_z \;\; \varepsilon_x \;\; \varepsilon_y \;\; \varepsilon_z \right]. \tag{5}$$

The SINS error dynamics model can be expressed as follows:

$$\dot{x} = Fx + w, \tag{6}$$

where w is a zeros-mean Gaussian white noise which is determined by the accuracy of the gyros and accelerometers. F is the state transformation matrix constructed according to equations (1)–(4). It can be given as follows:

$$F = \begin{bmatrix} F_{11} & F_{12} & C_b^n & 0_{3\times3} \\ F_{21} & F_{22} & 0_{3\times3} & -C_b^n \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix}, \tag{7}$$

where

$$F_{11} = \begin{bmatrix} \dfrac{v_D}{R_N + h} & -2\omega_{ie}\sin L - 2\dfrac{v_E tgL}{R_E + h} & \dfrac{v_N}{R_N + h} \\ 2\omega_{ie}\sin L + 2\dfrac{v_E tgL}{R_E + h} & \dfrac{v_D + v_N tgL}{R_E + h} & 2\omega_{ie}\cos L + \dfrac{v_E}{R_E + h} \\ -2\dfrac{v_N}{R_N + h} & -2\omega_{ie}\cos L - \dfrac{v_E}{R_E + h} & 0 \end{bmatrix}, \tag{8}$$

whereR_E and R_N are the transverse radius and meridian radius of the Earth, respectively. L is the latitude. h denotes the altitude. ω_{ie} is the Earth's rotation rate.v_N, v_E, and v_D are the velocity in the north, east, and down direction, respectively.

$$F_{12} = \begin{bmatrix} 0 & -f_D & f_E \\ f_D & 0 & -f_N \\ -f_E & f_N & 0 \end{bmatrix}, \tag{9}$$

where f_N, f_E, and f_D are the specific force in the navigation frame. Also,

$$F_{21} = \begin{bmatrix} 0 & \frac{1}{R_E + h} & 0 \\ -\frac{1}{R_N + h} & 0 & 0 \\ 0 & \frac{tgL}{R_E + h} & 0 \end{bmatrix}, \tag{10}$$

$$F_{22} = \begin{bmatrix} 0 & -\omega_{ie} \sin L - \frac{V_E tgL}{R_E + h} & \frac{V_N}{R_N + h} \\ \omega_{ie} \sin L + \frac{V_E tgL}{R_E + h} & 0 & \omega_{ie} \cos L + \frac{V_E}{R_E + h} \\ -\frac{V_N}{R_N + h} & -\omega_{ie} \cos L - \frac{V_E}{R_E + h} & 0 \end{bmatrix}. \tag{11}$$

2.2. DVL Error Model. The velocity of DVL can be transformed into the navigation frame as follows:

$$\tilde{v}_d^n = \tilde{C}_b^n C_d^b v_d, \tag{12}$$

where v_d is the velocity of DVL in the Doppler instrumental frame d. C_d^b is the alignment matrix which presents the transformation of the Doppler instrumental frame to the body frame b; it can be presented as follows:

$$C_d^b = I_{3\times3} - \theta\times, \tag{13}$$

where $I_{3\times3}$ is a 3×3 identity matrix and $\theta\times$ denotes the skew matrix of the misalignments between the DVL and the SINS. It can be obtained by calibration [7, 8]. $\tilde{C}_b^n$ is the direction cosine matrix in error; it can be obtained by

$$\tilde{C}_b^n = [I_{3\times3} - \phi\times] C_b^n, \tag{14}$$

where $\phi\times$ denotes the skew matrix of the attitude error ϕ. Substituting (14) into (12), it yields

$$\begin{aligned} \tilde{v}_d^n &= [I_{3\times3} - \phi\times] C_b^n C_d^b v_d, \\ &= C_b^n C_d^b v_d + \left[\left(C_b^n C_d^b v_d\right)\times\right]\phi, \\ &= v^n + \delta v_d. \end{aligned} \tag{15}$$

The velocity of the SINS in error can be modeled as

$$\tilde{v}_{\text{SINS}}^n = v^n + \delta v^n. \tag{16}$$

Differencing the velocity of the SINS and DVL in the navigation frame, there exists

$$\begin{aligned} \tilde{v}_{\text{SINS}}^n - \tilde{v}_d^n &= \left(v^n + \delta v^n\right) - \left(v^n + \delta v_d^n\right) \\ &= \delta v^n - \left[\left(C_b^n C_d^b v_d\right)\times\right]\phi. \end{aligned} \tag{17}$$

It can be represented by the following equation:

$$z = \tilde{v}_{\text{SINS}}^n - \tilde{v}_d^n = Hx + \eta, \tag{18}$$

where η is a zeros-mean Gaussian white noise which is determined by the accuracy of the DVL. H can be obtained from equation (17) as follows:

$$H = \left[\, I_{3\times3} \quad \left(C_b^n C_d^b v_d\right)\times \quad 0_{3\times6} \,\right]. \tag{19}$$

2.3. Standard Kalman Filter. The integrated navigation model described by equation (6) and equation (18) can be discretized as follows:

$$\begin{cases} X_k = \Phi_{k,k-1} X_{k-1} + W_k, \\ Z_k = H_k X_{k-1} + V_k, \end{cases} \tag{20}$$

where X_k is the state vector, Z_k is the measurement vector, $\Phi_{k,k-1}$ is the state transition matrix, and H_k is the measurement matrix. W_k and V_k are uncorrelated zeros-mean Gaussian white noise and their conversances can be given as follows:

$$E[W_k W_i] = \begin{cases} Q_k & k = i, \\ 0 & k \neq i, \end{cases} \tag{21}$$

$$E[V_k V_i] = \begin{cases} R_k & k = i, \\ 0 & k \neq i, \end{cases} \tag{22}$$

$$E[V_k W_i] = 0, \quad \forall k, i. \tag{23}$$

Kalman filter is able to estimate the state vector from the measurement vector in an optimum way. It is consisted of time update and measurement update. If there are no DVL measurements, time update is employed to predict the state vector as follows [11]:

$$\widehat{X}_{k,k-1} = \Phi_{k,k-1}\widehat{X}_{k-1}, \tag{24}$$

$$P_{k,k-1} = \Phi_{k,k-1}P_{k-1}\Phi_{k,k-1}^{T} + Q_{k-1}, \tag{25}$$

where $\widehat{X}_{k,k-1}$ is the predicted state vector and $P_{k,k-1}$ is the predicted covariance matrix of $\widehat{X}_{k,k-1}$. Once the DVL measurement is available, the measurement update can be accomplished as follows:

$$\widehat{X}_k = \widehat{X}_{k,k-1} + K_k\left[Z_k - H_k\widehat{X}_{k,k-1}\right], \tag{26}$$

$$K_k = P_{k,k-1}H_k^T\left[H_kP_{k,k-1}H_k^T + R_k\right]^{-1}, \tag{27}$$

$$P_k = \left[I - K_kH_k\right]P_{k,k-1}\left[I - K_kH_k\right]^T + K_kR_kK_k^T, \tag{28}$$

where K_k is the Kalman Filter gain. Then, the navigation solutions of the SINS/DVL integration can be updated by the estimation of the state vector $\widehat{X}_k$.

3. Neural Network-Based Integrated Navigation Structure

As mentioned in Section 1, the DVL may fail to provide velocity in some situations. To deal with this problem, a neural network-based approach is proposed. The main idea is to forecast the velocity measurements by neural network during DVL malfunction. And hence the subsequent integrated navigation can be conducted with the predictions. Since the velocity of the DVL is in the Doppler instrumental frame d, it should be transformed into the body frame by the alignment matrix C_d^b before integration as follows:

$$v^b = C_d^b v^d. \tag{29}$$

The neural network will predict the body frame velocity directly. In this paper, a nonlinear autoregressive exogenous neural network is employed [24]. That is, the current value of a time series can be forecasted by two series: the previous values of the same series and the current and previous values of the driving series. The driving series is the series that influences the series of interest.

The NARX model can be described as follows [25]:

$$y_{k+1} = M\begin{pmatrix} y_k & y_{k-1} & \cdots & y_{k-m} & x_{k+1} \\ x_k & x_{k-1} & \cdots & x_{k-m} & \end{pmatrix}, \tag{30}$$

where $M(\cdot)$ is the mapping function of the neural network, x_k is the value of the driving series at time k, and y_k is the value of the target series at time k. Obviously, the body frame velocity sequence is the target time series. Considering that the body frame velocity increment sequence of SINS can also indicate the change of the velocity, it is chosen as the driving series. The velocity increment of the SINS in n-frame u^n can be obtained by the two-sample iteration algorithm [26]. The velocity increment of the SINS in b frame during each DVL update cycle can be given as follows:

$$\Delta v_k^b = \sum_{i=1}^{N} C_n^b\left(t_{k-1} + i \times T\right)u^n\left(t_{k-1} + i \times T\right), \tag{31}$$

where T is the SINS update cycle and N denotes the amount of SINS update during each DVL update cycle.

The variables of the mapping function, which are initially stochastic, can be fine-tuned in the training process. The multilayer perceptron network is employed to perform the approximation and its configuration is shown in Figure 2. Considering that the training of the neural network requires a lot of data, a light-weight neural network model is employed. Therefore, the employed model is easy for training and fine-tuning, and no massive training data is required. The network is composed of three layers: input layer, hidden layer, and output layer. Both the body frame velocity and the velocity increment have three dimensions, so there are six attributes of the input. The velocities of the previous two steps are also adopted as the input of the network. There are ten neurons in the hidden layer. The final output of the network is the predicted body frame velocity which is a three-dimensional vector. The Levenberg–Marquardt (LM) optimization method is adopted as the training function for the network [27].

The structure of the improved SINS/DVL integration is shown in Figure 3. When the DVL is available, the SINS is integrated with the DVL to get navigation solutions. Meanwhile, the neural network is trained by the velocity and velocity increment in b frame. The driving series can be obtained continuously by the SINS:

$$x_k = \Delta v_k^b,\ k = 1, 2, \cdots, m, \tag{32}$$

where the velocity increment Δv_k^b can be obtained from equation (31). The target series can be acquired from equation (29) with DVL measurements as follows:

$$y_k = v_k^b,\ k = 1, 2, \cdots, m. \tag{33}$$

Once the DVL fails, the velocity in b frame can be predicted by the well trained network. However, with the DVL measurements, the network is able to predict only one step ahead. Therefore, the subsequent integrated navigation is conducted as follows:

Step 1: predict the velocity in b frame one step ahead with the input time series of current and the previous two steps.

Step 2: update the SINS navigation solutions and the velocity increment in b frame with the SINS raw data.

Step 3: conduct the SINS/DVL integrated navigation and update navigation solutions. Update the driving series with the velocity increment obtained by Step 2. Update the target series with the prediction obtained by Step 1.

Step 4: go to Step 1 until the integrated navigation is finished

4. Experimental Results and Discussion

4.1. Test Configuration. The ship-mounted experimental data were collected to evaluate the performance of the proposed algorithm. The trial was carried out in the Yangtze River. The equipped instruments include a navigation grade

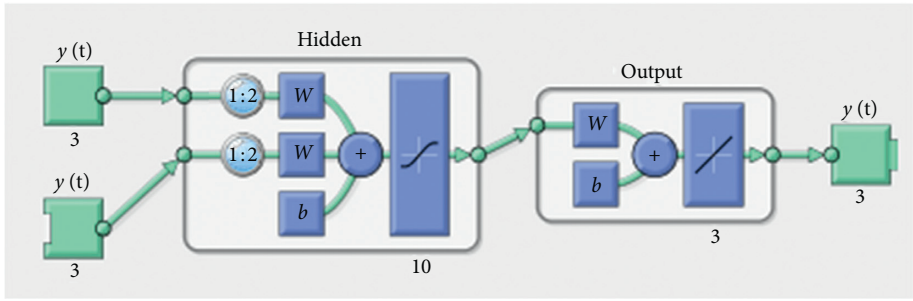


Figure 2: NARX generated by MATLAB R2014a.

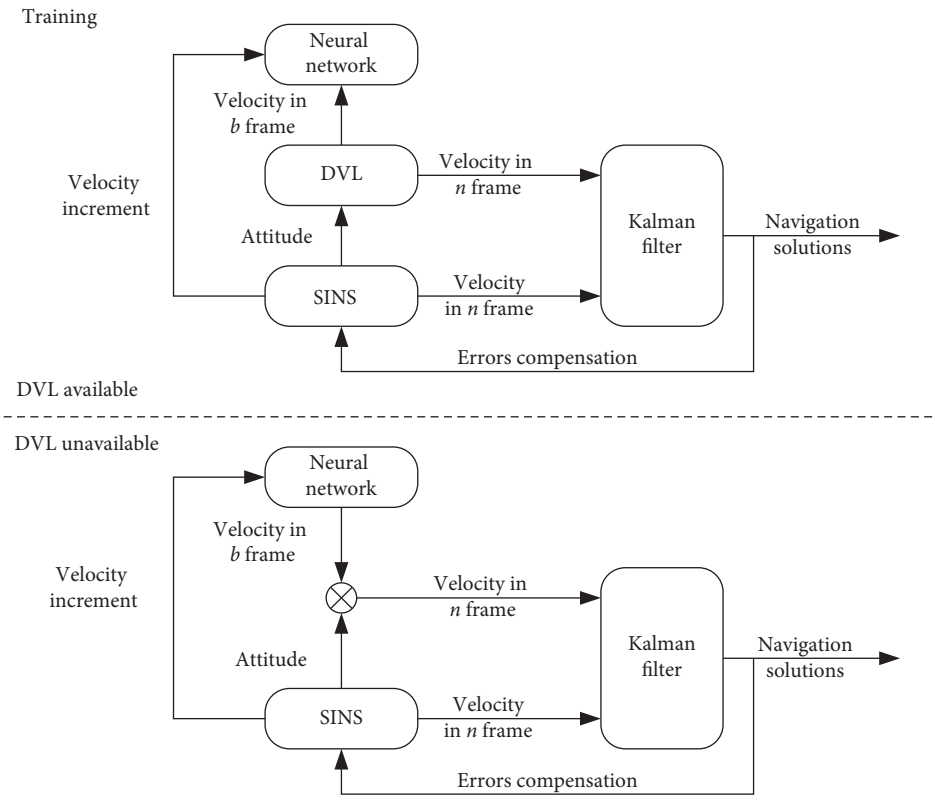


Figure 3: Structure of the improved SINS/DVL integration.

SINS, a bottom-locked DVL, and a NovAtel GPS receiver. The specifications of the SINS and DVL are shown in Tables 2 and 3, respectively.

4.2. Performance Evaluation of the SINS/DVL Integration. During the test, the ship sailed up to about 30 kilometers at the speed of around 9 knots (approximately 6600 s). From the experimental data, it is shown that the GPS receiver was able to provide credible position and velocity measurements during the whole test. Therefore, the integrated navigation results, including position, velocity, and attitude from SINS/GPS, can be used as the benchmark to compare with the SINS/DVL integration. Before integrated navigation, an initial alignment of SINS was conducted. The trajectories obtained from SINS/GPS and SINS/DVL are shown in

Table 2: SINS specifications.

Parameter item	Gyroscope	Accelerometer
Bias	$<0.02^{\circ}/h(1\sigma)$	$<5\times10^{-5}$g (1σ)
Update rate	200 Hz	200 Hz
Bias stability	<30 ppm	<50 ppm
Dynamic range	$\pm200^{\circ}$(s)	±15 g

Table 3: DVL specifications.

Parameter item	
Velocity accuracy	±0.5%V ± 0.5 cm/s
Update rate	1 Hz
Operating rate	300 kHz
Bottom-locked range	300 m
Dynamic range	−10 knot~20 knot

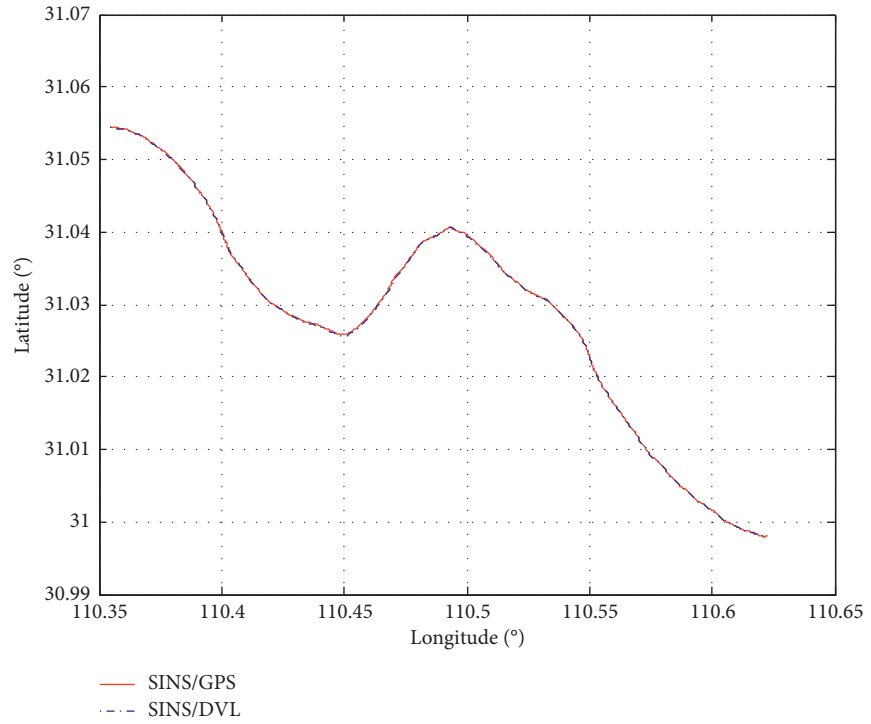


Figure 4: Trajectories obtained from the SINS/GPS and the SINS/DVL.

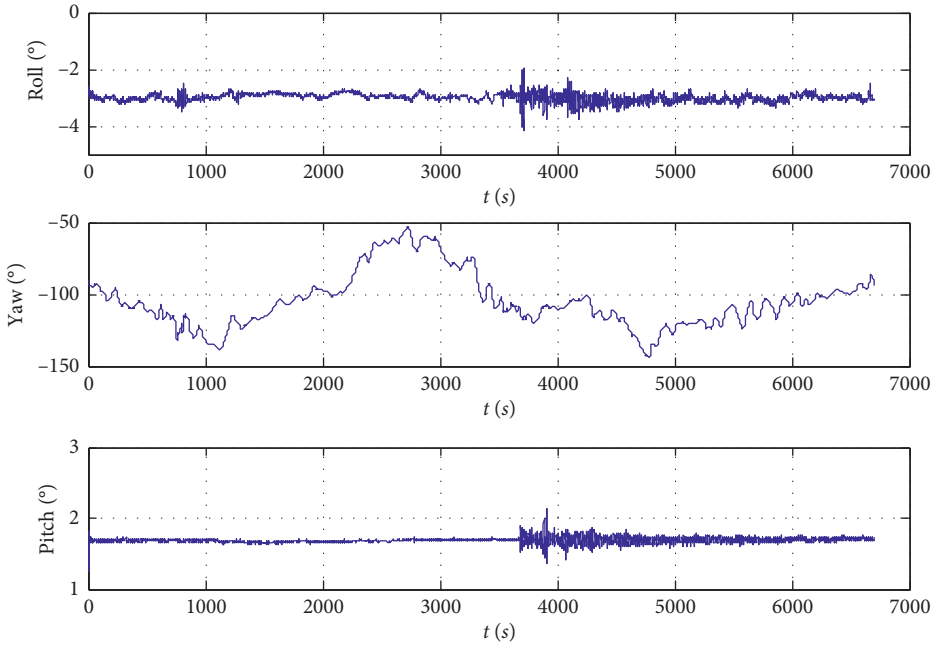


Figure 5: The attitude of the ship.

Figure 4. Figure 5 shows the changing of attitude. Figures 6 and 7 show the velocities in the navigation frame and Doppler instrumental frame, respectively. As can be seen from the figures, the roll and the pitch of the ship remain around −3° and 1.7°, respectively. The heading and the velocity of the ship vary slowly. The AUVs usually operate within a low dynamics range. It may fit the real motion of the AUVs, such as submarines. As can be seen from Figure 4, the two trajectories closely match each other. The error curve of the position is shown in Figure 8. And the final position error of the SINS/DVL integration is about 28 meters, which is less than 1% of the distance travelled. Predictably, as long as external position observations are unavailable, the position error will accumulate with time.

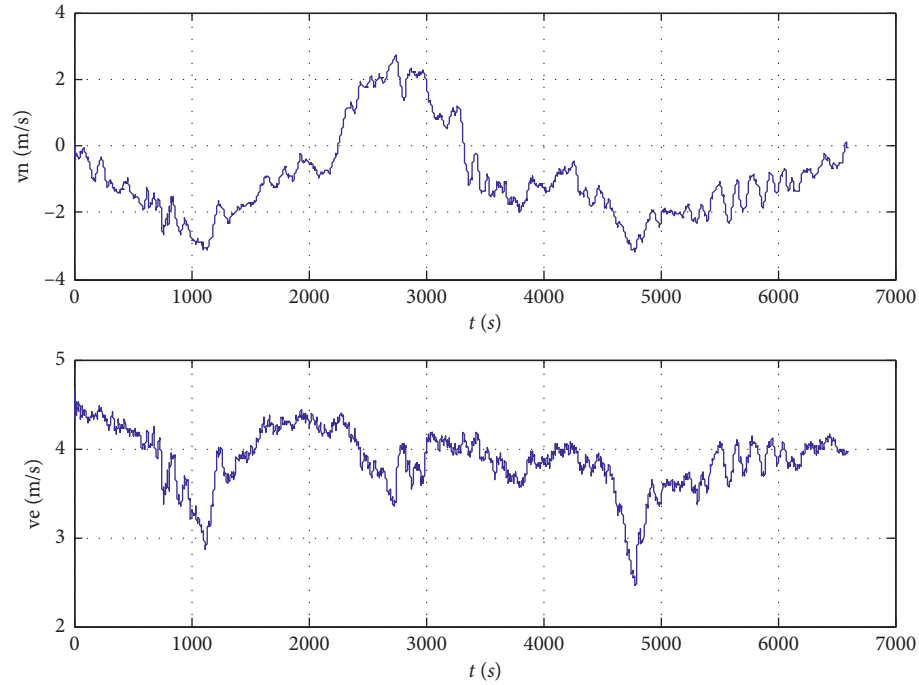


Figure 6: The velocity of the ship in the navigation frame.

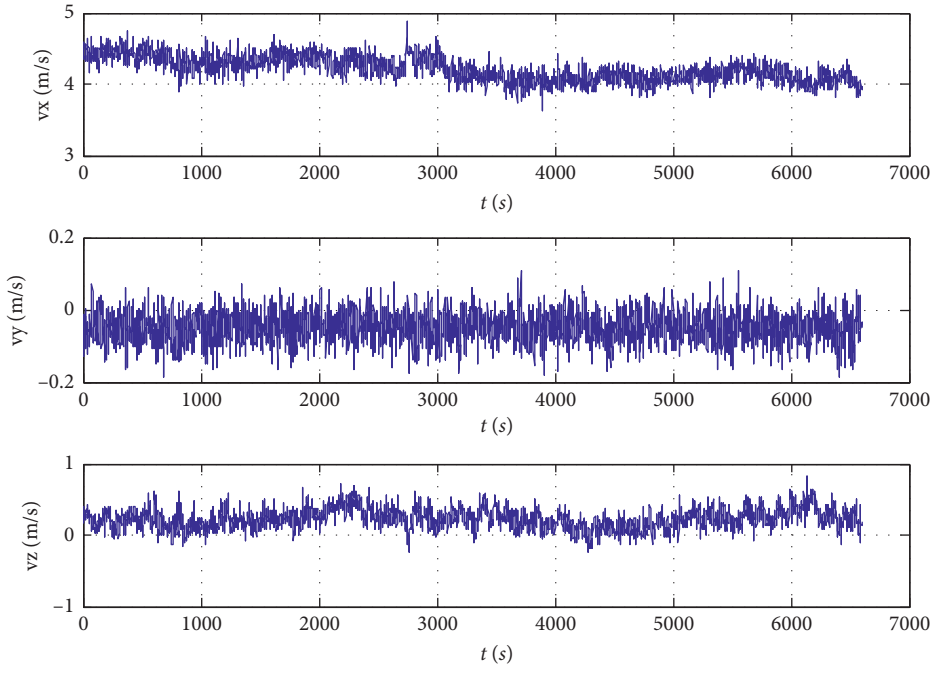


Figure 7: The DVL measurements during the experiment.

As the DVL provides the velocity measurements in the Doppler instrumental frame but not geodetic frame, it leads to the slow convergence of the heading error and INS measurement errors (the gyro biases and the accelerometer biases). In addition, the AUVs usually operate within a low dynamics range, which makes this problem more serious. Both the gyro and the accelerometer biases in the up direction cannot be estimated as reliable. Therefore, the open loop is adopted to estimate the INS biases [28]. Misalignments of the SINS/DVL are shown in Figure 9. The misalignments converge with time. However, the heading error converged slower than the roll error and the pitch error. It oscillates at the beginning and then converges. From the partial magnification of Figure 9, the heading error remains

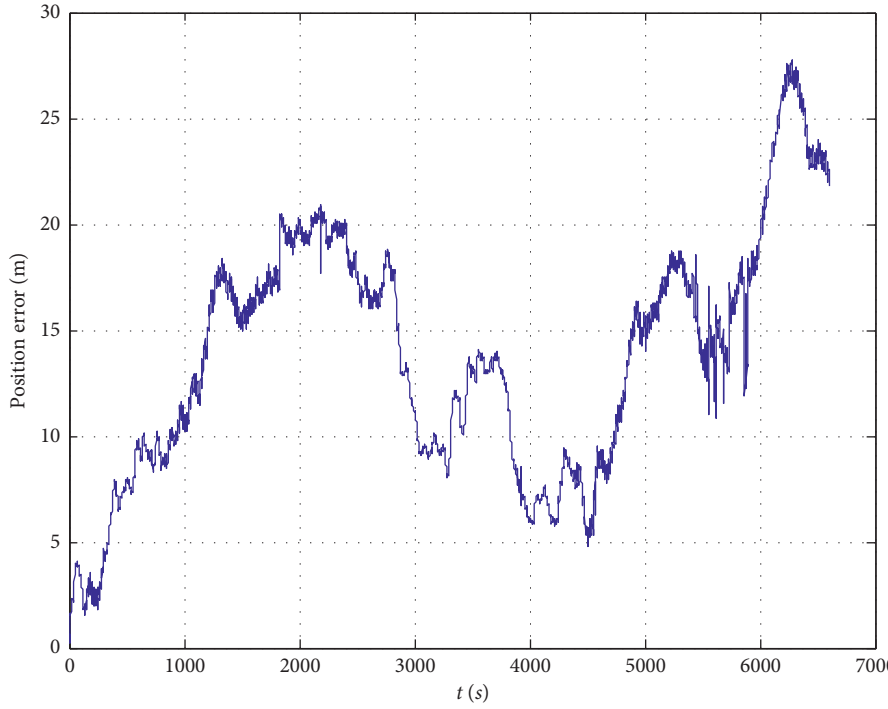


Figure 8: Position error of the SINS/DVL.

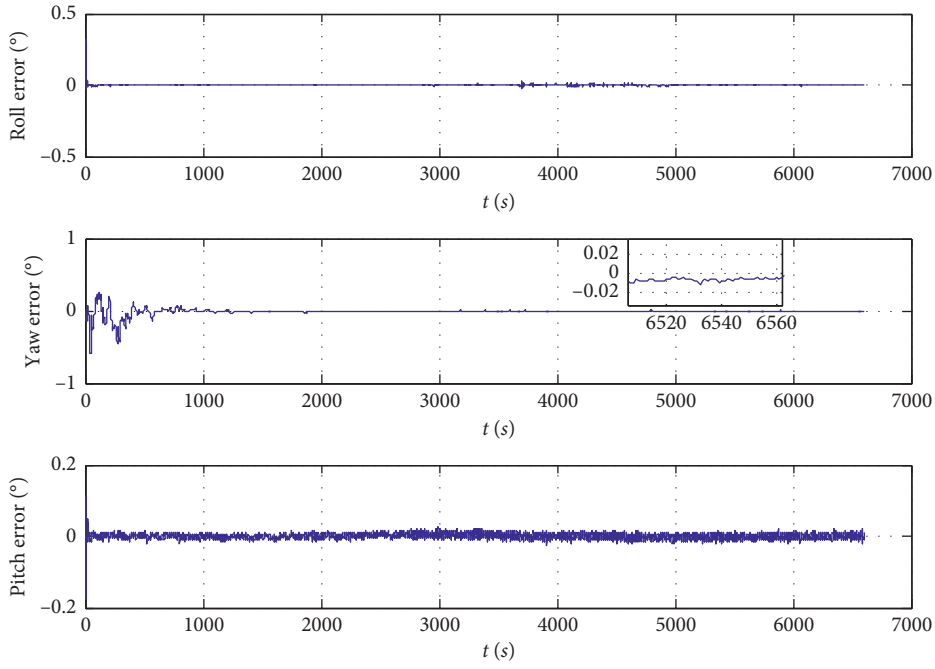


Figure 9: Misalignments of the SINS/DVL.

around 0.01° finally. It is known that DVL provides velocity measurements in the Doppler instrumental frame. And it leads to the slow convergence of the heading. Figure 10 shows the velocity error in the north and east direction, respectively. As a velocity-aided integration, the velocity errors of SINS/DVL show fast convergence.

4.3. Validation of the Proposed Neural Network-Based Algorithm. A test was designed to evaluate the performance of the proposed neural network-based algorithm. In practical use, the flag denoting whether the DVL is available has been set. As the DVL is available during the whole mission, it is assumed that DVL is unavailable after 4000 s. And the

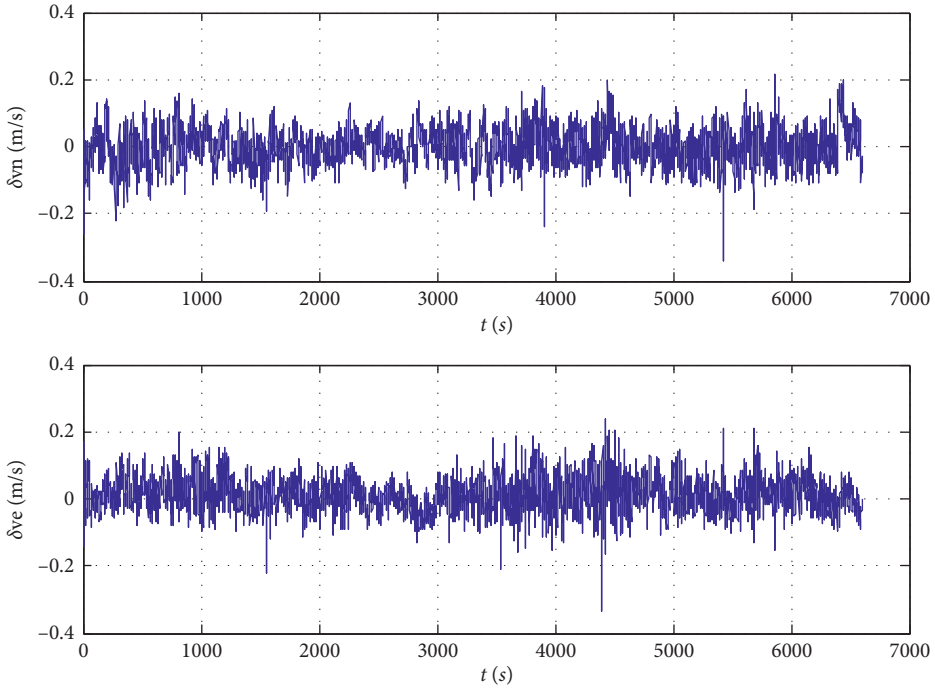


Figure 10: Velocity errors of SINS/DVL in the navigation frame.

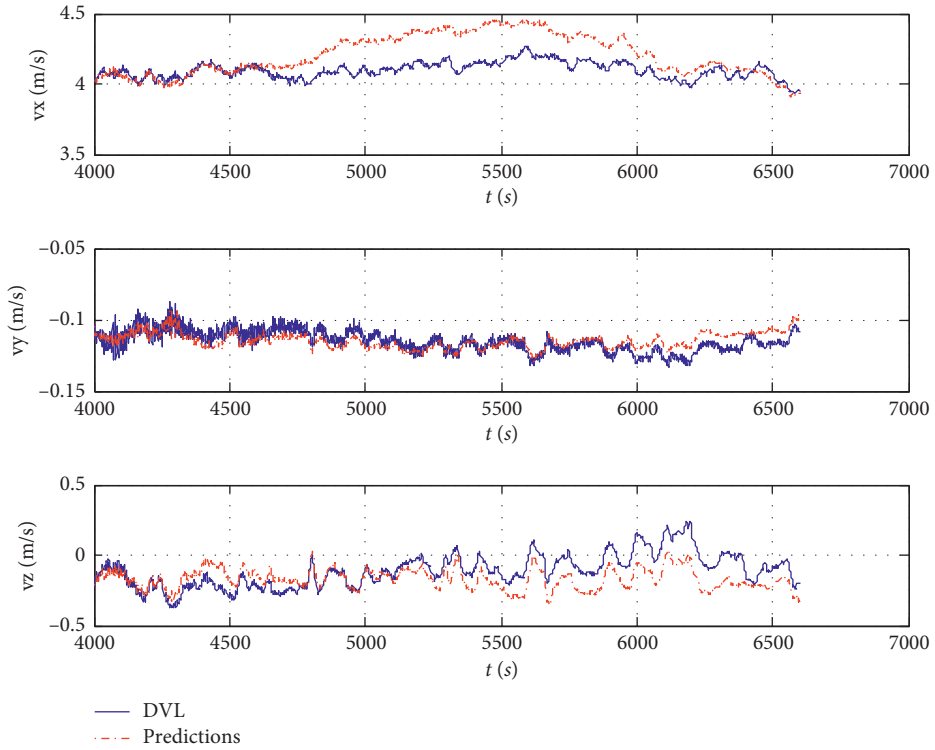


Figure 11: Body frame velocity obtained from the DVL and the neural network.

former data was employed to train the neural network. Therefore, the network is able to forecast the body frame velocity which can be used in the subsequent integrated navigation.

The body frame velocity obtained by DVL can be regarded as the actual value. Figure 11 compares the body frame velocity from the DVL and the network. As can be seen from the figure, the predicted values match well with

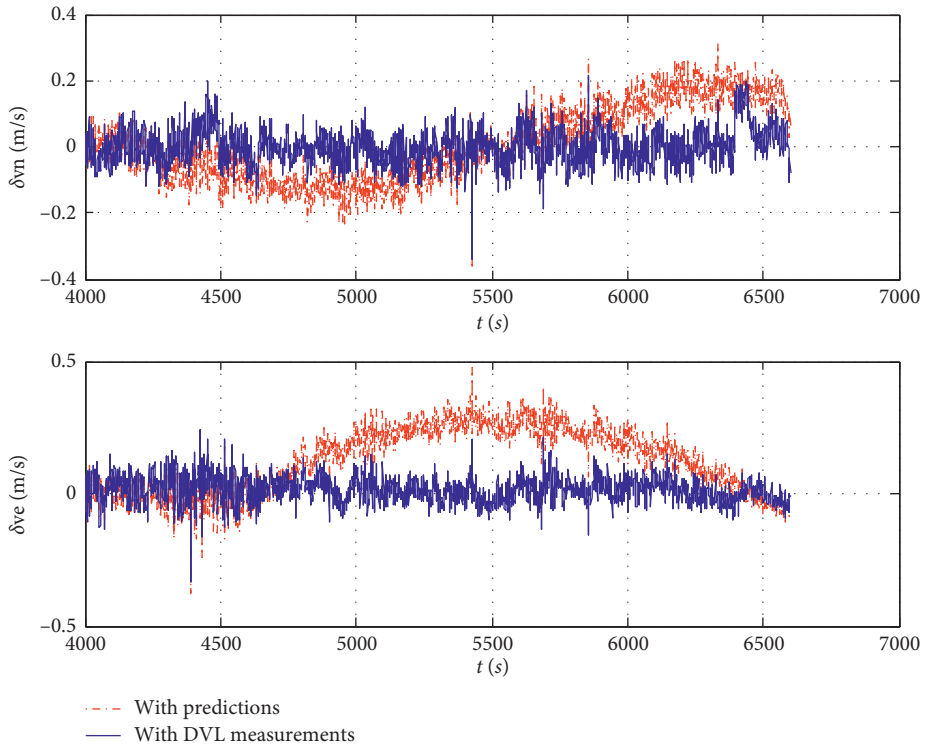


Figure 12: Velocity errors of SINS/DVL in the navigation frame.

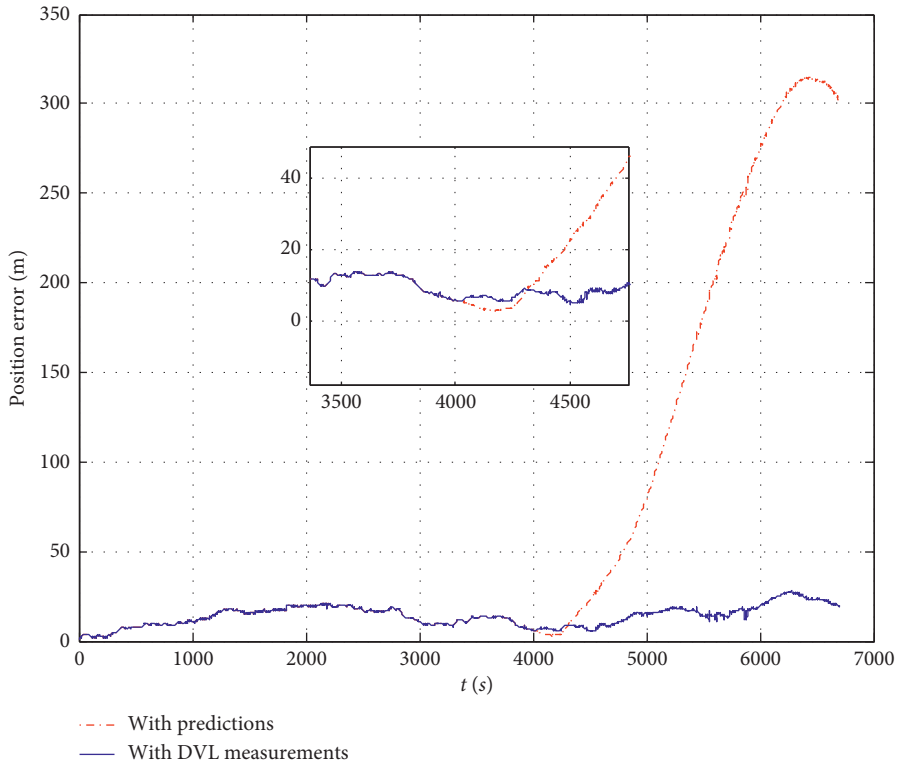


Figure 13: Position errors of the SINS/DVL.

the real value at the beginning, especially the first 200 s–300 s. However, with the increase of time, the difference of the velocities becomes larger.

Compared with the experimental results that DVLs are always available, Figures 12 and 13 show the velocity errors and the position errors of the integrated navigation,

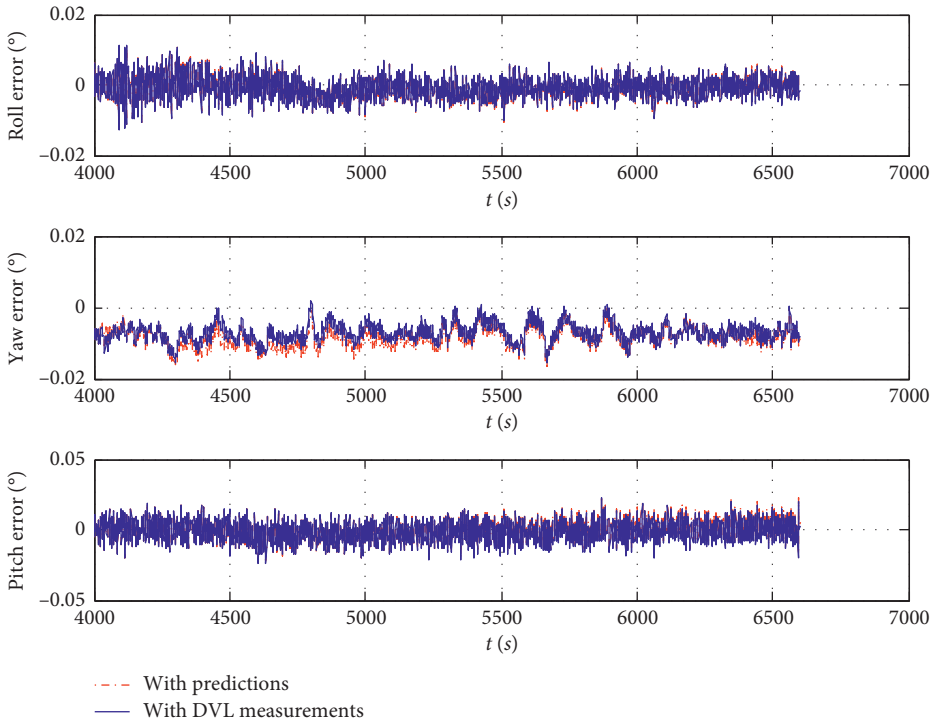


Figure 14: Attitude errors of the SINS/DVL.

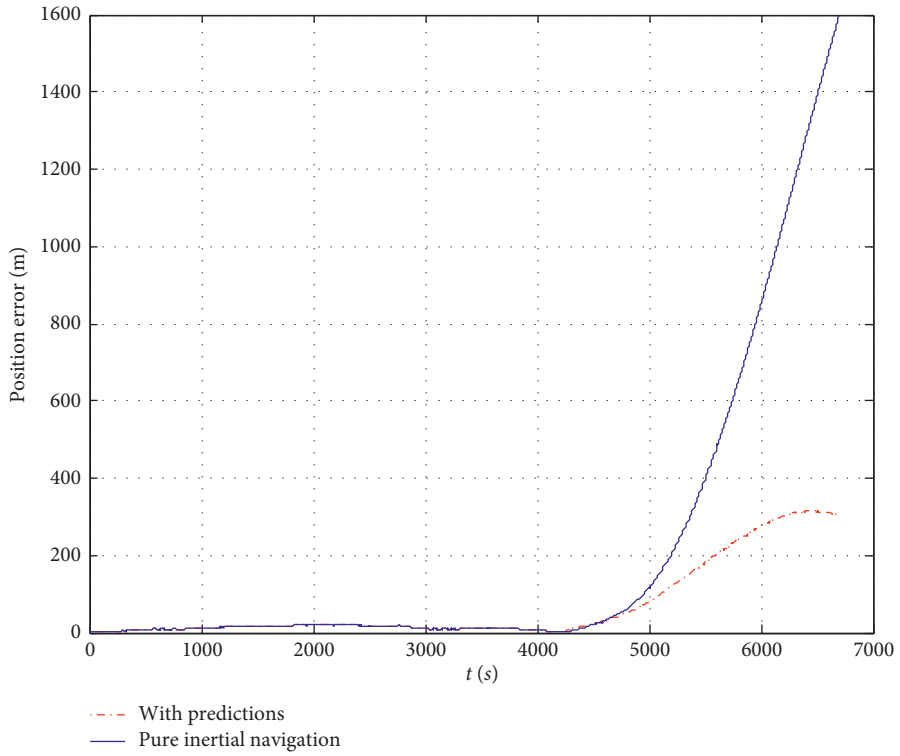


Figure 15: Comparison of the position errors.

respectively. It is shown in Figure 12 that the velocity error of the SINS/DVL with DVL measurements remains around zero. And the velocity error of the SINS/DVL with the predictions also remains around zero at the beginning. However, it diverges with the increase of time. Obviously, the divergence of the velocity error will lead to the growth of the position error. As can be seen from the partial magnification of Figure 13, the positioning error of the SINS/DVL

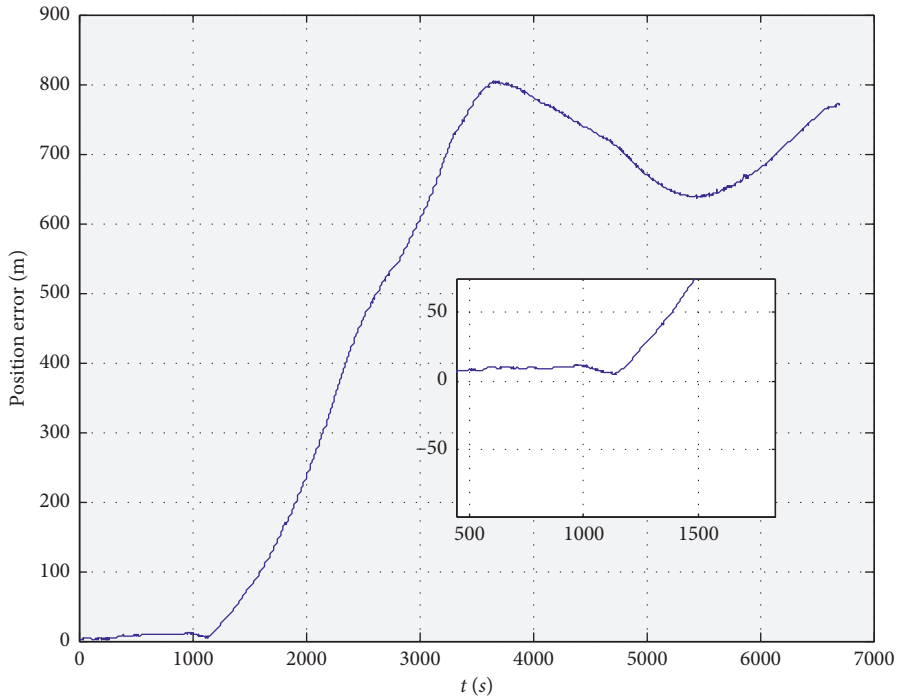


Figure 16: Position error of the SINS/DVL with 1000 s training data.

with predictions is extremely close to that with DVL measurements during the first 200 s–300 s and then increases with time. From the experimental results presented above, it is clearly shown that the proposed neural network-based approach is able to deal with short-term (approximately 200 s–300 s) malfunction of the DVL. However, as the predicting error of the neural network becomes larger, the positioning error of SINS/DVL will increase gradually.

Figure 14 shows the attitude errors of the SINS/DVL. As can be seen from the figure, the influence of the inaccurate velocity predictions on the attitude is not that obvious. The error curves still match well with each other. The final heading error of the SINS integrated with the predictions is around 0.01°. However, it is expected that the difference of the attitude errors will become larger if a low accuracy SINS is employed.

In practical use, when the DVL output is unavailable, the SINS work solely. A comparison has been done by comparing the position error of pure inertial navigation and SINS integrated with DVL predictions. It is clearly shown in Figure 15 that the SINS/DVL integration with DVL predictions is able to reduce the error growth while the position error of the pure inertial navigation increases rapidly. This provides another confirmation to the superiority of the proposed method.

Experiments have also been done by intentionally reducing the amount of the training data. Figure 16 shows the position error of the SINS/DVL with 1000s training data. From partial magnification of the figure, it is found that the SINS/DVL integration is able to maintain accuracy for only about 100 s–150 s. It is shown that, once the amount of the training data is reduced, the accuracy of the predictions will be reduced, too.

5. Conclusions

To deal with the problem of DVL malfunction in SINS/DVL integrated navigation, a neural network-based approach is proposed. When the DVL is available, the measurements from the SINS and DVL are employed to train the network. Once the DVL fails, the well trained network is able to forecast the velocity which can be used in the subsequent integrated navigation. From the experimental results, the following conclusions can be drawn.

(1) With a navigation grade SINS and a DVL, the trained neural network is able to provide credible velocity predictions. But the error of the prediction will increase with time gradually.

(2) With the predicted velocity series from the well trained network, the SINS/DVL integration is able to maintain accuracy for about 200 s–300 s. Therefore, the proposed approach is capable of dealing with DVL short-term malfunction.

(3) If the amount of training data is reduced, the accuracy of DVL predictions may decrease, too.

In this paper, a navigation grade SINS is employed in the experiment. Further investigation is still needed to determine whether this approach is suitable for low cost inertial navigation systems.

Data Availability

The data used to support the findings of this study is owned by the Information Engineering University and so cannot be made freely available. Access to these data will be considered by the author upon request, with the permission of the

Research Management Department of Information Engineering University.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was supported in part by the National Natural Science Foundation of China (Nos. 41674027 and 41604011).

References

[1] D. H. Won, E. Lee, M. Heo, S. Sung, J. Lee, and Y. J. Lee, "GNSS integration with vision-based navigation for low GNSS visibility conditions," *GPS Solutions*, vol. 18, no. 2, pp. 177–187, 2014.

[2] T. Zhang, L. Chen, and Y. Li, "AUV underwater positioning algorithm based on interactive assistance of SINS and LBL," *Sensors*, vol. 16, no. 42, 2016.

[3] J. C. Kinsey and L. L. Whitcomb, "In situ alignment calibration of attitude and Doppler sensors for precision underwater vehicle navigation: theory and experiment," *IEEE Journal of Oceanic Engineering*, vol. 32, no. 2, pp. 286–299, 2007.

[4] A. Karmozdi, M. Hashemi, and H. Salarieh, "Design and practical implementation of kinematic constraints in inertial navigation system-Doppler velocity log (INS-DVL)-based navigation," *Navigation*, vol. 65, no. 4, pp. 629–642, 2018.

[5] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: a review," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 1, pp. 131–149, 2014.

[6] X. Pan and Y. Wu, "Underwater Doppler navigation with self-calibration," *Journal of Navigation*, vol. 69, no. 2, pp. 295–312, 2016.

[7] W. Li, L. Zhang, F. Sun, L. Yang, M. Chen, and Y. Li, "Alignment calibration of IMU and Doppler sensors for precision INS/DVL integrated navigation," *Optik*, vol. 126, no. 23, pp. 3872–3876, 2015.

[8] W. Li, L. Yang, L. Zhang, M. Chen, and K. Tang, "A robust method for alignment calibration of an inertial measurement unit (IMU) and Doppler sensors," *Lasers in Engineering*, vol. 34, pp. 93–106, 2016.

[9] L. Chang, Y. Li, and B. Xue, "Initial alignment for a Doppler Velocity Log-aided Strapdown Inertial Navigation System with limited information," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 329–338, 2016.

[10] "Workhorse navigator Doppler Velocity Log (DVL) ," 2013, http://www.teledynemarine.com/Lists/Downloads/navigator_datasheet_lr.pdf.

[11] L. Zhao, X. Liu, L. Wang, Y. Zhu, and X. Liu, "A pretreatment method for the velocity of DVL based on the motion constraint for the integrated SINS/DVL," *Applied Sciences*, vol. 6, no. 79, 2016.

[12] Q. Wang, X. Cui, Y. Li, and F. Ye, "Performance enhancement of a USV INS/CNS/DVL integration navigation system based on an adaptive information sharing factor federated filter," *Sensors*, vol. 17239 pages, 2017.

[13] Y. Yao, X. Xu, Y. Li, and T. Zhang, "A hybrid IMM based INS/DVL integration solution for underwater vehicles," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5459–5470, 2019.

[14] P. Liu, B. Wang, Z. Deng, and M. Fu, "INS/DVL/PS tightly coupled underwater navigation method with limited DVL measurements," *IEEE Sensors Journal*, vol. 18, no. 7, pp. 2994–3002, 2018.

[15] D. Wang, X. Xu, Y. Yao, T. Zhang, and Y. Zhu, "A novel SINS/DVL tightly integrated navigation method for complex environment," *IEEE Transactions on Instrumentation and Measurement*, In press, 2019.

[16] R. Eliav and I. Klein, "INS/Partial DVL measurements fusion with correlated process and measurement noise," in *Proceedings of 5th International Electronic Conference on Sensors and Applications*, vol. 4, Amsterdam, Netherlands, November 2019.

[17] T. Yoo, M. Kim, S. Yoon, and D. Kim, "Performance enhancement for conventional tightly coupled INS/DVL navigation system using regeneration of partial DVL measurements," *Journal of Sensors*, vol. 2020, Article ID 5324349, 15 pages, 2020.

[18] A. Tal, I. Klein, and R. Katz, "Inertial navigation system/Doppler velocity log (INS/DVL) fusion with partial DVL measurements," *Sensors*, vol. 17, no. 415, 2017.

[19] Y. Zhu, X. Cheng, J. Hu, L. Zhou, and J. Fu, "A novel hybrid approach to deal with DVL malfunctions for underwater integrated navigation systems," *Applied Sciences*, vol. 7, no. 759, 2017.

[20] Z. Wu and W. Wang, "INS/magnetometer integrated positioning based on neural network for bridging long-time GPS outages," *GPS Solutions*, vol. 23, no. 3, 88 pages, 2019.

[21] G. Wang, X. Xu, Y. Yao, and J. Tong, "A novel BPNN-based method to overcome the GPS outages for INS/GPS system," *IEEE Access*, vol. 7, pp. 82134–82143, 2019.

[22] Y. Yao, X. Xu, C. Zhu, and C.-Y. Chan, "A hybrid fusion algorithm for GPS/INS integration during GPS outages," *Measurement*, vol. 103, pp. 42–51, 2017.

[23] Q. Y. Wang, K. Liu, Z. Sun, and M. Zhang, "Research into the high-precision marine integrated navigation method using INS and star sensors based on time series forecasting BPNN," *Optik*, vol. 172, pp. 494–508, 2018.

[24] E. Cadenas, W. Rivera, R. Campos-Amezcua, and C. Heard, "Wind speed prediction using a univariate ARIMA model and a multivariate NARX model," *Energies*, vol. 9109 pages, 2016.

[25] Z. Boussaada, O. Curea, A. Remaci, H. Camblong, and N. Mrabet Bellaaj, "A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation," *Energies*, vol. 11620 pages, 2018.

[26] L. Wang, L. Fu, and M. Xin, "Sculling compensation algorithm for SINS based on two-time scale perturbation model of inertial measurements," *Sensors*, vol. 18282 pages, 2018.

[27] Z. Ye and M. K. Kim, "Predicting electricity consumption in a building using an optimized back-propagation and Levenberg-Marquardt back-propagation neural network: case study of a shopping mall in China," *Sustainable Cities and Society*, vol. 42, pp. 176–183, 2018.

[28] W. Li, W. Wu, J. Wang, and M. Wu, "A novel backtracking navigation scheme for autonomous underwater vehicles," *Measurement*, vol. 47, pp. 496–504, 2014.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 6352539, 12 pages
https://doi.org/10.1155/2020/6352539

Research Article

Optimization of Order-Picking Problems by Intelligent Optimization Algorithm

Zhong-huan Wu,[1] **Hong-jie Chen**,[2] **and Jia-jia Yang**[3]

[1]*Department of Business Administration, Huashang College Guangdong University of Finance & Economics, Guangzhou 510000, China*
[2]*School of Business Administration, South China University of Technology, Guangzhou 510000, China*
[3]*Department of Public Policy, King's College London, London WC2R 2LS, UK*

Correspondence should be addressed to Hong-jie Chen; raphel888@gmail.com

Received 29 April 2020; Accepted 3 June 2020; Published 23 July 2020

Guest Editor: Wen-Tsao Pan

To improve the efficiency of warehouse operations, reasonable optimization of picking operations has become an important task of the modern supply chain. For the purpose of optimization of order picking in warehouses, a new fruit fly optimization algorithm, particle swarm optimization, random weight, and weight decrease model are used to solve the mathematical model. Further optimization is achieved through the analysis of the warehouse shelves and screening of the optimal solution of the picking time. In addition, simulation experiments are conducted in the MATLAB environment through programming. The shortest picking time is found out and chosen as an optimized method by taking advantage of the effectiveness of these six algorithms in the picking optimization and comparing the data obtained under the simulation. The result shows that the optimization capacity of RWFOA is better and the picking efficiency is the best.

1. Introduction

As a part of the logistics, the efficiency of the automated warehouse is largely dependent on the efficiency of order picking. Therefore, the picking plays an important role in the automated warehouse for improving the efficiency of picking operation. Although the automated stereoscopic warehouse provides more orderly and standardized management and the error rate is also small, for small batch warehouse, frequent warehousing, and warehouses with various products, logistics storage becomes more stringent, and requirements for the efficiency of the logistics are higher, and thus, the efficiency of picking needs to be improved.

There are a lot of algorithms for optimization of order picking, all of which have made minor or major contributions to the optimization of order picking [1–6]. At present, the ant colony algorithm [7], genetic algorithm [8] and multipopulation fruit fly optimization algorithm [9], which have been used to solve the picking operation problems, yielded good results. Based on the existing research, we will use the new fruit fly optimization algorithm and particle swarm optimization to solve the mathematical model.

The main structure of this paper is as follows: Section 1 introduces the motive and purpose of this study; Section 2 presents the literature review; Section 3 introduces research methods—original FOA, original particle swarm algorithm (PSA), random weight algorithm, weight decrease, and related literature; Section 4 introduces case description; Section 5 presents results and discussion; and Section 6 puts forward the research conclusions and suggestions.

2. Literature Review

2.1. Order Picking. Across the various operations in a warehouse, order picking is the most time-consuming operation in general [10] and accounts for around 55–75 percent of total warehousing costs [11]. Therefore, order

picking has the highest priority for productivity improvement [12].

Order picking is a particular case of the traveling salesman problem (TSP). This problem, introduced by Dantzig et al. [13], is one of the most studied problems in operations research. Efficient algorithms have been designed for the TSP [14]. Therefore, in order to improve the performance of order picking, reducing travel time is critical. Since the travel distance is proportional to travel time for picker-to-parts system [15], minimizing the travel distance (total or average) of a picking tour is often considered as an imperative factor to reduce travel time and consequently improve warehouse operation efficiency [12]. There are four methods to reduce the travel distance of an order picker [12]: storage location assignment, warehouse zoning, order batching, and pick-routing methods. And this paper focuses on the pick-routing methods.

To most order-picking research studies, optimization algorithms are still the center of routing studies [16]. To pursue the optimal order-picking route in a typical rectangular, the order-picking routing problem is considered as the STSP (Steiner traveling salesman problem) [12, 17]. There are two general methods to solve the STSP: the first method is to reformulate an STSP into the classic TSP by computing the shortest paths between every pair of required nodes (e.g., Renaud and Ruiz [17]) and the second one for the solution of a STSP is by using dedicated algorithms (e.g., Lucie Pansart [4]). The latter method is preferred to the former.

The dedicated algorithms include dynamic programming, integer programming, and branch and bound method. Although this kind of algorithm can get the exact solution, the calculation time is long and it is seldom used in the practical application [18]. The common approximation algorithms are the insertion algorithm, the r-opt algorithm, and the nearest neighbour algorithm. Although this kind of algorithm can quickly get a feasible solution to the optimal solution, the degree of its close to the optimal solution is not satisfactory [19]. Intelligent optimization algorithm is a more effective algorithm to solve this problem in recent years [20]. These algorithms are mainly genetic algorithms [8], ant colony algorithm [7], particle swarm optimization [21], and modified FOA, e.g., MSFOA [22] and IFOA4WSC [23].

2.2. Automated Stereoscopic Warehouse Model. At present, the shelves of the automatic stereoscopic warehouse are mainly fixed shelves, and each row of shelves in the warehouse is equipped with a stacker, which is responsible for picking up a cargo on the shelf. This paper takes some of the shelves in the warehouse as the object of study. In an automated warehouse, the stacker enters from the entrance, performs order picking, and chooses goods according to the programmed procedure. Assuming that there are k cargo spaces on each shelf, the stacker can only get one cargo space for each picking, removes the cargo from the shelve and transports it to the exit, and then returns to the shelves to pick, and the above steps were repeated. As the order of picking is not the same, the time required for each picking is not the same. We need to set the optimal sorting order, so as to minimize the time on picking and to improve the efficiency of order picking.

This study refers to the high-level rack model designed by Professor Ning and Hu [9], and the formulas from (1) to (8) are also proposed by them [9]. The structure is shown in Figure 1.

Where, the position of the column x and tier y can be set to (i, j), the position at the entrance is set to $(0, 0)$, the length of the shelf is D, and the height of the shelf is G. Assuming a shelf has x columns and y tiers, the goods allocation is (Y, j) $(m = 1, 2, 3, \ldots, k)$, and position is (α, β), so

$$\alpha m = \frac{D \cdot i}{X} - \frac{D}{2X}, \tag{1}$$

$$\beta m = \frac{G(j-1)}{Y}. \tag{2}$$

Assuming that the velocity in the horizontal direction of the stacker is v_1 and the velocity in the vertical direction is v_2, and the time spent by two adjacent cargo spaces m (i, j) and $m+1$ (p, q) in the process of picking up the goods in the horizontal direction is t_1 and t_2; the operating equation used is as follows:

$$t_1 = \frac{|\alpha_m - \alpha_{m+1}|}{v_1} = \frac{|i-p|D}{X \cdot v_1}, \tag{3}$$

$$t_2 = \frac{|\beta_m - \beta_{m+1}|}{v_2} = \frac{|j-q|G}{Y \cdot v_2}. \tag{4}$$

Since the horizontal and vertical movements of the stacker occur at the same time, the time of operation at the adjacent cargo space is t_m, and the maximum value for running speed t_1 in the horizontal direction and the running speed t_2 in the vertical direction is given by

$$t_m = \max\{t_1, t_2\}. \tag{5}$$

Then, the k cargo positions are selected and the total running time T_z used by the stacker is given by

$$T_z = \sum_{m=0}^{n} t_m. \tag{6}$$

If the picking time spent by the stacker is the same each time, that is, t_s, then the running time of all the cargo spaces is T_s, as shown in the following equation:

$$T_s = \sum_{m=1}^{n} t_s = k t_s. \tag{7}$$

Thus, the total time T of k cargo spaces is as follows:

$$T = T_z + T_s. \tag{8}$$

Under above circumstances, we will ask for the total time of operation of the automated warehouse stacker and the minimum value T. Six intelligence algorithms including original particle swarm, particle swarm weight decrease, particle swarm random weight, original FOA, fruit fly weight decrease, and fruit fly random weight are used to evaluate the minimum value of T.

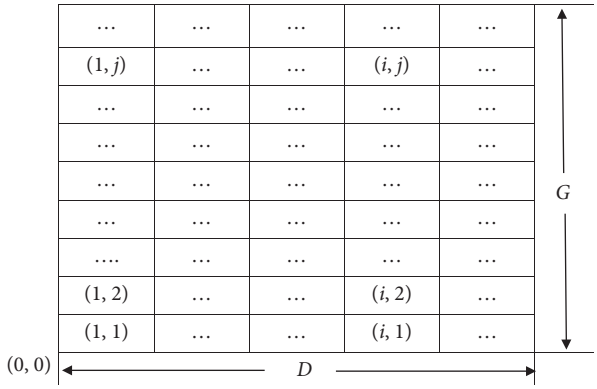


Figure 1: Structural model of high-level rack.

3. Research Methods

Particle swarm optimization is a new algorithm in recent years, which solves the TSP problem, and a good result is obtained [20]. And the fruit fly optimization algorithm (FOA) is a newly developed bioinspired algorithm. The continuous variant version of FOA has been proven to be a powerful evolutionary approach to determining the optima of a numerical function on a continuous definition domain [24]. The FOA and PSO are also easy to program and can be modified to other practical applications. Due to these advantages, they have been used to solve a wide range of optimization problems, including prediction and classification problems [25–27]. However, the FOA and PSO must be modified in order to effectively manage the discrete variables associated with optimization issues. Therefore, RW and WD were integrated into FOA and PSO to improve its advantage and to look for the better optimal order-picking time.

3.1. Fruit Fly Optimization Algorithm (FOA). The original FOA was invented by Professor Pan [24], and the FOA is highly accurate. Many studies will use the FOA to solve the optimization problem. FOA can be used in any field, such as military, engineering, medical science, management, and financial and other fields. It can also be combined with other algorithms, complementing each other. FOA is a new method of global optimization derived from foraging behaviors of fruit flies. Because a fruit fly itself is superior to other animals in perception, it comes close to the food using its olfactory organ, knowing where the food and partners gathered, and then fly to the destination. Following is the original fruit fly algorithm:

(1) Set initial location of fruit flies at random (x and y are two coordinate axes, initial position on coordinates):

$$\begin{aligned} &\text{Init}X \text{ axis}, \\ &\text{Init}Y \text{ axis}. \end{aligned} \tag{9}$$

(2) Random directions and distance of fruit flies searching for food relying on good sense of smell, which is equivalent to the initial location of the fruit flies plus random flight distance:

$$\begin{aligned} Xi &= X \text{ axis} + \text{Random Value}, \\ Yi &= Y \text{ axis} + \text{Random Value}. \end{aligned} \tag{10}$$

(3) As the location of food cannot be obtained, estimate the distance (Di) to the origin first, and then calculate the decision value of Smelli (Si), and this value is the reciprocal of Di:

$$\begin{aligned} Di &= \sqrt{X_i^2 + Y_i^2}, \\ Si &= \frac{1}{Di}. \end{aligned} \tag{11}$$

(4) Substitute decision value of Smelli (Si) into the above function to get the Smelli of location of fruit flies:

$$\text{Smelli} = \text{Function}(Si). \tag{12}$$

(5) Locate the fruit fly with the best Smelli from fruit flies (max):

$$[\text{bestSmellbestIndex}] = \max(\text{Smelli}). \tag{13}$$

(6) Retain the smell best and X-axis and Y-axis, and the fruit flies will fly to this position.

$$\begin{aligned} \text{Smellbest} &= \text{bestSmell}, \\ X \text{ axis} &= X(\text{bestIndex}), \\ Y \text{ axis} &= Y(\text{bestIndex}). \end{aligned} \tag{14}$$

(7) Enter into iterative optimization, repeat steps 2–5, and judge whether the Smelli is superior to the Smelli of the previous iteration, if yes, execute step 6.

The foraging process of a fruit fly group is illustrated in Figure 2 [25].

In view of the optimization of picking in this paper, we know that the range of search distance of the original fruit fly in the coordinates is limited, which leads to the weak optimal performance. If the weight is added to the original FOA, the search range of fruit flies will be enlarged, which will greatly enhance the optimization ability of fruit flies.

3.2. Particle Swarm Optimization (PSO). Particle swarm algorithm [28] is a kind of random search algorithm, which is a new intelligent optimization technique, and can converge on the global optimal solution with larger probability. PSO is derived from the study of predatory behavior of birds: a group of birds randomly search for food in a region, all birds know how far they are away from the food, and then the simplest and most effective strategy is to search the surrounding area of birds that is closest to the food. Inspired by this model, it is applied to solve the optimization problem. The basic PSO is as follows:

(1) Suppose in a D-dimensional target search space, N particles form a community, where the i-th particle is expressed as a D-dimensional vector:

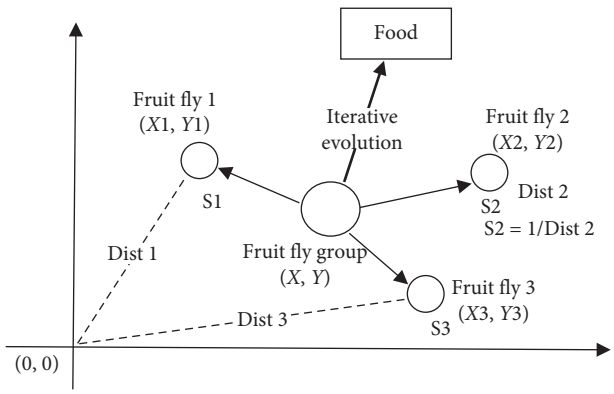


Figure 2: Foraging process of a fruit fly group.

$$Xi = (xi1, xi2, \ldots, xiD), \quad i = 1, 2, \ldots, N. \tag{15}$$

(2) The "flying" velocity of the i-th particle is also a D-dimensional vector, denoted as follows:

$$Vi = (vi1, vi2, \ldots, viD), \quad i = 1, 2, \ldots, N. \tag{16}$$

(3) The optimal position of the i-th particle searched so far is called the individual extremum, denoted as follows:

$$P\text{best} = (pi1, pi2, \ldots, piD), \quad i = 1, 2, \ldots N. \tag{17}$$

(4) The optimal position of the whole particle swarm searched so far is called the global extremum, denoted as follows:

$$g\text{best} = (gi1, gi2, \ldots giD). \tag{18}$$

When these two optimal values are found, the particles will update their speed and position according to the following two formulas:

$$\begin{aligned} Vij(t+1) &= w * vij(t) + c1r1(t)[pij(t) - xij(t)] \\ &\quad + c2r2(t)[pgj(t) - xij(t)], \\ Xij(t+1) &= xij(t) + vij(t+1). \end{aligned} \tag{19}$$

where $c1$ and $c2$ are learning factors, also known as acceleration constants; $r1$ and $r2$ are uniform random numbers within the scope [0, 1], $i = 1, 2, \ldots, D$; vij is the velocity of the particle, $vij \in [-v_{\max}, v_{\max}]$, in which $v_{\max}$ is a constant, and the speed of the particle is set by the user. $r1$ and $r2$ are random numbers between 0 and 1, which increases the randomness of particle flight. W refers to the extent to retain the original speed the greater of the w is, the stronger ability of global convergence and weak ability of local convergence, and the reverse is also true.

The foraging process of a particle swarm group is illustrated in Figure 3 [28].

3.3. Weight Decrease (WD). In this paper, we refer to the weight decrease and random weight algorithm mentioned by Gao [29], and the WD is based on the original PSO and FOA. The larger weighting factor is beneficial to jump out of the local minimum point and is convenient for global search, and the smaller inertia factor is beneficial to the accurate local search of the current search area, which is better for algorithm convergence. Therefore, for the phenomenon that PSO and FOA are easy to get premature and the algorithms are easy to oscillate near the global optimal solution at a later stage, the weight of linear change can be used to reduce the inertia weight linearly from the maximum $\omega_{\max}$ to the minimum $\omega_{\min}$. The formula for the number of iterations with the algorithm is $\omega = \omega_{\max} - (t * (\omega_{\max} \times \omega_{\min}))/t_{\max}$, where $\omega_{\max}$, $\omega_{\min}$, respectively, represent the maximum and minimum values of ω, t indicates the current number of iterations, and $t_{\max}$ indicates the maximum number of iterations.

The weight decrease method can adjust the global and local search capabilities of PSO and FOA, but it still has two shortcomings: first, the local search ability of early iterations is relatively weak, even if the initial particles are close to the global optimal point, it will be missed, and the global search ability will become weak at the later stage, so the program is caught in the local optimal value; second, the maximum number of iterations is difficult to predict, which will affect the adjustment function of the algorithm [30].

3.4. Random Weight (RW). The random weight algorithm is based on the original PSO and FOA. In this paper, the RW refers to taking ω value randomly, so that the impact of the historical speed of particles on the current speed is random. In order to accord with a random number that is randomly distributed ($N(\mu, \hat{\sigma 2})$), the shortcomings of ω linear decrease can be overcome from two aspects. In addition, we can apply the random direction and distance of fruit flies in FOA to increase its global search ability. If the evolution is close to the most power consumption at the beginning of evolution, the linearity of ω decreases, so the algorithm will not converge to the best point, and the random generation of ω can overcome this limitation. ω is calculated as follows:

$$\begin{aligned} \omega &= \mu + \sigma * N(0, 1), \\ \mu &= \mu\min + (\mu\max - \mu\min) * \text{rand}(0, 1), \end{aligned} \tag{20}$$

where N (0, 1) represents the random number of the standard normal distribution, and rand (0, 1) represents a random number between 0 and 1. Researches show that RW-based PSO and FOA algorithm can avoid the local optimum to a certain extent.

4. Case Description

Suppose the length of the shelf is 80 m, the height is 8 m, and a complete shelf has 40 rows and 5 tiers. The lateral movement speed Va of the stacker is 1 m/s and longitudinal velocity Vb is 0.2 m/s. The picking time of each cargo space is assumed to be 10 s. According to the above optimization algorithms, Popsize1 = 5 and Popsize2 = 10, that is, the number of all populations is Popsize1 × Popsize2 = 50. The largest number of iterations of six algorithms is 1000 times. In terms of FOA parameter, the random initial position of a

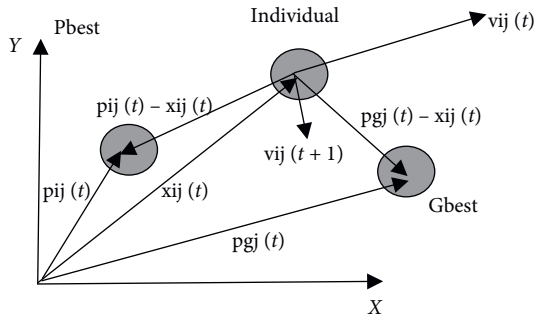


Figure 3: Process of a Particle swarm group.

fruit fly swarm is [−5, 5], fruit flies searching for food randomly, and the distance interval is [−50, 50]; in terms of PSO parameter, $C1$ and $C2$ are set to be 1.49445, $V_{\max}$ and $V_{\min}$ are set to be 1, $\text{pop}_{\max}$ is set to be 50, and $\text{pop}_{\min}$ is set to be −50; six algorithms are run independently of 20 times.

We apply the RW and WD mathematical model to FOA and PSO and take the individual position as the encoding object, and the length of the code is a randomly generated cargo space number. We then assume that the number of subpopulations is Popsize1, the number of individuals in each population is Popsize2, and the number of individuals in all populations is Popsize1 × Popsize2, and then the population quantity is Popsize1 × Popsize2. If m cargo spaces are randomly generated, then the coding scheme of No. b fruit flies in No. a subpopulation is shown in Table 1.

In order to check the optimization capability of the proposed FOA and PSO, two groups of 10 cargo spaces and 20 cargo spaces are randomly generated, as shown in Tables 2–5.

Table 1: Coding scheme of fruit flies randomly generated.

Cargo space	1	2	...	7	8	$m-1$	m
Tier	xab1	xab2	...	xab7	xab8	xab $(n-1)$	xabn
Row	yab1	Yab2	...	yab7	yab8	yab $(n-1)$	yabn
Smell	Sab1	Sab2	...	Sab7	Sab8	Sab $(n-1)$	Sabn

Table 2: 10 cargo spaces in Group 1.

Tier	24	32	40	26	17	12	38	15	7	29
Row	4	1	4	5	3	2	2	3	5	1

Table 3: 10 cargo spaces in Group 2.

Tier	22	32	12	28	40	12	25	34	17	27
Row	2	1	3	3	1	5	3	2	2	4

Table 4: 20 cargo spaces in Group 1.

Tier	20	32	42	35	22	6	19	43	18	38
Row	3	2	3	5	1	3	2	1	4	5
Tier	25	10	44	16	41	17	28	3	7	15
Row	1	5	2	4	4	2	1	3	5	4

Table 5: 20 cargo spaces in Group 2.

Tier	8	20	22	6	12	13	28	14	34	4
Row	4	3	2	4	2	1	3	6	3	1
Tier	33	11	32	3	36	27	40	4	22	25
Row	4	2	2	5	3	1	3	4	2	6

5. Results and Discussion

The results (subfigures) are shown below in proper order: PSO (upper left), WDPSO (center left), RWPSO (lower left), FOA (upper right), WDFOA (center right), and RWFOA (lower right).

5.1. Iteration Verification of 10 Cargo Spaces in Group 1. According to the data of Figure 4, the optimal search time of PSO, WDPSO, and RWPSO is 243 s, 235 s, and 234 s, and the optimal search time of FOA, WDFOA, and RWFOA is 236 s, 228 s, and 226 s.

According to the data of Table 6, the optimal average search time of PSO is 237 s, the optimal search time of FOA is 230 s, and the optimization of FOA is better. The average optimal search time of the original, WD, and RW is 234.5 s, 232 s, and 231 s, respectively, and the optimization of RW is better. Thus, RWFOA is the best.

From the standard deviation in Table 7, RWFOA is the smallest, better than the other five. Therefore, PSO algorithm is featured with good accuracy and speed, but its optimization performance is worse than FOA. For six different algorithms, the optimization of RWFOA is relatively good.

The optimal picking time of 10 cargo spaces is 226 s, and the corresponding picking order is as follows: 8–5–6–9–2–1–10–7–4–3.

5.2. Iteration of 10 Cargo Spaces in Group 2. According to the data of Figure 5, the optimal search time of PSO, WDPSO, and RWPSO is 216 s, 214 s, and 212 s, and the optimal search time of FOA, WDFOA, and RWFOA is 209 s, 208 s, and 207 s.

According to the data of Table 8, the optimal average search time of PSO is 214 s, the optimal search time of FOA is 208 s, and the optimization of FOA is better. The average optimal search time of the original, WD, and RW is 212.5 s, 211 s, and 209.5 s, respectively, and the optimization of RW is better. Thus, RWFOA is the best.

From the standard deviation in Table 9, RWFOA is the smallest, better than the other five. Therefore, PSO algorithm is featured with good accuracy and speed, but its optimization performance is worse than FOA. For six different algorithms, the optimization of RWFOA is relatively good.

The optimal picking time of 10 cargo spaces is 207 s, and the corresponding picking order is as follows: 3–2–1–8–5–7–6–10–9–4.

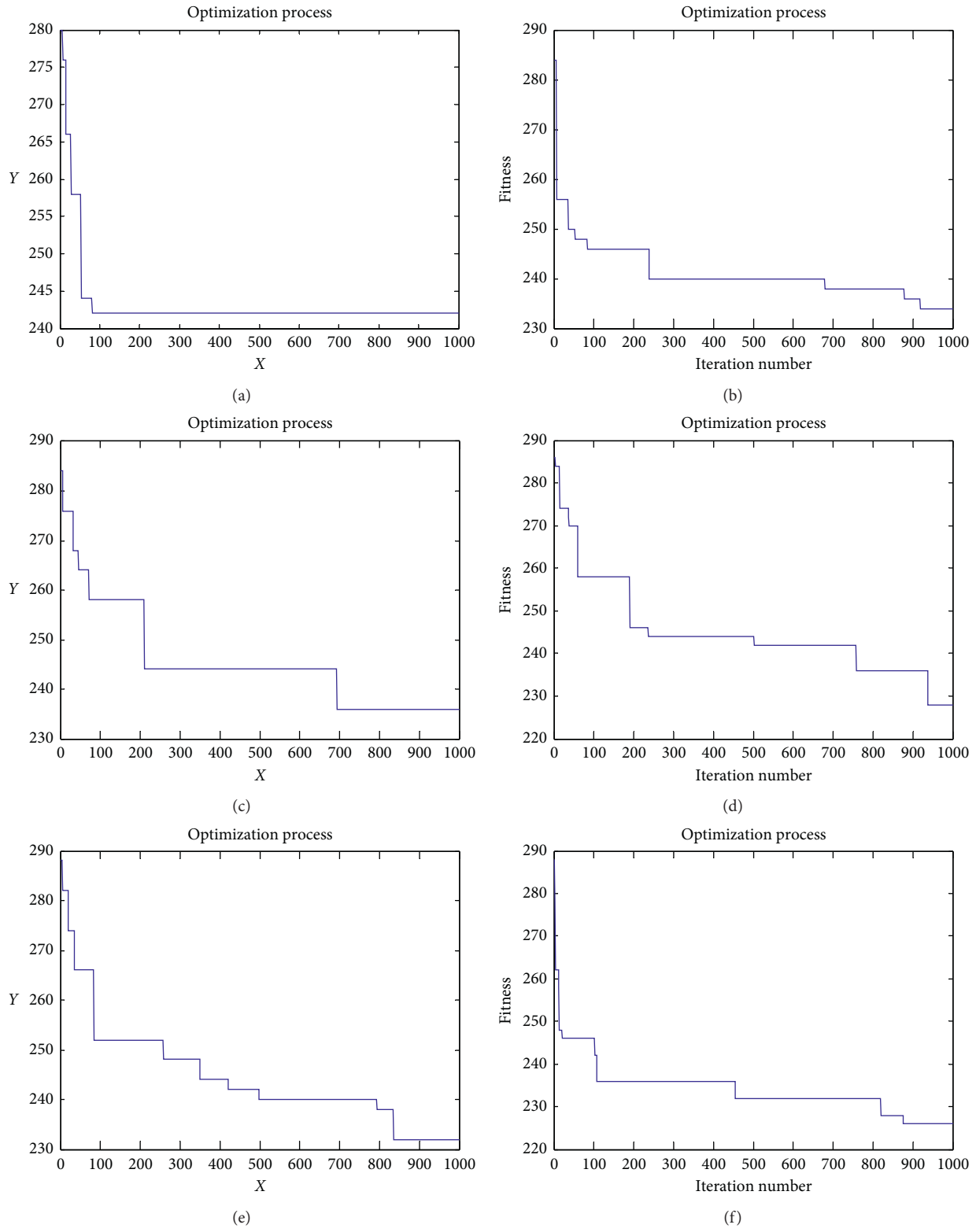


Figure 4: Iterative changes of 10 cargo spaces in Group 1: (a) PSO; (b) FOA; (c) WDPSO; (d) WDFOA; (e) RWPSO; (f) RWFOA. Note: X is the total time of iteration and Y is 1000 iterations.

5.3. Iteration of 20 Cargo Spaces in Group 1. According to the data of Figure 6, the optimal search time of PSO, WDPSO, and RWPSO is 553 s, 550 s, and 549 s, and the optimal search time of FOA, WDFOA, and RWFOA is 545 s, 543 s, and 541 s.

According to the data of Table 10, the optimal average search time of PSO is 550 s, the optimal search time of FOA is 543 s, and the optimization of FOA is better. The average optimal search time of the original, WD, and RW is 549 s,

Table 6: Average of picking time of 10 cargo spaces in Group 1.

Algorithm	Original (s)	WD (s)	RW (s)	Average (s)
PSO	243	235	234	237
FOA	236	228	226	230
Average	234.5	232	231	

Table 7: Standard deviation of picking time of 10 cargo spaces in Group 1.

	Algorithm SD		
Algorithm	Original	WD	RW
PSO	6.6	6.4	6.2
FOA	4.9	3.8	3.7

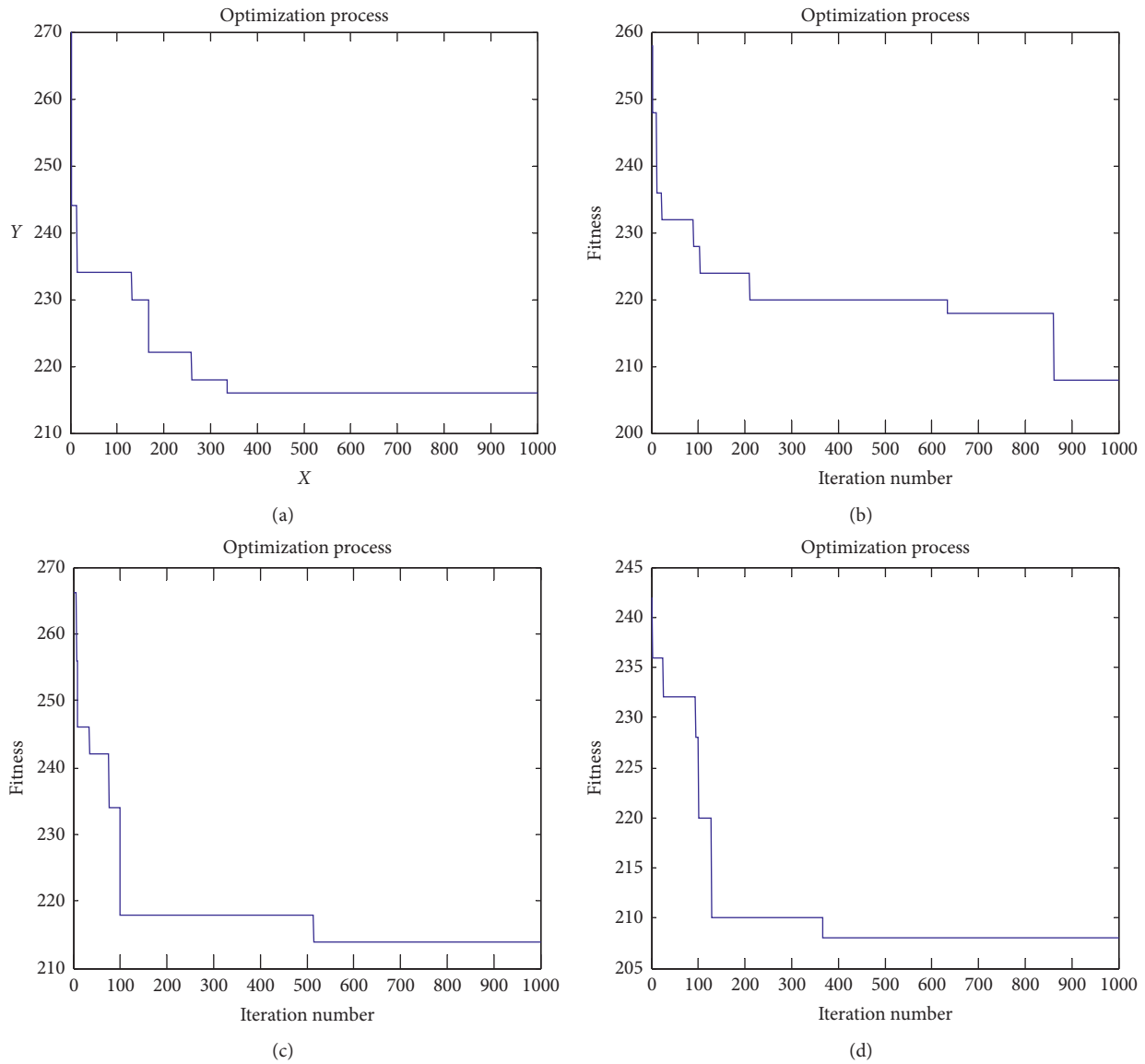
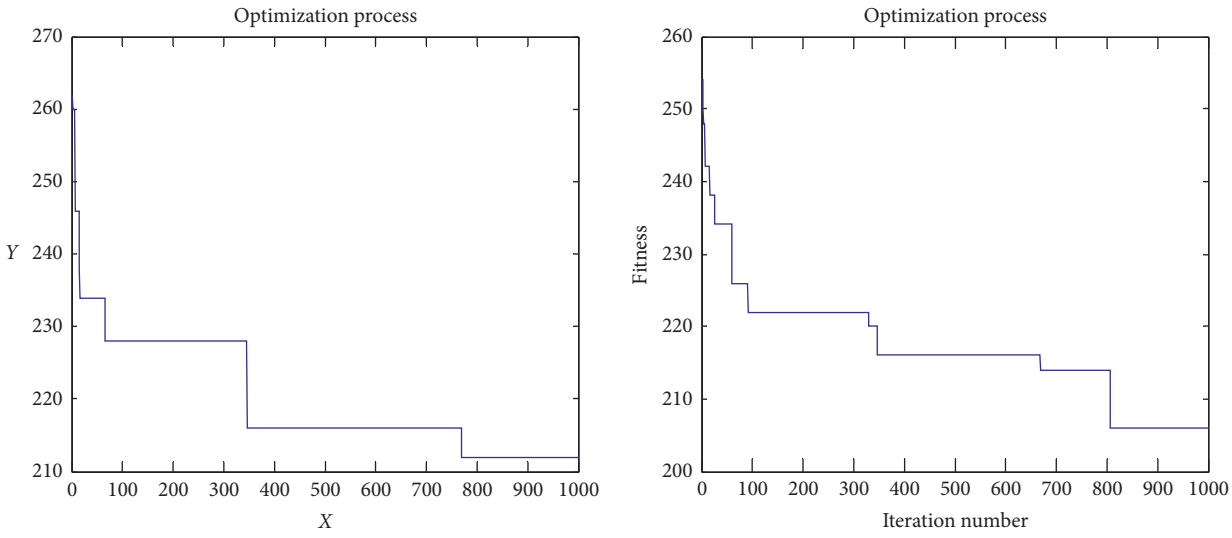


Figure 5: Continued.



(e) (f)

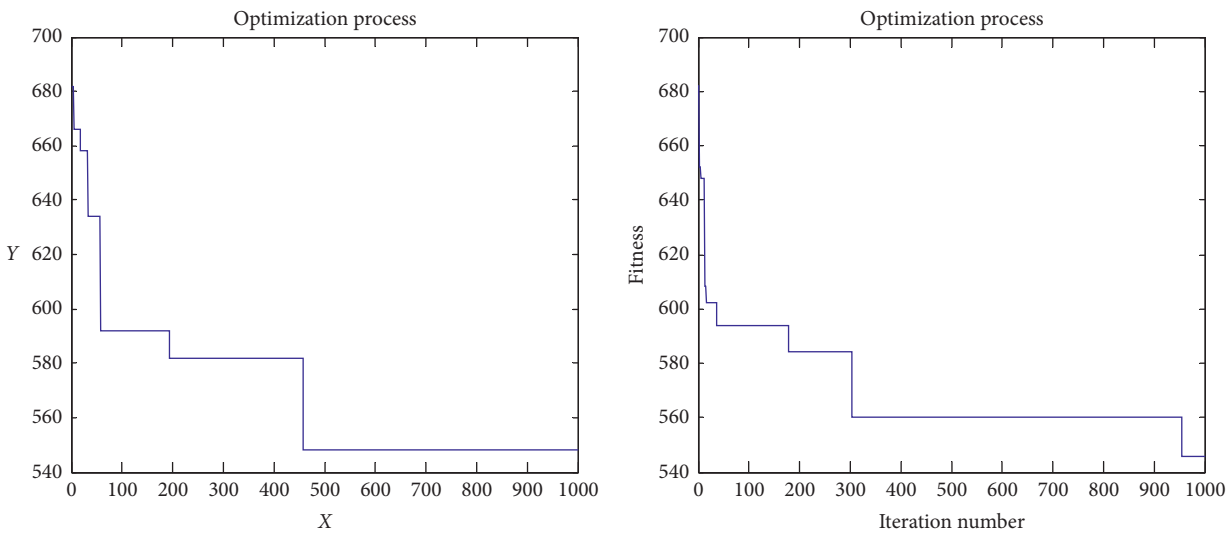
Figure 5: Iterative changes of 10 cargo spaces in Group 2: (a) PSO; (b) FOA; (c) WDPSO; (d) WDFOA; (e) RWPSO; (f) RWFOA. Note: X is the total time of iteration and Y is 1000 iterations.

Table 8: Average of picking time of 10 cargo spaces in Group 2.

Algorithm	Original (s)	WD (s)	RW (s)	Average (s)
PSO	216	214	212	214
FOA	209	208	207	208
Average	212.5	211	209.5	

Table 9: Standard deviation of picking time of 10 Cargo spaces in Group2.

	Algorithm SD		
Algorithm	Original	WD	RW
PSO	6.3	5.9	5.5
FOA	4.8	4.7	4.6



(a) (b)

Figure 6: Continued.

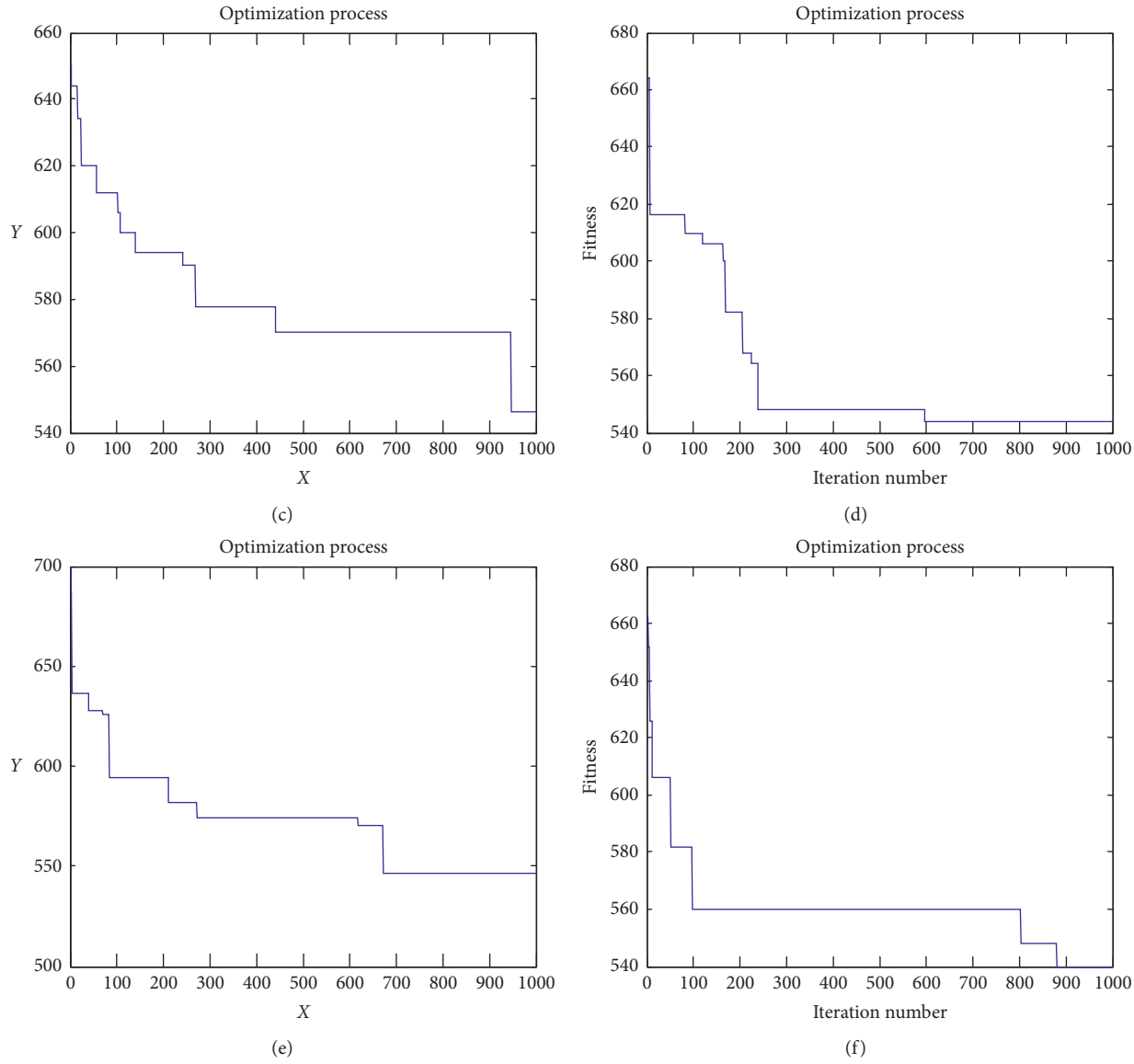


Figure 6: Iterative changes of 20 cargo spaces in Group 1: (a) PSO; (b) FOA; (c) WDPSO; (d) WDFOA; (e) RWPSO; (f) RWFOA. Note: X is the total time of iteration and Y is 1000 iterations.

546.5 s, and 545 s, respectively, and the optimization of RW is better. Thus, RWFOA is the best.

From the standard deviation in Table 11, RWFOA is the smallest, better than the other five. Therefore, PSO algorithm is featured with good accuracy and speed, but its optimization performance is worse than FOA. For six different algorithms, the optimization of RWFOA is relatively good.

The optimal picking time of 20 cargo spaces is 541 s, and the corresponding picking order is as follows: 9–12–19–13–20–15–4–17–8–1–10–2–16–5–14–3.

5.4. Iteration of 20 Cargo Spaces in Group 2. According to the data of Figure 7, the optimal search time of PSO, WDPSO, and RWPSO is 544 s, 542 s, and 540 s, and the optimal search

Table 10: Average of picking time of 20 cargo spaces in Group 1.

Algorithm	Original (s)	WD (s)	RW (s)	Average (s)
PSO	553	550	549	550
FOA	545	543	541	543
Average	549	546.5	545	

Table 11: Standard deviation of picking time of 20 cargo spaces in Group 1.

	Algorithm SD		
Algorithm	Original	WD	RW
PSO	18.4	15.3	15.0
FOA	15.6	14.3	11.2

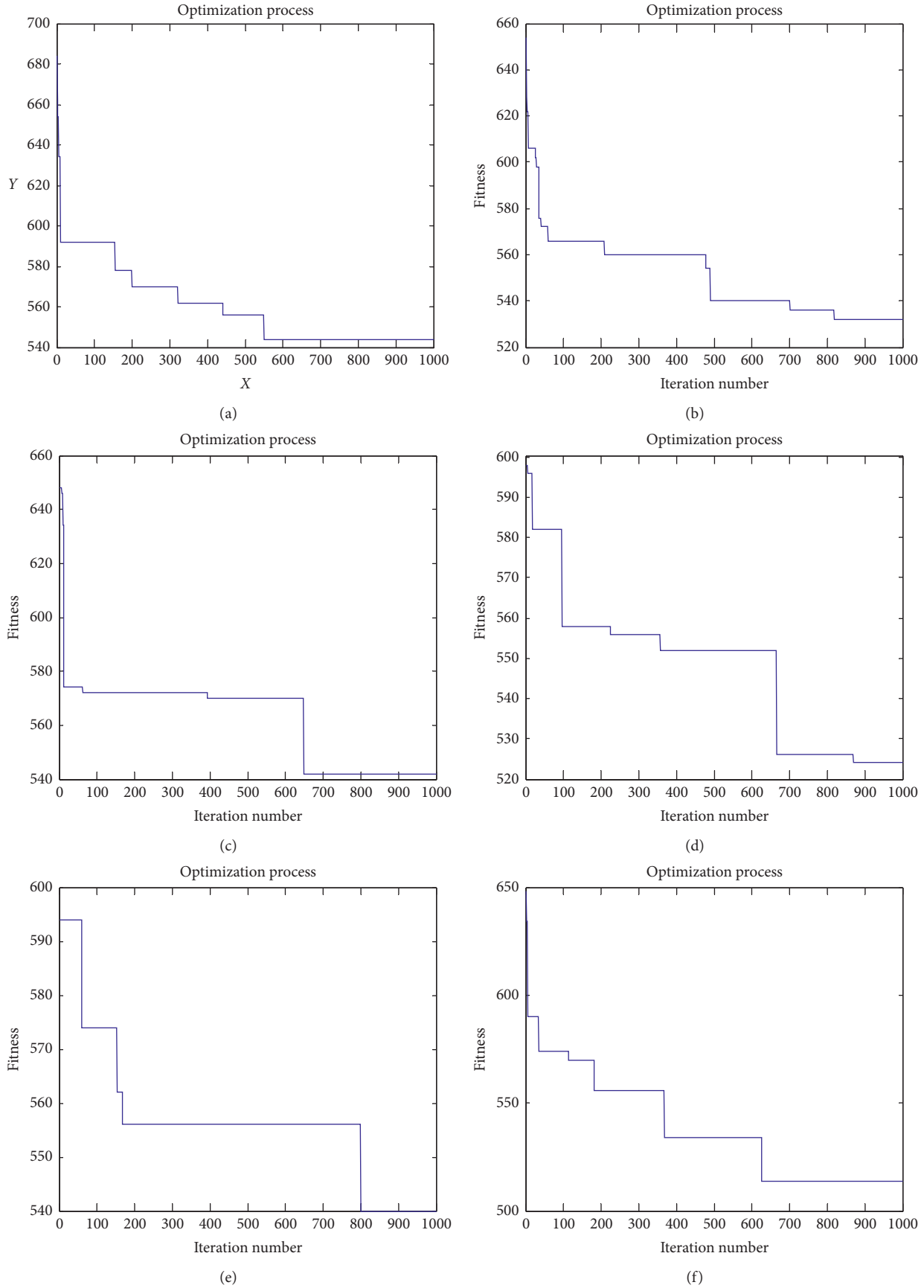


Figure 7: Iterative changes of 20 cargo spaces in Group 2: (a) PSO; (b) FOA; (c) WDPSO; (d) WDFOA; (e) RWPSO; (f) RWFOA. Note: X is the total time of iteration and Y is 1000 iterations.

Table 12: Average of picking time of 20 cargo spaces in Group 2.

Algorithm	Original (s)	WD (s)	RW (s)	Average (s)
PSO	544	542	540	542
FOA	531	528	514	524
Average	537	535	527	

Table 13: Standard deviation of picking time of 20 cargo spaces in Group 2.

	Algorithm SD		
Algorithm	Original	WD	RW
PSO	15.7	14.5	13.2
FOA	21.3	15.5	14.7

time of FOA, WDFOA, and RWFOA is 531 s, 528 s, and 514 s.

According to the data of Table 12, the optimal average search time of PSO is 542 s, the optimal search time of FOA is 524 s, and the optimization of FOA is better. The average optimal search time of the original, WD, and RW is 537 s, 535 s, and 527 s, respectively, and the optimization of RW is better. Thus, RWFOA is the best.

From the standard deviation in Table 13, RWFOA is the smallest, better than the other five. Therefore, PSO algorithm is featured with good accuracy and speed, but its optimization performance is worse than FOA. For six different algorithms, the optimization of RWFOA is relatively good.

The optimal picking time of 20 cargo spaces is 514 s, and the corresponding picking order is as follows: 8–18–19–4–5–1–12–2–10–6–16–15–20–14–11–7–9–3–13–17.

6. Conclusion

With the increasing pursuit of efficiency in logistics warehousing, order picking has also become an important research, and it is constantly proposed to apply a variety of different algorithms to optimize picking time. This paper assumes a model of automated warehouse shelves. By referring to previous studies, the study is designed to set the picking route to get the optimal picking time so as to improve the efficiency of order picking. It has been widely used in various industries, including electronic appliances, pharmaceutical logistics, tobacco logistics, machinery automation, and food industry.

A new FOA, PSO, RW, and WD are used to improve FOA and PSO and to look for the optimal order picking time. The result shows that the optimization capacity of RWFOA is better and the picking efficiency is the best. Therefore, it can be applied to the order picking in automated warehouses, thereby improving warehouse operation efficiency and reducing the time cost of order picking.

RWFOA is a more effective local search method which can be used in future work. The proposed RWFOA could be applied to other variations of the TSP; for example, fixed edges are listed that are required to appear in each solution to the problem, path problem, or vehicle routing problem etc. Therefore, future work could focus on the development of adaptive algorithms with the implementation of other problem-specific features that could improve the performance of the RWFOA.

This study also has certain limitations. For example, the paper assumes that the stacker is moving at a constant speed, but the speed in the actual operating conditions is uncertain. Secondly, this paper takes part of the shelves as the object of study instead of shelf-to-shelf, which means it is the local optimal in the warehouse rather than the global optimal.

Data Availability

The data used to test the algorithm are randomly generated, readers need to pay more attention to intelligent algorithms. Anyway, the data used to support the findings of this study are available from all the authors upon request.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was financially supported by the 2018 Social Science Planning Project of Guangzhou "Research on the Construction and Development of Guangzhou Smart International Shipping Center Based on the One Belt One Road Strategy" (Grant no. 2018GZGJ169) and 2016 Humanities and Social Sciences Research Projects of Universities in Guangdong Province "Construction of key disciplines in business administration" (Grant no. 2015WTSCX126).

References

[1] R. L. Daniels, J. L. Rummel and R. Schantz, L. J and R. Schantz, "A model for warehouse order picking," *European Journal of Operational Research*, vol. 105, no. 1, pp. 1–17, 1998.

[2] H.-I. Jeong, J. Park, and R. C. Leachman, "A batch splitting method for a job shop scheduling problem in an MRP environment," *International Journal of Production Research*, vol. 37, no. 15, pp. 3583–3598, 1999.

[3] W. Lu, McF. Duncan, V. Giannikas, and Q. Zhang, "An algorithm for dynamic order-picking in warehouse operations," *European Journal of Operational Research*, vol. 248, no. 1, pp. 107–122, 2016.

[4] L. Pansart, N. Catusse, and H. Cambazard, "Exact algorithms for the order picking problem," *Computers & Operations Research*, vol. 100, pp. 117–127, 2018.

[5] J. Bolaños Zuñiga, J. A. Saucedo Martínez, T. E. Salais Fierro, and J. A. Marmolejo Saucedo, "Optimization of the storage location assignment and the picker-routing problem by using mathematical programming," *Applied Sciences*, vol. 10, no. 2, 2020.

[6] H. Hwang, W. J. Baek, and M.-K. Lee, "Clustering algorithms for order picking in an automated storage and retrieval system," *International Journal of Production Research*, vol. 26, no. 2, pp. 189–201, 1988.

[7] M. J. LI, X. B. Chen, and C. Q. Liu, "Solution of order picking optimization problem based on improved ant colony algorithm," *Computer Engineering*, vol. 35, no. 3, pp. 219–221, 2009.

[8] C.-M. Hsu, K.-Y. Chen, and M.-C. Chen, "Batching orders in warehouses by minimizing travel distance with genetic algorithms," *Computers in Industry*, vol. 56, no. 2, pp. 169–178, 2005.

[9] X. Ning and H. Hu, "Multiple-population fruit fly optimization algorithm for scheduling problem of order picking operation in automatic warehouse," *Journal of Lanzhou Jiaotong University*, vol. 33, no. 3, pp. 108–113, 2014.

[10] K. J. Roodbergen and R. Koster, "Routing methods for warehouses with multiple cross aisles," *International Journal of Production Research*, vol. 39, no. 9, pp. 1865–1883, 2001.

[11] D. M.-H. Chiang, C.-P. Lin, and M.-C. Chen, "The adaptive approach for storage assignment by mining data of warehouse management system for distribution centres," *Enterprise Information Systems*, vol. 5, no. 2, pp. 219–234, 2011.

[12] R. De Koster, T. Le-Duc, and K. J. Roodbergen, "Design and control of warehouse order picking: a literature review," *European Journal of Operational Research*, vol. 182, no. 2, pp. 481–501, 2007.

[13] G. Dantzig, R. Fulkerson, and S. Johnson, "Solution of a large-scale traveling-salesman problem," *Journal of the Operations Research Society of America*, vol. 2, no. 4, pp. 393–410, 1954.

[14] M. Hahsler and K. Hornik, "TSP infrastructure for the traveling salesperson problem," *Journal of Statistical Software*, vol. 23, no. 1, pp. 1–21, 2007.

[15] C. G. Petersen and G. Aase, "A comparison of picking, storage, and routing policies in manual order picking," *International Journal of Production Economics*, vol. 92, no. 1, pp. 11–19, 2004.

[16] C. Theys, O. Bräysy, W. Dullaert, and B. Raa, "Using a TSP heuristic for routing order pickers in warehouses," *European Journal of Operational Research*, vol. 200, no. 3, pp. 755–763, 2010.

[17] J. Renaud and A. Ruiz, "Improving product location and order picking activities in a distribution centre," *Journal of the Operational Research Society*, vol. 59, no. 12, pp. 1603–1613, 2007.

[18] C. G. Petersen II, "The impact of routing and storage policies on warehouse efficiency," *International Journal of Operations & Production Management*, vol. 19, no. 10, pp. 1053–1064, 1999.

[19] X. Yu, X. Liao, W. Li, X. Liu, and T. Zhang, "Logistics automation control based on machine learning algorithm," *Cluster Computing*, vol. 22, no. 6, pp. 14003–14011, 2019.

[20] X. H. Shi, Y. C. Liang, H. P. Lee, C. Lu, and Q. X. Wang, "Particle swarm optimization-based algorithms for TSP and generalized TSP," *Information Processing Letters*, vol. 103, no. 5, pp. 169–176, 2007.

[21] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43, Nagoya, Japan, October 1995.

[22] Y. Zhang, G. Cui, J. Wu, W.-T. Pan, and Q. He, "A novel multi-scale cooperative mutation Fruit Fly Optimization Algorithm," *Knowledge-Based Systems*, vol. 114, pp. 24–35, 2016.

[23] Y. Zhang, G. Cui, S. Zhao, and J. Tang, "IFOA4WSC: a quick and effective algorithm for QoS-aware servicecomposition," *International Journal of Web and Grid Services*, vol. 12, no. 1, pp. 81–108, 2016.

[24] W.-T. Pan, "A new fruit fly optimization algorithm: taking the financial distress model as an example," *Knowledge-Based Systems*, vol. 26, pp. 69–74, 2012.

[25] H.-z. Li, S. Guo, C.-j. Li, and J.-q. Sun, "A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm," *Knowledge-Based Systems*, vol. 37, pp. 378–387, 2013.

[26] S.-M. Lin, "Analysis of service satisfaction in web auction logistics service using a combination of fruit fly optimization algorithm and general regression neural network," *Neural Computing and Applications*, vol. 22, no. 3-4, pp. 783–791, 2013.

[27] D. Shan, G. Cao, and H. Dong, "LGMS-FOA: an improved fruit fly optimization algorithm for solving optimization problems," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–9, 2013.

[28] T. A. M. Abdel, M. B. Abdelhalim, and S. E.-D. Habib, "Efficient multi-feature pso for fast gray level object-tracking," *Applied Soft Computing*, vol. 14, pp. 317–337, 2014.

[29] F. Gao, *Matlab Intelligent Algorithm Super Learning Manual*, Post & Telecom Press, Beijing, China, 2014.

[30] Y. Shi and R. C. Eberhart, "A modified particle swarm optimizer," in *Proceedings of the Congress on Evolu-Tionary Computation*, pp. 79–73, Washington DC, USA, July 1998.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 7259724, 14 pages
https://doi.org/10.1155/2020/7259724

Research Article

CSR Image Construction of Chinese Construction Enterprises in Africa Based on Data Mining and Corpus Analysis

Yaoping Zhong,[1,2] Wenzhong Zhu ,[3] and Yingying Zhou [3]

[1]*School of English for International Business, Guangdong University of Foreign Studies, Guangzhou 510420, China*
[2]*College of Foreign Studies, Guangxi Normal University, Guilin 541006, China*
[3]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510006, China*

Correspondence should be addressed to Wenzhong Zhu; wenzhongzhu2020@163.com

Received 3 May 2020; Accepted 16 June 2020; Published 15 July 2020

Guest Editor: Shianghau Wu

Since there is negative coverage of some western media on the business activities of Chinese overseas enterprises, which has adverse impact on the image of Chinese enterprises and even the national image of China, this study aims to detect the corporate social responsibility image (hereafter CSR image) of Chinese construction enterprises in Africa (hereafter CCEA) through analyzing the coverage of Financial Times (hereafter FT) from the UK and The Wall Street Journal (hereafter WSJ) from the US and dig up the motives behind their coverage. Octopus is first applied to mine and collect the reports data on CCEA from 2011 to 2019 by the two media. Two small corpora including the reports are then built. NVivo is next used to do the statistical analysis and clustering analysis of the keywords in two corpora as a whole and AntConc is finally utilized to do the statistics of high-frequency evaluative adjectives and nouns modified by evaluative adjectives as well as the concordance of the low-frequency words but closely relevant to corporate social responsibility (hereafter CSR) in two corpora, respectively. The results of the detailed analyses of the keywords are combined to unveil the CSR image of CCEA, which is followed by a discussion about the motives behind the coverage and finally some suggestions are put forward to improve the CSR image of CCEA. Theoretically, the present study promotes the interaction among data science, management, communications, and linguistics; practically it offers some advice to CCEA to elevate their CSR image.

1. Introduction

Since the Belt and Road Initiative (hereafter BRI) was put forward in 2013, China has been broadening its cooperation network. As of the end of January 2020, China has signed up to 200 BRI cooperation documents with 138 countries and 30 international organizations [1]. The African continent, with great development potential, is an important destination and foothold for China to push the BRI towards the west. By the end of January 2020, a total of 44 African countries have signed the BRI cooperation documents with China [1]. The construction advanced by BRI brings unprecedented opportunities as well as huge challenges to China-Africa cooperation. China's investment in Africa is mainly concentrated in the construction, mining, manufacturing, finance, and technology services. Among them, infrastructure construction in the construction industry is one of the highlights [2]. Thus it can be seen that China's overseas contracted project is one of the main directions of investment in Africa.

However, the construction industry is rather sensitive to the issues like labor safety, production cost, and environmental protection. With more Chinese enterprises investing and trading in Africa, such negative remarks as "neocolonialism" and "resource predator" spread in all kinds of news reports, which prevail in the west and even in some African countries. A survey on "Africans' Perceptions of Chinese Business in Africa" displays that 58% of 1,056 Africans from 15 African countries held a negative view that Chinese enterprises did not have a good reputation in their countries [3]. The root cause of this negative speech is that China-Africa

cooperation is becoming more pragmatic in the new century, while the direct cause is some Chinese enterprises' inadequacy or failure of fulfilling their social responsibilities in Africa. As the representative entities of operating abroad, Chinese enterprises are not only economic actors, but also "social responsibility bearers." CSR has a positive effect on corporate image and on customer citizenship behavior [4]; therefore, it is urgent for CCEA to polish their CSR image.

Although researches on CSR drivers or influence factors [5–8] and the relations between CSR and corporate performance [9–12] are prominently offered, only a few studies concentrate on the impact of media coverage on CSR [13–15] and CSR activities or themes of construction industry [16–18] and it is even rare to see the studies on CSR image of Chinese overseas construction enterprises. Therefore, the current study aims to probe into the CSR image of CCEA through answering the following questions based on the analyses of the news reports of two mainstream media, namely, Financial Times from UK and The Wall Street Journal from the US.

(1) What is the CSR image of Chinese construction enterprises in Africa under the coverage of Financial Times and The Wall Street Journal?

(2) What are the motives behind the two media's coverage of such an image?

(3) How can Chinese construction enterprises in Africa improve their CSR image?

2. Literature Review

2.1. Definition of CSR. "Social Responsibilities of the Businessman" written by Howard R. Bowen was published in 1953, which marks the beginning of the study of CSR in modern literature. In his book, Bowen [19] defined social responsibility as "the obligations and responsibilities of businessmen, that is, businessmen have the obligation to determine policies, make decisions, and take actions according to the goals and values of the society." Even though businesses are mainly driven by profits, CSR is not just about economic responsibilities of the enterprises. Some scholars have discussed the dimensions of CSR. Carroll [20, 21] conceptualized the CSR model by four responsibilities: economic responsibility, legal responsibility, ethical responsibility, and philanthropic responsibility. Dahlsrud [22] identified five major dimensions of CSR based on the analysis of 37 definitions of CSR since 1980, namely, economic dimension, social dimension, environmental dimension, stakeholder dimension, and voluntariness dimension.

Research on CSR in China starts late compared with the west. In China CSR movement was brought in on a large scale in 1990s, especially since China implemented the strategies of "Two Resources and Two Markets" and "Going Global," and there has been a growing demand for enterprises to fulfill their social responsibilities internationally [23]. Liu [24] proposed a broad sense and a narrow sense regarding CSR. The broad view holds that CSR is the responsibility of enterprises to all stakeholders, including shareholders. The narrow sense is divided into two views. One refers to enterprises' pursuit of profit maximization and the other points to the responsibility to all stakeholders excluding shareholders. The connotation of CSR has become increasingly rich with the wide spread of the concept in China and the enterprises' deepening understanding of it. The Chinese government holds that companies should fulfill their social responsibilities and abide by the relevant laws, regulations, and business ethics. To be more specific, while pursuing economic interests, an enterprise is responsible for stakeholders, employees, consumers, suppliers, communities, and other interested parties as well as environmental protection.

There are also some definitions of CSR posed by international organizations. Yang [25] discussed several definitions of CSR in the summary of International Symposium on "Corporate Social Responsibility and Development in Africa." For example European Commission points out that "CSR is a concept in which a company voluntarily decides to contribute to a better society and a cleaner environment as well as a concept of an interactive relationship between a company and its stakeholders on a voluntary basis by combining its operations with society and the environment." According to the World Bank, CSR is the commitment and obligation of enterprises to contribute to sustainable economic development and promote their own and society's development by working together with their employees and their families, local communities, and even all the people of the society. World Business Council for Sustainable Development defines CSR as the continuous commitment of enterprises to make contributions to economic development and operate by law while improving the life quality of workers and their families, local communities, and the majority of the society. Business for Social Responsibility holds that CSR is the business activities conducted on the basis of meeting or exceeding the public expectations on morality, law, and commerce. Wu et al. [26] proposed a definition in their paper based on an overview of CSR research development; that is, CSR means an enterprise is responsible for all of its stakeholders in its operations and activities to balance the needs of stakeholders and the profits of the company by considering the impact of their operations and activities on labor, the environment, communities, customers, and even their competitors when making decisions.

Even though there are different versions of the definition of CSR, most of the versions share great commonalities. To conclude, CSR in this study refers to the responsibility and obligation that enterprises should fulfill to such stakeholders as shareholders, employees, consumers, and their communities during the process of pursuing economic interests.

2.2. Definition of CSR Image. CSR image is closely relevant to both CSR and corporate image. Corporate image is the result of an aggregate process by which the public compares and contrasts the various attributes of firms [27]. To be more specific, corporate image refers to the impression and evaluation of the public and employees on an enterprise,

which consists of product image, media image, and the like, while CSR image is less complex than corporate image. Liu [24] defined CSR image as people's impression of an enterprise's attitude and behavior towards stakeholders excluding shareholders and a wide range of social issues, or people's attitude and evaluation of the enterprise based on its CSR concept and behavior. Thus it can be said that CSR image is the direct reflection of CSR behavior, which is formed on the basis of the social attributes of the enterprise and the social responsibility it undertakes.

2.3. Significance of CSR. At the early beginning, many enterprises regarded CSR as charitable activities like donation, and even a burden for the enterprises which still struggle to survive. Fortunately, both western and Chinese enterprises have a deeper understanding of CSR nowadays and realize its importance after their experience of the four-stage development, namely, start, development, maturity, and steady development or suffering of crisis. Externally, to ensure the healthy and orderly operation, an enterprise must respect and satisfy the interests of stakeholders and establish a harmonious relationship with them, which is conducive to maintaining and even expanding its profits. Internally, fulfilling CSR benefits the enterprise in improving employees' sense of mission and satisfaction, enhancing an enterprise's innovation ability, building a positive corporate culture, and strengthening its cohesion.

In the era of economic globalization, fulfilling CSR is not only beneficial to economic development and social progress, but also helpful to the long-term development of enterprises, for they can promote and maintain economic growth, reduce poverty, and create prosperity by providing investment, technology, jobs, and training of labor skills. When pursuing economic benefits, enterprises produce products and services that meet the needs of the society, thus promoting the growth of the economy [25]. More researchers argue the benefits of fulfilling CSR to the elevation of corporate image [24, 28]. Liu [24] pointed out that a positive CSR image is a shortcut for enterprises to obtain strategic advantage in the information age. As to the promotion of China's soft power in Africa, Song [28] stated that it is of strategic significance for Chinese enterprises in Africa to strengthen CSR building under the background of the implementation of "Going Global" strategy and China's imbalance development of soft and hard power in Africa.

2.4. Research on CSR of Chinese Enterprises in Africa. The western media have reported much on the economic activities of Chinese enterprises in Africa, but few western scholars pay attention to the CSR implementation of Chinese enterprises in Africa. On the contrary, Chinese scholars have been keeping their enthusiasm in this field. Some of them agree that fulfilling CSR is of great significance for Chinese enterprises to operate in Africa [29, 30]. Some Chinese enterprises in Africa, however, do not attach enough importance to CSR, thus giving chances to the western countries to smear Chinese overseas companies, which seriously damages the image of Chinese enterprises in Africa and even the image of China [28, 30]. Fortunately, more and more Chinese researchers have realized the poor CSR performance of some Chinese enterprises in Africa and proposed solutions to the problems after analyzing the causes behind this phenomenon [31–34]. In recent years, some researchers have conducted in-depth research on the understanding and implementation of CSR in China through case analysis and field investigation [23, 29, 31, 35], and their findings reveal that most firms have shouldered their responsibilities, with the big state-owned enterprises performing better than the small private ones [36]. One typical example is illustrated by Feng's [37] study which analyzed the Mombasa-Nairobi Standard Gauge Railway Project in Kenya constructed by the China Road and Bridge Corporation.

Some limitations of the abovementioned studies should not be neglected. In terms of research methods, most of the abovementioned studies adopt qualitative research methods, which are mostly limited to summary of experience or lessons. Case study is widely used in these qualitative researches and although the results are persuasive, the generalization of their conclusions is limited by the number of samples. As to the source of data, second-hand data is collected in most of the researches, and only a minority of the data are collected through face-to-face interviews and field investigation. News data are easy to collect in great quantity, yet only few studies are conducted to explore CSR through the analysis of news reports. Thus the present study will first apply Octopus to collect the news data related to CCEA from FT and WSJ, then NVivo to do the clustering analysis [38] to classify the keywords of the news reports, and next AntConc to do the word frequency statistics. This is an interdisciplinary study involving data science, management, communications, and linguistics, which innovates the research perspectives and methods of CSR study.

3. Methodology

3.1. Sample and Data Collection. Although social media plays an increasingly important role in people's life, word-of-mouth still plays a large part in people's communication and it is often beyond enterprises' ability to control the dissemination of news through this nonmass media. Therefore, many enterprises still prioritize mass media to promote their corporate image [39]. In this study, we choose FT and WSJ as two representative mainstream media based on their huge influence in providing business news. FT is a world famous international financial media. As of October 2019, its print circulation and digital circulation reach 168,958 and 740,000, respectively [40]. WSJ is a comprehensive newspaper featured by financial reports, focusing on financial and business coverage and has a broad influence on the international community. As of August 2019, its daily circulation reaches 2,834,000 [41].

Since we focus on the CSR image portrayed by the western media in this study, news data is collected from the ProQuest ABI/INFORM business information database which was released by ProQuest, an American company, 40 years ago. Guangdong University of Foreign Studies where

the authors study or work has purchased this database as part of its library resources. The database has long been a gold standard for business information storage, providing a wealth of information resources in economics, management, business, and related fields. Africa has witnessed some progress in understanding CSR since 2011, which pushes CSR construction into a consensus stage [42]. In view of this, we set the period of news collection from January 1, 2011, to December 31, 2019, a total of nine years. The data is collected in four steps, which is illustrated as follows.

Step 1: the publication type was set to newspaper, the document type to news, and the language to English. Keyword search was conducted by entering Chin* (including China and Chinese) construction compan* (including company and companies) and Afric* (including Africa and African). 3813 news reports were obtained after the retrieval.

Step 2: only FT and WSJ were focused, and 585 news reports from the former and 937 news reports from the latter were gained, respectively, after the screening.

Step 3: the websites of the database were mined to collect the screened news reports by Octopus and saved in xls format, with the titles, sources, and content of the news reports remained only.

Step 4: 19 news reports from the FT and 14 news reports from WSJ were further chosen for the current study after two researchers' browsing of the titles and content of the news reports and their consensus on the selected reports. Finally the news reports in xls format were transferred into txt format for the convenience of subsequent processing of the data.

3.2. Data Analysis. Clustering analysis is conducted to classify the keywords of the news reports in this study. It is a technique for analyzing statistical data and is widely used in many fields, including machine learning, data mining, pattern recognition, image analysis, and biological information. Clustering is a static classification of similar objects into different groups or more subsets so that each member in the same subset shares some similar properties, such as shorter spatial distance in a coordinate system. This analytical method is now widely used in public opinion research and enterprise perfection.

NVivo is utilized to first retrieve the 33 news reports from the two media, then extract the information of the news reports to count the word frequencies, and finally select the first 30% of the keywords as high-frequency ones which are related to the topics about CSR of CCEA. These high-frequency keywords are presented in a table in terms of serial number, content, frequency, percentage, and cumulative percentage. The software is also used to generate a word cloud and do the clustering analysis. We interpret the clusters to detect the relationship between the high-frequency keywords and reveal the structure of these research focuses.

As a complement to the clustering analysis, the corpus tool AntConc is first used to conduct word frequency of the 33 news reports and pick out the top 15 high-frequency lemmas to reveal which words the two media use to report CCEA. Then, we analyze the news headlines to see what events the two media focus on. Next, we count the evaluative adjectives and nouns modified by evaluative adjectives which are relevant to CSR performance of these Chinese companies to elicit the overall CSR image of CCEA.

4. Results and Discussion

4.1. Statistics of Keywords. In this study a total of 3,766 keywords are extracted, the first 30% of which are processed to generate a word cloud (see Figure 1). The bigger the word in the word cloud, the higher frequency it appears in the coverage. It is clear to see that the FT and WSJ mainly focus on Chinese companies' infrastructure construction, projects, and investments in Africa, which is supported by the Chinese government.

More details of the coverage by the two media can be revealed in terms of the frequency and cumulative percentage of the first 30% high-frequency keywords which are the focus of the present study. We are able to reveal the themes of the coverage of the two media regarding CCEA by referring to these high-frequency keywords. It can be seen from Table 1 that there are 42 high-frequency keywords in the coverage, with a total frequency of 3844 times, accounting for 29.82% of the total.

4.2. Clustering Analysis of Keywords. Keywords embody the topics of a news report, so clustering analysis of the 42 keywords is conducted to highlight the focuses of the coverage of CCEA as well as the consistency of the topics. Figure 2 demonstrates the tree graph of high-frequency keywords, in which the vertical distance within each cluster at the horizontal level means the average distance between two words. The distance indicates their intimate relations, so the shorter the distance, the closer their connotation, which suggests a higher frequency of their concurrence in the same report and in turn reflects greater consistence of the topic. Figure 2 demonstrates that there are four clusters which refer to four CSR focuses on CCEA. The four coverage hotspots include (1) Chinese state-owned construction enterprises' investment in infrastructure and trading with African countries; (2) China's economic assistance to Africa by financing projects and China's business activities involving military action; (3) China's investment in power plants construction and officials' involvement in construction activities; and (4) China's international construction companies' contribution to the world development.

4.3. Statistics of High-Frequency Lemmas. To seek more evidence for the revelation of the themes of the coverage about CCEA, high-frequency content words like nouns and verbs are counted (see Table 2). The statistics only focus on lemmas, which are the words derived from the same stem, so that different word forms of a stem can be included in the frequency of the same lemma. Different from the frequency statistics in Table 1 of section 4.1, the frequency statistics in

Figure 1: The word cloud of the news reports of FT and WSJ.

Table 1: High-frequency keywords in the news reports of FT and WSJ.

No.	Words	Frequency	Weighted percentage (%)
1	china	357	1.66
2	chinese	327	1.52
3	africa	231	1.08
4	projects	156	0.73
5	companies	144	0.67
6	countries	135	0.63
7	african	124	0.58
8	investments	111	0.52
9	beijing	112	0.52
10	state	112	0.52
11	governments	110	0.51
12	construction	108	0.5
13	railway	95	0.44
14	building	87	0.41
15	power	89	0.41
16	development	85	0.4
17	infrastructure	81	0.38
18	kenya	74	0.34
19	years	71	0.33
20	economic	66	0.31
21	economy	67	0.31
22	president	65	0.3
23	billions	65	0.3
24	capital	60	0.28
25	world	58	0.27
26	officials	59	0.27
27	military	59	0.27
28	trading	55	0.26
29	national	55	0.26
30	region	53	0.25
31	according	53	0.25
32	internationally	51	0.24
33	first	51	0.24
34	people	50	0.23
35	including	47	0.22
36	business	48	0.22
37	nigeria	45	0.21
38	continent	45	0.21
39	financial	45	0.21
40	london	46	0.21
41	ghana	46	0.21
42	ethiopia	46	0.21

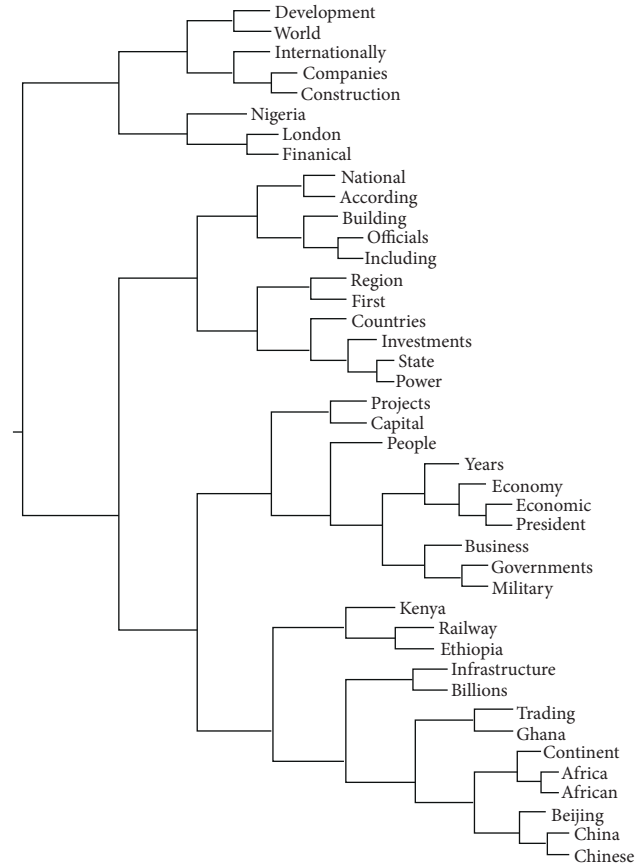


Figure 2: The tree graph of high-frequency keywords in the news reports of FT and WSJ.

Table 2 provides respective details of the focuses of FT and WSJ. We can see that the lemmas like *China, Chinese, Africa, African, project,* and *Beijing* appear in the two corpora, only with different frequencies. In addition, the lemma *state* appears in FT and *government* appears in WSJ, which indicates that the activities of CCEA are supported by the Chinese government. In general, the two media are focused on the projects carried out by Chinese construction enterprises, with the difference lying in FT concerning railway construction while WSJ linking China's military activities to infrastructure construction and highlighting the economic scale of these projects.

4.4. Headline Analysis. The headline represents the main idea of a news report, so headlines of the news reports of the two media are compared below to dig more details about the themes of their coverage. It is clear to see from Table 3 that the headlines of FT news are obviously shorter than those of WSJ. As to the content of the headlines, FT uses relatively objective titles which mainly focus on Chinese companies helping African countries build railways and roads, while WSJ tends to use negative titles. For example, the headline of article 3 from WSJ argues that most African countries are caught in a debt trap designed by China, and China's growing economic and military influence on the continent has raised concerns in the west.

Table 2: The top 15 high-frequency lemmas in the news reports of FT and WSJ.

	FT corpus	Frequency	WSJ corpus	Frequency
1	China	204	China	153
2	Chinese	197	say	153
3	say	156	Chinese	131
4	Africa	137	Africa	94
5	company	99	year	85
6	project	75	new	81
7	railway	75	project	80
8	African	74	country	68
9	more	72	government	64
10	state	72	billion	59
11	construction	69	military	57
12	year	68	economy	55
13	country	67	more	52
14	power	65	African	50
15	Beijing	64	Beijing	48

Table 3: Statistics of headlines of the new reports of FT and WSJ.

	FT corpus	WSJ corpus
1	Chinese builders target contracts in EU and US	Soft power: China backs Egypt's new $45 billion. Capital; deal to build new capital near Cairo is part of Beijing's plans to boost business in emerging markets
2	Construction companies stage second act overseas	Violence imperils Kenya port project; land disputes, terrorism threaten plan to turn Coastal Region into economic hub
3	China's plan to connect the world	More of Africa finds itself in China's debt; Beijing's widening economic and military footprint on continent raises concerns; president. Xi Jinping secures host of deals on weeklong visit
4	Chinese investors herald second golden age of east Africa rail building	After winning Re-Election, Kenya's president faces economic, social challenges; Uhuru Kenyatta's tasks: keeping economy ticking, cutting debt, and uniting a nation divided along tribal lines
5	China's testing ground	Egypt's Sisi clamped down on political Opposition--next up is the economy; the military has amassed a growing business empire under the former general-turned-president, leading to renewed popular resentment
6	Beijing relations raise the emotional heat	China builds first overseas military outpost; naval facility under construction in Djibouti shows Beijing's ambitions to be a global maritime power and protect its expanding interests abroad
7	Greater Chinese presence built on a longstanding relationship	Beijing spins a web of Chinese infrastructure; 'New silk road' projects keep the spirit of globalization alive
8	Kenyan railway highlights sharper focus on affordability	China takes wary steps into new Africa deals; premier pledges more business during trip, but Beijing shows more caution
9	Beijing denies five-year hack of African union HQ in Ethiopia	In Africa, those who bet on China face fallout; economic slowdown in China exacerbates strain for trading partners in Africa
10	Chinese loans keep Kenya railway project rolling	Ethiopia set to gain transport lifeline to sea with new $4 billion railway; project backed by China will connect Addis Ababa to seaport at Djibouti, aiding Ethiopian agriculture exporters

Judging from the above twenty randomly selected headlines, we can see that the Chinese construction enterprises have fulfilled their economic responsibility through helping the African countries build infrastructure. Both news media, however, cover the events mainly from the perspective of China's seeking its own interests, and WSJ links China's assistance to its outward expansion and even claims that China should be responsible for the debt woes of African countries. Apparently the headlines of the coverage of the two media reflect their political stance and this will cause a negative impact on the image of CCEA and even China's national image in Africa and beyond.

4.5. Targeted Keywords Analysis. Adjectives are evaluative in nature, such as good, great, bad, or terrible, and are often used to express the positive or negative effects of a proposition [43], so adjectives are one of the most prominent ways to express the values of different communities [44]. Biber et al. [43, 45] listed three semantic categories of adjectives: emotion, attitude, and evaluation. In addition, other researchers have also studied evaluative adjectives, whose function can be divided into subcategories in a particular language class or discourse community. For example, Hewings [46] analyzed evaluative adjectives, whose propositions can be divided into nine semantic categories, including interest, accuracy, importance, adequacy,

intelligibility, character, certainty, emotion, and judgment. Nelson [47] classified the adjectives in a business English corpus of 1.5 million words, including size or speed, place, positive, negative, neutral, work or business, currency, technology, and time [48]. Considering this study focuses on the CSR image construction of CCEA which contains emotions, feelings, or attitudes, these previous studies can provide references for the current study. Therefore detailed analysis of the adjectives of the news reports will be made to achieve the aim of the study.

4.5.1. Evaluative Adjectives. There are two types of adjectives and both predicative and attributive adjectives can be included in the analysis of the adjectives for the study; however, only the nouns in the adjective + noun pattern were included for a targeted analysis of the entities immediately adjacent to the evaluative adjectives [48]. So we select the attributive adjectives from the word list generated by AntConc according to their frequencies and then check by concordance to see whether they are used to describe the events related to CCEA. Table 4 shows that the evaluative adjective lists of both media share great similarities in containing such adjectives as *new*, *local*, *state-owned*, *big*, *first*, *international*, *overseas*, *economic*, and so forth. Judging from the two lists, it can be seen that both FT and WSJ focus on Chinese state-owned construction enterprises' construction of big projects, Chinese companies' development of new overseas markets, and their bringing commercial benefits to the local people. Thus, we can conclude that Chinese construction enterprises have fulfilled their economic responsibilities in Africa.

Yet there is one detail that should not be ignored. The word *military* appears 28 times and the other word *political* appears 20 times in WSJ, respectively. The high frequencies of the two words convey such a message that China aids Africa in constructing infrastructure with an intention of building its military bases to expand its political influence on the continent, which portrays China as a major rival to the US in Africa. In general the coverage of the two media on CCEA concentrates on the economic activities of CCEA, and their reports embed their ideology.

Table 4: High-frequency evaluative adjectives in the new reports of FT and WSJ.

	FT corpus	Frequency	WSJ corpus	Frequency
1	new	34	new	38
2	local	22	economic	34
3	state-owned	18	military	27
4	other	17	other	21
5	big	16	political	20
6	first	15	foreign	17
7	international	15	first	12
8	global	13	western	13
9	financial	13	local	13
10	overseas	13	many	12
11	economic	12	private	12
12	regional	12	electoral	11
13	private	11	international	10
14	huge	9	regional	10
15	domestic	8	state-owned	9
16	foreign	8	senior	9
17	political	8	public	9
18	western	8	former	8
19	many	8	major	8
20	commercial	7	big	8

4.5.2. Nouns Modified by Evaluative Adjectives. Entities, such as nouns are often applied to describe the activities of enterprises, so statistics of high-frequency nouns allow readers to learn more about the CSR performance of enterprises. Table 5 lists the top 20 nouns related to the CSR performance of CCEA, which are modified by evaluative adjectives. Judging from the statistical results, we can learn that the coverage of two media is fairly consistent, which is mainly about the business activities of CCEA. It is crystal to see that these companies have performed well in fulfilling their economic social responsibilities by constructing infrastructure, thus benefiting local people as well as the enterprises themselves. Furthermore, the economic activities of CCEA involve the officials of the China and African countries, which suggests that the mutual cooperation is largely driven by the Sino-Africa governments. There are, however, still some differences between the coverage of the two media. When taking a look at the concordance of the word *loans*, we see that FT propagated the negative impact of Chinese loans with additional conditions on Africa, while WSJ highlighted Chinese and African leaders' efforts to bolster infrastructure building for the benefit of their respective countries and claimed that many of China's projects served oil exploration and military activities when we referred to the concordance of the words *power, oil, and officers*. On the surface the reports of the two media focus on China's economic activities in Africa, but in fact they overinterpret China's aid to Africa as neocolonialism or resources plunder. In view of the influence of the two media in the UK and the US, such coverage may mislead the African people and even the people from other countries to misjudge Chinese overseas companies as well as China, which is definitely harmful to the overseas activities of Chinese enterprises and even China's diplomatic activities.

4.5.3. Concordance of Low-Frequency Keywords. The statistics of the high-frequency adjectives and nouns reveal the economic social responsibilities of CCEA, but the other social responsibilities are seldom concerned. Feng and Wei [49] pointed out that in 1979 Carroll was the first to summarize the connotation of CSR and constructed a "four-layer pyramid" model, namely, economic responsibility, legal responsibility, ethical responsibility, and discretionary responsibility, which was revised to "economy-law-ethic-charity" in 1991. This model shows that while enterprises are committed to maximizing profits for shareholders, they should also pay attention to laws, regulations, ethics, and morality and vigorously carry out philanthropy. Therefore, concordance is conducted to explore words of low-frequency but closely related to CSR in the specific context so as to further analyze the CSR image of CCEA portrayed by FT and WSJ.

Table 5: High-frequency nouns in the news reports of FT and WSJ.

	FT corpus	Frequency	WSJ corpus	Frequency
1	power	22	capital	14
2	projects	12	forces	13
3	railway	10	growth	12
4	investment	10	infrastructure	11
5	companies	9	economy	11
6	infrastructure	9	investment	9
7	development	9	projects	8
8	investments	9	countries	8
9	countries	7	lines	7
10	rail	7	ties	7
11	company	7	project	6
12	trade	7	election	6
13	markets	7	development	6
14	line	6	economies	6
15	construction	5	power	6
16	grid	5	officers	6
17	loans	5	oil	5
18	state	5	port	5
19	country	5	trade	5
20	people	5	investors	5

From the word list generated in the two corpora, the word *environmental* appears in the coverage of both media, and the word *illegal* appears in FT reports while the word *labor* appears in WSJ reports. Abiding by law is the legal responsibility in CSR, while protecting the environment and the rights and interests of employees are the ethical responsibilities of CSR. These two responsibility dimensions are often discussed in CSR as well as economic responsibility. FT covered that the projects aided by China to Africa, such as dam building, had disastrous effects on the local lakes and people's lives (see Figure 3). Some reports even claimed that China did not take environmental protection seriously and some Chinese companies even took their bad habits to Africa. Besides, some African projects which faced difficulties in financing due to environmental concerns had been allowed to proceed with loans from Chinese banks.

The word *environmental* appears in WSJ only two times less than in FT, but the focus here is not the frequency but the context in which the word is used. Figure 4 indicates that WSJ reports on the environmental protection problems caused by CCEA, among which these companies were forced to sell their shares of some projects after a halt due to the environmental problems and some Chinese companies had caused adverse effects on the local environment because of their ignoring or breaching the local environmental standards.

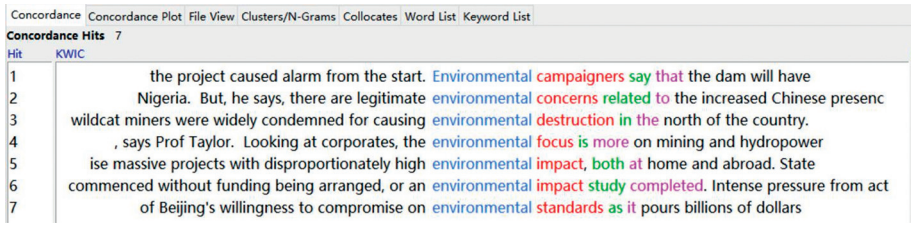
It is a bit surprising to see only a few reports on the fulfillment of legal responsibilities of CCEA from both media. Yet FT covered that some Chinese companies illegally harvested fish and shellfish and some of them even smuggled ivory and rhino horns across the continent (see Figure 5). These illegal behaviors have obviously damaged the local ecological environment and also violated the laws to protect wildlife, which incurs complaints and protests from the local residents.

Labor disputes are commonly seen in the reports on Chinese enterprises' operations in Africa, which are mainly about local workers' complaints about unfair pay and even striking to defend their legitimate interests. Figure 6 shows that Chinese construction companies often hire Chinese workers to do their projects, which leads to no significant increase of job opportunities for local workers and may have a negative impact on the local economy. In addition, some reports mentioned that some projects were forced to halt due to labor disputes and ended up with share transfer. Unfortunately this coverage did not trace the causes of these problems but simply listed the problems. Apparently, such coverage is unfair to CCEA.

As is discussed in this section, both FT and WSJ focus on the economic responsibility of CCEA, and these reported Chinese construction enterprises are mainly state-owned enterprises, whose activities, as the media claim, are sponsored by the Chinese government, and their activities form a significant part of China's BRI. Moreover it is said that China's massive lending to Africa for infrastructure construction has thrown the continent into a debt trap so that China can get better access to Africa's natural resources in its negotiation with African countries. Besides, the two media released a few reports on the legal and ethical responsibilities of CCEA, but these reports are all negative, and there is no coverage on the charity responsibility. This may contribute to Chinese companies' priority in pursuing economic returns, or indeed Chinese companies' inadequacy in fulfilling other responsibilities or just the deliberate neglect of the two media. No matter what reasons are behind it, the motives of the two media's coverage are worth further discussion.

4.6. The Motives Behind the Coverage of FT and WSJ. From the above analyses, it can be seen that the coverage of the two media is greatly consistent, which points to the Chinese construction enterprises' construction of large-scale infrastructure and investments in Africa with the support of Chinese government, but the economic activities of these Chinese companies have caused harm to local environment and residents. A certain ideology is embedded in such coverage, which is detrimental to Chinese companies' overseas operations and even to China's diplomatic activities. To cope with such a challenge, it is crucial for Chinese companies to understand the motives behind such biased coverage so that they can better regulate their own behaviors and fulfill corresponding CSR. The motives behind the coverage of FT and WSJ can be roughly inferred on the basis of the results of the previous analyses.

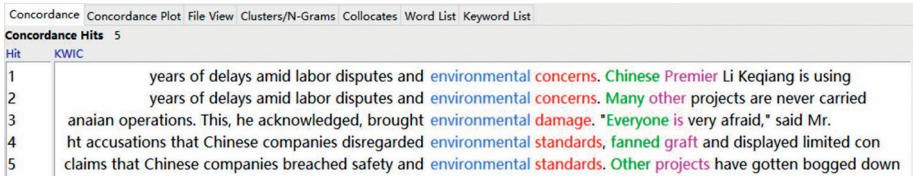
4.6.1. Motive One. Western countries' rejection of China's subsidies to its state-owned enterprises: in China's national economic system, state-owned enterprises are the mainstay of the country's economic lifeline. These companies are leaders and strong competitors in exploiting overseas markets. In the eyes of the western countries, however, Chinese state-owned construction enterprises have gained an absolute advantage with the support of their government, and that is the reason why western construction companies often lose the contracts in the competition with their


Concordance | Concordance Plot | File View | Clusters/N-Grams | Collocates | Word List | Keyword List

Concordance Hits 7

Hit	KWIC
1	the project caused alarm from the start. Environmental campaigners say that the dam will have
2	Nigeria. But, he says, there are legitimate environmental concerns related to the increased Chinese presenc
3	wildcat miners were widely condemned for causing environmental destruction in the north of the country.
4	, says Prof Taylor. Looking at corporates, the environmental focus is more on mining and hydropower
5	ise massive projects with disproportionately high environmental impact, both at home and abroad. State
6	commenced without funding being arranged, or an environmental impact study completed. Intense pressure from act
7	of Beijing's willingness to compromise on environmental standards as it pours billions of dollars

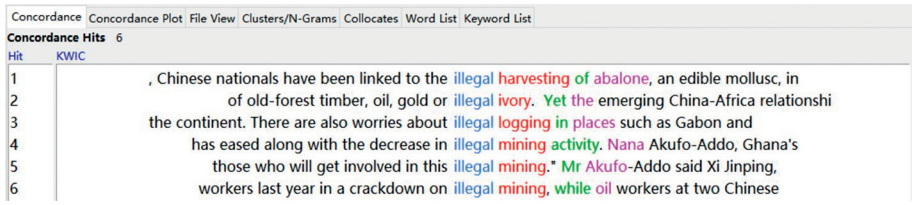
Figure 3: Concordance of environmental in the news reports of FT.


Concordance | Concordance Plot | File View | Clusters/N-Grams | Collocates | Word List | Keyword List

Concordance Hits 5

Hit	KWIC
1	years of delays amid labor disputes and environmental concerns. Chinese Premier Li Keqiang is using
2	years of delays amid labor disputes and environmental concerns. Many other projects are never carried
3	anaian operations. This, he acknowledged, brought environmental damage. "Everyone is very afraid," said Mr.
4	ht accusations that Chinese companies disregarded environmental standards, fanned graft and displayed limited con
5	claims that Chinese companies breached safety and environmental standards. Other projects have gotten bogged down

Figure 4: Concordance of environmental in the news reports of WSJ.


Concordance | Concordance Plot | File View | Clusters/N-Grams | Collocates | Word List | Keyword List

Concordance Hits 6

Hit	KWIC
1	, Chinese nationals have been linked to the illegal harvesting of abalone, an edible mollusc, in
2	of old-forest timber, oil, gold or illegal ivory. Yet the emerging China-Africa relationshi
3	the continent. There are also worries about illegal logging in places such as Gabon and
4	has eased along with the decrease in illegal mining activity. Nana Akufo-Addo, Ghana's
5	those who will get involved in this illegal mining." Mr Akufo-Addo said Xi Jinping,
6	workers last year in a crackdown on illegal mining, while oil workers at two Chinese

Figure 5: Concordance of illegal in the news reports of FT.


Concordance | Concordance Plot | File View | Clusters/N-Grams | Collocates | Word List | Keyword List

Concordance Hits 5

Hit	KWIC
1	for mining projects, which has involved importing labor and equipment, minimizing the boost for local
2	managers. "You maybe can use some local labor, but you definitely will need foremen and
3	Gabon after five years of delays amid labor disputes and environmental concerns. Chinese Prem
4	Gabon after five years of delays amid labor disputes and environmental concerns. Many other p
5	tax-exempt status and access to cheap labor in the form of conscripted soldiers. But

Figure 6: Concordance of labor in the news reports of WSJ.

Chinese counterparts in the African market. Admittedly, Chinese state-owned enterprises enjoy a greater competitive advantage in subsidies, credit, government relations, policy support, and so forth. The former deputy secretary of the US, Robert D. Hormats, said that more and more state-owned enterprises and sovereign funds entered the international market after the financial crisis with a competitive edge in the American market and in third countries, but the comparative advantages were provided by their governments instead of their own productivity and innovation. Therefore, he argued that American firms without government support were at a disadvantaged position when competing with these companies [50, 51]. It should not be ignored, however, that China adopts socialist market economy system, so state-owned enterprises dominate the pillar industries in China's national economy and the Chinese government has launched an increasingly comprehensive and thorough reform of state-owned enterprises since 2012 guided by the core document entitled "Guiding Opinions on Deepening the Reform of State-Owned Enterprises," which was published in 2015, along 22 supplementary files [52].

4.6.2. Motive Two. Western countries' restlessness and unwillingness to see the world center shift from the west to the east: since the reform and opening up, the world has witnessed China's economic miracle. China's desire to become a global economic powerhouse is thus deemed a counterweight to US hegemony in the international system [53]. With China's peaceful rising, it is emerging as a much more active world player. Since the postcrisis era, China has contributed with over 30% of the world economic growth, making it one of the most dynamic countries in the globe. For example, the promotion of BRI brings more friends to China, yet China's expanding of the African market will

cause a shock to its western counterparts who have settled in the resource-abundant continent for many years. The interpretation of Africa's renewed geoeconomic position in the global setting leads to a perception that China's growing geopolitical and economic influence in Africa encroaches upon what western powers has seen as its traditional cauldron of interest and influence [54]. So some western countries may interpret China's overseas development as a challenge to the existing international political and economic order lasted for decades. Then, it can be said that "west-centrism" is the root of misjudging China-Africa cooperation [55]. In fact, China has reiterated that it never seeks to challenge the current international order, nor does it seek hegemony.

4.6.3. Motive Three. Western countries' distorting Chinese aid to Africa as neocolonialism: since China began its aid and investment in Africa, its action is interpreted as earlier colonial investments to ensure access to raw materials [54, 56]. At the present stage of development, Chinese construction enterprises' labor export and outbound investment are a natural outcome of China's economic expansion. China has elevated to the second largest economy and possesses its advantages of low cost but high efficiency in infrastructure construction, coupled with overcapacity in domestic industrial products, such as cement and steel, and at the same time Africa is in urgent need for large-scale infrastructure construction, all of which push Chinese construction enterprises to conduct construction and investments in Africa. With BRI put forward and implemented deeply, however, some western media and politicians are extremely worried about China and desperately vilify and distort the initiative, playing up the so-called "Chinese neocolonialism" fantasy, thus denigrating China-Africa economic cooperation as neocolonialism [57, 58].

4.6.4. Motive Four. Western countries' stereotype of bad behaviors of Chinese overseas companies: in the early days of Chinese enterprises' entry to the overseas market, such problems as rampant fake goods, back pay, and environmental damage occurred from time to time due to these companies' limited strength, excessive pursuit of economic benefits, and insufficient understanding of CSR. As a result, an adverse corporate image of Chinese overseas enterprises was highlighted through the overwhelming reports by the western media, which gradually becomes a stereotype of Chinese multinational corporations [59–62]. While Chinese overseas enterprises have been aware of the importance of CSR in recent years, some, especially the small private enterprises which struggle for survival, are still taking chances and conducting illegal activities. Once such adverse events are exposed, the western media will naturally bring up the old story of Chinese companies' misconducts. Even if most of Chinese overseas enterprises have complied with the local laws nowadays, the western media still tend to exaggerate the bad behaviors of a single case.

To sum up, any news media has its own stance, so the coverage of mainstream media often represent the opinions of its sponsors. Once the motives behind the negative coverage of western media are revealed, the CCEA and even other Chinese overseas enterprises can better handle the prejudiced or even unreal reports.

4.7. Suggestions for CCEA to Improve Their CSR Image. In view of the above findings, the following suggestions are proposed to help the CCEA improve their CSR image.

First, establish a comprehensive CSR operation pattern: although enterprises are the main body of fulfilling CSR, it does not mean other stakeholders can be bystanders. For the overseas businesses, CSR image is not only related to their own development but also to the image of their home country, so all the stakeholders, including government, enterprises, and nongovernmental organizations, should be involved in the implementation of CSR. It is noted that in 2014 EthicsSA [3] had conducted a questionnaire survey on Chinese companies' operations in Africa by inquiring 1,056 respondents from 15 Sub-Saharan African countries and the survey results show that only 6% of respondents from South Africa and 7% of respondents from Nigeria believed that Chinese companies had considered the interests of the society and the community when making business decisions, and in Kenya 23% of the respondents shared the same view. This demonstrates a lack of localization of the Chinese enterprises in Africa and their focus is on the communication with African officials and the upper class [63]. Therefore, CCEA should strengthen their dialogue with the local people so as to win more understanding and support when making business decisions.

Second, deepen the understanding of the strategic significance of CSR. Even though Chinese companies are increasingly aware of the importance of fulfilling CSR, many of them still have a shallow understanding on it due to the short history of CSR in China and their excessive emphasis on immediate economic benefits. Wu et al. [26] pointed out that Chinese international contractors' understanding on CSR is distorted and incomplete, which is caused by their insufficient attention to the health and safety management, imperfect professional ethics standards, and neglect of CSR in decision-making and their construction activities. In this regard, it is suggested that China international construction corporations pay more attention to labor conditions, environmental protection, anticorruption, and ameliorate professional ethics so as to establish common values with local enterprises and society. In addition, CCEA should also enhance their understanding of the strategic significance of CSR under the background of China's vigorous promotion of BRI. Carrying out CSR in Africa has received great support from Chinese President Xi Jinping because helping African countries improve their capacity will be beneficial to their self-development, thus advancing a new type of China-Africa strategic partnership and pushing the mutual cooperation [42].

Third, balance the economic responsibility and other social responsibilities. Although economic responsibility comes first, legal, ethical, and philanthropic responsibilities are no less important. CCEA should fulfill the noneconomic

responsibilities within their power, which is helpful to their long-term interests. According to the survey on "Africans' Perceptions of Chinese Business in Africa" conducted by EthicsSA [3] in 2014, 22.7% of the respondents gave positive evaluation, while 55.9% of the respondents offered negative evaluation in respect of the quality of Chinese products and services. As for the feedback on the environmental responsibility of Chinese enterprises in Africa, 53.9% of the evaluations were negative and only 11.1% were positive. Besides these disappointing numbers, the CSR performance of the Chinese companies in other aspects was equally unsatisfactory [63]. These figures indicate that there is still much space for Chinese overseas companies to improve their CSR image.

Fourth, create an exclusive communication mechanism. Under the influence of Chinese traditional culture, Chinese enterprises advocate keeping a low profile, thus seldom publicizing what they have done well. Liao et al. [64] conducted a comparative analysis of CSR communication of the selected contractors among Asia, EU, US/Canada, and China and found that European contractors tended to present the highest levels of CSR communication at both dimensional and issue levels, whereas Chinese contractors ranked the lowest. Keeping a low key may work in China, but for companies aiming to operate overseas, it could be the opposite because if a company keeps a low key, it means it gives up the opportunity of making itself known to the public. It is vital for a company to communicate with the outside world when there is negative news about it. Deng and Zhang [65] found that although there is no significant correlation between the quantity of media reports on a company and the public's awareness of it, negative coverage on an enterprise can significantly affect the public's negative perception of it. So publicity is indispensable to transnational corporations. They should disclose CSR information regularly and strengthen the dissemination of CSR activities: for example, releasing English CSR reports annually, enhancing communication with stakeholders, creating CSR websites, and cohosting meetings and forums on CSR with international organizations.

Fifth, strengthen the construction of CSR culture. CSR culture refers to the corporate culture guided by the concept of CSR. It reflects the awareness and sense of CSR and the corporate thinking and behavior norms based on the concept of CSR. Freeman [66] stated that a company is not only owned by the shareholders, but also by many other parties, including employees, customers, suppliers, communities, governments, trade unions, and even competitors. Therefore, an enterprise should consider the responsibilities of stakeholders other than shareholders while pursuing profits, seeking the coordinated development of the interests of the enterprise and stakeholders [24]. CSR culture plays a key role in promoting enterprises to discretionarily fulfill their social responsibilities and its function is more effective and more lasting than the mandatory role of law. To construct CSR culture is to internalize the CSR consciousness into the enterprises as a value, thus forming their inner impetus. It includes two tasks: one is to transform CSR into the consciousness and action of employees through the normative function of culture and the other is to influence external stakeholders to accept CSR culture through the radiation and cohesion of the culture.

5. Conclusion

This study finds that both FT and WSJ focus on the economic responsibilities of CCEA, with much less attention on the other social responsibilities. The overall CSR image of CCEA portrayed by the two media is that Chinese state-owned construction enterprises have conducted large-scale infrastructure construction and investments in Africa, which not only wins economic benefits for Chinese companies but also promotes the economic development of Africa. However, the CCEA are mainly state-owned enterprises, and their business activities reflect the will of their home country, thus constituting an important part of China's BRI. These activities are interpreted as significant measures of China to expand its influence in Africa. Both media, especially WSJ, pointed out that China had provided a large number of loans for African countries, and this behavior has thrown African countries into debt traps so that China is able to gain the upper hand in its negotiations with African countries for more natural resources. Moreover, the two media also reported the misconducts, like environmental damage and labor disputes, of the Chinese construction enterprises. In general both FT and WSJ portray a negative CSR image of CCEA, which coincides with such speech as "neocolonialists," "resource predators," and "ecological disruptors" prevalent in the west.

The main motives behind the biased coverage of the two media are related to their political stance, ideology, and their stereotype of the early misconducts of Chinese overseas enterprises. The Chinese companies involved in African infrastructure construction are mainly state-owned enterprises, so the western media often misinterpret the competitive advantages of these companies as China's improper subsidies to them. Moreover, China's promotion of BRI brings vast opportunities to Chinese enterprises as well as the companies of those countries which have joined the initiative. Unfortunately, this proposal is distorted by the western media as China's colonization of Africa and such misinterpretation is not only prevalent in the west, but also stirs up concerns in Africa. Besides, western media's rigid impression of the early inappropriate behaviors of Chinese overseas enterprises, like labor disputes and environmental destruction, also leads to their prejudiced coverage.

Even though the coverage of the two media contains subjective elements, it also provides a warning for CCEA to reflect on their business behaviors in the international markets. This study thus puts forward some suggestions for CCEA to perfect their CSR image, which involve comprehensive participation of stakeholders in CSR construction, better understanding of CSR, balance of four levels of CSR, and so forth. These proposals are also applicable to other Chinese overseas companies.

The present study innovates the research methods in exploring CSR image of enterprises by utilizing data mining

and corpus analysis from an interdisciplinary perspective, namely, an integration of data science, communications, management, and linguistics. However, it also has its own limitations. Only two media are included and the quantity of news samples are confined to thirty-three, which may affect the generalization of the results. Therefore, the future study is recommended to cover more media to increase the quantity of samples and a comparison of the coverage of African media and Chinese media is also worth trying.

Data Availability

The original report data of Financial Times and The Wall Street Journal are obtained from the ProQuest ABI/INFORM business information database which was released by ProQuest. Guangdong University of Foreign Studies, where the authors study or work, has purchased this database as part of the resources of its school library. The data related to statistics of word frequency and concordance of key words used to support the findings of this study are included within the article. The complete key word list generated by NVivo and the complete word lists generated by AntConc used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the 2019 Project of the National Social Science Foundation of China: On the Overseas CSR Driving Forces and Influencing Mechanism for Chinese Enterprises (19BGL116).

References

[1] "List of countries that have signed the Belt and Road Initiative cooperation document with China," 2020, https://www.yidaiyilu.gov.cn/xwzx/roll/77298.htm.

[2] A. P. Zeng, "The status quo, challenges and countermeasures of China's investment and financing in Africa," 2020, https://baijiahao.baidu.com/s?id= 1637201381073798869&wfr=spider&for=pc.

[3] S. Geerts, N. Xinwa, and D. Rossouw, *Africans' Perceptions of Chinese Business in Africa, Geneva: Globethics.net/Hatfield: Ethics Institute of South Africa*, Ethics Institute of South Africa, Hatfield, Pretoria South Africa, 2014.

[4] M. J. Kim, X. M. Yin, and G. M. Lee, "The effect of CSR on corporate image, customer citizenship behaviors, and customers' long-term relationship orientation," *International Journal of Hospitality Management*, vol. 88, Article ID 102520, 8 pages, 2020.

[5] P. C. Tang, S. X. Yang, and S. W. Yang, "How to design corporate governance structures to enhance corporate social responsibility in China's mining state-owned enterprises?," *Resources Policy*, vol. 66, Article ID 101619, 11 pages, 2020.

[6] P. S. Hofman, J. Moon, and B. Wu, "Corporate social responsibility under authoritarian capitalism: dynamics and prospects of state-led and society-driven CSR," *Business & Society*, vol. 56, no. 5, pp. 651–671, 2017.

[7] Q. Zhang, B. L. Oo, and B. T. H. Lim, "Drivers, motivations, and barriers to the implementation of corporate social responsibility practices by construction enterprises : A review," *Journal of Cleaner Production*, vol. 210, pp. 563–584, 2019.

[8] Z. Chen and T. Hamilton, "What is driving corporate social and environmental responsibility in China? An evaluation of legacy effects, organizational characteristics, and transnational pressures," *Geoforum*, vol. 110, pp. 116–124, 2020.

[9] B. Xiong, W. Lu, M. Skitmore, K. W. Chau, and M. Ye, "Virtuous nexus between corporate social performance and financial performance: a study of construction enterprises in China," *Journal of Cleaner Production*, vol. 129, pp. 223–233, 2016.

[10] Q. Zhu, J. Liu, and K.-H. Lai, "Corporate social responsibility practices and performance improvement among Chinese national state-owned enterprises," *International Journal of Production Economics*, vol. 171, pp. 417–426, 2016.

[11] J.-M. Byun and J. M. Oh, "Local corporate social responsibility, media coverage, and shareholder value," *Journal of Banking & Finance*, vol. 87, pp. 68–86, 2018.

[12] Y. Zhang and M. Cui, "The impact of corporate social responsibility on the enterprise value of China's listed coal enterprises," *The Extractive Industries and Society*, vol. 7, no. 1, pp. 138–145, 2020.

[13] F. Li and T. Morris, "An analysis of the impact of corporate visibility in print media and its effects on corporate social responsibility," *International Journal of Global Business*, vol. 11, no. 1, pp. 39–66, 2018.

[14] S. F. Cahan, C. Chen, L. Chen, and N. H. Nguyen, "Corporate social responsibility and media coverage," *Journal of Banking & Finance*, vol. 59, pp. 409–422, 2015.

[15] X. Du, H. Pei, Y. Du, and Q. Zeng, "Media coverage, family ownership, and corporate philanthropic giving: evidence from China," *Journal of Management & Organization*, vol. 22, no. 2, pp. 224–253, 2016.

[16] B. Xia, A. Olanipekun, Q. Chen, L. Xie, and Y. Liu, "Conceptualising the state of the art of corporate social responsibility (CSR) in the construction industry and its nexus to sustainable development," *Journal of Cleaner Production*, vol. 195, pp. 340–353, 2018.

[17] W. Jiang and J. K. W. Wong, "Key activity areas of corporate social responsibility (CSR) in the construction industry: a study of China," *Journal of Cleaner Production*, vol. 113, pp. 850–860, 2016.

[18] Z.-Y. Zhao, X.-J. Zhao, K. Davidson, and J. Zuo, "A corporate social responsibility indicator system for construction enterprises," *Journal of Cleaner Production*, vol. 29-30, no. 29-30, pp. 277–289, 2012.

[19] H. R. Bowen, *Social Responsibilities of the Businessman*, Harper & Row, New York, NY, USA, 1953.

[20] A. B. Carroll, "A three-dimensional conceptual model of corporate performance," *Academy of Management Review*, vol. 4, no. 4, pp. 497–505, 1979.

[21] A. B. Carroll, "The pyramid of corporate social responsibility: toward the moral management of organizational stakeholders," *Business Horizons*, vol. 34, no. 4, pp. 39–48, 1991.

[22] A. Dahlsrud, "How corporate social responsibility is defined: an analysis of 37 definitions," *Corporate Social Responsibility and Environmental Management*, vol. 15, no. 1, pp. 1–13, 2008.

[23] C. Y. An, "China's CSR in Africa: a case study of Zambia-China economic and trade cooperation zone," *Asia & Africa Review*, vol. 2, pp. 92–130, 2014.

[24] Z. F. Liu, *Corporate Social Responsibility and Corporate Image Building*, China Financial & Economic Publishing House, Beijing, China, 2008.
[25] B. R. Yang, "Summary of international symposium on corporate social responsibility and development in africa," *West Asia and Africa*, vol. 8, pp. 73–75, 2010.
[26] C.-L. Wu, D.-P. Fang, P.-C. Liao, J.-W. Xue, Y. Li, and T. Wang, "Perception of corporate social responsibility: the case of Chinese international contractors," *Journal of Cleaner Production*, vol. 107, pp. 185–194, 2015.
[27] N. Nguyen and G. Leblanc, "Corporate image and corporate reputation in customers' retention decisions in services," *Journal of Retailing and Consumer Services*, vol. 8, no. 4, pp. 227–236, 2001.
[28] G. B. Song, *Study on CSR Building of Chinese Enterprises in Africa, Shanghai*, Shanghai Institutes for International Studies, Shanghai, China, 2012.
[29] J. Fang, "Some considerations about Chinese enterprises' social responsibility in Africa," *Petroleum & Petrochemical Today*, vol. 26, no. 10, pp. 42–47, 2018.
[30] Q. M. Guo, "On image of China and China's MNCs in Africa," *China Opening Journal*, vol. 4, pp. 37–40, 2012.
[31] X. H. Chen, T. Jia, and Y. Zhou, "Chinese corporate social responsibility in Africa: present state and proposals," *China Development Review*, vol. 12, no. 1, pp. 75–83, 2010.
[32] X. S. Han, "The development of Chinese enterprises in Africa and their social responsibilities," *Journal of International Economic Cooperation*, vol. 7, pp. 18–20, 2007.
[33] J. C. Shu, "Small and medium sized enterprises' CSR in Africa: status quo and some recommendations," *Asia & Africa Review*, vol. 5, pp. 54–62, 2011.
[34] H. W. Zhong and X. K. Yang, "Social responsibility: a problem that Chinese overseas companies must handle," *China WTO Tribune*, vol. 8, pp. 50–52, 2007.
[35] X. L. Wang, "A study on the social responsibility of Chinese enterprises in Africa - A case study of the Mombasa-Nairobi Railway," *Economic Forum*, vol. 12, pp. 139–141, 2016.
[36] S. M. Zhan, "Fulfilling corporate social responsibility promotes the common development of China and Africa-investigation report on 'Chinese enterprises fulfilling social responsibilities in Africa'," *West Asia and Africa*, vol. 7, pp. 63–66, 2008.
[37] J. Feng, "Chinese contractor involvement in wildlife protection in Africa: case study of mombasa-nairobi standard Gauge railway project, Kenya," *Land Use Policy*, vol. 95, Article ID 104650, 7 pages, 2020.
[38] Y. X. Tan and C. C. Hua, "Analysis of cluster on the inversion problem of network public opinion events," *Journal of Yunnan University*, vol. 41, no. S1, pp. 16–20, 2019.
[39] L. H. Liu, *Research on National, Regional and Corporate Communication: A Perspective of Discourse Analysis*, People Daily Press, Beijing, China, 2016.
[40] Wikipedia, "Financial Times," 2020, https://en.m.wikipedia.org/wiki/Financial_Times.
[41] Wikipedia, "The Wall Street Journal," 2020, https://en.m.wikipedia.org/wiki/The_Wall_Street_Journal.
[42] G. F. Yin, Y. Wang, Z. S. Guan et al., "The development course and characteristics of CSR in Africa," *China WTO Tribune*, vol. 11, pp. 17–26, 2016.
[43] D. Biber and University language, *A Corpus-Based Study of Spoken and Written Registers*, John Benjamins, Amsterdam & Philadelphia, 2006.
[44] M. J. Luzon, "'Your argument is wrong': a contribution to the study of evaluation in academic weblogs," *Text & Talk*, vol. 32, no. 2, pp. 145–165, 2012.
[45] D. Biber, "Stance in spoken and written university registers," *Journal of English for Academic Purposes*, vol. 5, no. 2, pp. 97–116, 2006b.
[46] M. Hewings, "An important contribution? or tiresome reading? A study of evaluation in peer reviews of journal article submissions," *Journal of Applied Linguistics*, vol. 1, no. 3, pp. 247–274, 2004.
[47] M. Nelson, "A corpus-based study of the lexis of business english and business English teaching materials," Unpublished thesis, University of Manchester, Manchester, England, UK, 2000.
[48] R. Poole, "New opportunities and Strong performance: evaluative adjectives in letters to shareholders and potential for pedagogically-downsized specialized corpora," *English for Specific Purposes*, vol. 47, pp. 40–51, 2017.
[49] M. Feng and J. Wei, *An Introduction to Corporate Social Responsibility*, Economic Science Press, Beijing, China, 2017.
[50] G. R. Hu, "The influence of competitive neutrality on state-owned enterprises in China and the legal response," *Science of Law*, vol. 6, pp. 165–172, 2014.
[51] L. Q. Bu, "Competitive neutrality and preliminary discussion on Chinese state-owned enterprise reform," *Economic Law Forum*, no. 2, pp. 95–121, 2017.
[52] K. J. Lin, X. Lu, J. Zhang, and Y. Zheng, "State-owned enterprises in China: a review of 40 years of research and practice," *China Journal of Accounting Research*, vol. 13, no. 1, pp. 31–55, 2020.
[53] P. R. Carmody and F. Y. Owusu, "Competing hegemons? Chinese versus American geo-economic strategies in Africa," *Political Geography*, vol. 26, no. 5, pp. 504–524, 2007.
[54] S. Naidu and D. Mbazima, "China-African relations: a new impulse in a changing continental landscape," *Futures*, vol. 40, no. 8, pp. 748–761, 2008.
[55] W. Zhou and F. Zhao, "Is China-Africa cooperation under the belt and road initiative neo-colonialism?" *Studies on Marxism*, vol. 1, pp. 129–152.
[56] H. Z. Ma and E. F. Cheng, "An analysis of 'new neocolonialism' in China-Africa economic and trade relations," *Shanghai Journal of Economics*, vol. 4, pp. 96–106, 2018.
[57] M. X. Sun, "The predicament of China's national image construction and and breakthrough," *Journal of the Central Institute of Socialism*, vol. 4, pp. 64–70, 2018.
[58] X. L. Han, "Reflections on Chinese economic engagement in Africa from the perspective of international law-Refuting western slander of neocolonialism through grafting and transplanting," *Journal of International Economic Law*, vol. 2, pp. 31–48, 2019.
[59] C. Peng, *A Study of the News Coverage on Chinese Privately-Owned Multinational Enterprises in the Main Stream Media of the USA and UK*, Shanghai International Studies University, Shanghai, China, 2014.
[60] C. X. Li, *'Image of Made in China' Constructed by American Media and Solution to the Risk*, Fudan University, Shanghai, China, 2010.
[61] J. Yan and J. Y. Shi, "Image of 'Made in China' depicted by western media-An content analysis of the Wall Street Journal in 2007 and 2008," *Henan Social Sciences*, vol. 18, no. 1, pp. 183–186, 2010.
[62] X. L. Wang and G. Han, "'Made in China' and national image: a content analysis of 30 years' coverage in American

mainstream media," *Chinese Journal of International Communication*, no. 9, pp. 49–55, 2010.

[63] P. F. Li, "Reputational risk of Chinese companies in Africa," *Sino Foreign Management*, vol. 9, p. 31, 2014.

[64] P.-C. Liao, N.-N. Xia, C.-L. Wu, X.-L. Zhang, and J.-L. Yeh, "Communicating the corporate social responsibility (CSR) of international contractors: content analysis of CSR reporting," *Journal of Cleaner Production*, vol. 156, pp. 327–336, 2017.

[65] L. F. Deng and N. Zhang, "Agenda-setting effects of business news on corporate reputation: research on the news of corporate social responsibility in China," *Modern Communication*, vol. 5, pp. 119–125, 2013.

[66] R. E. Freeman, *Strategic Management: A Stakeholder Approach*, Pitman, Boston, MA, USA, 1984.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 1847194, 7 pages
https://doi.org/10.1155/2020/1847194

Research Article

The Impact of Online Media Big Data on Firm Performance: Based on Grey Relation Entropy Method

Hai-qing Qin,[1] **Zhen-hui Li**,[1] **and Jia-jia Yang**[2]

[1]*School of Economics and Management, Communication University of China, Beijing 100024, China*
[2]*Department of Public Policy, King's College London, London, WC2R 2LS, UK*

Correspondence should be addressed to Hai-qing Qin; alex_qin@cuc.edu.cn

Received 4 May 2020; Accepted 15 June 2020; Published 14 July 2020

Guest Editor: Wen-Tsao Pan

The study uses the grey relation entropy method to explore the impact of online media big data on firm performance, based on 17 randomly selected Chinese A-share listed companies during the period from 2012 to 2017. It shows that the media big data, especially the negative media coverage, is highly associated with both short-term and long-term firm performance. Then, this study employs the system GMM method to testify how negative media coverage affects firm performance. It indicates that negative media coverage may be a damage crisis for the focal firm in the short term, but a favorable chance for change in the long run. These findings not only enrich the research on the influence of online media big data but also provide some references for enterprise managers.

1. Introduction

Past research has highlighted the "4V" features of big data [1], which garnered significant attention among scholars and managers. Online media coverage contains a huge amount of information and spreads quickly nowadays. There are more than 60 million pieces of original news and 8.3 million pieces of news of listed companies by the mid-2019 calculated by the Chinese Research Data Services Platform. Moreover, online media coverage also has characteristics such as relevance, scalability, and authenticity. Meanwhile, the value of online media coverage data resides in the fact that it contains multiple aspects of information [2, 3]. As such, we theorize that online media coverage with more massive volume, higher velocity, and more variety can be defined as online media big data.

With the development of China's media convergence process, media big data has become an essential factor influencing the development of enterprises. As a vital information infomediary [4], the media plays several major functions or roles as they report on firm issues [5]. It provides message about firm issues that curtails the information asymmetry and also frames issues through "persistent patterns of cognition, interpretation, and presentation, of selection, emphasis, and exclusion" [6, 7]. Thus, media coverage can affect firm's impression, reputation, and legitimacy [8]. Since the Internet drastically scales down the participation cost of information receivers, online media allows a huge number of users to communicate mutually [9], and public attention can be quickly focused [10]. Online media big data, especially online financial news, is a vital channel for stakeholders to obtain a focal firm's information [11]. The tones of financial news convey how journalists perceive the issue or the firm, which may form, enhance, or even alter the audience's impression of the focal firm.

A flood of positive media coverage of the film *Lost in Russia* released online for free watching during the 2020 Chinese New Year helped boost the share price of Huanxi Media Group Limited. Institutional experts regard this phenomenon as legitimation processes, in which firms can gain stakeholders preference. Pollock and Rindova (2003) provided substantial evidence by analyzing 225 IPOs and found that the positive tenor of media coverage has a positive relationship with underpricing [7]. In organizational literature, researchers emphasized the role of media in serving as an effective external governance mechanism. They hold that

the media can help ensure the interests of stakeholders, curtail information asymmetry, and offer overall strategic guidance [8]. Since July of 2018, the "fake doctor" scandal of Meinian Healthcare has pushed it to the focus of public attention, resulting in performance declining as well as market value shrinking. Meinian Healthcare survived the winter after having made many changes during the past two years. In this case, the media served as a social arbiter, which has increased the reputational cost of its misbehavior [12], but a firm change driver who helped executives realize the potential threat and take actively strategic change. However, these results are mainly driven by traditional media (e.g., newspaper) [13], and the conclusion is not yet consistent. Nevertheless, to date, research has not evaluated the different aspects of media coverage. The role of online media should be further explored, considering the immense volume of information and audiences' limited attention.

In this paper, by analyzing how online media big data affects firm performance, two methods are applied. To explore the complicated relationship between online media big data and firm performance, we first use the grey relation entropy method to measure what is the essential aspect and then conduct a dynamic panel-data model to further unfold whether it affects firm performance positively or not.

It is shown that the salience, the tones, and the overall tenor of media big data all matter in framing firm reputation. The impact is more pronounced when the media information spreads to a larger area. The positive tenor of media coverage may provide firms with chances to achieve sustained superior profit outcomes. What unexpected we found is that the negative media coverage has a long-lasting influence on firm performance. The negative media coverage may damage a focal firm's current financial performance. However, it can serve as an efficient external governance mechanism in the long run as well, which may provide executives with strategic guidance to focus on long-term firm performance.

2. Literature Review

Accordingly, there are two primary roles that the media plays [14]. Firstly, the media provides a platform to disseminate information, which can help reduce the information asymmetry between internal and external stakeholders of a focal firm [15]. Stakeholders can experience even distant events about a firm [16], especially in the case that the online media has broken the boundary of time and space [17]. In this espousing process, the firm can increase its familiarity with stakeholders, such as investors. Thus, the media has assisted firms in gaining public attention [7] which depends on the volume of media coverage. Signal theory holds that repeated signals increase the impression of the audience [18]. As demonstrated by Dyck et al., the impact of media is more pronounced when the media reach a larger number of relevant groups [12].

Secondly, the media can frame firm issues in positive or negative terms, which conveys the approval or disapproval judgements of firms and their actions [7, 19]. Therefore, the media can affect cognition and behavioural preferences of stakeholders, the reputation of firms [15], and the reputation of managers [12]. Past research proposes that positive media coverage is beneficial for firm reputation [15] and legitimacy [7]. These benefits above can improve the availability of resources the firm needs [20], which is conducive to the improvement of firm's profit outcomes [21]. Besides, it can also bring higher personal fame to managers and other internal stakeholders [22]. As an external incentive, it is beneficial to improve employees' work enthusiasm to a certain extent [23].

Online media coverage for adverse events, on the other hand, provides a field of public opinion, which may accelerate the fermentation of negative events, and therefore, damage the company reputation and legitimacy. The higher the volume of adverse online media coverage, the stronger the negative signals will be transmitted, which will cause deleterious online public opinions and lead to greater market punishment for enterprises [24]. For example, stock price crashes [20]. For enterprise managers and the internal staff, it can also cause negative emotions, resulting in a decline in confidence [12], and decrease work enthusiasm, which is not conducive to enterprise performance improvement.

However, the study believes that negative media reports have specific supervisory effects on enterprises. Dyck and Zingales first embark from the reputation mechanism and emphasize on the media as a social mediator of vital force to protect the interests of the investors [25]. Dyck et al. argue that the media plays a significant role in exposing corporate scandals and is able to stop unreasonable allocation of resources timely [26].

3. Data Sources and Variables

Data for corporate performance and corporate financial data are derived from the China Stock Market and Accounting Research (CSMAR) database. The data for online media big data comes from the database of China Services (CNRDS), which provides reliable support for various studies in the Chinese market [27]. The CNRDS platform gathered online financial media coverage articles posted by online media financial accounts, including 20 mainstream online financial media in China, such as Hexun.com, Sina Finance, and China Economic Net. All articles are coded by the sentiment of the title and content as positive, neutral, or negative by using artificial intelligence algorithms [28].

Moreover, we used a stratified random sampling method to obtain sample firms that cover almost every industry according to the two-digit industry codes of China Securities Regulatory Commission. Due to the particularity of the financial industry, it was not included in the sample. Finally, 17 Chinese A-share listed companies during the period from 2012 to 2017 were studied.

The variables for the grey relation entropy method include different aspects of online financial news. Salience, the boldness of the articles, was calculated based on the total number of articles with the firm named in the title. Volume was calculated by the total number of articles of a focal firm. Positive, which represents that there is a favorable tone of the

article, was measured by the total number of articles with a positive tone. Negative was measured by the total number of articles on an adverse tone. Besides, we also use the total number of original ones to see if there is any difference. Tenor, which means that the overall tenor of media coverage, was measured by the Janis–Fader coefficient of imbalance [7, 15, 29]. This measure was calculated using the formula:

$$\text{Tenor} = \frac{e^2 - ec}{t^2}, \quad \text{if } e > c; 0, \text{if } e = c, \\ \frac{ec - c^2}{t^2}, \quad \text{if } c > e, \tag{1}$$

where e is the number of positive articles about a firm, c is the number of negative articles about it, and t is the total volume of articles about it, including articles that are neutral in tenor. The range of this variable is −1 to 1, where −1 equals "all negative presses" and 1 equals "all positive presses." Following the past literature, we use ROA (return on assets) and ROE (return on equity) to measure firm performance.

4. Mathematical Model and Data Analysis

4.1. Grey Relation Entropy Method. Grey relational analysis approach is a method to measure the degree of correlation among factors according to the degree of similarity or difference of the development trend among factors. However, because the average value of grey correlation coefficient conceals many sparse features of grey management, it is not possible to make full use of the rich information provided by the coefficient of point management. Based on the grey correlation analysis, the grey relation entropy method is introduced to ameliorate the lack of grey correlation analysis [30]. In this paper, we use the grey relation entropy method to decide which aspect or dimension of media big data is the most vital one and how the correlation may change from time to time. Some firms receive more media attention while others do not (see Table 1), which implies that media environment differs from firm to firm.

Firstly, we set the reference sequence and comparison sequence as below.

The reference sequence:

$$A_0 = \left[X_{01}, X_{02}, X_{0j}, \ldots, X_{0n}\right]. \tag{2}$$

The comparison sequence:

$$A_i = \left[X_{i1}, X_{i2}, X_{ij} \ldots, X_{in}\right], \quad i = 1, 2, 3, \ldots, m, \tag{3}$$

where X_{0j} refers to firm performance and X_i refers to different aspects of media big data.

First, we initialize the data:

$$X_i = \frac{\max X_i - X_i}{\max X_i - \min X_i}. \tag{4}$$

Calculate the grey correlation distance Δ_{0ij}, where Δ_{0ij} is the distance between each comparison sequence and the reference sequence. The formula is

$$\Delta_{0ij} = \left|X_{0j}^* - X_{ij}^*\right|. \tag{5}$$

Then, calculate the grey relation coefficient:

$$R_{oij} = \frac{\Delta_{\min} + \rho\Delta_{\max}}{\Delta_{0ij} + \rho\Delta_{\max}}, \tag{6}$$

where ρ is the discrimination coefficient $[0, 1]$, usually $\rho = 0.5$. Results (part of them) are shown in Table 2.

The grey relation coefficient (R) distribution map value P:

$$P_h = \frac{R_h}{\sum_1^n R_h}, \; P_h \in P_i, \quad h = 1, 2, 3, \ldots, n. \tag{7}$$

The grey relation entropy value (see Table 3) is

$$H\left(R_i\right) = -\sum_1^n \left(P_h \ln\left(P_h\right)\right). \tag{8}$$

According to the law of entropy, when the grey correlation entropy of sequence X is the largest, it means that the influence of X points on the reference sequence is balanced. In other words, the grey correlation entropy reaches its maximum value when the distribution map values of each grey correlation coefficient are equal.

The maximum value is

$$\text{Hm}\left(X\right) = \ln n. \tag{9}$$

The grey relation entropy correlation degree is

$$E_j\left(X_i\right) = \frac{H\left(R_i\right)}{H_m}. \tag{10}$$

The GREM results (see Table 4) show that online media big data is highly associated with firm performance. Although there are subtle differences between several aspects of media big data, it can tell that the impact differs. Generally, the negative of online media coverage is highly associated with firm performance followed by the tenor of media coverage.

The positive (original) online media coverage is highly associated with both short-term and long-term firm performance, which indicates that audiences may prefer to believe credible sources when the news about a focal firm is framed in a positive tone.

Total online media coverage is highly associated with firm performance than the original ones when it comes to negative articles, which hints that the impact of negative media coverage may rely on the scope it spreads.

The correlation between the tenor of media coverage and firm performance has not changed much over time.

4.2. System GMM Method for Dynamic Panel Data. Following consideration of the negative media coverage is the highest correlated factor with firm performance, we conduct an econometric model to testify how it affects firm performance. Before the estimation, we use the Winsorize method to avoid the influence of extreme values on data analysis. Furthermore, there are some endogenous problems, for as much as firm performance and negative media coverage may have mutual causality and firm performance may be affected by previous firm performance. To reduce the aforementioned endogeneity, this study employs the system GMM method that fits a linear dynamic panel-data model

Table 1: The summary statistics of the variables for GREM.

Category	Variables	Mean	Std. dev.	Min	Max
A_{01}	ROA_t	0.033	0.069	−0.453	0.189
A_{02}	ROE_t	0.052	0.146	−1.067	0.243
A_{03}	ROA_{t+1}	0.023	0.121	−0.927	0.168
A_{04}	ROE_{t+1}	0.039	0.210	−1.433	0.248
A_1	Salience	161.029	137.830	1.000	656.000
A_2	Positive (total)	157.804	149.220	0.000	631.000
A_3	Positive (original)	47.029	46.525	0.000	216.000
A_4	Negative (total)	143.990	219.920	3.000	1588.000
A_5	Negative (original)	33.549	42.166	2.000	187.000
A_6	Tenor	0.057	0.154	−0.563	0.423

Table 2: The grey relation coefficient (e.g., A_{01} and A_1).

	A_{01}					
	[1]	[18]	[35]	[52]	[67]	[84]
	0.4722399	0.3956641	0.6797184	0.6525302	0.9256974	0.3996716
	0.5199922	0.4790806	0.4826101	0.4245426	0.8051225	0.3954566
	0.4540571	0.4029553	0.5769253	0.4296573	0.5136660	0.4099845
	0.4644591	0.4384371	0.7501023	0.4039934	0.4277149	0.3730879
	0.4388676	0.4658275	0.9753574	0.4119134	0.4921814	0.3956314
	0.4667398	0.4642610	0.5936379	0.4228632	0.4241844	0.3719485
	0.5463972	0.3794781	0.7400184	0.4766334	0.5625606	0.3896975
	0.7584622	0.4569635	0.5776415	0.7615225	0.5072221	0.4068271
A_1	0.4649920	0.4524507	0.3870351	0.3999047	0.4044883	0.4350132
	0.8396081	0.4531167	0.4089473	0.5365372	0.5481666	0.4579653
	0.5533175	0.5787409	0.4168465	0.4600016	0.4574930	0.3823617
	0.4453939	0.4358607	0.4730191	0.5835118	0.5072272	0.4256498
	0.4091393	0.4039758	0.4267023	0.6889290	0.5207435	0.4436271
	0.4476427	0.4698516	0.3714691	0.6078893	0.5436337	0.4825398
	0.4114355	0.4960617	0.5466753	0.4476838	0.4137933	0.5029635
	0.7126215	0.5398761	0.5837759	0.4117442	0.4247983	0.4509281
	0.4321090	0.6619091	0.9465222	0.6797493	0.4398294	0.4500233

Table 3: The grey relation entropy results.

Category	Variables	ROA_t	ROE_t	ROA_{t+1}	ROE_{t+1}
H_1	Salience	4.595743	4.59178	4.589593	4.586525
H_2	Positive (total)	4.589219	4.592326	4.591505	4.589894
H_3	Positive (original)	4.594324	4.597193	4.595985	4.593646
H_4	Negative (total)	4.615016	4.609228	4.608330	4.605134
H_5	Negative (original)	4.585639	4.582943	4.583577	4.581322
H_6	Tenor	4.610332	4.605310	4.607550	4.606146

Table 4: The entropy correlation degree.

Category	Variables	ROA_t	ROE_t	ROA_{t+1}	ROE_{t+1}
E_1	Salience	0.9936800	0.9928231	0.9923504	0.9916868
E_2	Positive (total)	0.9922695	0.9929412	0.9927636	0.9924154
E_3	Positive (original)	0.9933731	0.9939936	0.9937323	0.9932265
E_4	Negative (total)	0.9978472	0.9965957	0.9964016	0.9957105
E_5	Negative (original)	0.9914953	0.9909125	0.9910495	0.9905619
E_6	Tenor	0.9968345	0.9957486	0.9962329	0.9959293

Table 5: Results of system GMM for dynamic panel data.

Model	Model1 (FP = ROA)			Model2 (FP = ROE)		
Variables	Coef.	Std. Err.	z	Coef.	Std. Err.	z
Constant	−0.017	0.189	−0.09	−0.364	0.526	−0.69
FP L1.	0.050	0.139	0.36	−0.040	0.034	−1.18
FP L2.	0.119	0.148	0.80	−0.082***	0.030	−2.74
Firm age	0.018	0.016	1.11	0.023	0.035	0.64
Firm size	0.000	0.008	0.01	0.021	0.023	0.92
Slack	−0.001	0.006	−0.15	0.004	0.021	0.20
MSh	0.065*	0.034	1.90	0.251**	0.113	2.23
CEO tenure	−0.050**	0.023	−2.14	−0.186**	0.087	−2.14
CEO duality	0.008	0.011	0.75	−0.009	0.030	−0.31
Negative	−0.011*	0.006	−1.85	−0.038**	0.018	−2.06
Negative L1.	0.009*	0.005	1.68	0.023**	0.010	2.33
Negative L2.	0.007	0.005	1.38	0.004	0.014	0.30
Wald chi2	126.21***			111.39***		

Note. Obvious number = 68. Industry dummy variable is controlled but not presented. $^{*}P < 0.10$; $^{**}P < 0.050$; and $^{***}P < 0.010$.

where the unobserved panel-level effects are correlated with the lags of the dependent variable, to evaluate the relationship between media coverage and firm performance [31].

4.2.1. Dependent Variables. Firm performance (FP) was measured as the total return on assets (ROA) in a given year and the total return on equity (ROE) in a given year.

4.2.2. Explanatory Variables. Negative online media coverage was calculated as the natural logarithm of the total number of online media articles reporting about each listed firm in negative tones plus one in a given year.

4.2.3. Control Variables. Consistent with the previous literature of corporate performance, we employ control variables of organizational characteristics, corporate governance, and CEO characteristics [5]. We select *firm age* (calculated as the natural logarithm of its duration from the listed year to the sample year), *firm size* (measured as the natural logarithm of total assets), *firm slack* (current asset/current debt), *MSh* (management shareholding ratio), *CEO tenure* (measured as the entire year the CEO served in the focal firm), and *CEO duality* (dummy variable). We also control for *industry* as a dummy variable.

In this paper, the model used for the analysis of the impact of online media coverage on firm performance is

$$\begin{aligned} \mathrm{FP}_{\mathrm{it}} &= \alpha + \beta_1 \mathrm{FP}_{\mathrm{it-1}} + \beta_2 \mathrm{FP}_{\mathrm{it-2}} + \beta_3 \mathrm{Negative}_{\mathrm{it}} \\ &\quad + \beta_4 \mathrm{Negative}_{\mathrm{it-1}} + \beta_5 \mathrm{Negative}_{\mathrm{it-3}} + \beta_6 \mathrm{Controls}_{\mathrm{it}} \\ &\quad + \varepsilon, \quad n = 0, 1, 2. \end{aligned} \tag{11}$$

The Arellano–Bover/Blundell–Bond model estimation results (see Table 5) show that the negative online media coverage has different impacts on short-term firm performance and long-term firm performance. The coef. of Negative in Mode1 (−0.011, $P < 0.01$) implies that negative media coverage is adversely associated with short-term firm performance. However, the coef. of Negative L.1 (0.001, $P < 0.01$) indicates a positive relationship between negative media coverage and long-term firm performance. Moreover, the influence of negative media coverage on firm performance gradually converged in the third year and began to become insignificant (see coef. of Negative L2. of Model1). The consequences of Model2, where firm performance is measured by ROE, are consistent with Model1, which suggests the findings are robust.

5. Results and Discussion

Results show that online media big data plays a vital role in firm performance through a variety of aspects.

It is suggested in the results of grey relation entropy that the heterogeneity in different aspects of online media big data is not apparent. However, the subtle difference shows that the positive media coverage is highly associated with short-term firm performance and by the same token it is highly associated with long-term firm performance. And the original positive media coverage is more pertinent to corporate performance than total positive ones. This may be due to the dependence of information credibility when it comes to affirmative reporting. On the contrary, unlike the highlighted beneficial impact of the positive media coverage in the previous literature, the negative media coverage is of the highest correlation with firm performance. Furthermore, the total negative media coverage is highly associated with firm performance followed by the original negative ones, which indicates the scope of negative information may play a more essential role in affecting firm performance. Additionally, the tenor of media big data is highly associated with firm performance as well, followed by the salience.

The system GMM results indicate that the negative media coverage may have an opposite impact in the short and long run. Although the negative media coverage may cause considerable damage to the focal firm in a short-term, it can be positively associated with long-term firm performance due to the social arbiter role it plays. This is consistent with prior research that demonstrated that negative media

coverage may offer managers strategic guidance [5] and promote the efficiency of internal corporate governance (e.g., board quality) [32].

6. Conclusion

Based on the GRE method and the system GMM method, this paper explored which is the highest correlated factor of online media big data for firm performance and how it affects firm performance. The results have both theoretical and practical implications.

The results expanded the research on media coverage and firm performance by considering different aspects of online media big data. It is found that not only the volume and tenor of media coverage serve as vital roles in exposing and framing firm issues but also the salience of the news is of great importance especially in the big data era.

There are some practical implications given the high correlation between media big data and corporate performance. Firstly, managers can use media big data to predict corporate performance and keep abreast of stakeholders' evaluation of the company. In particular, the external supervisory role of media big data should be emphasized. Secondly, executives should incorporate online media coverage into decision-making reference. Specifically, managers should not indulge in a favorable media environment and be overconfident. On the contrary, they should reflect on and make corresponding adjustments in the face of negative public opinions given the vigorous impact of stakeholders' perception of the company. Timely and accurate response may help the company stop losses to a certain extent. Thirdly, managers can make full use of online media and timely disclose information through the network platform, such as the enterprise official accounts, to curtail information asymmetry with stakeholders. Audiences may take the initiative to disclose negative information as a manifestation of the company's responsibility and choose to trust the focal firm again. Thus, it would be better to learn firm information from official and credible sources rather than informal sources, especially negative information. Lastly and importantly, investors and consumers can also have a more comprehensive understanding of the company's information through online media coverage and avoid the potential loss caused by information asymmetry.

There are also several limitations to this study, which directs our suggestions for future research. An initial limitation is the rough category of the aspects of online media big data. Our analysis basically measures different aspects by the number of articles, but does not examine the content characteristics. It would be interesting to use the text analysis method to further explore the impact of specific words or phrases based on artificial intelligence algorithms. Furthermore, the current study mainly concentrated on the supervisory role of negative media coverage following the majority of previous studies. Nonetheless, it is supposed to delve into how negative media coverage serves as an external governance mechanism. It does highlight the lack of consideration of overall aspects of online media big data, while this in itself is not an apparent limitation. Future research may benefit from further unfolding the overall impacts of online media big data on firm performance through the use of various data mining methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was financially supported by the Asian Media Research Center Project of China-"Research on the Improvement Path of China's International Communication Ability in the New Media Era" (Grant nos. AMRC2018-3), National Natural Science Foundation of China-"Research on factors and process of successful academic entrepreneurship within the perspective of institutional complexity" (Grant no. 71772167), and National Radio and Television Administration Ministerial Social Science Research of China-"Research on China's Media Convergence Ecosystem in 5G" (Grant no. GDT1912).

References

[1] A. McAfee, E. Brynjolfsson, T. H. Davenport et al., "Big data: the management revolution," *Harvard Business Review*, vol. 90, no. 10, pp. 60–66, 2012.

[2] M. Nathan and J. Warren, *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*, Manning Publications Co., Shelter Island, NY, USA, 2015.

[3] J. von Bloh, T. Broekel, B. Özgun et al., "New (s) data for entrepreneurship research? An innovative approach to use Big Data on media coverage," *Small Business Economics*, pp. 1–22, 2019.

[4] D. L. Deephouse and P. P. M. A. R. Heugens, "Linking social issues to organizational impact: the role of infomediaries and the infomediary process," *Journal of Business Ethics*, vol. 86, no. 4, pp. 541–553, 2009.

[5] M. K. Bednar, S. Boivie, and N. R. Prince, "Burr under the saddle: how media coverage influences strategic change," *Organization Science*, vol. 24, no. 3, pp. 910–925, 2013.

[6] T. Gitlin, "The world is watching: mass media in the making and unmaking of the new left," 1980.

[7] T. G. Pollock and V. P. Rindova, "Media legitimation effects in the market for initial public offerings," *Academy of Management Journal*, vol. 46, no. 5, pp. 631–642, 2003.

[8] R. V. Aguilera, K. Desender, M. K. Bednar, and J. H. Lee, "Connecting the dots: bringing external corporate governance into the corporate governance puzzle," *The Academy of Management Annals*, vol. 9, no. 1, pp. 483–573, 2015.

[9] L. Salter, *Democracy, New Social Movements, and the Internet: A Habermasian Analysis*, Cyberactivism. Routledge, Abingdon, UK, 2013.

[10] X. R. Luo, J. Zhang, and C. Marquis, "Mobilization in the Internet age: Internet activism and corporate response,"

Academy of Management Journal, vol. 59, no. 6, pp. 2045–2068, 2016.

[11] C. E. Carroll and M. McCombs, "Agenda-setting effects of business news on the public's images and opinions about major corporations," *Corporate Reputation Review*, vol. 6, no. 1, pp. 36–46, 2003.

[12] A. Dyck, N. Volchkova, and L. Zingales, "The corporate governance role of the media: evidence from Russia," *The Journal of Finance*, vol. 63, no. 3, pp. 1093–1135, 2008.

[13] T. Nokelainen and J. Kanniainen, *Coverage Bias in Business News: Evidence and Methodological Implications*, Management Research Review, Bingley, UK, 2018.

[14] B. J. Bushee, J. E. Core, W. Guay, and S. J. W. Hamm, "The role of the business press as an information intermediary," *Journal of Accounting Research*, vol. 48, no. 1, pp. 1–19, 2010.

[15] D. L. Deephouse, "Media reputation as a strategic resource: an integration of mass communication and resource-based theories," *Journal of Management*, vol. 26, no. 6, pp. 1091–1112, 2000.

[16] J. M. McLeod, G. M. Kosicki, and Z. Pan, *On Understanding and Misunderstanding Media Effects*, Mass Media and Society, London, UK, 1991.

[17] M.-L. Lau and B. Wydick, "Does new information technology lower media quality? The paradox of commercial public goods," *Journal of Industry, Competition and Trade*, vol. 14, no. 2, pp. 145–157, 2014.

[18] J. J. Janney and T. B. Folta, "Signaling through private equity placements and its impact on the valuation of biotechnology firms," *Journal of Business Venturing*, vol. 18, no. 3, pp. 361–380, 2003.

[19] K. Lamertz and J. AC. Baum, "The legitimacy of organizational downsizing in Canada: an analysis of explanatory media accounts," *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, vol. 15, no. 1, pp. 93–107, 1998.

[20] A. V. Shipilov, H. R. Greve, and T. J. Rowley, "Is all publicity good publicity? The impact of direct and indirect media pressure on the adoption of governance practices," *Strategic Management Journal*, vol. 40, no. 9, pp. 1368–1393, 2019.

[21] S. Kiousis, C. Popescu, and M. Mitrook, "Understanding influence on corporate reputation: an examination of public relations efforts, media coverage, public opinion, and financial performance from an agenda-building and agenda-setting perspective," *Journal of Public Relations Research*, vol. 19, no. 2, pp. 147–165, 2007.

[22] J. R. Nofsinger, "The impact of public information on investors," *Journal of Banking & Finance*, vol. 25, no. 7, pp. 1339–1366, 2001.

[23] M. L. A. Hayward, V. P. Rindova, and T. G. Pollock, "Believing one's own press: the causes and consequences of CEO celebrity," *Strategic Management Journal*, vol. 25, no. 7, pp. 637–653, 2007.

[24] S. A. Hawkins and S. J. Hoch, "Low-involvement learning: memory without evaluation," *Journal of Consumer Research*, vol. 19, no. 2, pp. 212–225, 1992.

[25] A. Dyck and L. Zingales, "The corporate governance role of the media the right to tell: the role of mass media in economic development," 2002.

[26] A. Dyck, A. Morse, and L. Zingales, "Who blows the whistle on corporate fraud?" *The Journal of Finance*, vol. 65, no. 6, pp. 2213–2253, 2010.

[27] J.-B. Kim, L. Li, Z. Yu, and H. Zhang, "Local versus non-local effects of Chinese media and post-earnings announcement drift," *Journal of Banking & Finance*, vol. 106, pp. 82–92, 2019.

[28] https://www.cnrds.com/Home/Index#/FeaturedDabase/DB/CFND/ViewName/Basic Information on Internet Financial News, in Chinese.

[29] I. L. Janis and R. Fader, "The coefficient of imbalance," in *Language of Politics*, H. Lasswell and N. Leites, Eds., pp. 153–169, MIT Press, Cambridge, MA, USA, 1965.

[30] Q. Zhang and J. Deng, "Grey relation entropy method of grey relation analysis," *Systems Engineering-Theory & Practice*, vol. 16, no. 8, pp. 7–11, 1996, in Chinese.

[31] F. Windmeijer, "A finite sample correction for the variance of linear efficient two-step GMM estimators," *Journal of Econometrics*, vol. 126, no. 1, pp. 25–51, 2005.

[32] J. R. Joe, H. Louis, and D. Robinson, "Managers' and investors' responses to media exposure of board ineffectiveness," *Journal of Financial and Quantitative Analysis*, vol. 44, no. 3, pp. 579–605, 2009.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 1283740, 14 pages
https://doi.org/10.1155/2020/1283740

Research Article

Fuzzy Comprehensive Evaluation of Decoupling Economic Growth from Environment Costs in China's Resource-Based Cities

Zhidong Li[1] **and Zhifan Zhou**[2]

[1]*School of Economics and Management, Hefei Normal University, Hefei, China*
[2]*Northwestern Polytechnical University, Xi'an, China*

Correspondence should be addressed to Zhidong Li; 502597394@qq.com

Received 26 February 2020; Accepted 2 May 2020; Published 10 June 2020

Guest Editor: Wen-Tsao Pan

Most of China's resource-based cities are recently threatened by the problems of environmental pollution, resource depletion, and even economic recession. There is an urgent need for these cities to decouple economic growth from environmental degradation through improving resource utilization efficiency and eco-efficiency. This paper quantitatively evaluates the decoupling trends between GDP and environmental damage in major Chinese resource-based cities using big data. To explore the decoupling trends in the development of 115 resource-based cities in China, we develops a model to evaluate the eco-efficiency between economic growth and environmental pollution. Industrial pollutants are used as indicators of environmental degradation and GDP as an index for economic growth. This study finds diverse decoupling levels among the cities and that nearly one-third are developing unsustainably. The cities have encountered serious environmental problems due to past economic policies which have encouraged extensive GDP growth for the past 4 decades. This study demonstrates that the urban development and environmental protection of Chinese resource-based cities are not on a good path. It will be more difficult for the cities with quicker GDP growth to achieve a satisfactory decoupling trend. Therefore, there is an urgent need for industries to undertake resource-conserving and environmental protection measures, and, particularly, endorse technology innovation and the green economy. In order to achieve an overall balanced decoupling in China's resource-based cities, it is essential for local governments to encourage environmentally friendly behaviors and enhance eco-efficiency when developing the country's economy. Finally, the findings illustrate significant unbalanced decoupling levels across Chinese resource-based cities.

1. Introduction

With the development of big data, the study of resource-based cities is more effective, and the research conclusions are more accurate using big data [1]. Resource-based cities have played an important role during China's ascendance to the position of the world's second largest economy by providing resources, energy, and their own continuous economic growth. However, Chinese resource-based cities face challenges due to the nonrenewable nature of natural resources, such as the minerals, energy, and forests, on which they depend [2]. They are far from achieving the sustainability goals of balanced development towards a green economy, ensuring ecological protection and social advancement. More broadly, the world, including China, will not be able to meet future demands for minerals, ores, fossil fuels, and biomass if economic growth does not decouple from the rate of natural resource consumption. The Chinese government and the leaders of resource-based cities recognize that the natural environment and resources should not be sacrificed for achieving fast economic growth and are looking for possible solutions. This paper investigates the issues that are related to the sustainable development of these resource-based cities, which, as a special economic group, are playing a key role in China's economic development.

Previous studies indicate that China's resource-based cities, mainly established in the 1950s and 1960s, have

contributed significantly to the fast economic growth and industrial expansion of the country [3, 4]. Before the 1960s, the resource industries played a constructive and significant role in achieving economic growth, with natural resources, such as energy and mining, being the main engine of industrialization [4–6]. However, after six decades of resource-based industrial development and boom, these cities are now confronted with issues, such as shrinking mining sites, structurally imbalanced economies, slowing economic growth, high unemployment, and the unsustainable use and overexploitation of natural resources [5, 7]. Recently, other scholars have also drawn attention to the imbalanced development of resource-based cities and ecological problems in China and have examined the cities' shrinking resources, slowing economic growth, serious industrial pollution issues, and declining ecological environment [6, 8]. However, most studies are mainly focused on development in individual resource-based cities. For example, Liu and Li [9] studied the relationship between industrial waste gas emissions and economic and population growth as well as industrial structure in Xuzhou, China, suggesting the waste gas discharged from resource industries of the city accounted for more than 90% of total emissions.

In fact, Chinese resource-based cities are responsible for many environmental damage problems. They have an unevenly distributed ecological footprint. Furthermore, during the expansion of the resource industries, they have contributed significantly to environmental pollution [10, 11] and resource depletion [12]. Recently, Xu and Tian [13] used DEA model to analyze the eco-efficiencies of 27 coal-sourced cities, selecting waste gas and wastewater pollution as indicators, and they suggest that overall eco-efficiency has increased. However, the study also illustrated imbalanced eco-efficiencies with the eastern region being higher than the western and the central regions. In addition, the life cycle of resource-based cities has also been explored by Wang and Liu [14], who examined the coupling between economic and population growth and the ecological environment in resource-based cities in central China. It was found that, since 2005, the relationship between urbanization and the ecological environment of most resource-based cities in central China has remained unchanged. The resource-based cities with deteriorating relationships are mainly located in the provinces of Anhui, Jiangxi, and Hunan, and the better performing resource-based cities are generally located in Shanxi. Huang et al. [15] have studied the spatial convergence of eco-efficiency in selected Chinese resource-based cities, using the Meta-US-SBM model to analyze the relationship between the cities' spatial distance and eco-efficiency [16]. Their research suggests that the closer the distance is between the resource-based cities, the smaller the difference is in eco-efficiency. This paper uniquely analyzes the eco-efficiency of China's resource-based cities as a whole, selecting key indicators, which have caused China's environmental problems, including SO_2, industrial soot, and wastewater to systematically analyze the eco-efficiency of the resource-based cities from the perspectives of both resource types and life cycle. Since the previous study did not cover many resource-based cities and lacked a comparative analysis of the resource-based cities, this study analyzes the eco-efficiency of 115 resource-based cities from two angles, GDP growth and environment deterioration, and proposes development suggestions for resource-based cities in China.

China's substantial demand for resources due to continuous economic expansion intensifies environmental degradation. Resource exploitation is a process which seriously damages the natural environment, including its geological balance, surface ecology, and air. The atmospheric ecological damages include air pollution and climate change, as well as droughts and floods, due to surface and geological changes. China's National Resource-based Sustainable Urban Development Plan 2013–2020 put a strong emphasis on transforming Chinese resource-based cities to green economies. To avoid economic development and environmental health being perceived as contradictory demands, decoupling between economic growth and environmental degradation is a necessary requirement to achieve sustainability. While there are quite a few studies on resource-based cities in China, such a decoupling relation is yet to be explored and quantified.

This study evaluates the decoupling trends between economic growth and environmental degradation in Chinese resource-based cities. Based on data availability, 115 prefectural level cities were selected and their eco-efficiencies were analyzed. This study examines why and how these cities rely on their industrial resources base and the potential for developing new industries. Firstly, a review of the existing literature is presented followed by an examination of environmental issues and a sustainable development path for resource-based cities in China. The life cycle theory of resource-based cities is explained afterwards, allowing the classification of China's prefecture-level resource-based cities. An evaluation of the cities' decoupling trends is then conducted leading to policy implications and recommendations. A major innovation of this study is that it establishes a new evaluation model for calculating eco-efficiency by combining both the OECD decoupling index and Tapio's decoupling elasticity index. Another innovation is the measurement of eco-efficiency for China's resource-based cities.

2. Resource-Based Cities in China

2.1. A Historical Perspective. Most resource-based cities in China were established in the 1950s as a result of planned development which allowed for heavy reliance on a limited number of industries and imbalanced development within and between these settlements [17, 18]. Achieving economic growth was the main objective at the time, and this was further reinforced with the economic reform of the 1970s. As it was difficult to establish modern industries without access to technology, resource development was seen as the only effective choice to improve living standards in China. The emergence of a large number of resource-based cities followed extensive utilization of a development model around low-level technology. This created overdependence on resource industries, which leads to overexploitation of natural reserves. As a result, the resources in many such cities

became exhausted too early and too quickly without proper replacement industries. For example, small industrial and mining cities in coal and mineral resource-rich Shanxi Province attracted attention with their unsustainable development model, which produced alarming waste creation, resource damage, serious ecological destruction, and environmental problems [19].

While resource-based cities in Eastern and Southern China have since undergone transformation [20], those in the central and western regions have been trapped in an unsustainable development mode. Since China's open-door policies of the 1970s, special economic zones were established in Eastern and Southern China with better access to modern technology, better management, and greater foreign direct investment. Many new industries developed quickly and the once resource-based cities in those regions changed dramatically. The resource-based cities in China's central and western regions lacked such supporting policies. Instead, they provided substantial raw materials and primary resource products for the eastern and southern cities. These different roles created a large development gap between China's resource-based cities [21]. After more than 60 years of development, many resource-based cities, mainly in Central and Western China, face resource exhaustion, including Yichun, Liaoyuan, Jiaozuo, Pingxiang, and Daye. Over 20 resource-based cities are in decline with poor economic prospects and serious pollution issues.

Since 2007, the Chinese government began to focus its attention on the resource-based cities. Its official policy, "Suggestions about Promoting Sustainable Development in Resource-based Cities" encouraged local governments to establish long-term mechanisms for ensuring the sustainable development of resources industries and improved nature conservation and environmental protection. In 2013, the State Council articulated "Sustainable Development Strategies for Resource-based Cities, 2013–2020" in a revised document, which set clear sustainability goals for further economic development, improvement in people's quality of life, resource conservation, and ecological protection [22]. These strategies incorporated for the first time the goals of establishing a diversified industrial system while effectively utilizing natural resources and enhancing people's well-being, which corresponds to the three sustainability pillars of enhancing economic prosperity, strengthening environmental protection, and ensuring social advancement. Furthermore, in 2014, the Chinese government fully implemented an action plan for air and water pollution prevention and drafted strategies for eliminating soil contamination. The current 13th Five-Year Plan also covers actions and plans for resource preservation. At the 19th Communist Party Congress, China declared its commitment to pursue a model of sustainable development and to maintaining green GDP growth, through upgrading industrial structures, improving production, and ensuring better living standards and a healthy ecological environment. Sustainability in resource-based cities should be achieved through an integrated progress involving all three aspects of socioeconomic and environmental development and green GDP growth. This will allow for improved, protected, and optimized use of the natural environment as well as the improvement of people's lives in an ecologically civilized society.

A necessary condition to achieve sustainable development is the decoupling between GDP growth and natural resource consumption [23]. The intensity of natural resource exploitation combined with poor technology and management has caused severe environmental degradation and ecological damage in China, particularly in its resource-based cities. These problems are magnified by the fast urban development and industrialization in all resource-based cities. Previous research on the sustainability of resource-based cities in Northeastern China indicates a declining overall trend: while the sustainability of the economic subsystem increased, that of the environmental subsystem decreased and the social system remained stable [24]. There appear to remain significant decoupling challenges.

2.2. Life Cycle Theory. Unlike other human settlements, resource-based cities go through a distinctive life cycle [6, 8, 25]. Lucas [26] described their development as a four-stage life cycle covering construction, development, transformation, and maturity. Given the end-of-life possibility, two additional stages were added later by Bradbury and St-Martin [27], namely, decline and closure (see Figure 1). At the construction stage, the exploitation of resources is at a small scale and the level of economic growth is slow due to limited resources being available during industrial development, and therefore, the ecological footprint is relatively good. During the growth stage, labor and capital are heavily invested, and the economy grows quickly due to the fast development of resources industries; however, there is progressive ecological deterioration. After a longer period of growth, the development of resource-based cities would face issues characteristic of the whole resource market with less exploitative reserves available and fluctuating demand. Because of unstable profits, enterprises would feel more competition pressure, and consequently, some would start to consider transforming their operations, and therefore, the city would enter the transformation stage.

There are two different paths after the transformation stage. On the one hand, the successfully transformed cities enter into a mature stage, changing their single-industry structure into a comprehensive industrial system. As a result, they are on a new track towards a cleaner ecological environment with less environmental pollution due to better prevention and control measures and less reliance on natural resources. On the other hand, the cities which are not able to transform successfully enter a decline stage where their economy quickly deteriorates, their population decreases because labour and capital move to new locations, and environmental pollution decreases, allowing the natural environment to gradually recover. Eventually, resource-based cities will enter into a closing stage because of declining industries and society should return to its state before resource exploitation and development [6].

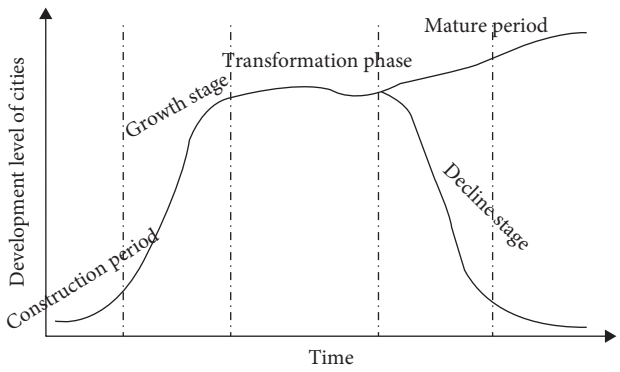


Figure 1: Life cycle of resource-based cities (source: based on Lucas [26] and Ding and Zhang [5]).

2.3. Classification and Sample Selection. In this study, "resources" refer to nonrenewable resources which are extracted and processed. When the natural reserves are exhausted, they become difficult to access or economically unviable, and the related resource industries would also decline. The definition of a resource-based city is one based on natural resources and resource industries. These cities have resource exploitation and related industries as economic pillars, which directly support a large labor force employed directly and indirectly for the resource industries.

From a development perspective, a city's prospects depend on the richness of natural resources and their exploitation. Resources may be found first and then a city developed, or the other way around—a city is developed and resources are discovered later [28]. In both cases, the discovery and utilization of resources play an important role in local development. Resources include not only oil, coal, ferrous and nonferrous metals, and mining reserves, but also forests, water, or other biological and environmental assets [23]. A city's resource utilization is important, including its eco-efficiency, as that is the link between environmental costs and environmental impacts for economic activities [29] and eco-industrial chains allowing a better resource utilization [30]. On the other hand, from a functional perspective, with resource commodities being the major local output, a city's function becomes the provision of these products to other places in the country.

According to the 2013–2020 National Resource-based Sustainable Urban Development Plan [18], there are 262 resource-based cities in China, 126 of which are at a prefecture level and the other 100 at a county level. The prefecture-level resource-based cities can be classified according to the dominant major industry as oil- and gas-based, metal-based, coal-based, and forest- and other resource-based (see Table 1). Furthermore, these cities can be categorized according to their life cycle stage, namely, developing, transforming, mature, and declining (see Table 1 and also Figure 1). This classification however does not provide any insights about the environmental performance of the resource-based cities and in particular, about their eco-efficiency. The analysis to follow is based on 115 of the 126 prefecture-level resource-based cities with 11 large prefectures (in italics in Table 1) excluded because of data unavailability.

Because the data cannot be obtained, we put 11 resource-based cities aside, which include Haixi Mongol and Tibetan Autonomous Prefecture, Altay Prefecture, Qiannan Buyei and Miao Autonomous Prefecture, Chuxiong Yi Autonomous Prefecture, Qianxinan Buyei and Miao Autonomous Prefecture, Ngawa Tibetan and Qiang Autonomous Prefecture, Bayingolin Mongol Autonomous Prefecture, Yanbian Korean Autonomous Prefecture, Puer, Liangshan Yi Autonomous Prefecture, and Da Hinggan Ling Prefecture; all of them are minority autonomous regions. Then, we have studied 115 resource-based cities in China.

3. Methodology

This study evaluates the eco-efficiency of the selected 115 resource-based cities by calculating the decoupling elasticity between GDP and pollutant discharge based on a fuzzy comprehensive analysis model of the eco-efficiency evaluation theory and Tapio's decoupling index. Eco-efficiency evaluation is one of the most important methods for assessing decoupling trends for sustainable development. All definitions include balancing the use of resources and environmental impact with human demands, expressed as the ratio between economic value added and added environmental impact. The Organization for Economic Co-operation and Development (OECD) defined eco-efficiency as "the efficiency with which ecological resources are used to meet human demands" [26]. According to the World Business Council for Sustainable Development, this is a management philosophy for environmental improvements leading to economic benefits. In other words, more human demands are met with less stress on the environment with environmental impacts and resource use kept within Earth's bearable level.

Decoupling describes the relationship between two upward curves—of economic growth and of the utilization of resources or environmental impacts. In environmental studies, decoupling assessment is used to measure the link between environmental pressure and GDP growth [33, 34]. A decoupling index represents the changing rate of an environmental indicator over the changing rate of an economic indicator. If the former is not growing with economic growth, the trend is considered decoupling [35, 36].

3.1. Two Decoupling Indexes. In order to assess the decoupling trends of China's resource-based cities, the indicator chosen is the separation degree between the environmental indicator curve and the economic indicator curve [32]. Two decoupling indexes are used, namely, the OECD decoupling index [37] and the decoupling elasticity index [38].

According to the OECD [39], the decoupling index is the separation degree between the environmental pressure (EP) and the development momentum (DM):

$$\text{decoupling index} = \frac{(\text{EP}/\text{DM})_{\text{beginning}}}{(\text{EP}/\text{DM})_{\text{end}}}. \quad (1)$$

Table 1: Classification of selected 126 resource-based cities at prefecture level in China.

	Oil and gas [13]	Metal [26]	Coal [47]	Forest and others [27]
Developing [22]	Yan'an Qingyang Longnan *Haixi Mongol and Tibetan Autonomous Prefecture* [4]	Songyuan Hezhou Liupanshui Wuwei *Altay Prefecture* *Qiannan Buyei and Miao Autonomous Prefecture* [6]	Shuozhou Hulun Buir Erdos Zhaotong Xianyang Yulin *Chuxiong Yi Autonomous Prefecture* *Qianxinan Buyei and Miao Autonomous Prefecture* [9]	Nanchong Bijie [2]
Transforming [17]	Panjin Nanyang [2]	Huludao Tangshan Baotou Anshan Tonghua Ma'anshan Zibo Luoyang Zhangye *Ngawa Tibetan and Qiang Autonomous Prefecture* [11]	Xuzhou [1]	Suqian Linyi Lijiang [3]
Mature [66]	Daqing Dongying Baoshan Karamay *Bayingolin Mongol Autonomous Prefecture* [5]	Benxi Longyan Ganzhou Laiwu Ezhou Hengyang Handan Chenzhou Shaoyang Loudi Hechi Panzhihua Dazhou Ya'an Baoji Jinchang [17]	Zhangjiakou Xingtai Datong Yangquan Changzhi Jincheng Linfen Xinzhou Yuncheng Lvliang Chifeng Heihe Jixi Huainan Hebie Pingdingshan Guangyuan Guang'an Anshun Qujing Weinan Pingliang *Yanbian Korean Autonomous Prefecture* [24]	Chengde Jinzhong Jilin Suzhou Mudanjiang Huzhou Bozhou Chuzhou Chizhou Xuancheng Nanping Sanming Yichun (Jiangxi Province) Jining Taian Sanmenxia Yunfu Baise Zigong Lincang *Puer* *Liangshan Yi Autonomous Prefecture* [18]
Declining [25]	Puyang [1]	Tongling Xinyu Huangshi Shaoguan Baiyin [5]	Wuhai Fuxin Fushun Liaoyuan Yichun (Heilongjiang Province) Hegang Shuangyashan Qitaihe Huaibei Zaozhuang Pingxiang Jiaozuo Luzhou Tongchuan Shizuishan [16]	Baishan Jingdezhen *Da Hinggan Ling Prefecture* [3]

Note: the number in parentheses is the quantity of the city, and the cities written in italics cannot be studied because of data lacking.

Environmental pressure can be represented by energy consumption, greenhouse gas emissions, and pollution indicators while development momentum is represented with GDP. The decoupling is breaking the link between "environmental bads" and "economic goods" [39], hence between growth in environmental pressure and the generation of economic benefits and services [37, 40]. Decoupling is considered to occur when the GDP growth rate is higher than environmental damage over a certain period [2].

Alternatively, according to Tapio [34], the decoupling elasticity index is the changing rate of an environmental indicator over the changing rate of economic development calculated according to the following equation:

$$\text{decoupling elasticity index (DEI)} = \frac{\Delta \text{EI}/\text{EI}}{\Delta \text{GDP}/\text{GDP}}, \quad (2)$$

where ΔEI/EI is an environmental indicator's change rate and ΔGDP/GDP is the change rate of GDP. Using the decoupling elasticity index, Tapio divides the relationship between environmental pollution and GDP growth into 8 types: strong positive decoupling, expansive weak decoupling, declining strong coupling, expansive coupling, declining coupling, expansive strong coupling, declining weak decoupling, and strong negative decoupling [40, 41] (see Table 2).

In Tapio's decoupling elasticity model, data for the start and end year of the time period are used to calculate the index, which does not show annual trends. Due to the wide yearly fluctuation of pollution in the Chinese cities, Tapio's model does not properly represent the real decoupling relationship between economic growth and environmental degradation [39]. According to the observation of the GDP, INS, INW, and SO_2 data of resource-based cities, it is found that the annual data fluctuate greatly. For example, consider two randomly selected cities—Karamay and Tangshan. Karamay is located in the Xinjiang Uygur Autonomous Region of China in the northwest areas and is a mature oil and gas resource-based city. Tangshan is located in Hebei Province, which is a metal resource-based city in the stage of transformation. They are both typical resource-based cities in China.

Karamay would show strong positive decoupling between GDP and industrial soot emissions only if the start and end of the 2004–2017 period are used (see Figure 2). However, due to considerable volatility in the data for industrial soot emissions, the real decoupling trend with GDP over the entire sample period is difficult to estimate. Similarly, for Tangshan, if only the start and end year of the 2004–2017 period are used, the city would also show a strong positive decoupling between GDP and industrial soot emissions (see Figure 3). In fact, a strong positive decoupling was observed only from 2003 to 2010. Furthermore, Tapio's eight decoupling trends (shown in Table 2) are not suitable for Chinese resource-based cities which experience fast economic growth.

3.2. Eco-Efficiency of China's Resource-Based Cities. This study adopts both the OECD decoupling index and Tapio's decoupling elasticity index, but also includes the middle year variations to establish a new evaluation model for calculating the eco-efficiency of the resource-based cities (see equation (3)):

$$\text{DEI}_{\text{EP,ED}} = \frac{\Delta \text{EP}/\text{EP}}{\Delta \text{ED}/\text{ED}} = \frac{(\text{EP}_t - \text{EP}_{t-1})/\text{EP}_{t-1}}{(\text{ED}_t - \text{ED}_{t-1})/\text{ED}_{t-1}}, \quad (t = 1, 2, \ldots, n). \quad (3)$$

$\text{DEI}_{\text{EP,ED}}$ is the decoupling elasticity between environmental pressure and economic growth; EP_t is the environmental pressure in year t of the study period; EP_{t-1} is the environmental pressure in year $t-1$ of the study period; ED_t is the economic growth in year t of the study period; ED_{t-1} is the economic growth in year $t-1$ of the study period.

Environmental pressure (EP) is represented by industrial wastewater (INW) discharge, industrial SO_2, and industrial soot emissions (INS) while economic development is represented by GDP growth.

In recent years, China is suffering from serious pollution, including air pollution, water pollution, and soil pollution. The pollution sources mainly come from domestic waste and industrial pollutants. While the high pollution industries of resource exploitation and resource and primary processing are the main industries of China's resource-based cities, industrial pollution is the main source of pollution in resource-based cities. In these pollutions, industrial soot, industrial wastewater, and SO_2 are the main pollutants that cause air pollution, water pollution, and soil pollution, so the 3 indicators of INS, INW, and SO_2 have been selected to represent environmental pressure.

Table 2 shows Tapio's decoupling-elastic model, which is suitable for evaluating the decoupling of economy and environment in mature economies. Because the value range of each stage is narrow in 8 decoupling trends of Tapio's model, meanwhile the trends of economic development and environmental degradation are relatively flat in mature economies, the narrower value interval can reflect the difference among trends [6]. While resource-based cities GDP increase rapidly with environmental deteriorating severely, then the narrower numerical area is not enough to show the decoupling trends between economy and environment [7]. Tables 3 is obtained based on Table 2, Figure 2 and 3.

Taking into consideration economic growth, environmental degradation, and decoupling indexes, the decoupling trends are categorized into five levels, namely, excellent, very good, satisfactory, poor, and unsatisfactory (see Table 3 and Figure 4). The decoupling trends are then evaluated using the fuzzy comprehensive evaluation method.

4. Empirical Results

The evaluation method is presented first using the illustrative example of Karamay. This is followed by the full sample results. There are 38 resource-based cities with data available for 2003–2014, 76 with data for 2004–2017, and one (Bijie) with data only for 2011 to 2017. This should not affect the evaluation results as yearly evaluations are used.

Table 2: Tapio's evaluation of eco-efficiency.

Decoupling trend	ΔEI/EI	ΔGDP/GDP	DEI	Trends
Strong positive decoupling	<0	>0	<0	GDP grows while pollution decreases
Expansive weak decoupling	>0	>0	0–0.8	GDP grows while pollution increases slowly
Declining strong decoupling	<0	<0	>1.2	GDP declines slowly while pollution decreases quickly
Expansive coupling	>0	>0	0.8–1.2	GDP grows while pollution increases
Declining coupling	<0	<0	0.8–1.2	GDP declines while pollution decreases
Expansive negative decoupling	>0	>0	>1.2	GDP grows slowly while pollution increases dramatically
Declining negative decoupling	<0	<0	0–−0.8	GDP decreases quickly while pollution decreases slowly
Strong negative decoupling	>0	<0	<0	GDP declines while pollution increases

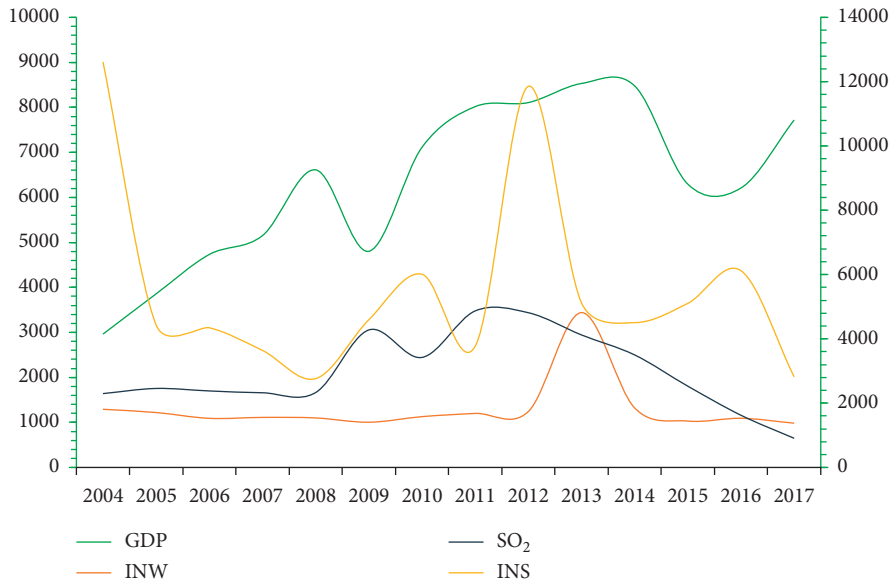


Figure 2: Pollution and GDP of Karamay.

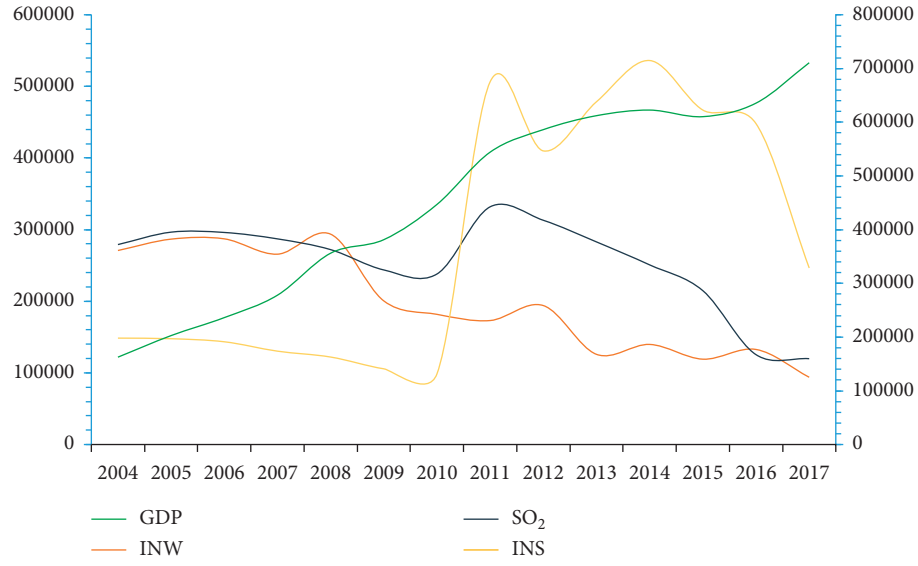


Figure 3: Pollution and GDP of Tangshan.

Table 3: Evaluation of eco-efficiency.

Decoupling trends	ΔEP	ΔED	DEI	Level	Trends
Strong positive decoupling	<0	>0	<−0.5	A: excellent	Economic growth increases while pollution decreases at a rate not less than half the rate of economic growth
Weak positive decoupling	<0	>0	>−0.5	B: very good	Economic growth increases while pollution decreases at a rate less than half the rate of economic growth
Expansive decoupling	>0	>0	<1	C: satisfactory	Economic growth increases while pollution increases at a rate not higher than the rate of economic growth
Declining decoupling	<0	<0	>1.2		Economic growth decreases while pollution decreases at a rate more than 1.2 times the rate of economic growth
Expansive negative decoupling	>0	>0	>1	D: poor	Economic growth increases while pollution increases at a rate higher than the rate of economic growth
Declining negative decoupling	<0	<0	<1.2		Economic growth decreases while pollution decreases at a rate less than 1.2 times the rate of economic growth
Strong negative decoupling	>0	<0	<0	E: unsatisfactory	Economic growth decreases while pollution increases

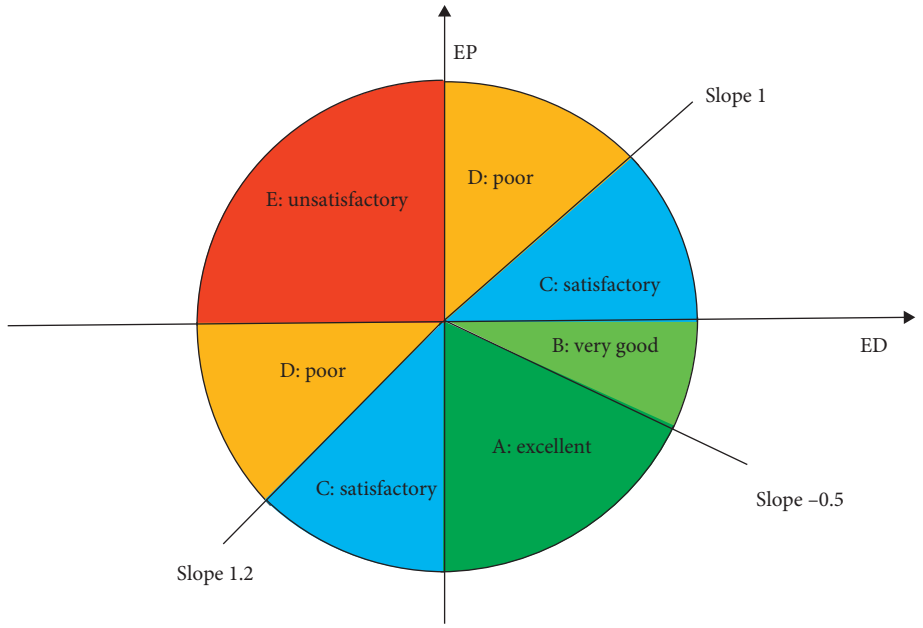


Figure 4: Evaluation areas of resource-based cities' eco-efficiencies.

4.1. Illustrative Calculation for Karamay. The following four steps are used for the evaluation model:

Step1: data processing–data for GDP, INW, SO_2, and INS from 2004 to 2017 (14 years) are used for Karamay.

Step 2: evaluating yearly indexes according to equation (3), divided into three pollutants:

$$\frac{(\mathrm{INW}_t - \mathrm{INW}_{t-1})/\mathrm{INW}_{t-1}}{(\mathrm{GDP}_t - \mathrm{GDP}_{t-1})/\mathrm{GDP}_{t-1}},$$

$$\frac{(\mathrm{SO}_{2t} - \mathrm{SO}_{2t-1})/\mathrm{SO}_{2t-1}}{(\mathrm{GDP}_t - \mathrm{GDP}_{t-1})/\mathrm{GDP}_{t-1}}, \quad (4)$$

$$\frac{(\mathrm{INS}_t - \mathrm{INS}_{t-1})/\mathrm{INS}_{t-1}}{(\mathrm{GDP}_t - \mathrm{GDP}_{t-1})/\mathrm{GDP}_{t-1}}.$$

Table 4 presents the calculation results for Karamay.

Step 3: evaluating according to the maximum subordinate degree principle, ΔEP, ΔED, and DEI are calculated and the eco-efficiency level is assigned by a computer program. The evaluation level is assigned based on the highest frequency level; if there are several similar frequency levels, the lowest eco-efficiency is assigned. Table 5 shows the eco-efficiency levels by pollutants for Karamay. As there are 9 records of "A," 8 of "B," 10 of "C," 7 of "D," and 5 of "E," according to the maximum subordinate degree principle, the comprehensive evaluation level for Karamay is "C."

4.2. Full Sample Results. Figure 5 shows the combined evaluation outcomes for all 115 cities by pollutants. The DEIs for industrial wastewater and industrial SO_2 are moderate, while that for industrial soot is even better. Table 6 presents

Table 4: Data of Karamay.

Year	ΔGDP	ΔINW	ΔSO_2	ΔINS	ΔINW/ΔGDP	$\Delta SO_2/\Delta GDP$	ΔINS/ΔGDP
2005	0.3022	−0.0558	0.0696	−0.6489	−0.1847	0.2302	−2.1470
2006	0.2269	−0.1060	−0.0315	−0.0192	−0.4670	−0.1390	−0.0846
2007	0.0885	0.0194	−0.0240	−0.1621	0.2187	−0.2714	−1.8322
2008	0.2836	−0.0106	0.0034	−0.2412	−0.0375	0.0119	−0.8504
2009	−0.2736	−0.0870	0.8359	0.6629	0.3180	−3.0551	−2.4228
2010	0.4811	0.1238	−0.1999	0.3085	0.2572	−0.4156	0.6413
2011	0.1270	0.0633	0.4227	−0.3730	0.4984	3.3286	−2.9377
2012	0.0113	0.0351	−0.0116	2.1462	3.1214	−1.0295	190.7563
2013	0.0523	1.7694	−0.1436	−0.5685	33.8288	−2.7459	−10.8681
2014	−0.0064	−0.6161	−0.1510	−0.1194	96.6737	23.7013	18.7408
2015	−0.2575	−0.2201	−0.2778	0.13296	0.85496	1.07889	−0.5164
2016	−0.0134	0.06033	−0.3581	0.19886	−4.5043	26.7310	−14.847
2017	0.2432	−0.1001	−0.4418	−0.5340	−0.4115	−1.8170	−2.2206

Data source: calculated from the originally collected data from the Development Research Centre Net (DRCNET, 2017), http://www.drcnet.com.cn/www/integrated/#.

Table 5: Evaluation of Karamay.

Year	INW	SO_2	INS
2005	B	C	A
2006	B	B	B
2007	C	B	A
2008	B	C	A
2009	D	E	E
2010	C	B	C
2011	C	D	A
2012	D	A	D
2013	D	A	A
2014	C	C	C
2015	D	D	E
2016	E	C	E
2017	B	A	A

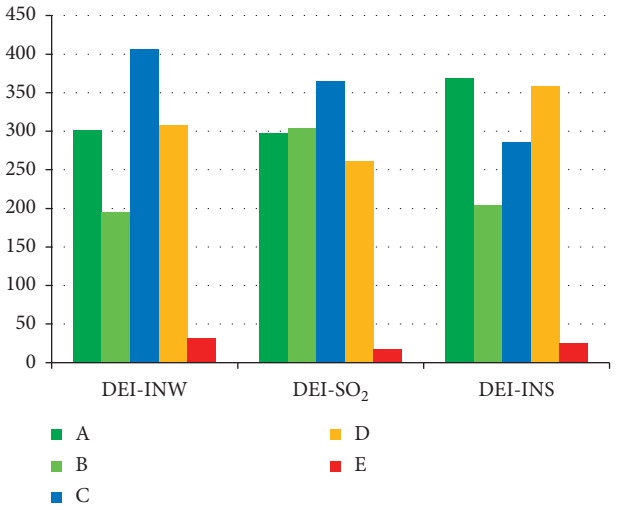


Figure 5: Evaluation of DEI of INW, SO_2, and INS.

the eco-efficiency evaluation results for each of the 115 resource-based cites while the map on Figure 6 shows their geographical location.

The eco-efficiency assessment shows that overall, there are 30 level A, 10 level B, 39 level C, 34 level D, and 2 level E cities (see Figure 7). There were several provinces with relatively low eco-efficiencies. Heilongjiang and Jilin provinces in Northeast China had no level A or B cities but two level E cities and the others at C or below. Hebei, Henan, Sichuan, Hunan, Hubei, and the Inner Mongolia Autonomous Region also had low eco-efficiencies with six cities at level D and one at level C. Jiangxi, Guizhou, and Guangxi were also not performing well.

4.3. Results by Commodity Type. Figure 8 and Table 7 show the distribution of the evaluation results according to resource types. Level A cities account for the highest percentage in the coal-based cities which is equal to that of level C cities. The coal cities also have overall the highest share of excellent and good eco-efficiencies. Level C cities also represent the highest percentage in oil and gas and forest and other categories while level D cities are the highest in metal-based resource-based cities. The two level E cities belong, respectively, to the coal and forest and other group. Overall, there is no significant difference among different commodity types. This suggests that the challenges for resource-based cities are not affected by the different price performance of different commodities.

Among the 115 resource-based cities, environmental pollution is most serious among oil and gas resource-based cities, with 90% of the cities being at levels C, D, and E. The forest and other resource-based cities are performing better in pollution control with nearly 60% of them being at levels C, D, and E. The coal and metal-based cities account for less cities at levels C, D, and E, 57% and 56%, respectively.

4.4. Results by Life Cycle Type. The results presented in Figure 9 and Table 8 show the distribution of the eco-efficiencies according to life cycle stages of the resource-based cities. Level A cities represent the highest percentage in the transforming cities which indicates that there are efforts to decrease the environmental impacts of economic activities. There are no level E cities in the two categories of developing and transforming cities, but this unsatisfactory eco-

Table 6: Evaluated eco-efficiencies of the 115 resource-based Chinese cities.

No.	City	Level	No.	City	Level	No.	City	Level
1	Anshan	D	40	Hulun Buir	C	79	Sanming	C
2	Anshun	D	41	Huzhou	C	80	Shaoguan	A
3	Baise	D	42	Jiaozuo	A	81	Shaoyang	A
4	Baishan	A	43	Jilin	E	82	Shizuishan	B
5	Baiyin	A	44	Jinchang	D	83	Shuangyashan	C
6	Baoji	B	45	Jincheng	B	84	Shuozhou	C
7	Baoshan	C	46	Jingdezhen	C	85	Songyuan	C
8	Baotou	A	47	Jining	D	86	Suqian	C
9	Benxi	A	48	Jinzhong	D	87	Suzhou	B
10	Bijie	D	49	Jixi	C	88	Taian	C
11	Bozhou	D	50	Karamay	C	89	Tangshan	A
12	Changzhi	C	51	Laiwu	D	90	Tongchuan	B
13	Chengde	D	52	Liaoyuan	C	91	Tonghua	D
14	Chenzhou	C	53	Lijiang	C	92	Tongling	B
15	Chifeng	B	54	Lincang	C	93	Weinan	A
16	Chizhou	D	55	Linfen	A	94	Wuhai	C
17	Chuzhou	C	56	Linyi	D	95	Wuwei	D
18	Daqing	C	57	Liupanshui	B	96	Xianyang	A
19	Datong	A	58	Longnan	D	97	Xingtai	A
20	Dazhou	A	59	Longyan	D	98	Xinyu	D
21	Dongying	D	60	Loudi	C	99	Xinzhou	D
22	Erdos	A	61	Luoyang	A	100	Xuancheng	A
23	Ezhou	B	62	Luzhou	A	101	Xuzhou	A
24	Fushun	A	63	Lvliang	C	102	Ya'an	D
25	Fuxin	D	64	Ma'anshan	C	103	Yan'an	C
26	Ganzhou	D	65	Mudanjiang	D	104	Yangquan	A
27	Guang'an	D	66	Nanchong	A	105	Yichun (Heilongjiang Province)	C
28	Guangyuan	A	67	Nanping	C	106	Yichun (Jiangxi Province)	C
29	Handan	C	68	Nanyang	A	107	Yulin	C
30	Hebei	C	69	Panjin	C	108	Yuncheng	A
31	Hechi	D	70	Panzhihua	C	109	Yunfu	C
32	Hegang	C	71	Pingdingshan	C	110	Zaozhuang	D
33	Heihe	D	72	Pingliang	D	111	Zhangjiakou	A
34	Hengyang	B	73	Pingxiang	D	112	Zhangye	D
35	Hezhou	C	74	Puyang	D	113	Zhaotong	D
36	Huaibei	C	75	Qingyang	D	114	Zibo	D
37	Huainan	C	76	Qitaihe	E	115	Zigong	A
38	Huangshi	A	77	Qujing	A			
39	Huludao	A	78	Sanmenxia	C			

Data source: calculated from the originally collected data from the Development Research Centre Net (DRCNET, 2017), http://www.drcnet.com.cn/www/integrated/#.

efficiency is demonstrated in the group of mature and even stronger among declining cities. Level C cities dominate developing and mature cities. The declining group is quite interesting as overall it has the highest share of excellent (A) and very good (B) eco-efficiency but also the highest number of cities with only satisfactory eco-efficiency. This could be interpreted as indicating that unless concerted efforts are made early in the life cycle of resource-based cities to conserve the natural resources on which they depend, eco-efficiency comes a stage when the environmental damage, including resource depletion, becomes irreversible. The improved eco-efficiency is not capable of stopping the overall decline.

Table 8 shows that, among the resource-based cities classified by life cycle, the cities in the developing stage have encountered the most serious ecological problems with 73% of such cities falling in categories C, D, and E. However, the transforming, mature, and declining cities are performing better with 60%, 68%, and 25% of cities falling in C, D, and E, respectively. Although the declining cities show the positive decoupling trend, there is a need for local governments to formulate proper strategies of transformation to revitalize their cities' economy.

5. Discussion

There are many factors influencing the evaluation results. The data analysis showed that from the perspective of city types, in coal-based cities, the decoupling between INW discharge and GDP was large while the decoupling between GDP and SO_2 or INS was smaller. The decoupling of environmental discharge and economic growth may be caused by the fast growth of coal production in the 2000s and increases in coal prices. The relatively better performance of

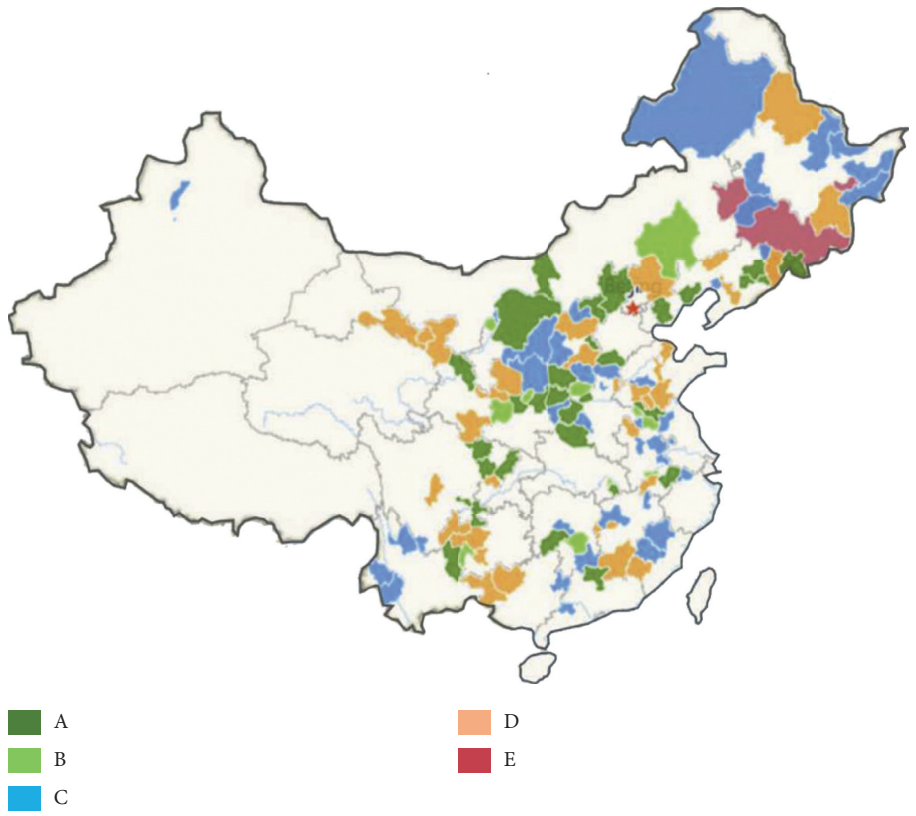


Figure 6: Location of the evaluated Chinese resource-based cities.

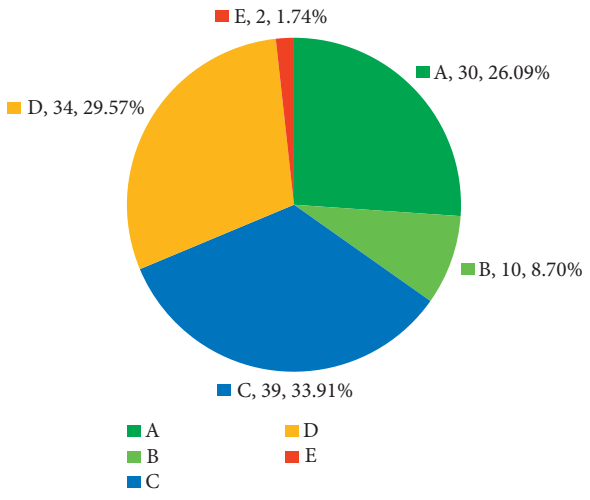


Figure 7: Eco-efficiency distribution of the 115 evaluated Chinese resource-based cities.

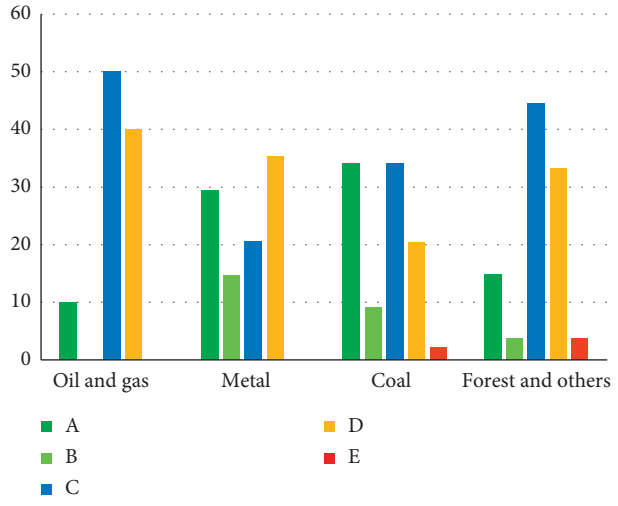


Figure 8: Eco-efficiency evaluation results by resource type (%).

wastewater compared to SO_2 and INS is likely due to improved regulations about INW discharge during the sample period. On the other hand, SO_2- and INS-related regulations were enforced before the sample period and there were no significant changes in both technology and policy environment. Therefore, the trend is stable.

In oil resource-based cities, GDP was increasing relatively slowly and the increase in pollution discharge was also small. A possible explanation is that oil-based cities faced the threat of resource depletion accompanied by decreasing resource exploitation. Most oil-producing cities, except those in Xinjiang, are no longer able to sustain their production.

Prior to 2010, GDP was growing quickly in metal-based cites while the pollution growth rate was lower than that of GDP. After 2010, the pollution growth rate decreased and was accompanied by decreased GDP growth rate. In addition, before 2010, the industries of the steel-based cites were booming, but after 2010 with the slowing of national

Table 7: Eco-efficiency evaluation results by resource type.

Number	A	B	C	D	E	Total
Oil- and gas-based cities	1	0	5	4	0	10
Metal-based cities	10	5	7	12	0	34
Coal-based cities	15	4	15	9	1	44
Forest and other resource-based cities	4	1	12	9	1	27
Total	30	10	39	34	2	115

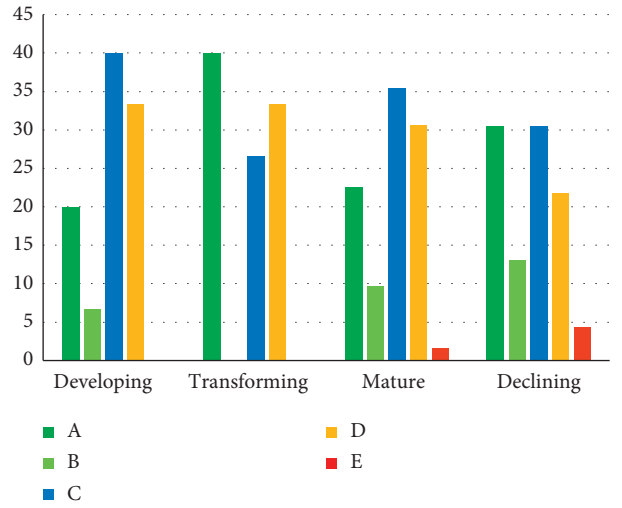


Figure 9: Eco-efficiency evaluation results by life cycle stage type (%).

Table 8: Eco-efficiency evaluation results by life cycle stage type.

Number	A	B	C	D	E	Total
Developing cities	3	1	6	5	0	15
Transforming cities	6	0	4	5	0	15
Mature cities	14	6	22	19	1	62
Declining cities	7	3	7	5	1	23
Total	30	10	39	34	2	115

economic growth, the industries which required large amounts of steel such as construction have been developing slowly. As a consequence, the GDP and pollution growth rates started to slow down.

In the forest and other resource-based cites, the GDP growth rate was relatively slow and the decoupling rate between GDP and pollution emission rates was also small. This was due to a lack of significant changes in either technology or regulation.

From a regional perspective, in the resource-based cities of Northeast China (Heilongjiang, Liaoning, and Jilin) and in Southwest China (Yunnan, Guizhou, Sichuan, and Guangxi) as well as in Shanxi Province, the GDP growth rate was relatively slower and their pollution emissions were decreased quickly. Moreover, in North China (Hebei and Mongolia) and Central China (Henan, Wuhan, Hunan, and Anhui), the GDP of the resource-based cities was increasing relatively fast; however, the pollution emission rate was lower than the GDP growth rate. In East China (Jiangsu, Zhejiang, and Jiangxi) and South China (Fujian and Guangdong), the resource-based cities experienced the fastest GDP growth rates and their pollution emissions were generally decreasing. Overall, East China presents the most desirable case of decoupling, which is likely due to technology progress allowing the economy to grow with less environmental damage. The decoupling in Northeast and Southwest China is also acceptable, even at the cost of lower economic growth. The North China pattern however needs to be monitored as environmental deterioration may continue with economic growth.

6. Conclusion and Policy Implications

This paper quantitatively assesses the decoupling between economic growth and environmental degradation in 115 Chinese resource-based prefectural level cities using an innovative method to calculate eco-efficiencies. The results suggest that, among the 115 resourced-based cities, 30 level A cities are performing well in decoupling between economic growth and environmental degradation, and the performance of 10 cities of level B has been slightly improved. However, 39 cities of level C have just maintained the decoupling and 36 cities of level D have been performing in deteriorating rate, particularly with last 2 cities of levelE being deteriorating seriously. The results indicate that about one-third of these cities fall in level D or E, which indicates unsustainable development. Moreover, there are very different decoupling levels among the cities located in different regions of the country. The extensive environmental degradation in China's resource-based cities has been accompanied by fast economic expansion. Yet, it is encouraging to see that the Chinese government is currently taking strict measures to protect resources and the natural environment. The imbalance in resource-based cities' decoupling trends should be able to allow the poorly performing cities to search for more development opportunities.

In order to protect the natural environment, there is an urgent need for the resource-based cities to decouple GDP growth from environmental degradation. Achieving sustainability in the resource-based cities should be taken from a concept to proper engagement with weakening the link between GDP growth and environmental degradation. This is essential for their future [38]. There is a need to summarise the key experiences from those cities, who have achieved levels A and B, and lessons from those who failed into levels D and E. Such a share of knowledge might help resource-based in China and the rest of the world to achieve sustainable development.

In addition, the urban transition and environmental performance of Chinese resource-based cities are not equally smooth [1, 37]. For the cities with fast GDP growth, it is harder to achieve a high decoupling level. As suggested by Mao [41] and Ru et al. [12], the urban transformation should be achieved through enhancing technology innovation, improving resource utilization, and upgrading the industrial structure, particularly in the cities in North and West China and this will help them increase their competitiveness. Local leaders should support green investment and encourage

industries to increase their eco-efficiency and avoid resource waste and utilization [42]. Most importantly, local governments need to take effective actions to achieve a balanced future development for resource-based cities, including improving local resource utilization and promoting the circular and green economies as the key solutions [12].

The extend of the decoupling between economic growth and environmental degradation changes considerably with different macroeconomic circumstances and policy adjustments as well as depending on local government efforts, the types of resources exploited, resource management abilities, and anticorruption and environmental regulation policies. This study believes that improving eco-efficiency and choosing the economically desired level of environmental and resource management as suggested by Schaltegger and Synnestvedt [31] can effectively enhance sustainability. This is particularly important for achieving a more effective decoupling and a more balanced development of the resource-based cities in China.

Data Availability

The evaluation data used to support the findings of this study can be accessed publicly at the website https://github.com/zhidongli98762020123/China-s-resource-based-cities/.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This research was supported by a grant from the University of Humanities and Social Sciences of Anhui Province (grant number: SK2019A0616).

References

[1] Y. Zhang, K. Wang, Q. He et al., "Covering-based web service quality prediction via neighbourhood-aware matrix factorization," *IEEE Transactions on Services Computing*, 2019, In press.

[2] J. Tan, P. Zhang, K. Lo, J. Li, and S. Liu, "The urban transition performance of resource-based cities in Northeast China," *Sustainability*, vol. 8, no. 10, p. 1022, 2016.

[3] H. Du, "Adjusting strategies and establishing institutions for ensuring the sustainable development of the resource-based cities," *Chinese Journal of Population Resources and Environment*, vol. 23, no. 2, pp. 88–93, 2013, in Chinese.

[4] Z. Li, D. Marinova, X. Guo et al., "Evaluating pillar industry's transformation capability: a case study of two Chinese steel-based cities," *PLoS One*, vol. 10, no. 9, Article ID e0139576, 2015.

[5] X. Ding and Y. Zhang, "Economic transition of resourced-based cities based on the life cycle theory," *World Regional Studies*, vol. 17, no. 3, pp. 70–76, 2008.

[6] S. Y. He, J. Lee, T. Zhou, and D. Wu, "Shrinking cities and resource-based economy: the economic restructuring in China's mining cities," *Cities*, vol. 60, pp. 75–83, 2017.

[7] W. Sun and G. Dong, "China's resource-based cities' efficiencies and changes, based on a DEA model," *Geographical Research*, vol. 29, no. 12, pp. 2155–2165, 2010, in Chinese.

[8] L. Mou, "Spatial economic analysis on the life cycle of resource-based cities," *Economic Forum*, vol. 476, no. 4, pp. 27–31, 2017.

[9] J. Liu and H. J. Li, "Factors of industrial waste gas emission in resource-based city based on LMDI mode—a case study on Xuzhou," *Resources & Industries*, vol. 2, pp. 1–6, 2016.

[10] X. Huang and T. Wang, "Eco-efficiency assessment for resource-based cities," *Science Research Management*, vol. 36, no. 7, pp. 71–78, 2015.

[11] H. Li, *Transformation Index of China's Resource-Based Cities: Transformation Evaluation of Prefectural-Level Cities*, Commercial Press, Beijing, China, 2016.

[12] X. Ru, S. Chen, and H. Dong, "An empirical study on relationship between economic growth and carbon emissions based on decoupling theory," *Journal of Sustainable Development*, vol. 5, no. 8, pp. 43–51, 2012.

[13] J. F. Xu and S. Y. Tian, "Studying the green development path of coal resource-based cities by analyzing the eco-efficiencies," Ph.D thesis, Anhui University, Vol. 6, Anhui, China, 2018.

[14] G. X. Wang and T. Liu, "Coupling relationship change between urbanization and eco-environment of resource-based cities in Central China," *China Population Resources and Environment*, vol. 27, no. 7, pp. 80–88, 2017.

[15] Y. Huang, Y. Fang, G. F. Gu, and J. S. Liu, "The evolution and differentiation of economic convergence of resource-based cities in Northeast China," *Chinese Geographical Science*, vol. 28, no. 3, pp. 495–504, 2018.

[16] Y. Zhang, G. Cui, S. Deng et al., "Efficient query of quality correlation for service composition," *IEEE Transactions on Services Computing*, p. 1, 2018, In press.

[17] R. Wang, "A multi agent-based approach for supply chain network," *Advanced Materials Research*, vol. 136, pp. 82–85, 2010.

[18] L. Li, Y. Lei, D. Pan, and C. Si, "Research on sustainable development of resource-based cities based on the DEA approach: a case study of Jiaozuo, China," *Mathematical Problems in Engineering*, vol. 2016, Article ID 5024837, 10 pages, 2016.

[19] G. Hong, Z. Kai, and Z. Hanwen, "Research on sustainable development of resource-based small industrial and mining cities—a case study of Yangquanqu Town, Xiaoyi, Shanxi province, China," *Procedia Engineering*, vol. 21, pp. 633–640, 2011.

[20] X. Zhu, "Transformation of resource-based cities in China," in *Proceedings of the Resilience and Sustainability 36th Annual Conference of the International Association for Impact Assessment*, IAIA, Nagoya Congress Center, Aichi-Nagoya, Japan, May 2016.

[21] Y. Wang, *Competitiveness Strategies*, China Financial and Economic Publishing House, Beijing, China, 2010.

[22] Y. Huang, Y. Fang, G. Gu, and J. Liu, "The evolution and differentiation of economic convergence of resource-based cities in Northeast China," *Chinese Geographical Science*, vol. 28, no. 3, pp. 495–504, 2018.

[23] Y. Chen, N. Zhang, Y. Zhang et al., "Energy efficient dynamic offloading in mobile edge computing for Internet of Things," *IEEE Transactions on Cloud Computing*, p. 1. 2018, In press.

[24] C. Lu, B. Xue, C. Lu et al., "Sustainability investigation of resource-based cities in Northeastern China," *Sustainability*, vol. 8, no. 10, p. 1058, 2016.

[25] R. A. Lucas, *Minetown Milltown Railtown: Life in Canadian Communities of Single Industry*, University of Toronto Press, Toronto, Canada, 1971.

[26] R. A. Lucas, *Minetown Milltown Railtown: Life in Canadian Communities of Single Industry*, University of Toronto Press, Toronto, Canada, 1971.

[27] J. H. Bradbury and I. St-Martin, "Winding down in a Quebec mining town: a case study of schefferville," *The Canadian Geographer*, vol. 27, no. 2, pp. 128–144, 1983.

[28] P. Wang, W. Liu, and D. Lu, "Study of the efficiency of science and technology resource distribution in the metropolitan areas of China: case studies of Beijing, Tianjing and Hebei province, Yangtze river delta and the pearl river delta," *Geographical Science and Development*, vol. 30, no. 10, pp. 1233–1239, 2011.

[29] F. Guo, K. Lo, and L. Tong, "Eco-efficiency analysis of industrial systems in the Songhua river basin: a decomposition model approach," *Sustainability*, vol. 8, no. 12, p. 1271, 2016.

[30] C. Yu, H. Li, X. Jia, and Q. Li, "Improving resource utilization efficiency in China's mineral resource-based cities: a case study of Chengde, Hebei Province," *Resources, Conservation and Recycling*, vol. 94, pp. 1–10, 2015.

[31] S. Schaltegger and T. Synnestvedt, "The link between "green" and economic success: environmental management as the crucial trigger between environmental and economic performance," *Journal of Environmental Management*, vol. 65, no. 4, pp. 339–346, 2002.

[32] Z. Lu, H. Wang, and Q. Yue, "Decoupling index: the quantitative expressive of resource consumption, waste discharge and economic growth," *Resources Science*, vol. 33, no. 1, pp. 2–9, 2011.

[33] M. Smith, K. C. Hargroves, and C. Desha, *Cents and Sustainability: Securing Our Common Future by Decoupling Economic Growth from Environmental Pressures*, Earthscan, London, UK, 2010.

[34] P. Tapio, "Towards a theory of decoupling: degrees of decoupling in the EU and the case of road traffic in Finland between 1970 and 2001," *Transport Policy*, vol. 12, no. 2, pp. 137–151, 2005.

[35] Y. Chen, N. Zhang, Y. Zhang et al., "TOFFEE: task offloading and frequency scaling for energy efficiency of mobile devices in mobile edge computing," *IEEE Transactions on Cloud Computing*, p. 1. 2009, In press.

[36] C. Wang, "Decoupling analysis of China's economic growth and energy consumption," *China Population Resources and Environment*, vol. 20, no. 3, pp. 35–37, 2010.

[37] S. Xu, T. Wang, and T. Wang, "On energy equity and China's policy choices," *Energy & Environment*, vol. 28, no. 3, pp. 288–301, 2017.

[38] J. D. Ward, P. C. Sutton, A. D. Werner, R. Costanza, S. H. Mohr, and C. T. Simmons, "Is decoupling GDP growth from environmental impact possible?" *PLoS One*, vol. 11, no. 10, Article ID e0164733, 2016.

[39] Y. Zhang, C. Yin, Q. Wu et al., "Location-aware deep collaborative filtering for service recommendation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems (TSMC)*, p. 12. 2020, In press.

[40] R. Sun, "Improving the calculation method of Tapio decoupling index and its implications," *Techno-Econ Management Resources*, vol. 8, pp. 7–11, 2014.

[41] J. Mao, "The path selections of resource-based cities transformation in China," *International Journal of Financial Research*, vol. 5, no. 2, pp. 171–174, 2014.

[42] J. He, T. Zhuang, and X. Xie, "Energy consumption, economic development and environmental improvement in China," *Energy & Environment*, vol. 25, no. 8, pp. 1345–1357, 2014.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 1971286, 15 pages
https://doi.org/10.1155/2020/1971286

Research Article

ROPPSA: TV Program Recommendation Based on Personality and Social Awareness

Nana Yaw Asabere and Amevi Acakpovi

Accra Technical University, Accra, Ghana

Correspondence should be addressed to Nana Yaw Asabere; yawasabere2005@yahoo.com

Received 2 February 2020; Revised 11 April 2020; Accepted 16 May 2020; Published 8 June 2020

Guest Editor: Weilin Xiao

The rapid growth of mobile television (TV), smart TV, and Internet Protocol Television (IPTV) content due to the convergence of broadcasting and the Internet requires effective recommendation methods to select appropriate TV programs/channels. Many previous methods have been proposed to address this issue. However, imperative factors such as the utilization of personality traits and social properties to recommend programs for TV viewers remain a challenge. Consequently, in this paper, we propose a recommender algorithm called Recommendation of Programs via Personality and Social Awareness (*ROPPSA*) for TV viewers. *ROPPSA* utilizes normalization and folksonomy procedures to generate group recommendations for TV viewers who have common similarities in terms of personality traits and tie strength with a Target TV Viewer (*TTV*). Therefore, *ROPPSA* improves TV viewer cold-start and data sparsity situations by utilizing their personality traits and tie strengths. We conducted extensive experiments on a relevant dataset using standard evaluation metrics to substantiate our *ROPPSA* recommendation method. Results of our experimentation procedure depict the advantage, recommendation accuracy, and outperformance of *ROPPSA* in comparison with other contemporary methods in terms of precision, recall, *f*-measure (*F*1), and arithmetic mean (AM).

1. Introduction

Globally, recent decades have witnessed the emergence of novel challenges regarding television (TV) content consumption. TVs are machine screens that broadcast signals and convert them into multimedia pictures and sounds for educational and entertainment purposes. The evident convergence of Internet and broadcasting has paved the way for the proliferation of different TV technologies such as Internet Protocol TV (IPTV), mobile TV, and smart TV which can be viewed on multiscreen and device ecosystems [1–6]. In particular, the content of mobile TV is watched through mobile phones and smartphones, which makes them digitally ubiquitous.

Additionally, online social networking sites such as Facebook, Netflix, YouTube, and Twitter have complemented TV viewing experience by creating a large number of multiscreen applications [1–3]. Internationally, due to habit changes in accordance with diversification of TV content, active viewers experience difficult times deciding which programs to watch among thousands of channels as shown in Figure 1. Such situations create difficulty for TV producers to predict viewers' interests and preferences in certain circumstances [1–6].

Consequently, the identification and selection of TV programs/channels have improved over the years. Electronic Program Guides (EPGs) have become available and more appropriate to use in comparison to historical and traditional methods such as browsing printed program guides and channel surfing [8–11]. Intelligent personalization techniques and systems can be innovatively utilized to significantly improve the effectiveness of searching and retrieving TV programs of interest.

Many researchers such as [1–6, 8–11] have utilized traditional recommender systems and techniques such as hybrid filtering (HF), collaborative filtering (CF), and content-based filtering (CBF) as shown in Table 1 to accomplish the task of searching, retrieving, and suggesting the

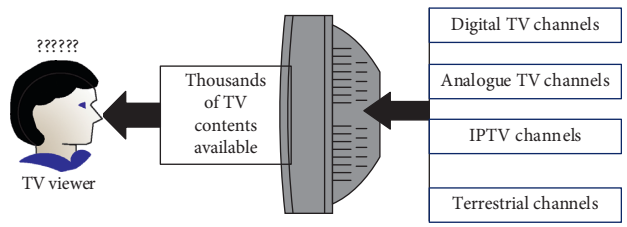


Figure 1: TV viewers often get confused with modern TV technologies.

right TV programs (items) for viewers through different recommender systems. Contextual information such as location, time, and physical conditions has also been applied in different recommendation scenarios [12].

Unfortunately, although the burden of information overload related to program consumption by TV viewers has been tackled to some extent, most of these existing personalized TV systems (e.g., [8–11]) did not exploit real-time social features/properties and personality traits of TV viewers to predict TV program interests/preferences. The incorporation of social (tie strength) and personality (traits) factors can help reduce the challenges of data sparsity (users rating a small proportion of items out of a larger number of available items) [13] and cold-start problems (new user and new item problems) [14] in TV recommender systems in order to significantly improve recommendation accuracy and quality of TV programs for viewers.

Consequently, this paper employs two main concepts, namely, personality and social properties/characteristics of TV users to generate effective and reliable group recommendations relating to TV program content. Social properties such as tie strength, closeness centrality, and degree centrality have been proved in literature [15–21] as very effective and important entities in social recommender systems. For instance, social ties and network centrality measures between attendees of a smart conference have positively been utilized to generate recommendations regarding conference attendees [15, 16], scholarly research papers [17, 18], and conference session venues [19–22].

Furthermore, some researchers who have worked on cold-start challenges in recommender systems [14, 23, 24] employed a wide variety of user profiles to generate recommendations without considering the personality of both new and existing users in different domains. Research has shown that personality is a stable and measurable psychological construct which can be used to explain various human behaviours using some traits [25]. This means that the concept of personality is capable of providing rich information. Moreover, Hu and Pu [26] proved that personality is the main and persistent factor which reflects the interests of people and predicts their behaviors. For example, the utilization of personality profiles can both provide recommendations for decision support and effectively solve cold-start issues [15, 16]. Therefore, it is appropriate to adopt new users' personality as their profiles [15, 16, 25, 26] to carry out a deeper study on TV program recommendations for viewers.

In order to decipher the enumerated research problem above, we propose a recommender algorithm called Recommendation of Programs via Personality and Social Awareness (*ROPPSA*) for TV viewers. Specifically, *ROPPSA* utilizes normalization and folksonomy procedures to generate recommendations for TV viewers in groups that have strong similarities in terms of personality traits and tie strength with a Target TV Viewer (*TTV*). *ROPPSA* considers different sources of information including (i) personal (TV program interests of viewers), (ii) social (tie strengths between a *TTV* and TV viewers), and (iii) user personality traits between a *TTV* and TV viewers.

1.1. Contributions. The main contributions in this work are summarized as follows:

(i) In our proposed recommendation method, we consider the social properties and personality traits of users by computing tie strengths and personality similarities between a *TTV* and TV viewers.

(ii) Our recommendation method uses personality rating similarities coupled with a normalization procedure to generate reliable groups for personality group recommendations. Furthermore, our recommendation method utilizes the idea of tie strength computations and folksonomies to generate reliable groups for social group recommendations.

(iii) The recommendation methodology developed in this paper reduces cold-start and data sparsity and is favourable for improving the social and personality awareness of smart and mobile TV viewers.

(iv) Our proposed algorithm enables effective and reliable group recommendations of TV programs to TV viewers who have high tie strength and similar personalities with a *TTV* based on TV program interests.

(v) Finally, our *ROPPSA* recommendation approach is evaluated and validated through an experimentation procedure, which utilized a relevant dataset for an experimental comparison with other contemporary recommendation techniques methods.

1.2. Structure of the Paper. The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 describes our proposed *ROPPSA* recommender algorithm and model. Section 4 presents our experimentation procedure and results to show the effectiveness of our proposed *ROPPSA* method. Finally, Section 5 concludes this paper.

2. Related Studies

This section presents associated studies and literature relating to our study. We focus on related work pertaining to (i) personalized TV program recommender systems, (ii) social group TV recommender systems, and (iii) personality group recommender systems.

Table 1: Categories of traditional recommender systems.

Recommender system	Brief description
CBF	The CBF procedure involves the utilization of related content information with items and users. In CBF, recommendation of new items to a user involves matching content with items known to be of interest to the user [7]
CF	Using a rating process, the CF method is based on the notion that a user would usually be interested in items preferred by other users with similar interests [7]
HF	HF merges CF and CBF as well as other recommender algorithms/systems to reduce challenges such as data sparsity and cold-start [7]

2.1. Personalized TV Recommender Systems. Globally, a lot of researchers have made considerable efforts to develop effective personalized TV program recommender systems. A brief report of such related studies is presented as follows.

Krstic and Bjelica [27] designed a personalized TV program guide that employs a one-class classifier tool with an autoencoder neural network. Similarly, Martinez et al. [9] proposed a personalized TV program recommender system called *queveo.tv*, which hybridizes collaborative filtering and content filtering.

Furthermore, Kim et al. [1] generated personalized TV program recommendations for a target user by utilizing CF to propose a TV program recommender algorithm. Through genre and similar channel preferences, a rank model was employed to generate TV program recommendations in [1]. Similar to [1], Pyo et al. [2] used sequential pattern mining such as viewing time and watching length of each TV program to develop an automatic recommender algorithm for personalized TV program scheduling.

Kwon and Hong [28] employed memory-based CF to propose a personalized program recommender algorithm for smart TVs. The proposed method in [28] improved recommendation performance of EPGs and smart TV recommender applications. Similar to [28], Song et al. [6] presented a novel solution for content recommendation through the customization of IPTV content in accordance with current user context and environment, thereby ensuring better user experience.

The fundamental concept of group recommendation tries to maximize the group satisfaction function, which is generally based on periodically repeating recommendations or a partial satisfaction function after the sequence of items [29]. Kim et al. [30] proposed a TV program recommendation method by merging multiple preferences. They used channels and genres of programs as features for their recommendation method. Similarly, Chaudhry et al. [31] considered a heterogeneous information network environment and proposed to extract the heterogeneous relationship information between the target viewer and different TV programs following different metapaths, to deliver good quality recommendation using the implicit feedback history data of the viewer.

2.2. Social Group TV Recommender Systems. Social TV is defined as a social media service whereby TV users share familiarities about TV programs in a social network that they are viewing [3, 32]. Effective social TV service is typically provided by considering two technicalities: (i) how to create social TV communities by grouping similar TV users? and (ii) how to recommend TV programs based on group and personal interests for personalizing TV? [3].

As a consequence of the above, Pyo et al. [3] employed two Latent Dirichlet Allocations (LDAs) and proposed a unified topic model, which recommends TV programs as a social TV service based on grouping similar TV users. Similar to [3], Wang et al. [33] established a group preference model by evaluating the external followees' impact on group interest using external experts' guidance for social TV group recommendation. Similar to [3, 33], Shin and Woo [34] proposed a socially aware TV program recommender for several viewers based on individual and group interests. In order to achieve this task, the authors merged user profiles by combining their common interests to generate social TV recommendations.

Shepstone et al. [35] presented a recommendation method which describes how audio analysis of age and gender in a group of TV viewers can be used to generate recommendations of TV programs for the group. Similarly, based on TV content and semantic reasoning techniques, Sotelo et al. [36] presented a method of TV program recommendation for groups of people.

2.3. Personality Group Recommender Systems. As a result of the fact that most methods model real-life conditions in group recommendation, the need for personality detection in a group is tremendously important [29]. In relation to personality inclusion in recommender systems, various types have been extensively proposed in [37–42].

In terms of personality group recommendations, Quijano-Sachnez et al. [43] described some new concepts of enhancing recommendations to people in different groups. Their approach maximized global group satisfaction by considering social associations and personality among each group. Similarly, Quijano-Sachnez et al. [44] ensured agreements within a group by defining alliances through personality and trust.

Similar to [43, 44], Quijano-Sachnez et al. [45] utilized three imperative features: past memory, trust, and personality to generate group recommendations. Similar to [45], Quijano-Sachnez et al. [46] proposed an entrusting-based recommendation method, which includes an analysis of personality in conflict situations and group characteristics such as size, structure, and trust between group members.

Recio-Garcia et al. [47] presented an innovative method of generating recommendations to groups based on CF and group personality composition. Similarly, through the utilization of personality and trust, Recio-Garcia et al. [48] proposed a Decision Support System for groups of people, where each user entrusts an agent that represents his/her interests. In order to generate personality group recommendations, Prada et al. [49] introduced diversity in the behaviour of social agents through a computational model of personality based on the Five-Factor Model (FFM).

In comparison with the existing TV program recommendation methods described above, our proposed *ROPPSA* method exhibits novelty and advantage of integrating social and personality awareness through the respective computations of tie strengths and personality similarities of TV users (*TTV* and other TV viewers) in order to establish strong ties and highly similar personalities between them for effective TV program recommendations.

Our method paves the way for TV users with strong ties (social relationships) and similar personalities to be grouped so that similar TV viewer groups are recommended similar TV programs as a result of their relationships and correspondences with one another. Consequently, personalized TV program recommendations are possible with the similar TV user groups related to social and personality awareness.

The fundamental recommendation process of *ROPPSA* is illustrated in Figure 2. Figure 2 involves TV viewers, various recommendation entities, and personality and social group recommendations of TV programs. The integration of individual user profiles represents a possible solution for recommending TV programs to multiple viewers. This process is referred to as group modeling or group recommendation [31].

To the best of our knowledge and with reference to the literature above, this is the first time tie strength and personality traits are being utilized to generate group recommendations for TV viewers.

In the next section, we present details on the design of our proposed *ROPPSA* recommendation method for TV viewers/users. Specifically, personality traits for TV recommendation are not available in the literature. This reason further corroborates our innovative objectives for proposing *ROPPSA* as a recommendation method for TV viewers in our society.

3. Algorithmic Design of *ROPPSA*

In this section, we present the fundamental concept and architecture of *ROPPSA*. Figure 3 illustrates that our *ROPPSA* recommendation model generates both social group and personality group TV program recommendations through initial computations of tie strengths and personality similarities among mobile TV viewers, respectively.

Referring to Figure 3, the *TV Viewer Crawler (TVC)* collects and sends personality trait ratings of individual TV viewers to the *Group Profile Organizer* (*GPO*). The *GPO* is responsible for creating group profiles in accordance with personality trait ratings and interests of participants using a normalization procedure. The *Similarity Profiler (SP)*

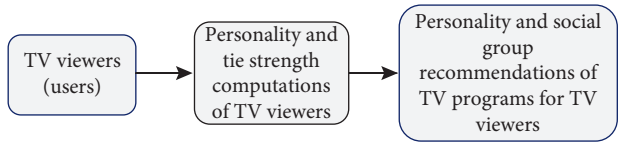


Figure 2: Fundamental recommendation procedure of *ROPPSA*.

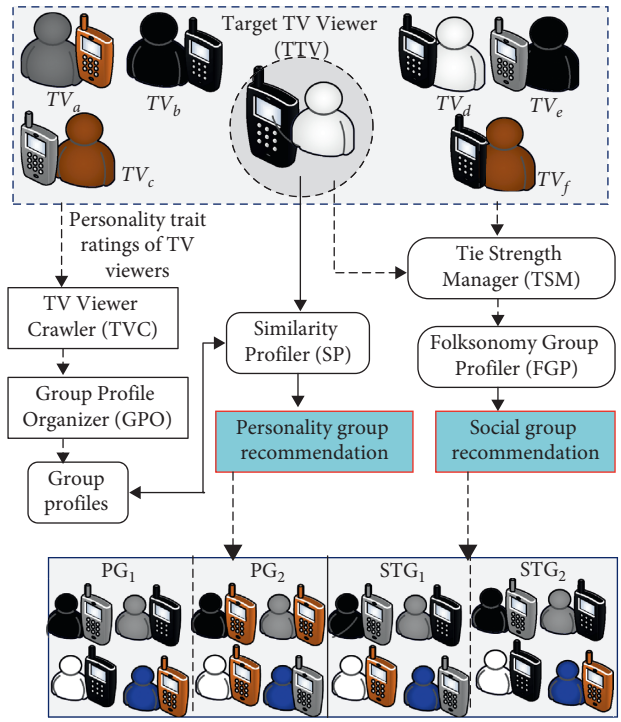


Figure 3: *ROPPSA* recommendation model.

computes similarities of personality trait ratings between the Target TV Viewer (*TTV*) and the created personality group profiles (pgp_b) in order to generate reliable and efficient personality group recommendations of TV programs based on the interests of the *TTV*.

The *Tie Strength Manager (TSM)* computes the contact frequencies and durations between the *TTV* and other TV viewers so that in accordance with their tie strengths, the Folksonomy Group Profiler (*FGP*) allocates groups for wider coverages of reliable and efficient social group recommendations based on program interests of the *TTV*.

The program interests of smart and mobile TV viewers can vary at any time as a result of modifications in TV schedules of a particular broadcasting station. Due to such variations, the recommendation process in *ROPPSA* depends on both mobile and smart TV viewer profiles in the current viewer situation. Due to the fact that *ROPPSA* operates on smart and mobile TVs, *ROPPSA* requires standard mobile and smart TVs with integrated operating systems, Internet, and web 2.0 features which use a web browser as a client supported by a programming language. We elaborate further on our *ROPPSA* recommendation framework below.

3.1. Capturing TV Program Preferences in ROPPSA. As stated earlier, in order to determine TV program preferences of TV viewers, *ROPPSA* initially utilizes and computes similar personalities between a *TTV* and other (grouped) TV viewers for its personality group recommendation service. Consequently, *ROPPSA* employs an explicit approach that allows TV viewers to input specific personality trait ratings from 1 to 5 using the Big Five Personality Traits (*BFPT*) [25]. *BFPT* consists of the following:

(i) *Extraversion*: talkative, outgoing, energetic assertive, etc.

(ii) *Agreeableness*: trusting, sympathetic, forgiving, generous, kind, appreciative, etc.

(iii) *Openness*: wide interests, innovative insightful, imaginative, curious, artistic, etc.

(iv) *Conscientiousness*: responsible, thorough, reliable, planful, organized, efficient, etc.

(v) *Neuroticism* (*emotional stability*): worrying, unstable, touchy, tense, self-pitying, anxious, etc.

ROPPSA further employs and computes tie strengths between a *TTV* and other (grouped) TV viewers (using their contact durations and frequencies) for its social group recommendation service.

3.2. Personality Group Profiling. Recommendation to groups usually involves the application of precise and appropriate characteristics in order to develop a compromise for different group semantics that validate disagreements and agreements among users [31, 44]. We organize our personality group profiling process illustrated in Figure 4 by adopting the arrangement of group recommendation in CF presented by Bobadilla et al. [7] and Asabere et al. [17]. Figure 4 demonstrates how our recommendation method combines TV viewer data and obtains data involving groups of TV viewers in four different steps.

The personality profiles of users (TV viewers) define their interests in terms of TV programs. Consequently, group profiles define common features used by individual TV viewers. We conformed to group profiling of individual TV viewers, by adopting the socially aware TV program recommender method used in [34]. As a consequence, we initially combine TV viewer profiles based on their personality trait ratings. The personality trait ratings of TV viewers are further arranged into groups by the *GPO* due to the fact that they have different traits and ratings.

As shown in Figure 4, the rating matrix of personality trait ratings by TV viewers is verified by TV_c in step 1. In steps 2 and 3, *GPO* generates neighbourhood intersections and group prediction phase (normalization), based on similar personality traits rated by TV viewers using (1). In the last step, *SP* decides which TV programs to recommend and generates a personality group recommendation based on the personality similarity between a *TTV* and the created personality group profiles (pgp_b).

Consequently, it is imperative to compute the weighted average associated with each personality trait rating using

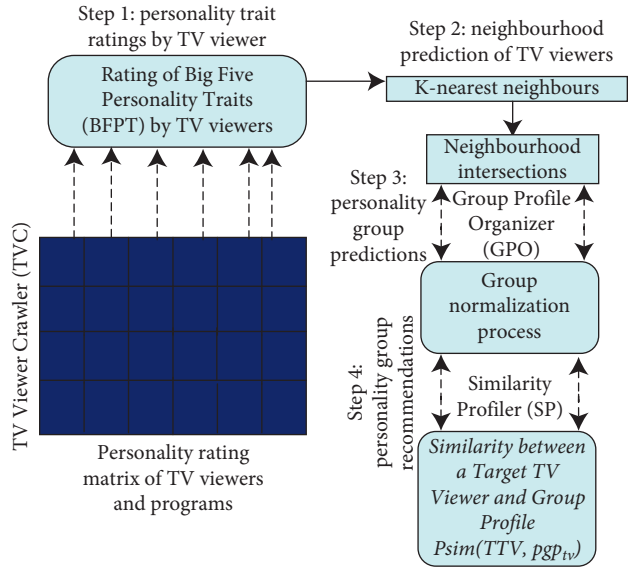


Figure 4: Personality group profiling and organization of TV viewers.

(2). This computation describes the program preferences of the TV viewers in order for the *GPO* to create a preference model for generating each group profile. As shown in Figure 4, *GPO* utilizes the *K-nearest neighbour* methodology by identifying unique personality trait ratings of TV viewers and categorizing them as a group [17, 31, 34]. Therefore, TV viewers with similar personality trait ratings are classified as nearest neighbours and grouped together.

Personality trait ratings of TV viewers are done on a scale of 1–5, where 1 and 5 are the lowest and the highest preference ratings, respectively. Let the TV program preferences of each group profile be $PGP = \{pgp_1, pgp_2, pgp_3, \cdots, pgp_b\}$. The normalization process by *GPO* interconnects common neighbours and predicts different groups of TV viewers [17, 18]:

$$\text{norm\_}TV_{a,b} = TV_{\min} + \frac{(TV_{\max} - TV_{\min}) \times TV_b}{PR_{\max} - PR_{\min}}. \quad (1)$$

We employ (1) to compute the normalized (n) values of personality trait ratings which have been rated differently by different TV viewers. In (1), $TV_{a,b}$ denotes individual personality trait ratings of the TV viewers, for a particular personality trait, where a is the notation for a particular TV viewer and b is his/her rating for a particular personality trait. The precise maximum and minimum ratings of the TV viewers for a particular personality trait are, respectively, represented as $TV_{\max}$ and $TV_{\min}$. In addition, $PR_{\max}$ and $PR_{\min}$ represent the overall highest and lowest personality trait ratings in the experimental dataset. It must be emphasized that, in (1), $PR_{\max}$ and $PR_{\min}$ are, respectively, equal to 5 and 1, representing the highest and lowest possible personality trait ratings in the dataset [17, 18]:

$$pgp_b = \frac{\sum_{a=1}^{TV} \text{norm\_}TV_{a,b}}{NTV}. \quad (2)$$

In (2), the values attained from (1) are totaled and divided by the number of TV viewers (*NTV*) who allotted ratings to a particular personality trait. Consequently, through (2), we compute and attain an average normalized value which designates the creation and allocation of a group profile for particular TV viewers. The TV viewers are allocated into their respective group profiles if their rating for a particular personality trait is greater than or equal to the corresponding average normalized value (pgp_b) of their group [17, 18].

With reference to Table 2, we demonstrate an example of our group profiling and normalization technique. In Table 2, there are four TV viewers with different personality trait ratings. For instance, in order to generate a group profile for the personality trait "Openness," the first step of our proposed recommender algorithm is to count the number of available ratings related to "Openness." Table 2 shows that there are four different ratings associated with "Openness" rated by each of the four TV viewers.

The normalized value of TV_a*'s* annotation on "Openness" with a rating of 3 is computed as follows:

$$\text{norm\_}TV_{a,3} = 2 + \frac{(4-2)\times 3}{5-1} = 2 + \frac{6}{4} = 3.5. \quad (3)$$

Similarly, the normalized value of TV_b*'s* annotation on "Openness" with a rating of 2 is computed as follows:

$$\text{norm\_}TV_{b,2} = 2 + \frac{(4-2)\times 2}{5-1} = 2 + \frac{4}{4} = 3. \quad (4)$$

Furthermore, the normalized value of TV_c*'s* annotation on Openness to experience with a rating of 4 is computed as follows:

$$\text{norm\_}TV_{c,4} = 2 + \frac{(4-2)\times 4}{5-1} = 2 + \frac{8}{4} = 4. \quad (5)$$

Additionally, the normalized value of TV_d*'s* annotation on Openness to experience with a rating of 2 is computed as follows:

$$\text{norm\_}TV_{d,2} = 2 + \frac{(4-2)\times 2}{5-1} = 2 + \frac{4}{4} = 3. \quad (6)$$

Consequently, the creation of a personality group profile (pgp_b) for the personality trait "Openness" in Table 2 is computed as follows:

$$pgp_b = \frac{\sum_{a=1}^{TV} norm\_TV_{a,b}}{NTV} = \frac{3.5+3+4+3}{4} = 3.4. \quad (7)$$

The above normalization and personality group profiling approach in our proposed *ROPPSA* recommendation model illustrates that allocation of TV viewers to a personality group profile associated with "Openness" ratings should be assigned to TV viewers with ratings of 3.4 or above. This implies that, as per the data on Table 2, only two TV viewers, namely, TV_a and TV_c, can be assigned to that particular personality group profile for personality group recommendation of TV programs.

We employ equation (8) to determine the Personality Similarity (*PSim*) Score between *TTV* and pgp_b, i.e., *PSim* (*TTV*, pgp_b). In (3), *TTV* and pgp_b are represented by *a* and *b*,

Table 2: Normalization and personality group profiling process.

TV viewers	Personality traits	Ratings (1–5)
TV_a	Extroversion	5
	Agreeableness	3
	Openness	3
	Conscientiousness	2
	Neuroticism	2
TV_b	Extroversion	2
	Agreeableness	2
	Openness	2
	Conscientiousness	3
	Neuroticism	5
TV_c	Extroversion	3
	Agreeableness	2
	Openness	4
	Conscientiousness	1
	Neuroticism	1
TV_d	Extroversion	5
	Agreeableness	4
	Openness	2
	Conscientiousness	2
	Neuroticism	1

respectively. The average of all personality trait ratings of *a* and *b* is symbolized by $\overline{P_a}$ and $\overline{P_b}$, respectively. Additionally, the personality trait ratings of *a* and *b* with one of the personality traits *x* are denoted by $P_{a,x}$ and $P_{b,x}$, respectively [15, 16, 26]:

$$PSim(a,b) = \frac{\sum_{x\in X}\left(P_{a,x}-\overline{P}_a\right)\left(P_{b,x}-\overline{P}_b\right)}{\sqrt{\sum_{x\in X}\left(P_{a,x}-\overline{P}_a\right)}\sqrt{\sum_{x\in X}\left(P_{b,x}-\overline{P}_b\right)}}. \quad (8)$$

Our algorithm utilizes and returns a *PSim* Score between −1 and 1, where 1 indicates that *TTV* and pgp_{tv} have precisely the same or quite similar ratings and −1 denotes otherwise [15, 16, 26]. As a result of the fact that our recommender algorithm generates multiple personality group profiles for recommendation of TV programs, it loops through different *TTV* and pgp_b transactions in accordance with the dataset using (8). Furthermore, with reference to the range for *PSim* Scores during experimentation, we set a threshold α in (9) for *PSim* computations between *TTV* and pgp_b as follows:

$$PSim(TTV, pgp_b) \geq \alpha. \quad (9)$$

3.3. Tie Strength. In a social network, tie strength is defined as the extent to which users are close to each other in terms of social/collaborative relations. A high tie strength normally depicts strong social relations such as friendship and familiarities among users in a social network. Users in a social network with low tie strengths have weaker relationships in terms of closeness. In order to acquire essential and important information from users in a social network, the concept of ties represents a major channel for collaboration and communication. Ties establish communication links in a social network, thereby improving and providing key

elements for social cooperation and collaboration among users [15–21].

The concept of social ties/tie strength has been applied in different socially aware recommender algorithms and techniques [15–21]. Available evidence in the literature shows that these algorithms have performed considerably well in the alleviation of issues regarding cold-start and data sparsity as well as improved social recommendation accuracy [15–21]. Using the following equation, we measure and evaluate the tie strength between *TTV* and another TV viewer, TV_a

$$\text{Tie\_Strength}_{TTV,TV_a}(t) = \frac{f_{TTV,TV_a} \times d_{TTV,TV_a}}{T}. \quad (10)$$

In (10), $f_{TTV,TVa}$ is the social contact frequency between *TTV* and TV_a in the time frame T and $d_{TTV,TVa}$ is their social contact duration within the same time frame T.

3.4. Social Group-Tie Profile Generation. The expansion of different social network communities substantiates innovative procedures and arrangements regarding the important role of self-organization principles. In the scope of this paper, we are interested in an algorithm which involves sharing of social resources, through the usage of a knowledge demonstration process called folksonomy [17, 50].

In order to generate social tie-group profiles in our *ROPPSA* recommendation model, we initially utilize the concept of folksonomies by connecting a *TTV* to other TV viewers in the dataset. In this paper, a folksonomy can further be described as a hypergraph $G = [TTV, TV, H]$, where *TTV* represents Target TV Viewers, TV represents other TV viewers, and $H = [h_1, h_2, h_3, \cdots, h_n]$ denotes the set of hyperedges which only exist among the nodes (TV viewers) in different sets [17, 50].

With reference to Figure 5, a *TTV* is connected to different TV viewers through hyperedges h_1 and h_2, respectively. TV viewers TV_a, TV_b, TV_c, and TV_d are connected to the *TTV* through hyperedge h_1. Furthermore, TV viewers TV_e, TV_f, TV_g, and TV_h are connected to the *TTV* through hyperedge h_2. We employ (10) to compute the tie strengths between *TTV* and all other TV viewers connected to hyperedges h_1 and h_2. After computing the tie strengths, we utilize (11) to compute a threshold for the social tie-group profile ($stgp_{tv}$) which is connected through hyperedge h_3:

$$stgp_{tv} = \frac{\text{sum of Tie\_Strengths}}{\text{number of Tie\_Strengths}}. \quad (11)$$

Using a time frame (T) of 660 minutes, we present an example of our social tie-group profile generation process in Table 3 and further explain how (10) and (11) are applied to generate social group recommendations for TV viewers.

Table 3 depicts the computation of tie strengths between *TTV* and other TV viewers using (10) which applies respective contact durations and frequencies. A close view of Table 3 depicts the tie strength for each computation of the *TTV* and other TV viewers. We employ (11) to compute the threshold of the social tie-group profile for the tie strength displayed in Table 3 as follows:

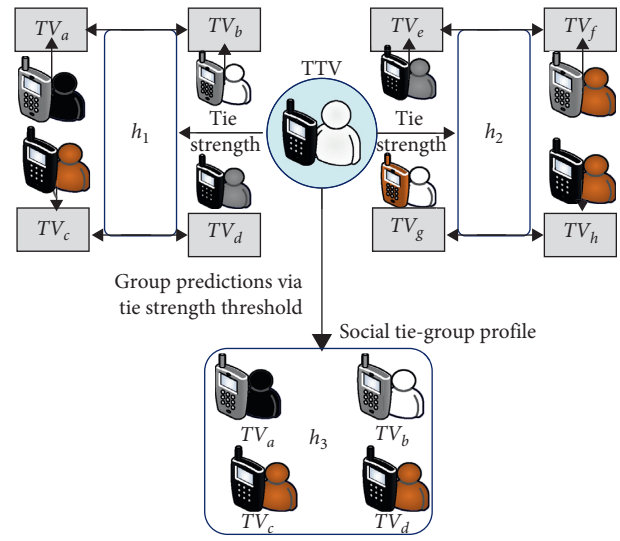


Figure 5: Social tie-group profiling and organization of TV viewers.

$$stgp_{tv} = \frac{0.6 + 0.6 + 0.3 + 0.1 + 0.2 + 0.2 + 0.3 + 0.2}{8} = \frac{2.3}{8} = 0.3. \quad (12)$$

The above social tie-group profiling approach in our proposed *ROPPSA* recommendation model demonstrates that the allocation of TV viewers to a social tie-group profile based on tie strength should be allotted to TV viewers who have a tie strength of 0.3 or above with a *TTV*. This implies that, as per the data on Table 3, only four TV viewers, namely, TV_a, TV_b, TV_c, and TV_g, can be assigned in that particular social tie-group profile to generate a social group recommendation of TV programs. This is depicted in the connections of hyperedge h_3 in Figure 5.

3.5. ROPPSA Recommender Algorithm. Our proposed *ROPPSA* recommender algorithm (Algorithm 1) displays the required inputs and outputs for personality and social group recommendation procedures. Variables in our recommender algorithm are declared in steps 2–4. In step 5, unique personality groups are identified. Steps 6–12 compute normalized personality groups using (1) and (2) and allocate TV viewers to their corresponding groups. Using (8), steps 13–16 compute the personality trait similarity between *TTV's* and personality group profiles (pgp_b) in order to recommend TV programs to personality group profiles based on personality trait similarity threshold utilized through (9).

Similarly, steps 18–22 compute the tie strength between *TTV*s and TV viewers using (10) and allocate TV viewers into social tie-group profiles using (11) as a threshold for social group recommendation of TV programs.

4. *ROPPSA* Experimentation Procedure

This section demonstrates our experimentation procedure for the evaluation of *ROPPSA*. Using an extended version of

Table 3: Social tie-group profiling process.

Target TV Viewer	TV viewers	Contact duration	Contact frequency	Tie strength
TTV	TV_a	5	7	0.6
	TV_b	3	6	0.6
	TV_c	3	5	0.3
	TV_d	2	4	0.1
	TV_e	2	3	0.2
	TV_f	2	4	0.2
	TV_g	2	6	0.3
	TV_h	2	5	0.2

the ATU dataset utilized in [21], we present our experimental results in comparison with other existing and relevant methods.

4.1. Experimental Dataset. In order to substantiate the procedure of our proposed recommender algorithm (*ROPPSA*), it was imperative that we utilize precise and relevant data for our experimentation process. Therefore, as stated above, we utilized and extended version of the Accra Technical University (ATU) dataset published in [21]. The ATU dataset in [21] needed to be extended because it did not contain personality trait ratings required for the experimentation in this paper. Consequently, we gathered more user, tie strength, and personality trait data from ATU students which we represented as TV viewers. Data were gathered from students in Higher National Diploma (HND) Marketing ($n = 2864$), HND Computer Science ($n = 720$), and HND Building Technology ($n = 812$).

Therefore, the ATU dataset utilized in this paper contains a total of 4396 TV viewers (2421 males representing 55.07% and 1975 females representing 44.93%). Tie strength details of the ATU dataset utilized in this paper are illustrated in Figures 6 and 7.

The personality data gathered in our ATU dataset involved personality trait ratings by the ATU students (designated as TV viewers) in the dataset. Among the *BFPT*, students rated each personality traits using a scale of 1 to 5. The total number of personality ratings for all the traits combined in the ATU dataset is 22,541. Figure 8 illustrates the personality trait data utilized in our ATU dataset for experimentation.

Additionally, in order to ensure and substantiate recommendation accuracy in our experimentation procedure, we gathered TV program interests of TV viewers in accordance with the following: movie/TV series—*Horror*, *Action*, *Comedy*, *Adventure*, and *Drama*; sports/games—*Football*, *Basketball*, *Hockey*, *Boxing*, and *Athletics*. Tables 4 and 5 illustrate data regarding TV program interests of TV viewers in the ATU dataset in accordance with their ratings.

In relation to the ATU dataset, the total time frame (T) which is 1140 minutes utilized in [21] remained the same. Consequently, using (10), we computed $Tie\_Strength_{TTV,TVa}(t) = (80 \times 7)/1440$ and achieved a result of 0.5 as the maximum (highest) Social Contact Score (SCS) between a *TTV* and other TV viewers for positive and effective social recommendations

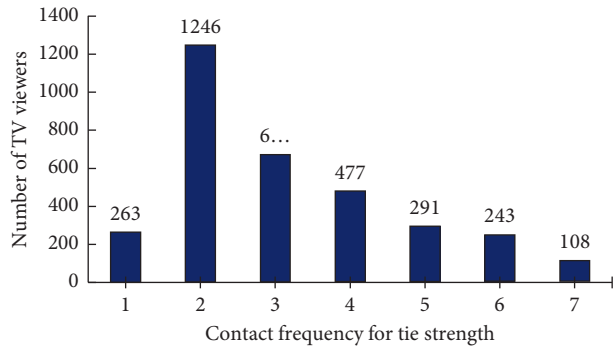


Figure 6: ATU dataset—contact frequency trends.

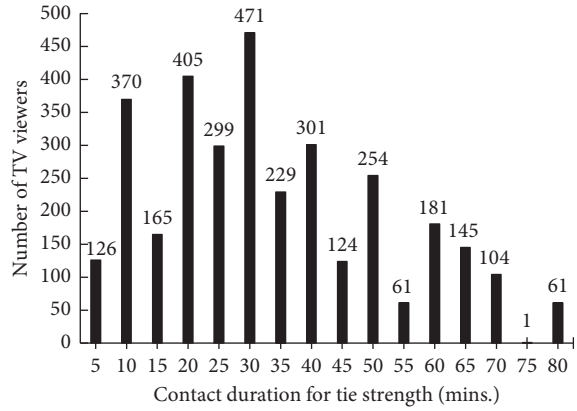


Figure 7: ATU dataset—contact duration trends.

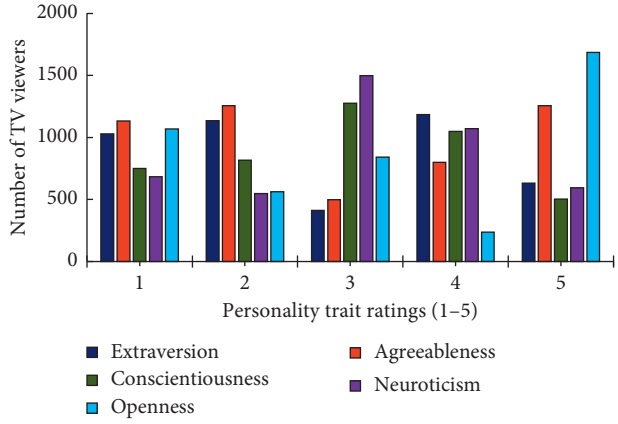


Figure 8: ATU dataset—personality data.

Input:
Target TV Viewer (*TTV*)
TV viewers (TV_s)
Personality trait ratings of *TTV* and TV_s
Tie strengths of *TTV* and TV_s
Output:
TTV personality and social group recommendations for TV_s

```
(1) procedure Personality and Social (Group) Recommendation of TV programs
(2) //Declare and Initialize Variables
(3) a and b //integer variables
(4)    personality_group_threshold, PSim_threshold and tie_strength_group_thresh; //floating variables
(5)    Identify unique personality trait ratings and classify as group
(6) //Compute normalized values of personality groups using (1) and (2)
(7) //Allocate TV viewers to respective groups
(8)    for every TV viewer do
(9)       if (TV viewer[a].Rating ≥ personality_group_threshold) then
(10)         Allocate TV viewer[a] in pgp[b];
(11)      end if
(12)   end for
(13) //Compute PSim of TTV and pgp_b using (8) and (9)
(14)      if (PSsim (TTV, pgp_b) ≥ PSim_threshold then
(15)      recommend TTV's TV programs to pgp_b;
(16)      end if
(17) for every TV viewer do
(18)   Compute tie_strength between TTV and TV_s using (10) and (11)
(19)   if (Tie_Strength (TTV, TV_s) ≥ tie_strength_group_thresh) then
(20)      Allocate TV viewer[a] in stgp[b] and recommend
             TTV's TV programs to stgp_b;
(21)      end if
(22)   end for
(23) end procedure
```

Algorithm 1: *ROPPSA* pseudocode for personality and social group recommendations.

Table 4: ATU dataset—movie/TV series program ratings.

Category	Ratings (1–5)				
	1	2	3	4	5
Horror	1890	2139	85	170	112
Comedy	167	1941	2039	146	1941
Action	114	108	1951	148	2075
Adventure	120	1477	525	1702	572
Drama	164	1805	129	2007	291
Total ratings	2456	7472	4732	4177	4996

[15–21]. In *ROPSAA*, the quality of social group recommendations for TV viewers is determined through the corresponding level of tie strength computation a *TTV* has with a social tie group.

Through the utilization of (11) in accordance with the ATU dataset, more effective social group recommendations in terms of quality and accuracy were generated with SCSs between 0.3 and 0.5. Therefore, the priority TV program recommendation list for a particular social contact group was established based on computed recommendation values that fell within this threshold. However, SCSs which fell out of the above threshold were considered as unsuccessful TV program recommendations. Consequently, we experimentally established $0.1 \leq SCS \leq 0.5$ as the range for TV program recommendation based on social contact (relations) and assigned a social group recommendation threshold of 0.3–0.5 in accordance with the dataset.

Typically, the computation of Rating Similarity Scores (RSSs) between two users/products is between −1 and 1. A value of 0 indicates that there is no association between two users or variables [51]. Consequently, using (8), we experimentally established $0.6 \leq RSS \leq 1$ as the range for TV program recommendation based on personality traits and assigned a personality group recommendation threshold of 0.8–1 using (9) according to the ATU dataset for successful recommendations in terms of prediction accuracy. Conversely, computed RSSs which fell out of the above threshold were considered as unsuccessful TV program recommendations.

4.2. Experimentation Setup and Evaluation Metrics. In our experimentation setup, we utilize a cross-validation technique to evaluate *ROPPSA* by initially partitioning our ATU dataset in training (80%) and test (20%) sets. The evaluation metrics employed in our experimentation of *ROPPSA* in the case of personality group recommendations focused on prediction accuracy. Therefore, we utilized root mean square error (RMSE), mean absolute error (MAE), and normalized mean absolute error (NMAE) in this regard. Lower values of

RMSE, MAE, and NMAE corroborate better performance. RSME is defined as follows [7]:

$$\text{RMSE} = \frac{\sqrt{\sum_{PR_{u,i} \in PR_{\text{test}}} \left(PR_{u,i} - \overline{PR}_{u,i}\right)^2}}{\left|PR_{\text{test}}\right|}. \quad (13)$$

Additionally, MAE is defined as follows:

$$\text{MAE} = \frac{\sum_{PR_{u,i}} \left|PR_{u,i} - \overline{PR}_{u,i}\right|}{\left|PR_{\text{test}}\right|}, \quad (14)$$

where $|PR_{\text{test}}|$ denotes the number of personality trait ratings in the test set. $PR_{u,i}$ and $\overline{PR}_{u,i}$, respectively, represent the real and prediction values of the personality rating in $|PR_{\text{test}}|$. As a result of the fact that different recommender systems/algorithms may use different numerical scales, we employed NMAE in our experiments to ensure that experimental errors can be fully expressed on a normalized scale. We therefore used (15) to compute NMAE. In (15), $PR_{\max}$ and $PR_{\min}$ are denoted as the upper and lower limits of user personality trait ratings in our ATU dataset, respectively. Consequently, in accordance with our ATU dataset, $PR_{\max} = 5$ and $PR_{\min} = 1$ [7]:

$$\text{NMAE} = \frac{\text{MAE}}{PR_{\max} - PR_{\min}}. \quad (15)$$

In relation to social group recommendations, we focused on recommendation quality and therefore employed accuracy, recall, and *f*-measure ($F1$). Higher values of accuracy (A), recall (R), and $F1$ substantiate better performance. We employed (16) to compute the preciseness of a recommender algorithm by utilizing the accuracy metric as follows [7]:

$$A = \frac{\text{Num}(N, tv)}{\text{Num}(TV)}. \quad (16)$$

Additionally, we utilized the recall metric to measure coverage of all good and successful TV programs using the following equation [7]:

$$R = \frac{\text{Num}(N, tv)}{\text{Num}(tv)}. \quad (17)$$

In (16) and (17), Num (N, tv) represents the number of relevant TV programs in the TV recommendation list for social group recommendation, Num (TV) in (16) denotes the number of all TV program recommendations and Num (tv) in (17) denotes TV programs that are relevant to a particular social group. Furthermore, in order to substantiate the results obtained through (16) and (17), we computed $F1$ using the following equation [7]:

$$F1 = \frac{2 \times A \times R}{A + R}. \quad (18)$$

4.2.1. Baseline Methods for Comparison. As a result of the fact that *ROPPSA* generates two main types of recommendations, namely, personality and social group recommendations, it was imperative to employ relevant methods for experimental comparison. Consequently, in order to validate our recommendation results in this paper, we compare *ROPPSA* to the following relevant methods described below.

In relation to personality group recommendations for TV programs, we utilized the work done by Kim et al. [30] and Chaudhry et al. [31] and represented them as $T1$ and $T2$, respectively, for comparison in our experiments. Both $T1$ and $T2$ are relevant personalized group TV recommendation methods which involved similar approaches of combining/merging individual user profiles/models to select TV programs to suit groups of TV viewers.

Furthermore, in relation to social group recommendations for TV programs, we exploited the work presented by Shin and Woo [34] and Wang et al. [33]. We represented them as $T3$ and $T4$, respectively, for comparison in our experiments. Both $T3$ and $T4$ are relevant social group TV recommendation methods which involved similar approaches regarding the establishment of group preference profiles/models from individual TV viewer profiles which relate to common social similarity and interests in terms of TV programs. The main goal of our experimentation procedure was to verify the effectiveness of *ROPPSA* and improve recommendation accuracy by alleviating issues of data sparsity and cold-start in TV program recommendation. The experimental results below present evidence on how the above goals were achieved.

4.3. Experimental Results and Analysis. In order to evaluate the experimental results effectively, we validated our experimentation procedure by answering the following questions:

(i) In comparison with $T1$, $T2$, $T3$, and $T4$, what was the overall assessment of *ROPPSA* in terms of TV program recommendation?

(ii) During our experimentation procedure, was there a reduction of cold-start and data sparsity in *ROPPSA*, compared to $T1$ and $T2$ for personality TV program group recommendations and $T3$ and $T4$ for social TV program group recommendations?

In terms of personality group recommendations, we utilized RMSE, MAE, and NMAE to evaluate *ROPPSA* based on the computations and measurements of similar personalities between a *TTV* and groups of TV viewers as described above.

As shown in Figure 9, at the maximum Rating Similarity Score (RSS) of 1.0, *ROPPSA* attained the lowest MAE (0.763) in comparison with that of $T2$ (0.824) and $T1$ (0.937). Similarly, in Figure 10, at the maximum RSS of 1.0, *ROPPSA* achieved the lowest NMAE (0.190) in comparison with $T2$ (0.205) and $T1$ (0.234). Furthermore, the RMSE value of *ROPPSA* in Figure 11 is 0.759 at an RSS of 1.0, in comparison with $T2$ (0.820) and $T1$ (0.933). Subsequent results of MAE, MMAE, and RMSE in Figures 9–11 corroborate the effectiveness of *ROPPSA*.

Table 6 summarizes the performance comparisons of the above recommendation methods on the ATU dataset in terms of MAE and RMSE, respectively. The experimental results show that *ROPPSA* outperforms $T1$ and $T2$ in terms

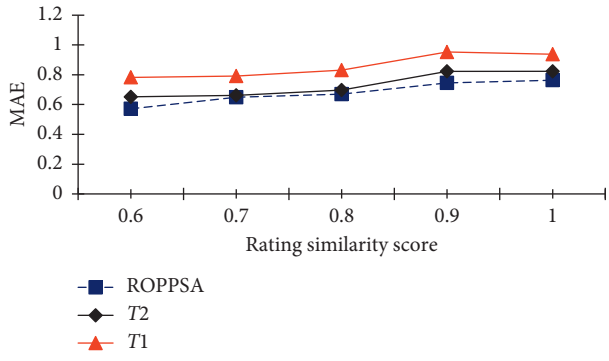


Figure 9: MAE performance on the ATU dataset.

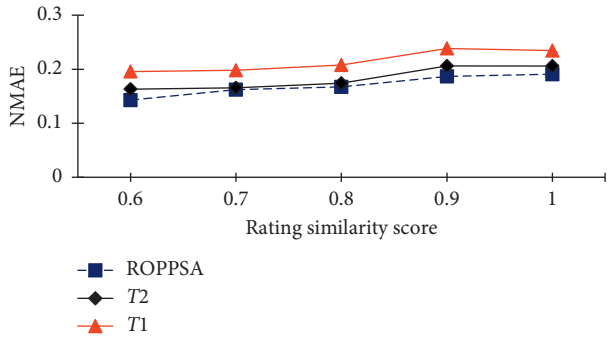


Figure 10: NMAE performance on the ATU dataset.

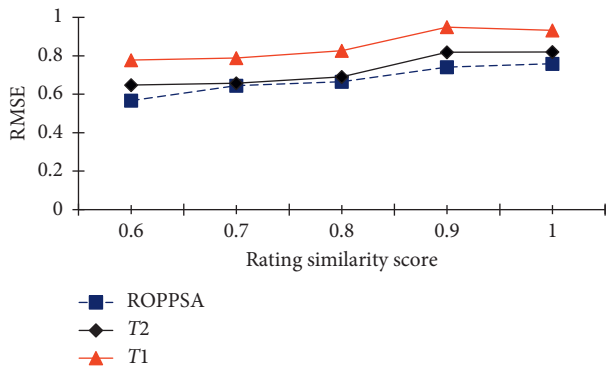


Figure 11: RSME performance on the ATU dataset.

Table 5: ATU dataset—sports TV program ratings.

Category	Ratings (1–5)				
	1	2	3	4	5
Football	1975	0	0	0	2421
Basketball	1381	594	2421	0	594
Hockey	2978	108	570	148	592
Boxing	1381	0	594	1483	938
Athletics	1381	2077	0	938	0
Total ratings	9097	2781	3588	2573	4550

of MAE and RMSE. This is because the use of personality trait information improves recommendation accuracy and performance by compelling TV viewers to be more similar in terms of personality. It can also be observed that $T2$ and *ROPPSA* outperform $T1$ in terms of RMSE and MAE due to the same utilization of personality information. This further proves and substantiates the necessity and effectiveness of integrating personality information into TV program recommendation methods for performance improvement.

The experimental results in Table 6 and Figures 9–11 reveal that the use of personality as a TV program recommendation entity for TV viewers can help to further improve prediction accuracy by determining more similar group TV viewers because direct personality information expresses similarities between TV viewers' program interests to some extent.

In our experimentation analysis, Figure 12 shows that at the maximum threshold value for social contact (0.5), *ROPPSA* achieved the highest accuracy value (0.237) when compared with $T3$ (0.177) and $T4$ (0.062). *ROPPSA* therefore generated more effective and successful TV program recommendations (high accuracy) than $T3$ and $T4$.

Higher recall specifies that a recommender algorithm returned the most relevant and effective results (TV recommendations) and exhibited minimal false-negative errors. According to Figure 13, the recall values of *ROPPSA* are higher than those of $T3$ and $T4$. Specifically, at the highest Social Contact Score of 0.5, *ROPPSA* achieved a higher recall (0.778) than $T3$ (0.569) and $T4$ (0.515). Therefore, *ROPPSA* outperformed $T3$ and $T4$ in terms of recall and hence generated more relevant and successful social group recommendation of TV programs (high recall).

In order to verify the transparency and authenticity of our experimental results for accuracy and recall in Figures 12 and 13, we computed $F1$ using (18). A critical view of Figure 14 shows that *ROPPSA* attained higher results (0.363) in terms of $F1$ in comparison with $T3$ (0.269) and $T4$ (0.112). This correspondence is due to the efficacy of information retrieval with respect to how *ROPPSA* devoted more importance to both accuracy and recall in terms of social group TV program recommendation. Furthermore, the positive transparency of $F1$ in social group recommendations is due to the outperformance of *ROPPSA* in accuracy and recall evaluation metrics in comparison with $T3$ and $T4$.

Additionally, in relation to the arithmetic mean (AM) results of accuracy and recall using (19), *ROPPSA* achieved higher results in comparison with $T3$ and $T4$. As shown in Table 7 and Figure 15, in terms of social group recommendation, at the highest tie strength (0.5), we attained AM results of 0.507, 0.373, and 0.289 for *ROPPSA*, $T3$, and $T4$, respectively:

$$\mathrm{AM} = \frac{1}{2} \times (A + R). \tag{19}$$

Our approach proves that the AM results attained for *ROPPSA* were the highest in comparison with the $F1$ (harmonic mean) results, which accordingly corroborates that AM should always be higher than harmonic mean regarding the retrieval effectiveness of a recommender system/algorithm [52].

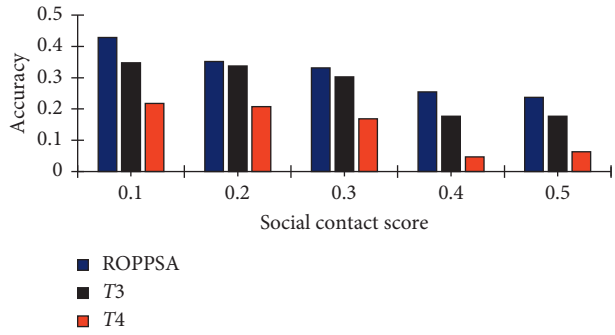


Figure 12: Performance evaluation for accuracy on the ATU dataset.

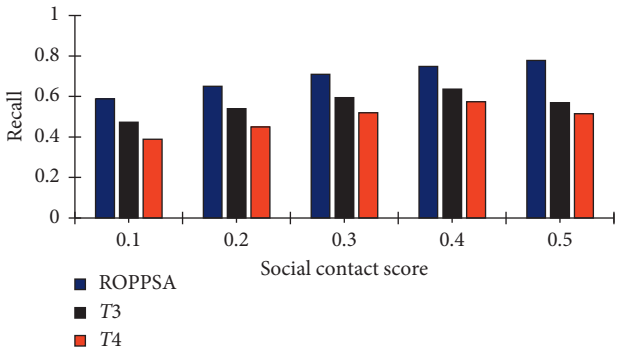


Figure 13: Performance evaluation for recall on the ATU dataset.

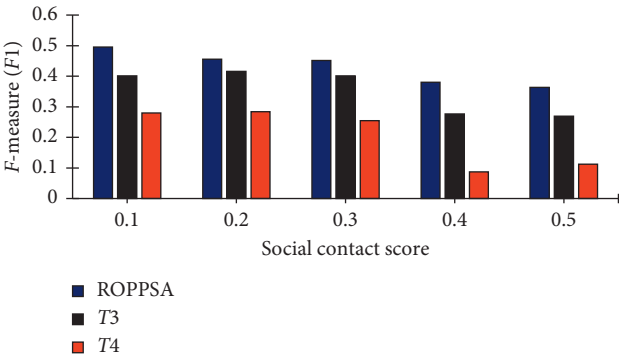


Figure 14: Performance evaluation for f-measure ($F1$) on the ATU dataset.

Table 6: RSME and MAE performance on the ATU dataset.

PSim	MAE performance			RMSE performance		
	*T*2	ROPPSA	*T*1	*T*2	ROPPSA	*T*1
0.8	0.696	**0.669**	0.830	0.690	**0.665**	0.826
0.9	0.823	**0.745**	0.953	0.819	**0.741**	0.949
1.0	0.824	**0.763**	0.937	0.820	**0.759**	0.933

4.4. Summary of Experimental Results. In summary, it can be verified from Figures 9–15, as well as Tables 6 and 7, that *ROPPSA* reliably attained more favourable results in all the utilized evaluation metrics. This observation indicates that our approach is more suitable and robust due to its capability

Table 7: Experimental comparison of AM and *F*1.

Method	Highest SCS	Accuracy	Recall	AM	*F*1
*T*3	0.5	0.177	0.569	0.373	0.269
ROPPSA	**0.5**	**0.237**	**0.778**	**0.507**	**0.363**
*T*4	0.5	0.062	0.515	0.289	0.112

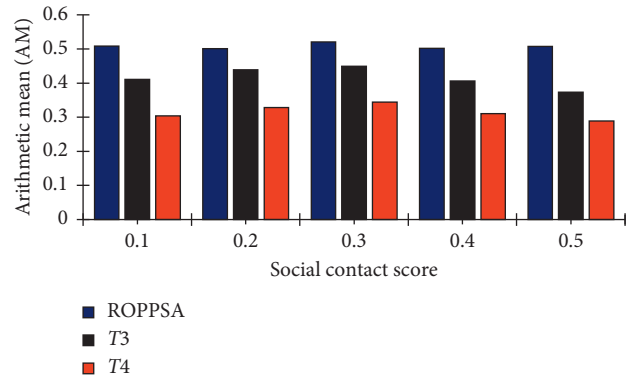


Figure 15: Performance evaluation for arithmetic mean (AM) on the ATU dataset.

to make use of tie strengths and personality trait ratings of TV viewers. Consequently, in comparison with the baseline methods, *ROPPSA* alleviated the challenges of cold-start and data sparsity due to the unlimited utilization of both personality and social features of TV viewers.

Our proposed *ROPPSA* recommender algorithm outperformed both *T*1 and *T*2 as well as *T*3 and *T*4 in terms of personality and social group recommendations, respectively. This exemplifies the importance of personality group profile normalization and social group relationships/influences in generating effective and efficient TV program recommendations for TV viewers.

5. Discussion and Concluding Remarks

With reference to our research findings and our proposed recommendation method (*ROPPSA*), it is extremely important and vital to note that, globally, the data protection law of every nation has an objective to protect individuals' fundamental rights to privacy and the protection of their personal data [53, 54]. Ghiglieri et al. [54] undertook an online survey which involved 200 participants to determine whether consumers were aware of smart TV-related privacy risks. The main findings in Ghiglieri et al. [54] depicted that generally participants are unwilling to disconnect the Internet from their smart TVs because they value the smart TV's Internet functionality more than their privacy [54].

In Germany, a sectoral investigation was upgraded into a concrete policy guidance for smart TV stakeholder/users in relation to privacy/data issues of television services in the nation's ecosystem. Additionally, the Netherlands has recently witnessed the implementation/enactment of three enforcement actions by the Dutch Data Protection Authority against leading providers of interactive television services [53]. These examples corroborate that the effective

implementation of our proposed *ROPPSA* method should be done in cognizance with the Data Protection Act/law of respective nations and such acts/laws should have content relating to privacy/personal data of smart TV stakeholder/users [53, 54].

Most existing TV program recommendation methods disregard the fact that people often exhibit similar personality traits and tie strengths in terms of different program interests or preferences. In this paper, we proposed an innovative recommendation method called *ROPPSA* for TV recommendation in social and personality groups. *ROPPSA* generates similar personalities and tie strengths among *TTV*s and TV viewers for improving group recommendation accuracy using two different procedures. We carried out an extensive experimentation procedure using a relevant ATU dataset. Experimental results show that our proposed *ROPPSA* method outperformed other relevant TV program recommendation methods in terms of recommendation accuracy and other utilized metrics.

In future, it is our aim to explore more social-based patterns pertaining to TV program recommendation by applying other social property features involving network centrality measures. Consequently, relevant TV viewers in such a method will be determined by analyzing their information pertaining to network centrality.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Materials

The dataset utilized in this research article involves relevant data collected from students in Accra Technical University (ATU), Ghana, and is an extension from a dataset used in [1]. Data were gathered from students in Higher National Diploma (HND) Marketing ($n = 2864$), HND Computer Science $n = 720$), and HND Building Technology ($n = 812$). Therefore, the ATU dataset utilized in this paper contains a total of 4,396 TV viewers (2,421 males representing 55.07% and 1,975 females representing 44.93%). The ATU dataset in this research article contains tie strength data which comprises computations of contact frequencies (Figure 1) and contact duration (Figure 2) of ATU students (users designated as TV viewers). The ATU dataset also comprises personality trait ratings (scale of 1 to 5) of all users (TV viewers) in accordance with the Big Five Personality Traits (BFPT). The total number of personality ratings for all the traits combined in the ATU dataset is 22,541. Figure 3 illustrates the personality trait data utilized in our ATU dataset for experimentation. Additionally, in order to ensure and substantiate recommendation accuracy in our experimentation procedure, we gathered TV program interest of TV viewers in accordance with the following: movie/TV series-horror, action, comedy, adventure, and drama; sports/games-football, basketball, hockey, boxing, and athletics. Tables 1 and 2 illustrate data regarding TV program interest of TV viewers in the ATU dataset in accordance with their ratings. Figure 1: ATU dataset—contact frequency trends. Figure 2: ATU dataset—contact duration trends. Figure 3: ATU dataset1—personality data. Table 1: ATU dataset—sports TV program ratings. Table 2: ATU dataset—movie/TV series program ratings. (*Supplementary Materials*)

References

[1] E. Kim, S. Pyo, E. Park, and M. Kim, "An automatic recommendation scheme of TV program contents for (IP)TV personalization," *IEEE Transactions on Broadcasting*, vol. 57, no. 3, pp. 674–684, 2011.

[2] S. Pyo, E. Kim, and M. Kim, "Automatic and personalized recommendation of TV program contents using sequential pattern mining for smart TV user interaction," *Multimedia Systems*, vol. 19, no. 6, pp. 527–542, 2013.

[3] S. Pyo, E. Kim, and M. Kim, "LDA-Based unified topic modeling for similar TV user grouping and TV program recommendation," *IEEE Transactions on Cybernetics*, vol. 45, no. 8, 2015.

[4] M. Bambia, R. Faiz, and M. Boughanem, "Context-awareness and viewer behavior prediction in social-TV recommender systems: survey and challenges," in *Communications in Computer and Information Science*, T. Morzy, P. Valduriez, and L. Bellatreche, Eds., vol. 539, pp. 52–59, Springer, Cham, Switzerland, 2015.

[5] D. Véras, T. Prota, A. Bispo, R. Prudêncio, and C. Ferraz, "A literature review of recommender systems in the television domain," *Expert Systems with Applications*, vol. 42, no. 22, pp. 9046–9076, 2015.

[6] S. Song, H. Moustafa, and H. Afifi, "Advanced IPTV services personalization through context-aware content recommendation," *IEEE Transactions on Multimedia*, vol. 14, no. 6, pp. 1528–1537, 2012.

[7] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, pp. 109–132.

[8] Z. Yu and X. Zhou, "TV3P: an adaptive assistant for personalized TV," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 393–399, 2004.

[9] A. Martinez, J. Arias, A. Vilas, J. Garcia Duque, and M. Lopez Nores, "What's on TV tonight? An efficient and effective personalized recommender system of TV programs," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 1, pp. 286–294, 2009.

[10] M. Bjelica, "Unobtrusive relevance feedback for personalized TV program guides," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 2, pp. 658–663, 2011.

[11] M. Krstic and M. Bjelica, "Context-aware personalized program guide based on neural network," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 4, p. 1301, 2012.

[12] K. Verbert, N. Manouselis, X. Ochoa et al., "Context-aware recommender systems for learning: a survey and future challenges," *IEEE Transactions on Learning Technologies*, vol. 5, no. 4, pp. 318–335, 2012.

[13] Y. Hu, Q. Peng, X. Hu, and R. Yang, "Time aware and data sparsity tolerant web service recommendation based on improved collaborative filtering," *IEEE Transactions on Services Computing*, vol. 8, no. 5, pp. 782–794, 2015.

[14] J. Zhu, J. Zhang, C. Zhang et al., "CHRS: cold start recommendation across multiple heterogeneous information networks," *IEEE Access*, vol. 5, pp. 15283–15299, 2017.

[15] F. Xia, N. Y. Asabere, H. Liu, Z. Chen, and W. Wang, "Socially aware conference participant recommendation with personality traits," *IEEE Systems Journal*, vol. 11, no. 4, pp. 2255–2266, 2017.

[16] N. Y. Asabere, A. Acakpovi, and M. B. Michael, "Improving socially-aware recommendation accuracy through personality," *IEEE Transactions on Affective Computing*, vol. 9, no. 3, pp. 351–361, 2017.

[17] N. Y. Asabere, F. Xia, Q. Meng, F. Li, and H. Liu, "Scholarly paper recommendation based on social awareness and folksonomy," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 30, no. 3, pp. 211–232, 2015.

[18] F. Xia, N. Y. Asabere, H. Liu, N. Deonauth, and F. Li, "Folksonomy based socially-aware recommendation of scholarly papers for conference participants," in *Proceedings of the 23rd International Conference on World Wide Web-WWW'14 Companion*, Seoul, Republic of Korea, April 2014.

[19] N. Y. Asabere, F. Xia, W. Wang, J. J. P. C. Rodrigues, F. Basso, and J. Ma, "Improving smart conference participation through socially aware recommendation," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 5, pp. 689–700, 2014.

[20] F. Xia, N. Y. Asabere, J. J. P. C. Rodrigues, F. Basso, N. Deonauth, and W. Wang, "Socially-aware venue recommendation for conference participants," in *Proceedings of the 110th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*, pp. 134–141, Vietri sul Mere, Italy, December 2013.

[21] N. Y. Asabere, B. Xu, A. Acakpovi, and N. Deonauth, "SARVE-2: exploiting social venue recommendation in the context of smart conferences," *IEEE Transactions on Emerging Topics in Computing*, 2018.

[22] M. Hornick and P. Tamayo, "Extending recommender systems for disjoint user/item sets: the conference recommendation problem," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 8, pp. 1478–1490, 2012.

[23] S. Sedhain, S. Sanner, D. Braziunas, L. Xie, and J. Christensen, "Social collaborative filtering for cold-start recommendations," in *Proceedings of the 8th ACM Conference on Recommender systems-RecSys'14*, pp. 345–348, New York, NY, USA, October 2014.

[24] S. Sedhain, A. Menon, S. Sanner, L. Xie, and D. Braziunas, "Low-rank linear cold-start recommendation from social data," in *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI-17)*, pp. 1502–1508, San Francisco, CA, USA, 2017.

[25] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 273–291, 2014.

[26] R. Hu and P. Pu, "Acceptance issues of personality-based recommender systems," in *Proceedings of the 3rd ACM Conference on Recommender Systems-RecSys'09*, pp. 221–224, New York, NY, USA, October 2009.

[27] M. Krstic and M. Bjelica, "Personalized program guide based on one-class classifier," *IEEE Transactions on Consumer Electronics*, vol. 62, no. 2, pp. 175–181, 2016.

[28] H.-J. Kwon and K.-S. Hong, "Personalized smart TV program recommender based on collaborative filtering and a novel similarity method," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1416–1423, 2011.

[29] M. Kompan and M. Bielikova, "Group recommendations: survey and perspectives," *Computing and Informatics*, vol. 33, no. 2, pp. 446–476, 2014.

[30] N. R. Kim, S. Oh, and J. H. Lee, "A television recommender system learning a user's time-aware watching patterns using quadratic programming," *Applied Sciences*, vol. 8, no. 8, pp. 1–14, 2018.

[31] M. U. Chaudhry, S. Oh, N. Kim, and J. H. Lee, "Heterogeneous information network based TV program recommendation," in *Proceedings of the 16th International Symposium on Advanced Intelligent Systems*, pp. 4–7, Mokpo, Republic of Korea, 2015.

[32] P. Cesar and D. Geerts, "Past, present, and future of social TV: a categorization," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC)*, pp. 347–351, Las Vegas, NV, USA, January 2011.

[33] X. Wang, L. Sun, Z. Wang, and D. Meng, "Group recommendation using external followee for social TV," in *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo*, pp. 37–42, Melbourne, Australia, July 2012.

[34] C. Shin and W. Woo, "Socially aware TV program recommender for multiple viewers," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 927–932, 2009.

[35] S. Shepstone, Z.-H. Tan, and S. Jensen, "Audio-based age and gender identification to enhance the recommendation of TV content," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 3, pp. 721–729, 2013.

[36] R. Sotelo, Y. Blanco-Fernandez, M. Lopez-Nores, A. Gil-Solla, and J. Pazos-Arias, "TV program recommendation for groups based on muldimensional TV-anytime classifications," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 1, pp. 248–256, 2009.

[37] A. Odic, M. Tkalcic, J. F. Tasic, and A. Kosir, "Personality and social context: impact on emotion induction from movies," in *Proceedings of the UMAP 2013*, Rome, Italy, June 2013.

[38] M. Elahi, M. Braunhofer, F. Ricci, and M. Tkalcic, "Personality-based active learning for collaborative filtering recommender systems," *Advances in Artificial Intelligence*, Springer, Berlin, Germany, pp. 360–371, 2013.

[39] B. Ferwerda, E. Yang, M. Schedl, and M. Tkalcic, "Personality traits predict music taxonomy preferences," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems-CHI EA'15*, pp. 2241–2246, Seoul, Republic of Korea, April 2015.

[40] A. Felfernig, L. Boratto, M. Stettinger, and M. Tkalčič, "Personality, emotions, and group dynamics," in *Group Recommender Systems. SpringerBriefs in Electrical and Computer Engineering*, Springer, Cham, Switzerland, 2018.

[41] M. Tkalčič, B. D. Carolis, M. de Gemmis, A. Odić, and A. Košir, "Introduction to emotions and personality in personalized systems," in *Emotions and Personality in Personalized Services. Human–Computer Interaction Series*, M. Tkalčič, B. De Carolis, M. de Gemmis et al., Eds., Springer, Cham, Switzerland, 2016.

[42] M. Tkalčič, "Emotions and personality in recommender systems," in *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 535–536, New York, NY, USA, 2018.

[43] L. Quijano-Sanchez, J. A. Recio-Garcia, and B. Diaz-Agudo, "Personality and social trust in group recommendations," in *Proceedings of the 22nd IEEE International Conference on Tools with Artificial Intelligence*, pp. 121–126, Arras, France, 2010.

[44] L. Quijano-Sanchez, J. A. Recio-Garcia, and B. Diaz-Agudo, "Using personality to create alliances in group recommender systems," in *Proceedings of the International Conference on Case-Based Reasoning*, Springer, London, UK, pp. 226–240, September 2011.

[45] L. Quijano-Sanchez, J. A. Recio-Garcia, and B. Diaz-Agudo, "Happymovie: a Facebook application for recommending movies to groups," in *Proceedings of the IEEE 23rd*

International Conference on Tools with Artificial Intelligence, pp. 239–244, Boca Raton, FL, USA, 2011.

[46] L. Quijano-Sanchez, J. A. Recio-Garcia, and B. Diaz-Agudo, "An architecture and functional description to integrate social behaviour knowledge into group recommender systems," *Applied Intelligence*, vol. 40, no. 4, pp. 732–748, 2014.

[47] J. A. Recio-García, L. Quijano, and B. Díaz-Agudo, "Including social factors in an argumentative model for group decision support systems," *Decision Support Systems*, vol. 56, pp. 48–55, 2013.

[48] J. A. Recio-Garcia, G. Jimenez-Diaz, A. A. Sanchez-Ruiz, and B. Diaz-Agudo, "Personality aware recommendations to groups," in *Proceedings of the Third ACM Conference on Recommender Systems-RecSys'09*, pp. 325–328, New York, NY, USA, October 2009.

[49] R. Prada, S. Ma, and M. A. Nunes, "Personality in social group dynamics," in *Proceedings of the 12th IEEE International Conference on Computational Science and Engineering*, pp. 607–612, Vancouver, Canada, August 2009.

[50] J. Wu, Y. Shi, and C. Guo, "A resource recommendation method based on user taste diffusion model in folksonomies," *International Journal of Knowledge and Systems Science*, vol. 3, no. 1, pp. 18–30, 2012.

[51] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," *Lecture Notes in Computer Science*, vol. 4231, pp. 291–324, Springer, Berlin, Germany, 2007.

[52] E. Baykan, M. Henzinger, L. Marian, and I. Weber, "A comprehensive study of features and algorithms for URL-based topic classification," *ACM Transactions on the Web*, vol. 5, no. 3, pp. 1–29, 2011.

[53] K. Irion and N. Helberger, "Smart TV and the online media sector: user privacy in view of changing market realities," *Telecommunications Policy*, vol. 41, no. 3, pp. 170–184, 2017.

[54] M. Ghiglieri, M. Volkamer, and K. Renaud, "Exploring consumers' attitudes of smart TV related privacy risks," in *Proceedings of the 5th International Conference on Human Aspects of Information Security, Privacy, and Trust*, pp. 656–674, Vancouver, Canada, July 2017.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 7069191, 11 pages
https://doi.org/10.1155/2020/7069191

Research Article

Research on the Technological Innovation Efficiency of China's Strategic Emerging Industries Based on SBM: NDEA Model and Big Data

Zuchang Zhong,[1] Fanchao Meng,[2] Yuanbing Zhu ,[3] and Gang Wang[2]

[1]*Guangdong Institute of International Strategy, Guangdong University of Foreign Studies, Guangzhou, China*
[2]*School of Business, Guangdong University of Foreign Studies, Guangzhou, China*
[3]*School of International Studies, Jinan University, Guangzhou, China*

Correspondence should be addressed to Yuanbing Zhu; zhuyuanbing@hotmail.com

Received 14 March 2020; Accepted 16 April 2020; Published 15 May 2020

Guest Editor: Wen-Tsao Pan

Characterized by large scale, variety, fast generation, and extremely high value but low density, big data can be used to mine effective information, provide users with auxiliary decision-making, and realize its own value. Based on the nonoriented SBM and the network DEA model, this paper systematically and objectively evaluates the technological innovation efficiency of strategic emerging industries in all provinces of China in 2002–2013. The study found the following. (1) The overall technological efficiency of China's strategic emerging industries is low. The average of comprehensive efficiency is 0.278; of 26 provinces, only 8 are above the average level. (2) The efficiency in the commercialization stage of scientific and technological achievements of strategic emerging industries in the whole country and most of the provinces is higher than that in the stage of knowledge innovation. The inefficiency of the knowledge innovation stage restricts the efficiency promotion of China's strategic emerging industries. (3) The overall innovation efficiency of strategic emerging industries has been increasing from 2002 to 2013. In comparison, the growth rate of pure technical efficiency is larger than that of scale efficiency. (4) The overall efficiency, the efficiency in the knowledge innovation stage, and the efficiency in the commercialization stage of scientific and technological achievements of the eastern region are higher than those of the central and western regions.

1. Introduction

After the international financial crisis in 2008, the world's major developed countries rerecognized the strategic significance of the real economy with the manufacturing industry as a mainstay, implemented the reindustrialization strategy, and pushed forward the development of strategic emerging industries with high technology, high risk, low carbon, and high research and development level to promote economic growth and achieve economic recovery. The strategic emerging industries have been regarded as the core industries to seize the commanding heights of the economy and the dominant force of a new round of international competition, and developing strategic emerging industries will certainly become the breakthrough for China's industrial structure optimization and upgrading and economic sustainable development. In recent years, the Chinese government has been attaching great importance to the development of strategic emerging industries. On October 18, 2010, the State Council issued the *Decision of the State Council on Accelerating the Fostering and Development of Strategic Emerging Industries*, formulated and proposed a series of plans to nurture and promote the development of strategic emerging industries, and defined seven major industrial areas: energy conservation and environmental protection, a new generation of information technology, new materials, new energy, biomedicine, high-end equipment manufacturing, and new energy automobiles. *Made in China 2025* issued by the State Council on May 20, 2015, clearly pointed out that through the government guidance

and integration of resources, five major projects including establishing a manufacturing innovation center, boosting intelligent manufacturing, strengthening industrial development at the grass-roots level, green manufacturing, and high-end equipment innovation, technological breakthroughs that constrains the development of manufacturing industries to enhance overall competitiveness of China's manufacturing and strategic emerging industries can be achieved. On December 19, 2016, the State Council published a *Guideline on Emerging Sectors of Strategic Importance During the 13th Five-Year Plan Period*, which adopted a comprehensive plan for development goals, key tasks, and policy measures of China's strategic emerging industries during the 13th five-year plan period. According to the guideline, by 2030, emerging sectors of strategic importance will become a leading force driving the sustained and healthy development of China's economy, China will become the world's important manufacturing center and innovation center of strategic emerging industries, and a number of innovative enterprises with global influence and leading role will spring up.

At present, China is the largest manufacturing economy in the world, and its manufacturing industry occupies a key position in the global manufacturing sector. Three economic regions, Yangtze River Delta, the Pearl River Delta, and Beijing-Tianjin-Hebei, have become competitive with industrial characteristics and the potential of becoming the world's leading industrial development base. However, the existing problem is that, in the labor system of international manufacturing sector, we have been lagging behind the global value chain controlled by the developed countries and in the low end of the global value chain for a long time [1, 2]. According to statistics, the average annual growth rate of R&D expenditure, full-time equivalent of research personnel, number of patent applications, and number of valid patents were 23.54%, 14.9%, 34.18%, and 40.13%, respectively, in 2012–2013, and the growth rate was the highest in the world, while the annual growth rate of sales of new products over the same period was 20.82%. This shows that China's patent application quantity and number of patents granted increased greatly, but at the same time many patents did not translate into the actual product sales. Therefore, compared with developed countries, the innovation efficiency of China's strategic emerging industries needs to be improved, especially the efficiency of transformation of scientific and technological achievements. Therefore, accurate and systematic assessment of the innovation efficiency of strategic innovation industries has important implications for optimizing the allocation of innovation resources, promoting strategic emerging industries to enhance the capacity of independent innovation and global value chain transitions, elevating domestic technical content of export products, and breaking the low-end lock of global value chain dilemma.

2. Literature Review

At present, there are two kinds of methods for evaluating the innovation efficiency of industries at home and abroad: one is the parametric method, which is mainly represented by stochastic frontier analysis (SFA); the other is nonparametric method, mainly represented by data envelopment analysis (DEA). In the study of innovation efficiency of strategic emerging industries, most scholars use DEA method. Raab et al. [3] used the input-oriented CCR model to evaluate the efficiency of high-tech industries in 50 states of the United States in 2002. Chen et al. [4] used the input-oriented CCR model to evaluate the comprehensive efficiency of six high-tech industries in Hsinchu Science Park, Taiwan, in 1991–1996. Lu et al. [5] used the DEA-Tobit model to evaluate the efficiency of 194 high-tech enterprises in Taiwan. Xiao and Xie [6] analyzed the influence of independent innovation, technology introduction, technological transformation, and digestive absorption on the innovation efficiency of China's strategic emerging industries and found that the independent innovation and technology introduction had significant impact on the innovation efficiency, and the impact of technological innovation and digestive absorption on innovation efficiency was not significant. Liu and Xia [7] estimated the innovation efficiency of 89 listed companies in 2007–2010 based on the SFA method and analyzed the influencing factors of innovation efficiency using the Tobit model. Li and Li [8] established stochastic frontier model (SFA) by using panel data of 10 Chinese LED listed companies in 2008–2010 as samples to calculate the innovation efficiency of LED strategic emerging industries and then put forward corresponding measures to improve innovation efficiency of China's emerging sectors. Lv and Sun [9] analyzed the technical efficiency and its influencing factors of the 19 categories of industries in China using SFA model. The results show that there are industrial heterogeneity and regional heterogeneity in the development of China's emerging industries. Huang and Zhang [10] adopted the Malmquist index decomposition method based on DEA model and estimated the Malmquist index of technical innovation in 28 provincial administrative regions and three major regions (eastern, central, and western regions) by using the provincial panel data of China in 2005–2012. Based on the index system in the Oslo manual, Zhang et al. [11] evaluated the innovation capacity of strategic emerging industries of 21 sample cities and explored the inherent evolution law and its application in the strategic emerging industries using gray fuzzy evaluation method and kernel density analysis. Liu et al. [12] adopted the DEA model to study the comprehensive efficiency, pure technical efficiency, and scale efficiency of China's strategic emerging industries using the panel data of 28 provinces and municipalities in China from 2007 to 2012. The results show that the pure technological efficiency of China's strategic emerging industries is at a low level, showing the process of rising first and then declining. Chen [13] used a three-stage DEA model to conduct a comparative study of its financing efficiency. The study found that the low financing efficiency of SMEs on the New Third Board was mainly affected by the low-scale efficiency caused by unreasonable input and output. Liu [14] measured the short-term and long-term innovation efficiency of five subsectors of strategic emerging

industries in Liaoning Province from 2000 to 2011 and discussed the impact of imitative innovation and independent innovation on innovation efficiency using DEA. Guo et al. [15] used the DEA model and the Malmquist index decomposition method to measure the technological efficiency, technological progress, and all factors in the process of high-tech industry innovation in 30 provincial administrative regions and three major regions (east, central, and western regions) of China's Productivity Growth and cluster analysis of regional high-tech innovation efficiency levels. It is found that, since 2009, the efficiency of technological innovation in China's regional high-tech industries has generally shown an upward trend, but it has been subject to large fluctuations. Zhang et al. [16] analyzed the achievements and efficiency of scientific and technological innovation in 31 provinces and cities of China and used the improved two-stage dynamic network DEA model to divide the innovation behavior of each region after the release of scientific and technological innovation policy into two stages. The first stage is from innovation input to intermediate innovation output stage, and the second stage is from intermediate output to achievement transformation stage, obtaining two-stage efficiency scores. Sun et al. [17] believed that there is a significant spatial correlation in the process of upgrading the urban industrial structure in China; the "U-shaped" relationship between the rationalization of the industrial structure and the urban total factor productivity; the performance between the advanced industrial structure and the urban total factor productivity out of the "inverted *U*" relationship. Han et al. [18] constructed a two-stage input-output index system of high-tech enterprise technology innovation based on the characteristics of high-tech enterprise technology innovation and considering the time lag effect, applied the dynamic two-stage DEA model to evaluate the performance of the two stages, and comprehensively analyzed the efficiency of high-tech enterprises in each province in the two stages of technology R&D and technology transformation. Zeng et al. [19] used the data envelopment analysis CCR model to calculate the comprehensive efficiency value under the condition of constant scale returns, and the BCC model calculates the pure technical efficiency value and the scale efficiency value under the conditions of variable scale returns. The efficiency is analyzed, and the trend of production efficiency is analyzed through the Malmquist model. Li et al. [20] used DEA model and Malmquist index decomposition method to measure the technological efficiency, technological progress, and total factor productivity growth in the process of high-tech industry innovation in 30 provincial administrative regions and three major regions of China by using the interprovincial R&D panel data of China's high-tech industry from 2009 to 2016 and made cluster analysis on the efficiency level of regional high-tech innovation. Guo and Li's [21] research believed that enterprise independent innovation and technology credit can have a significant positive impact on manufacturing innovation efficiency, and there are significant regional differences in the impact of technology finance on manufacturing innovation efficiency.

Based on research results of innovation efficiency of strategic emerging industries by domestic and foreign scholars, this paper argues that there are two shortcomings in the following two aspects. First, it focuses on the evaluation of single-stage innovation efficiency and the influencing factors. Most literatures mainly consider the innovation process of strategic emerging industries as a black box, regardless of internal structure and internal operating mechanism of the industrial innovation system. Second, when using DEA method to evaluate the innovation efficiency of strategic emerging industries, the traditional input- or output-oriented CCR and BCC models are mainly adopted, which ignore the information of slack variable; that is, the improvement is not reflected in the measurement of innovation efficiency value, which may cause a deviation in the efficiency measure. At the same time, the overall comparison of the efficiency of decision-making unit becomes more difficult when the input factor is increased and the slack output is considered. In order to make up for the shortcomings of the existing literature, this study divides the innovation system into the stage of knowledge innovation and the stage of commercialization of scientific and technological achievements. The regional differences and variation trends of innovation efficiency of China's strategic emerging industries are investigated using provincial panel data with network DEA (NDEA) model proposed by Tone et al. [22] and SBM model to accurately study innovation of strategic emerging industries and to provide references for regional innovation policy adjustment.

3. Two-Stage NDEA Model of Strategic Emerging Industries Based on SBM

3.1. Two-Stage Chain Process Division of Technological Innovation of Strategic Emerging Industries. The process of technological innovation is a process of transforming knowledge, skills, and materials into customer-satisfaction products. It is the evolution of knowledge generation, creation, and application. From the perspective of value chain, the process of enterprise technological innovation involves a series of value-added processes such as innovation resource input, innovation idea generation, research and development, technical management and organization, engineering design and manufacturing, user participation, and marketing. In the process of innovation, these value-added activities are interrelated and sometimes are operated in parallel. The entire technical innovation system is a community of the interaction between technology and economic activities. Specifically, in the process of technological innovation, once the new technology or technique is invented, new products and new more advanced production mode will be created through new technology, new process applications, and industrialization process to meet customer needs, so that the innovation resources are value-added. In the application of new technology, new technique, and industrialization process, the innovation system will timely transfer the customer's expectations, changes in demand, technical trends, and others to the development and application of new technologies and techniques through its

own feedback mechanism, thus promoting continuous improvement of technological innovation system.

Based on the above analysis, innovation output includes scientific and technological achievements and economic benefits based on the industry level. Scientific and technological achievements (technological output), as an intermediate output, are both the result of upfront innovation resource investment and the premise of later-stage new product development and technology improvement. Therefore, the process of technological innovation forms two innovative stages linked by technological output and input. The first stage is the stage of knowledge innovation from resource input to technological output, which produces intellectual output mainly through R&D funds and personnel input, etc., such as patented technology and nonpatented technology; the second stage is the commercialization stage of scientific and technological achievements from the technical output to the economic benefits, which improve business sales income, especially the sales of new products, and then the economic benefits mainly through new product development and new technological transformation. The process of the two stages is shown in Figure 1.

3.2. Two-Stage DEA Model considering Slack Variables. Data envelopment analysis (DEA) is a method of empirically measuring productive efficiency of decision-making units (or DMUs) of multiple inputs and outputs introduced by Charnes et al. in [23]. SBM model is the DEA model based on the slack variable measurement proposed by Tone [24]. It belongs to the nonradial DEA model. Compared with the traditional radial DEA models, CCR and BCC models, SBM model also considers the slackness of input and output. The noneffective degree of DMUs is measured by the average proportion of input (output) reduction (increase). The DMUs in the model do not have to follow the ray direction proportionately, so SBM model can improve the input and output to the maximum extent, but SBM model still considers the DMUs as a black-box system in the process of efficiency evaluation, and the estimated efficiency value does not reflect the internal structure and operation mechanism of the evaluation object system. Fare and Grosskopf [25] have laid the foundation for the development of the network DEA model and constructed the network DEA framework model. The model shows two basic characteristics of the chain network system: first, it contains two or more subsystems; second, the subsystems are linked by intermediate variables. Since then, many scholars such as Lewis et al. [26], Kao and Hwang [27], and Tone and Tsutsui [22] improved models on the basis of the model. Among them, Tone and Tsutsui [22] proposed the network DEA model taking slacks into account, that is, SBM-NDEA (Slacks-Based Measure Network DEA) model, which effectively combines SBM model and the traditional network DEA model, ensuring the accuracy of the efficiency measurement. The SBM-NDEA model is expressed as follows.

Given n decision-making units DMUj ($j = 1, 2, \ldots, n$), each decision-making unit consists of K nodes, and m_k and r_k refer to the numbers of input indicators and output indicators, respectively; $X_j^k = (X_{1j}^k, X_{2j}^k, \ldots, X_{m_k j}^k)^T$ represents the input vector of the DMUi at node k, $Y_j^k = (Y_{1j}^k, Y_{2j}^k, \ldots, Y_{m_k j}^k)^T$ represents the output vector of the DMUi at node k, $Z_j^{(k,h)} = (Z_{1j}^{(k,h)}, Z_{2j}^{(k,h)}, \cdots, Z_{t(k,h)j}^{(k,h)})^T$ represents the middle vector of the DMUi from node k to node h, (k, h) indicates from node k to node h, t (k, h) represents the intermediate index number of (k, h), L stands for a collection of intermediate indicators, and the production possibility set of $\{(X^k, Y^k, Z^{(k,h)})\}$ under CRS is expressed as

$$\begin{aligned} X^k &\geq \sum_{j=1}^{n} X_j^k \lambda_j^k \quad (k = 1, \ldots, K), \\ Y^k &\leq \sum_{j=1}^{n} Y_j^k \lambda_j^k \quad (k = 1, \ldots, K), \\ Z^{(k,h)} &\geq \sum_{j=1}^{n} Z_j^{(k,h)} \lambda_j^{(k,h)} \quad (\text{as the output of } k), \\ Z^{(k,h)} &\geq \sum_{j=1}^{n} Z_j^{(k,h)} \lambda_j^{(k,h)} \quad (\text{as the input of } h). \end{aligned} \tag{1}$$

For VRS, add constraint condition $\sum_{j=1}^{n} \lambda_j^k = 1$ ($\lambda_j^k \geq 0$, $k = 1, \ldots, K$).

DMU_{jo} ($j0 = 1, 2, \ldots, n$), (X_o, Y_o), $s_0^{k-} = (s_{1o}^{k-}, s_{2o}^{k-}, \ldots, s_{m_k o}^{k-})^T$ represents the input slack vector of DMU_{jo} at node k, $s_o^{k+} = (s_{1o}^{k+}, s_{2o}^{k+}, \ldots, s_{m_k o}^{k+})^T$ corresponding input-output meets

$$\begin{aligned} x_o^k &= \sum_{j=1}^{n} x_j^k \lambda_j^k + s_o^{k-} \quad (k = 1, \ldots, K), \\ y_o^k &= \sum_{j=1}^{n} y_j^k \lambda_j^k - s_o^{k+} \quad (k = 1, \ldots, K), \\ s_o^{k-} &\geq 0, s_o^{k+} \geq 0, \lambda^k \geq 0 \quad (k = 1, \ldots, K). \end{aligned} \tag{2}$$

The intermediate index under slack network DEA model satisfies the following formula:

$$z_o^{f(k,h)} = \sum_{j=1}^{n} z_j^{f(k,h)} \lambda_j^k + s_o^{f(k,h)} \quad (k = 1, \ldots, K), \tag{3}$$

where $\sum_{j=1}^{n} z_j^{f(k,h)} \lambda_j^k = \sum_{j=1}^{n} z_j^{f(k,h)} \lambda_j^h$ indicates that the intermediate variable satisfies the output of the previous node equal to the input of the next node. $Z_j^{f(k,h)} = (Z_{1j}^{f(k,h)}, Z_{2j}^{f(k,h)}, \ldots, Z_{t(k,h)j}^{f(k,h)})^T$ $(k = 1, \ldots, K)$ stands for free intermediate index, f (k, h) stands for free intermediate index number, and $s_o^{f(k,h)} = (s_{1o}^{f(k,h)}, s_{2o}^{f(k,h)}, \cdots, s_{t(k,h)}^{f(k,h)})^T$ indicates the slack vector of (k, h) free intermediate index and is unrestraint.

Definition 1. θ^* refers to overall efficiency of DMU_{jo}. When $\theta^* = 1$, DMU_{jo} is valid; when $\theta^* < 1$, DMU_{jo} is invalid. The overall efficiency of DMU_{jo} under different orientations is

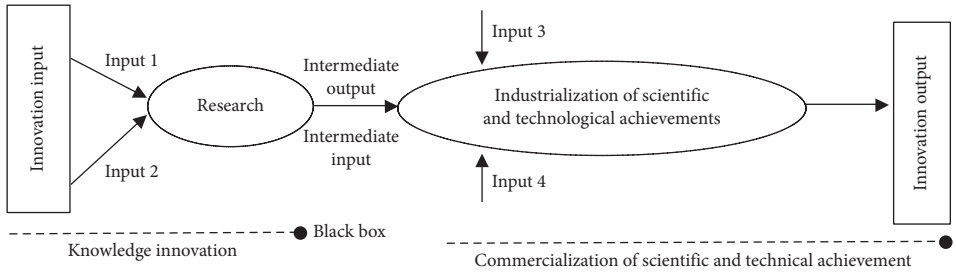


Figure 1: Value chain process of technical innovation activities in two stages.

$$\begin{aligned}
&\text{Input-oriented,} \quad \theta^* = \min_{\lambda^k, s^{k-}} \sum_{k=1}^{k} w^k \left[1 - \frac{1}{m_k}\left(\sum_{i=1}^{m_k} \frac{s_i^{k-}}{x_{io}^k}\right)\right],\\
&\text{Output-oriented,} \quad \theta^* = \min_{\lambda^k, s^{k+}} \frac{1}{\sum_{k=1}^{k} w^k \left[1 + 1/r_k\left(\sum_{r=1}^{r_k}\left(s_r^{k+}/y_{ro}^k\right)\right)\right]},\\
&\text{Non-oriented,} \quad \theta^* = \min_{\lambda^k, s^{k-}, s^{k+}} \frac{\sum_{k=1}^{k} w^k \left[1 - 1/m_k\left(\sum_{i=1}^{m_k}\left(s_i^{k-}/x_{io}^k\right)\right)\right]}{\sum_{k=1}^{k} w^k \left[1 + 1/r_k\left(\sum_{r=1}^{r_k}\left(s_r^{k+}/y_{ro}^k\right)\right)\right]},
\end{aligned} \tag{4}$$

where $1 - 1/m_k\left(\sum_{i=1}^{m_k} s_i^{k-}/x_{io}^k\right)$ indicates the average reduction rate of the DMU at node k; $[1 + 1/r_k\left(\sum_{r=1}^{r_k} s_r^{k+}/y_{ro}^k\right)]$ represents the average increase rate of node k, w^k 为 refers to the weight of node k, and $\sum_{k=1}^{k} w^k = 1$.

Definition 2. s_i^{k-*} and s_r^{k+*} are slacks of input and output, respectively, when the optimum efficiency is assigned for DMU_{jo} under different orientations, and θ^k refers to the efficiency of node k. Under different orientations, the efficiency of DMU_{jo} at node k is

$$\begin{aligned}
&\text{Input-oriented,} \quad \theta^k = 1 - \frac{1}{m_k}\left(\sum_{i=1}^{m_k} \frac{s_i^{k-*}}{x_{io}^k}\right),\\
&\text{Output-oriented,} \quad \theta^k = \frac{1}{1 + 1/r_k\left(\sum_{r=1}^{r_k}\left(s_r^{k+*}/y_{ro}^k\right)\right)},\\
&\text{Nonoriented,} \quad \theta^k = \frac{1 - 1/m_k\left(\sum_{i=1}^{m_k}\left(s_i^{k-*}/x_{io}^k\right)\right)}{1 + 1/r_k\left(\sum_{r=1}^{r_k}\left(s_r^{k+*}/y_{ro}^k\right)\right)}.
\end{aligned} \tag{5}$$

When $\theta^k = 1$, DMU_{jo} is valid at node k; when $\theta^k < 1$, DMU_{jo} is invalid at node k.

As defined by Definitions 1 and 2, the necessary and sufficient conditions for the overall effectiveness of DMU_{jo} are that all nodes are valid.

3.3. Setting of Input and Output Indicators

3.3.1. Input and Output Indicators of Knowledge Innovation Stage. As the first stage of technological innovation activities, knowledge innovation refers to new knowledge and new technology developed mainly through R&D activities. From the perspective of input indicators, scholars usually choose R&D investment and R&D personnel investment. The R&D funds are expressed by R&D internal fund expenditure and development costs of new products; and R&D personnel are expressed by full-time equivalent of research personnel. Specifically, we use the R&D capital stock and the new product development capital stock index to replace the indicator. For the measurement of R&D capital stock, the perpetual inventory method is used. 2002 is intended to be the base period, the 2002 R&D stock is R&D internal expenditure of strategic emerging industries in 2002 divided by the depreciation rate and average growth rate of several years after the base period, that is, $k_{2002} = I_{2002}/(\delta + \varsigma)$, where I_{2002} is traffic data of R&D internal expenditure of China's strategic emerging industries, and δ refers to the depreciation rate. Referring to the results of most of the researches, the depreciation rate is set to be 15%; ζ refers to average growth rate of R&D internal expenditure of China's strategic emerging industries. Based on this, it is calculated using perpetual inventory method, that is, R&D stock in t year = R&D stock in t-1 year X (1–15%) + R&D traffic in t year.

In terms of output index, the output of knowledge innovation stage mainly presents as patent and nonpatented technology. Because the nonpatented technology is enterprise business secret and is difficult to be accurately measured, the patented technology is major consideration. The patent application quantity and effective patents are taken as two indicators to measure the output of the knowledge innovation phase, which is consistent with most of the studies.

3.3.2. Input and Output Indicators of Commercialization Stage of Scientific and Technological Achievements. The commercialization stage of scientific and technological achievements is the second stage of technological innovation. The main purpose of this stage is to transform the first-stage intellectual output into economic benefit. In terms of input indicators, we take technology, capital, and personnel into consideration. As the two stages of technological innovation are closely related, the technical input mainly comes from the first stage of output; the number of patent applications and the number of valid invention patents are both output of the first stage and input of the second stage, so they are the common intermediate variables of the two stages of technological innovation. In terms of capital, the capital expenditure of technological transformation is taken as the capital investment in the commercialization of scientific and technological achievements. In terms of personnel investment, in consideration of important role of scientific and

technical personnel in the application and popularization of new technologies, scientific and technical personnel are taken as personnel investment at this stage. In terms of output indicators, the output at this stage is mainly reflected in the economic benefits. Learning from the practice of Peng et al. [28], the new product sales revenue and exports are selected to reflect the results of technological innovation. See Table 1 for input-output assessment indicator system.

3.4. Sample and Data Selection. According to the *China National Economy Classification Codes* (GB/T4754-2012), strategic emerging sectors are divided into the following industries: chemical manufacturing; traditional Chinese medicinal materials, and Chinese patent medicine processing industry; biological products manufacturing; aircraft manufacturing and services; spacecraft manufacturing; communication equipment manufacturing; radar and corollary equipment manufacturing; radio and television equipment manufacturing; electronic components manufacturing; electronic device manufacturing; home audio-visual equipment manufacturing; other electronic equipment manufacturing; electronic computer manufacturing; office equipment manufacturing; medical equipment and device manufacturing; and instrumentation manufacturing industry.

This paper is based on the data of strategic emerging industries of 26 provinces (municipalities and autonomous regions in China), and data in Tibet, Xinjiang, Hainan, Inner Mongolia, Qinghai, Hong Kong, Macao, and Taiwan are missing and eliminated. During the period from 2002 to 2013, due to the hysteretic nature of the two stages of technological innovation, referring to the research results of Guan and Liu [29], the final lag period of R&D investment is set to be 2 years; that is, there are a one-year lag period between innovation resources investment and scientific and technological achievements and a one-year lag period of the commercialization of scientific and technological achievements. Therefore, a lag of one year is adopted. The first-stage innovation resource investment, intermediate output and input, the final output use data of year t, year t+1, and year t+2, respectively, that is, input and output data of technological innovation of the first stage, are from 2002–2011 and 2003–2012, and the input and output data of the second stage are from 2003–2012 and 2004–2013, respectively. All data are derived from the 2002–2014 *China Statistical Yearbook on Science and Technology*, *China Statistical Yearbook on High Technology Industry*, *China Statistical Yearbook*, and provincial statistical yearbook.

4. Analysis and Discussion of Empirical Results

Based on the SBM-NDEA model of Variable Return to Scale (VRS), this paper uses the software MaxDEA6.4 to analyze the technological innovation efficiency of the strategic emerging industries in 26 provinces in China in 2002–2003. The results are shown in Tables 2 and 3. Based on the results of innovation efficiency, this paper will analyze the overall technical efficiency characteristics of strategic emerging industries, the regional differences of innovation efficiency and the convergence analysis, and the two-stage technological innovation efficiency.

4.1. Feature Analysis of Technological Innovation Efficiency. Nationwide, the overall technological efficiency of China's strategic emerging industries is low, the average overall efficiency is 0.278, and the minimum is 0.041 (Heilongjiang). The top three provinces are Tianjin (0.878), Beijing (0.778), and Shanghai (0.716) successively. Of 26 provinces in the study, only 8 are better than average, accounting for 30.76%; most of the provinces' overall efficiency of innovation is lower than the national average, indicating that most of provinces' innovation-driven development level is still low and resource utilization in the process of innovation is low, and the phenomenon of wasting resources is very serious. Further research shows that the pure technical efficiency value of most provinces is smaller than the scale efficiency value; the former is 0.488, and the latter is 0.576, which shows that the improvement of technological innovation efficiency of China's strategic growth industries is mainly dependent on scale economic effects, and the contribution of pure technical efficiency is relatively small.

In terms of the efficiency of the two-stage innovation, the average overall efficiency value of the 26 provinces in the knowledge innovation stage is 0.247, and the efficiency value of 18 provinces is below the average efficiency value, accounting for 69.24%. The top three are Tianjin (0.757), Beijing (0.677), and Guangdong (0.588), and the efficiency value of 17 provinces of knowledge innovation is less than 0.2. This shows that, in the period of knowledge innovation, the overall efficiency of China strategic emerging industries is low, the waste of innovation resources is serious, and the gap between innovation resource allocation and cultivation of innovative talents is increasing. It has restricted the improvement of overall innovation efficiency of China's strategic innovation industries. At the commercialization stage of scientific and technological achievements, the average overall efficiency value of the 26 provinces in China is 0.435, 11 provinces of which are higher than the national average. The top three provinces are Tianjin (1), Jiangsu (0.934), and Shanghai (0.897); Heilongjiang (0.129) is the lowest. It is noteworthy that the efficiency of scientific and technological achievements of Tianjin in 2002–2013 is effective, much higher than other provinces. In general, the efficiency of the commercialization state of scientific and technological achievements is higher than the efficiency of the knowledge innovation stage. This may be due to the fact that the Chinese government has made a substantial investment in the innovation resources in recent years, and the number of patent applications increased significantly, but the number of effective invention patents is lower than the number of patent applications. In addition, with the implementation of China's innovation-driven strategy, some provinces, especially in the eastern region, have established specialized agencies serving transformation of sci-tech achievements drawing on the advanced experience of developed countries, which effectively promoted the

Table 1: Assessment indicator system of two-stage technological innovation efficiency of strategic emerging industries.

Initial innovation input (**X**)	Fund input	R&D internal expenditure: billion yuan (X_1^1)
		New product development expenditure: billion yuan (X_2^1)
	Personnel input	Full-time equivalent of R&D personnel: (X_3^1)
Technological output (**MY**)	Knowledge innovation	Patent application quantity: (Y_1^1 or X_1^2)
		Effective patent application: (Y_2^1 or X_2^2)
Innovation resource input (**MX**)	Technological commercialization	Expenditure on technological transformation: yuan (X_3^2)
		Number of scientific and technical personnel (X_4^2)
Final output (**Y**)	Economic benefits	Sales of new products: yuan (Y_1^2)
		Exports: yuan (Y_2^2)

X_i^1 and Y_i^1 refer to input and output indicators of knowledge innovation at note 1; similarly, X_i^2 and Y_i^2 refer to input and output indicators of commercialization of scientific and technological achievements at node 2. At the same time, because technological output is both the output index of the first stage and the input index of the second stage, it has two marks.

Table 2: Comparative analysis of average efficiency value of different regions in 2002–2013.

Regions	Overall efficiency			Knowledge innovation stage			Commercialization stage of scientific and technological achievements		
	Technical efficiency	Pure technical efficiency	Scale efficiency	Technical efficiency	Pure technical efficiency	Scale efficiency	Technical efficiency	Pure technical efficiency	Scale efficiency
Anhui	0.141	0.239	0.508	0.155	0.233	0.556	0.327	0.592	0.597
Beijing	0.778	0.822	0.944	0.677	0.762	0.876	0.895	0.899	0.992
Chongqing	0.346	0.417	0.664	0.309	0.377	0.677	0.527	0.683	0.710
Fujian	0.451	0.516	0.882	0.395	0.486	0.824	0.862	0.878	0.977
Gansu	0.081	0.632	0.128	0.081	0.647	0.128	0.208	0.929	0.249
Guangdong	0.644	1.000	0.644	0.588	1.000	0.588	0.735	1.000	0.735
Guangxi	0.162	0.618	0.225	0.146	0.612	0.198	0.241	0.814	0.279
Guizhou	0.077	0.267	0.295	0.065	0.259	0.349	0.141	0.337	0.392
Hebei	0.096	0.179	0.485	0.091	0.154	0.506	0.205	0.425	0.529
Henan	0.201	0.272	0.636	0.191	0.277	0.624	0.327	0.440	0.676
Heilongjiang	0.041	0.127	0.296	0.033	0.087	0.355	0.129	0.419	0.345
Hubei	0.143	0.191	0.674	0.128	0.159	0.704	0.256	0.335	0.729
Hunan	0.173	0.273	0.563	0.172	0.248	0.595	0.283	0.481	0.612
Jilin	0.129	0.338	0.329	0.133	0.323	0.376	0.204	0.492	0.381
Jiangsu	0.527	0.894	0.604	0.418	0.804	0.563	0.934	1.000	0.934
Jiangxi	0.108	0.174	0.553	0.108	0.162	0.583	0.295	0.482	0.605
Liaoning	0.197	0.244	0.793	0.154	0.175	0.862	0.469	0.541	0.855
Ningxia	0.114	0.990	0.114	0.115	0.988	0.115	0.318	1.000	0.318
Shangdong	0.261	0.294	0.902	0.239	0.285	0.834	0.495	0.528	0.952
Shanxi	0.389	0.991	0.394	0.389	1.000	0.389	0.465	0.991	0.470
Shaanxi	0.058	0.082	0.685	0.057	0.067	0.762	0.201	0.248	0.794
Shanghai	0.716	0.907	0.791	0.547	0.815	0.689	0.897	1.000	0.897
Sichuan	0.176	0.208	0.847	0.155	0.195	0.818	0.442	0.465	0.948
Tianjin	0.878	0.912	0.963	0.757	0.825	0.912	1.000	1.000	1.000
Yunnan	0.165	0.895	0.186	0.176	0.889	0.197	0.156	0.877	0.181
Zhejiang	0.168	0.194	0.868	0.150	0.183	0.816	0.298	0.308	0.972
Average	0.278	0.488	0.576	0.247	0.462	0.573	0.435	0.660	0.659

commercialization of scientific and technological achievements.

Table 3 shows the results of technological innovation efficiency of the strategic growth industries in 2002–2003. It can be seen from Table 3 that the overall innovation efficiency of strategic emerging industries has been on the rise from 2002 to 2013, increasing from 0.194 in 2002 to 0.311 in 2011, and reached the highest value of 0.436 in 2010. In contrast, the growth rate of pure technical efficiency is greater than that of scale efficiency, which shows that the development of China's strategic emerging industries has gradually evolved from the traditional low-end development relying on scale advantages to integrate into global value chain, and then embed in global value chain, ultimately leading to the development of global value chain, and the technical level of strategic emerging industries has been greatly improved. The overall efficiency of the knowledge innovation phase is greater than the commercialization stage of scientific and technological achievements, which depends largely on the rapid growth of pure technical efficiency in the knowledge innovation phase. The efficiency value of the overall efficiency, the knowledge innovation stage, or the

Table 3: Comparative analysis of average innovation efficiency in 2002–2013.

Year	Overall efficiency			Knowledge innovation stage			Commercialization stage of scientific and technological achievements		
	Technical efficiency	Pure technical efficiency	Scale efficiency	Technical efficiency	Pure technical efficiency	Scale efficiency	Technical efficiency	Pure technical efficiency	Scale efficiency
2002–2004	0.194	0.394	0.510	0.168	0.379	0.515	0.405	0.660	0.584
2003–2005	0.179	0.417	0.452	0.151	0.387	0.428	0.374	0.620	0.633
2004–2006	0.191	0.417	0.458	0.131	0.375	0.409	0.337	0.573	0.564
2005–2007	0.172	0.416	0.453	0.163	0.424	0.461	0.296	0.652	0.458
2006–2008	0.248	0.457	0.543	0.185	0.417	0.430	0.384	0.734	0.526
2007–2009	0.314	0.523	0.614	0.312	0.480	0.732	0.507	0.698	0.738
2008–2010	0.313	0.512	0.660	0.327	0.525	0.724	0.460	0.627	0.762
2009–2011	0.419	0.585	0.736	0.389	0.562	0.739	0.613	0.715	0.853
2010–2012	0.436	0.602	0.746	0.393	0.569	0.743	0.594	0.738	0.804
2011–2013	0.311	0.553	0.586	0.252	0.500	0.547	0.380	0.584	0.664

commercialization stage of scientific and technological achievements decreased sharply after 2011. It may be related to the sluggish recovery of world economy after international financial crisis and implementation of the "reindustrialization" strategy by the developed countries.

4.2. Regional Difference and Convergence Analysis of Technological Innovation Efficiency. In order to examine the differences in innovation efficiency among regions, 26 provinces in China are divided into three regions: the eastern provinces include Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, and Guangdong; the central region includes 8 provinces: Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hubei, and Hunan provinces; and the western region includes Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, and Ningxia. From the results of innovation efficiency in subregions of Table 4, there is a large gap among the overall efficiency and efficiency in the knowledge innovation stage and the commercialization stage of scientific and technological achievements in the eastern, central, and western regions. The overall efficiency and efficiency in the knowledge innovation stage and the commercialization stage of scientific and technological achievements in the eastern region are higher than those of the central and western regions, indicating that the eastern region occupies the high-end segments of the domestic value chain, while the central and western regions are still in the low-end position. In both stages of knowledge innovation and the transformation of scientific and technological achievements, the pure technical efficiency in the western region is higher than that in the central region, which mainly benefited from continuous advancement of the western development strategy. The strategic emerging industries in the eastern region have been transferred to the western region, which has driven up the technological level of the entire western region. However, the scale efficiency of the western region in the stages of knowledge innovation and the transformation of scientific and technological achievements is lower than those in the eastern and central regions, which indicates that the scale effect of the strategic emerging industries in the western region has not yet appeared and the scale of industrial development should be further expanded. Chinese government should increase input in strategic emerging industries and constantly enhance the status of the industries in the domestic value chain. The central region should increase the use of innovation resources to reduce the waste of innovation resources, integrate more into the domestic value chain and the global value chain, and continuously upgrade the level of technological development of industrial sectors.

In order to further examine the variation trend of innovation efficiency gap, a convergence test is carried out. According to Barro et al. [30], convergence includes σ-convergence and β-convergence, and β-convergence includes absolute β-convergence and conditional β-convergence. σ-convergence means that the dispersion of real per capita income tends to decline over time, and β-convergence means that poor economies grow faster than rich economies. If the per capita income or output of each economy can achieve exactly the same steady-state level, it is absolute β-convergence, but if each economy approaches different steady-state level, it should be conditional β-convergence. We only take σ-convergence as an example in this paper. σ-convergence generally reflects the variation trend of the gap through the standard deviation or variable coefficient of national or regional level. Figure 2 gives variable coefficients of SBM innovation efficiency. It can be seen from Figure 2 that the variable coefficient of the innovation efficiency of strategic emerging industries among the provinces shows a significant downward trend, whether it is the overall efficiency, efficiency value of the knowledge innovation stage, or efficiency value of the transformation stage of the scientific and technological achievements. It indicates that σ-convergence exists and the gap of innovation efficiency among provinces is gradually narrowing. This also reflects the fact that, with the acceleration of international economic integration and regional economic integration, strategic emerging industries of the world and eastern regions have shifted to the central and western regions, and interregional economic activities and innovation activities have been more closely linked. The positive overflow effect of innovation

Table 4: Comparison of technological innovation efficiency value of strategic emerging industries in different regions.

Regions	Overall efficiency			Knowledge innovation stage			Commercialization stage of scientific and technological achievements		
	Technical efficiency	Pure technical efficiency	Scale efficiency	Technical efficiency	Pure technical efficiency	Scale efficiency	Technical efficiency	Pure technical efficiency	Scale efficiency
Central region	0.472	0.596	0.788	0.402	0.549	0.747	0.679	0.758	0.884
Eastern region	0.166	0.326	0.494	0.164	0.311	0.523	0.286	0.529	0.552
Western region	0.147	0.514	0.393	0.138	0.504	0.406	0.279	0.669	0.484
National average	0.278	0.488	0.576	0.247	0.462	0.573	0.435	0.660	0.659

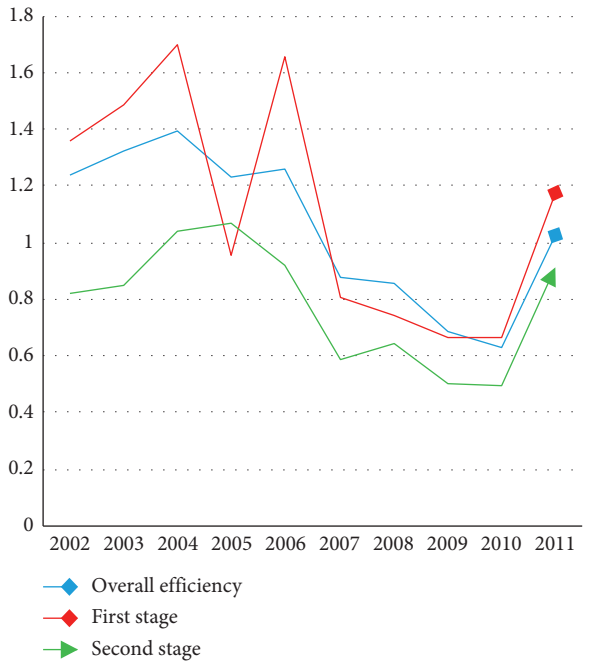


Figure 2: Variable coefficient trend of innovation efficiency of strategic emerging industries in 2002–2011.

Figure 3: Distribution diagram of technical efficiency in two stages of provinces and cities in 2002–2013.

activities among provinces is becoming more and more obvious. The central and western regions are improving technical level of strategic emerging industries by opening wider to the outside world, optimizing the industrial innovation environment, and introducing and absorbing advanced technologies.

4.3. Analysis of Technological Innovation Efficiency in Two Stages. In terms of innovation performance of strategic emerging industries in different stages, 26 provinces can be divided into four areas based on average efficiency. Class *A* area is an area of highly efficient and intensive technological innovation, which is mainly characterized by high efficiency of knowledge innovation and transformation of high-tech achievements. Such area boasts high level of economic development with good innovation environment and certain policy advantages and thus attracts high-end innovation resources and elements. Industries in this area are at medium and high positions of the domestic value chain or global value chain. Class *B* area is characterized by high efficiency of knowledge innovation and low efficiency of transformation of high-tech achievements. This region has abundant innovative talents, but lacks effective industrial base and market environment. Class *C* area is characterized by low innovation efficiency and high efficiency of transformation of high-tech achievements, which is difficult to attract high-end innovation elements due to restricts of innovation environment, innovation policies, and industrial base, leading to inefficient knowledge innovation. Class *D* area is featured by extensive growth and inefficient technological innovation, which is characterized by low knowledge innovation efficiency and inefficient transformation of scientific and technological achievements. The innovation resource allocation is inefficient, and it also lacks good environment for innovation and commercialization of scientific and technological achievements in this area. Its industries are also at the lowermost position of domestic value chain and the global value chain. As shown in Figure 3, Class *A* area mainly consists of Tianjin, Beijing, Guangdong, Shanghai, Jiangsu, Fujian, Shanxi, and Chongqing, and most of these areas belong to the eastern developed regions; and no provinces belong to the Class *B* area, which demonstrates that basic research and innovation ability of China's strategic emerging industries is weak but also shows a more serious polarization phenomenon; area *C* consists of Shandong,

Sichuan, and Liaoning provinces, and the remaining provinces belong to Class *D* area. It indicates that the strategic emerging industries of the majority of China's provinces are still at the low end of the domestic value chain and global value chain with extensive industrial development model, and the competitiveness needs to be improved.

5. Conclusions and Suggestions

This paper divides the technological innovation process of strategic emerging industries into knowledge innovation stage and commercialization stage of scientific and technological achievements. Based on the nonoriented SBM model and the network DEA model, this paper systematically and objectively evaluates the technological innovation efficiency of strategic emerging industries in all provinces of China in 2002–2013. The study found the following. (1) The overall technological efficiency of China's strategic emerging industries is low. The average of comprehensive efficiency is 0.278 and the minimum is 0.041 (Heilongjiang). The top three provinces are Tianjin, Beijing, and Shanghai. Of 26 provinces in the study, only 8 are better than average, accounting for 30.76%. (2) From the efficiency estimation results of the two-stage innovation, the efficiency in the commercialization stage of scientific and technological achievements of strategic emerging industries in the whole country and most of the provinces is higher than that in the stage of knowledge innovation. This shows that, in the period of knowledge innovation, the overall efficiency of China strategic emerging industries is low, the waste of innovation resources is serious, and the gap between innovation resource allocation and cultivation of innovative talents is increasing. It has restricted the improvement of overall innovation efficiency of China's strategic innovation industries. (3) By stages, the overall innovation efficiency of strategic emerging industries has been increasing from 2002 to 2013. In comparison, the growth rate of pure technical efficiency is larger than that of scale efficiency. (4) Throughout the region, the overall efficiency, the efficiency in the knowledge innovation stage, and the efficiency in the commercialization stage of scientific and technological achievements of the eastern region are higher than those of the central and western regions. The variable coefficient of the innovation efficiency of strategic emerging industries among the provinces shows a significant downward trend.

Based on the above research conclusions, this article puts forward the following policy recommendations. (1) To improve the efficiency of technological innovation in strategic emerging industries, we should focus on advantageous resources. From the perspective of the cost-benefit ratio of resources, we should rationally design resource allocation, improve resource use efficiency, and comprehensively promote diffusion and transformation of technological innovations. (2) Strengthening cross-provincial horizontal communication and cooperation and the establishment of interprovincial information and resource exchange and interoperability platforms all contribute to the promotion of innovative activities and innovative exchanges between provinces and cities, which is fundamental to solve the problem of insufficient intermediate output sexual measures. (3) Focus on the promotion of superior experience and promote balanced development between regions. Comprehensively promote the experience of policy support, development concepts, technology development, and production models in advantageous regions. Through technical cooperation, experience dissemination, industrial alliances, personnel training, etc., promote the development of regions with low technological innovation efficiency and comprehensively enhance China's high technology. (4) The succession among policies should be strengthened, so that the latter policies become the booster of the previous policies, and then the performance of China's regional science and technology innovation policies should be improved.

Data Availability

The data used to support the results of this study have copyright issues and so cannot be made freely available. Requests for access to these data should be made to Zuchang Zhong, zhongzuc@163.com.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was supported by the National Natural Science Foundation of China (Grants nos. 71673064 and 71974039), National Natural Science Foundation of Guangdong (Grant no. 2019A1515011475), innovation team project (Humanities and Social Sciences) of Universities in Guangdong (Grant no. 2017WCXTD003), and Regional and National Projects of Guangdong University of Foreign Studies in 2018 (Grant no. 6).

References

[1] Z. Liu and J. Zhang, "The formation, breakthroughs and countermeasures of captive network in developing countries under global OEM systems," *Chinese Industrial Economy*, no. 5, pp. 39–47, 2007.

[2] Z. Liu, "Transformation and upgrading and countermeasure of export-oriented economy in the eastern coastal region of China," *China Economic Studies*, no. 1, pp. 15–22, 2010.

[3] R. A. Raab and P. Kotamraju, "The efficiency of the high-tech economy: conventional development indexes versus a performance index," *Journal of Regional Science*, vol. 46, no. 3, pp. 545–562, 2006.

[4] C.-J. Chen, H.-L. Wu, and B.-W. Lin, "Evaluating the development of high-tech industries: taiwan's science park," *Technological Forecasting and Social Change*, vol. 73, no. 4, pp. 452–465, 2006.

[5] Y. H. Lu, C. C. Shen, C. T. Ting et al., "Research and development in productivity measurement: an empirical investigation of the high technology industry," *African Journal of Business Management*, vol. 4, no. 13, pp. 2871–2884, 2010.

[6] X. Xiao and Xie l, "An empirical analysis of innovation efficiency of China's strategic emerging industries," *Economic Management*, vol. 33, no. 11, pp. 26–35, 2011.

[7] Z. Y. Liu and Y. M. Xia, "Innovation efficiency and impact factors of China's strategic emerging industries," *Journal of Emerging Trends in Economic and Management Sciences*, vol. 3, no. 5, pp. 547–552, 2012.

[8] H. Li and S. Li, "Study on innovation efficiency evaluation of strategic emerging industries - an empirical analysis of LED," *Industry Journal of Central University of Finance and Economics*, no. 4, pp. 75–80, 2013.

[9] Y. Lv and H. Sun, "Technical efficiency and its influencing factors of China's strategic emerging industries," *The Journal of Quantitative and Technical Economics*, no. 1, pp. 128–143, 2014.

[10] H. Huang and Z. Zhang, "Technological innovation efficiency of china's strategic emerging industries- dea-based malmquist exponential index," *Model Technical Economy*, vol. 34, no. 1, pp. 21–28, 2015.

[11] Z. Zhang, J. Pan, and L. Peng, "Evaluation, evolution and rules of innovation ability of strategic emerging industries," *Science Research Management*, vol. 36, no. 3, pp. 1–11, 2015.

[12] H. Liu, Y. Liu, Q. Han, and Y. Hu, "Research on the efficiency of technological innovation in China's strategic emerging industries," *Systems Engineering-Theory and Practice*, vol. 35, no. 9, pp. 2296–2303, 2015.

[13] S. Chen, "Research on SMEs 'financing efficiency in the new Third board market——research on directional issuance based on three-stage DEA model," *Audit and Economic Research*, vol. 32, no. 3, pp. 78–86, 2017.

[14] Y. Liu, "Empirical study on innovation efficiency of strategic emerging industries- analysis based on DEA method," *Mathematics in Practice and Theory*, vol. 46, no. 10, pp. 16–25, 2016.

[15] S. Guo, M. Tong, J. Guo, and Y. Han, "Inter-provincial real environmental efficiency measurement and influencing factors analysis based on three-stage DEA model," *China Population · Resources and Environment*, vol. 28, no. 3, pp. 106–116, 2018.

[16] Y. Zhang, H. Ying, and B. Yan, "Research on the evaluation of regional innovation input-output and the improvement path of science and technology innovation policy performance based on Two-stage DEA Model -- Analysis Based on science and technology innovation policy information," *Information journal*, vol. 37, no. 1, pp. 198–207, 2018.

[17] X. Sun, Z. Wang, and G. Zhang, "Structural dividends and spatial spillover effects in the improvement of total factor productivity," *Economic Review*, no. 3, pp. 46–58, 2018.

[18] B. Han, Y. Su, L. Tong, and M. Wan, "Research on technological innovation performance of high-tech enterprises based on Two-stage DEA," *Scientific research management*, vol. 39, no. 3, pp. 11–19, 2018.

[19] Y. Zeng, L. Cai, W. Sun, G. Zhang, B. Lin, and Y. Fang, "Analysis of the operation efficiency of public hospitals in China based on the DEA model," *China Health Statistics*, vol. 35, no. 1, pp. 47–51, 2018.

[20] P. Li, L. Tong, and Y. Liu, "Research on evaluation of regional high-tech industry innovation efficiency based on DEA and malmquist index," *Industrial Technology and Economy*, vol. 38, no. 1, pp. 27–34, 2019.

[21] Y. Guo and H. Li, "Empirical research on the impact of investment in science and technology finance on manufacturing innovation efficiency——based on Chinese provincial panel data," *Industrial Technology and Economy*, vol. 38, no. 2, pp. 29–35, 2019.

[22] K. Tone, M. Tsutsui, and D. E. A. Network, "Network DEA: a slacks-based measure approach," *European Journal of Operational Research*, vol. 197, no. 1, pp. 243–252, 2009.

[23] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, pp. 429–444, 1978.

[24] K. Tone, "A slacks-based measure of efficiency in data envelopment analysis," *European Journal of operational Research*, vol. 130, no. 3, pp. 498–509, 2001.

[25] R. Fare and S. Grosskopf, "Network DEA," *Socio-Economic Planning Sciences*, vol. 34, no. 1, pp. 35–49, 2000.

[26] H. Lewis and T. Sexton, "Network D E A.: Efficiency analysis of organizations with complex internal structure," *Computers and Operations Research*, vol. 9, pp. 1365–1410, 2004.

[27] C. Kao and S.-N. Hwang, "Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan," *European Journal of Operational Research*, vol. 185, no. 1, pp. 418–429, 2008.

[28] W. Peng and Y. Yongze, "The output efficiency and influencing factors of China's national high-tech zones," *Journal of Xi'an Jiaotong University (Social Science Edition)*, vol. 30, no. 5, pp. 16–23, 2010.

[29] J. Guan and S. Liu, "Comparing regional innovative capacities of PR China based on data analysis of the national patents," *International Journal of Technology Management*, vol. 32, no. 3/4, pp. 225–245, 2005.

[30] R. J. Barro, Xavier S.-I. Martin, O. J. Blanchard, and R. E. Hall, "Convergence across states and regions," *Brookings Papers on Economics Activity*, vol. 1991, no. 1, pp. 107–182, 1991.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 5609272, 17 pages
https://doi.org/10.1155/2020/5609272

Research Article

Anticipating Corporate Financial Performance from CEO Letters Utilizing Sentiment Analysis

Siqi Che,[1] **Wenzhong Zhu**,[2] **and Xuepei Li**[3]

[1]*School of English for International Business, Guangdong University of Foreign Studies, Guangzhou 510420, China*
[2]*School of Business, Guangdong University of Foreign Studies, Guangzhou 510420, China*
[3]*School of Economics and Trade, Guangdong University of Foreign Studies, Guangzhou 510420, China*

Correspondence should be addressed to Wenzhong Zhu; 743106832@qq.com

Received 15 January 2020; Accepted 19 February 2020; Published 13 May 2020

Guest Editor: Weilin Xiao

With the emergence and tremendous growth of text mining, a computer-assisted approach for capturing sentiment viewpoints from textual data is gradually becoming a promising field, particularly when researchers are increasingly facing the problem of filtering bunches of useless information without capturing the essence in the big data era. This study aims at observing and classifying the sentiment orientation in CEO letters, digging the main corporate social responsibility (CSR) themes, and examining the effectiveness of CEO letters' sentiment on forecasting financial performance. A specific sentiment dictionary has been proposed to identify and classify the sentiment orientation in CEO letters by utilizing the appraisal theory. Additionally, the qualitative data analysis software NVivo is applied to explore the CSR topics. Furthermore, a modified Altman's Z-score model and machine-learning approach are employed to predict financial performance. The results of preliminary evaluations validate that approximately 62.14% of the texts represent positive polarity even when companies are not in a promising economic situation. The CSR themes mainly focus on business ethical responsibility, particularly ethical activities. Among various machine-learning approaches, the logistic regression approach is appropriate for predicting financial performance with the state-of-the-art accuracy of 70.46 %. The encouraging results indicate that the sentiment information inCEO letters is a vital factor for anticipating financial performance. This work not only offers a new analytic framework for associating linguistic theory with computer science and economic models but will also improve stakeholders' decision-making.

1. Introduction

Due to researchers' unrivalled and explosive expansion in data mining, big data, and artificial intelligence, natural language processing (NLP) in handling bunches of textual data becomes an explosive and prevalent field with great future prospects. The high-tech novel technique of sentiment analysis offers a more efficient and accurate way for text processing, and its amazing pace of innovation, low costs, and scalability make it a highly attractive and alternative approach.

Sentiment analysis, also known as "opinion/sentiment mining" or "subjectivity analysis," uncovers a prominent interdisciplinary field of mixing computational linguistics and computer science, which attempts to extract subjective opinions, feelings, and attitudes contained in the text and analyzes how to use language to deliver subjectivity and viewpoints on a particular topic [1].

Correspondingly, corporate social responsibility report (CSRR), containing abundant sentiments, is crucial for reflecting companies' sustainable standpoints on its operating ideas, strategies, and methods. For this reason, CSRR can be a valuable source for investors' decision-making. For instance, CEO letter contains the ideas that corporations credibly desire to convey some information about themselves to their potential stakeholders, and these ideas may be important in determining investment and lending decisions because they cover valuable background information about interpreting and explaining financial performance, which may not be covered in companies' financial statements [2]. This obviously boosts the requirements of sentiment analysis and turns into extremely precious resources for

stakeholders' decision-making [3]. China, as one of the biggest developing countries in the world, has developed CSR rapidly, especially the Chinese government has paid much more attention to sustainable development. However, fewer researchers have exerted empirical evidence on the sentiment analysis of CSR in China, particularly in the case of information asymmetry.

To our knowledge, prior studies have proposed manifold sentiment analysis techniques towards extracting useful content from massive social media [4, 5]; some scholars have categorized sentences into positive/negative [4], while other researchers have divided sentences into critical/emotional-based on researchers' intuition. The taxonomy results, however, are inconsistent and incomparable as various self-designed frameworks have been applied in diverse studies. Although sentiment analysis of CSR or nonfinancial information appears to be a possible trend for predicting financial performance and assisting investors to make future investment decisions, little research has been done in this field.

As an attempt to make up the deficiency in the above research field, this paper is divided into three dimensions: firstly, in the microlevel analysis, it determines the sentiment polarity and sentiment attributes in letters to shareholders by utilizing appraisal theory; secondly, in the mesolevel analysis, it identifies the prominent CSR themes by employing NVivo 12 plus software; thirdly, in the macrolevel analysis, it explores the sentiment elements so as to anticipate financial performance by using the expression of Z-score.

The following parts of the paper are organized as follows: in the part of literature review, the related CSR, CEO letters, primary sentiment analysis approaches, appraisal theory, and forecasting financial performance through textual information will be reviewed; the part of research methodology will introduce the annotation study and the experimental procedures; in the part of experimental results and analysis, the experimental results will be stated and analyzed; in the conclusion part, the research results, contribution, limitations, and suggestions will be provided.

2. Literature Review

2.1. CEO Letters in CSR Report. The CEO letter (hereafter, shareholders' letter or letter to shareholders) is the most widely read part of the CSRR and a valuable communication source for stakeholders' decision-making. Generally, letter to shareholders is widely regarded as a promotional genre, which tries to provide companies' subjective sentiment standpoints and aims to portray a positive image [6] such as what stakeholders desired to know about identifying the last year's performance, tracking important events, displaying corporate social responsibility, and predicting the future vision of the companies. Thus, the letter stands in a vital important position to deliver company's competitive advantage [7]. Although a number of scholars recognized the significance of CEO letters, for instance, Kohut and Segars [7] characterize the effectiveness of letter to shareholders and how this information can benefit the company. Patelli and Pedrini [8] indicate that, even under the tough economic conditions, companies still sincerely disclose nonfinancial information with stakeholders. Surprisingly, little research has focused on how companies try to construct their images and showed to readers through sentiment analysis.

2.2. Sentiment Analysis. In artificial intelligence (AI), text mining is an effective and efficient way to process a large number of textual contents through extracting the sentiment polarity based on natural language processing. In particular, sentiment analysis has gradually become a popular technique. A considerable amount of literature has concentrated on analyzing documents' sentiment polarity [9]. Automatic sentiment classification has been extensively applied to reviews of products [10], movies [11], books [12], shopping [13], social networks [14], and students' course evaluation comments [15, 16], which is a common application of classifying positive or negative reviews. Most recent work has involved in extracting the textual information in the financial reports, as the text may contain more information than the numerical part in an annual report [17, 18]. The final results indicate that sentiment analysis is an important technique for forecasting financial performance and, thus, can be used to support the decision-making process of potential stakeholders [19, 20]. Researchers, therefore, have realized the potential value of the textual information analysis of financial reports for predicting financial performance and managing risks.

In fact, the challenge of detecting sentiment from text has been tackled from various perspectives. Nonetheless, previous approaches to spot affect have been categorized into two main approaches: lexicon-based approach and machine-learning approach [21], which are explicitly depicted in Figure 1.

2.2.1. Lexicon-Based Approach. Lexicon-based methods are rule-based, requiring a predefined word list and polarity to identify viewpoints and sentiments [4, 22], containing computing orientation for a text from the semantic orientation of words or phrases [23]. In other words, when a new text has been selected, the words inside the text have to match with the words in the sentiment dictionary and then various algorithms have been employed to aggregate values. The total aggregation of positive and negative values of the words assembles the sentiment orientation of the whole text. An idealized operating mechanism is specified in Figure 2.

Table 1 enumerates some related work in the lexicon-based approach for sentiment analysis and illustrates the various types of objectives along with the associated models used and the experimental results produced. We employed precision, recall, and F-measure as evaluation metrics; they are common in information retrieval and document classification research. Precision measures the number of correctly classified items out of the total classified by a classification technique, and recall measures the amount of correctly classified items of those manually classified as the gold standard. The F-measure is the harmonic mean of precision and recall, which offers a better measure than the arithmetic mean of precision and recall.

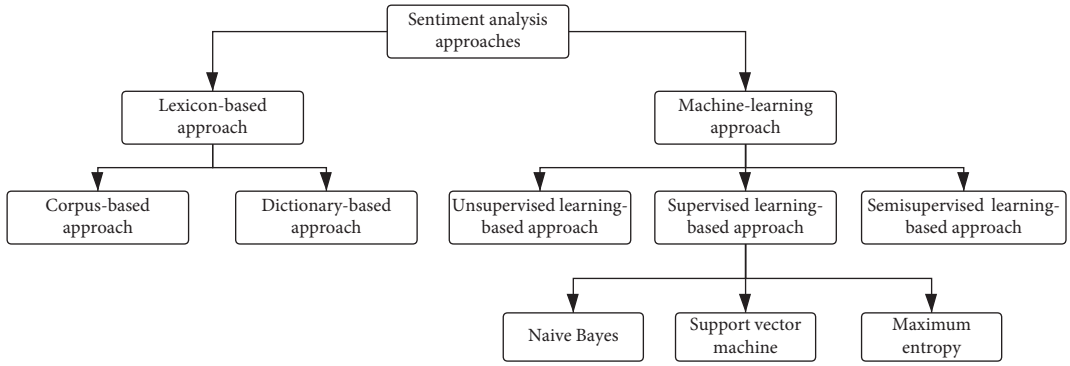


Figure 1: Sentiment classification methods.

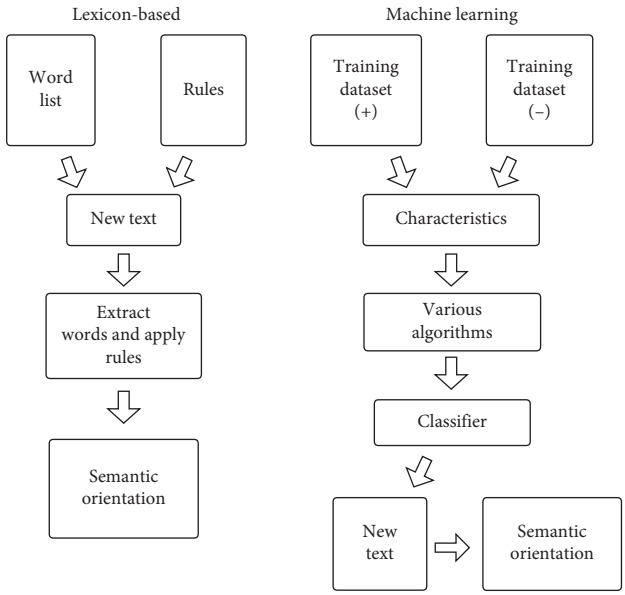


Figure 2: Machine-learning and lexicon-based approaches of sentiment analysis.

The merits of employing lexicon-based approach are that, across diverse domains, lexicon-based methods do not require to alter dictionaries [4]. Similarly, Brooke, Tofiloski, and Taboada [29] claimed that, compared with machine-learning, rule-based approach is not a cumbersome task because scholars do not need to take a large amount of time and effort to annotate the training data in a specific domain.

The lexicon-based approach, however, has its own limitations. Sometimes, the individual word extraction may miss important meanings in the text. Since the existing sentiment dictionary cannot meet the specific context of letter to shareholders, for example, in the dictionary, words like *lower* and *decrease* simply do not have a negative meaning in the CSR. Hence, in order to alleviate the dispute, sentiment identification requires a more comprehensive and proprietary sentiment dictionary for the specific CSR domain.

2.2.2. Machine-Learning Approach. The machine-learning approaches are presented with training classifiers, such as support vector machine (SVM), Naive-Bayes (NB), logistic regression (LR), and maximum entropy (ME), to assess the contents' positive or negative orientations based on an initial coding ten test upon the dataset to see if the sentiment is indeed captured [30]. To put it another way, human annotators code a small sample of dataset and then the machine takes over once it "learns" what kinds of words resemble these positive and negative sentiments [31]. A detailed operating mechanism is specified in Figure 2.

Generally speaking, the machine-learning approach can be further divided into supervised, unsupervised, and semisupervised [32]. Normally, the training objective is to be able to classify or distinguish instances. The main difference is that the supervised machine learning-based approach selects labeled instances to construct the model. The unsupervised approach is used for data mining to spot what is inside the unlabeled data. Semisupervised technique is halfway between supervised and unsupervised learning. This technique is about to use unlabeled data to learn a function improving the classification performance.

Table 2 summarizes a great diversity of relevant literature on machine-learning approach for sentiment identification, illustrating that most existing studies analyze text by creating a large dataset to measure subjectivity.

The advantages of the machine-learning approach are that, once the trained dataset is available, the classifier can rapidly determine the text's polarity. For instance, Yang et al. [35] applied a naive Bayes classifier and class association rules to identify the sentiment of consumer reviews and proved to achieve a satisfactory result. Additionally, Troussas et al. [36] classified the people's feelings and attitudes about certain topics and stated how sentiment analysis using naive Bayes can efficiently assist language learning. Specifically, Yuan et al. [39] manifested that the classification accuracy has been improved extremely based on the sentiment analysis of CSR in financial reports through the SVM approach. This provides the inspiration and basis for the further research of this study.

While the machine-learning approach has long been used in topical text classification with good results because of the efficiency and accuracy, it sometimes encounters with shortcomings as a result of high-level human manual intervention. That is to say, the classifiers necessitate sizable human annotated datasets for training and testing, especially amid the big data era, which is extremely costly and time-consuming

Table 1: Related work of lexicon-based approach in sentiment analysis.

Authors	Objectives	Models	Data source	Evaluation method	Data set	Accuracy (percent)	Precision (percent)	Recall	F_1
Hatzivassiloglou and Mckeown [24]	Assign adjective ±	Nonhierarchical clustering	WSJ corpus	N/A	657 adj (+) 679 adj (−)	78.1–92.4	N/A	N/A	N/A
Turney [23]	Assign docs sentiments	PMI-IR	Automobile, bank, movie, travel reviews	N/A	240 (+) 170 (−)	65.8–84	N/A	N/A	N/A
Turney and Littman [25]	Assign docs sentiments	PMI LSA	AV-ENG corpus AV-CA corpus TASA corpus	N/A	1614 (+) 1982 (−)	82.8–95	N/A	N/A	N/A
Taboada et al. [26]	Assign adjectives ±	SO-PMI	Reviews	N/A	521 adj	49.5–56.75	N/A	N/A	N/A
Ding et al. [27]	Assign adjectives, adverbs, verbs, and nouns ±	Opinion Observer	445 customer reviews of products from amazon.com	Macroaveraging (macroaveraging means given a set of confusion tables, a set of values are generated, and each value represents the precision or recall of an automatic classifier for each category)	N/A	N/A	91	90	90
		No context dependency				N/A	92	83	87
		Without using equation				N/A	90	85	87
		FBS				N/A	92	74	82
Taboada et al. [4]	Assign adjectives, verbs, adverbs, nouns ±	SO-CAL	Reviews	Across various domains	N/A	80	N/A	N/A	N/A
Dey et al. [28]	Assign docs sentiments	SO-CAL, Senti-N-Gram	Movie, books, cars, cookware, phones, hotels, music, and computers reviews	3-fold cross validation	Books	68	66	76	71
					Cars	84	76	100	86
					Computers	80	73	96	83
					Cookware	74	68	92	78
					Hotels	74	67	96	79
					Movies	66	64	72	68
					Music	64	62	72	67
					Phones	86	85	88	86

[31]. Moreover, sometimes, the training datasets are unavailable. Previous attempts at categorizing movie reviews and book reviews used the machine-learning approach with limited success. Because the technique may not suitable for various texts, facing the obstacle of domain specificity, the accuracy of analysis has reduced greatly because the thorough custom-made training dataset may not generalize well to texts from other domains [40, 41]. Consequently, in this study, the particular characteristics should be captured to perfectly satisfy the specific domain of letters to shareholders.

2.3. Forecasting Financial Performance by Using Text Information. Recently, sentiment analysis has been widely explored in understanding the relationship between text information and corporate financial performance. Through evaluating authors' opinions, attitudes, and sentiment polarity, the future financial performance could be predicted. A great number of scholars have recently engaged in analyzing textual information in annual reports, newspapers, and other documents to forecast stock return or financial performance. Table 3 lists some literature to indicate that the sentiment of documents can significantly correlate with financial performance.

From the above literature, our study will try to use letters to shareholders in CSRR to test whether the sentiment in the letter can predict the financial performance or not by employing various kinds of machine-learning methods. Furthermore, in order to have a better comparison of the sentiment attribute categorization and construct a suitable

Table 2: Related work of machine-learning approach in sentiment analysis.

Authors	Objectives	Models	Data source	Evaluation method	Data set	Accuracy (percent)	Precision (percent)	Recall	F_1
Bruce and Wiebe [33]	Assign sentences subjectivity using 1 to 4 scale	LCA	14 articles in Wall Street Journal Treebank Corpus	10-fold cross validation	486 subjective, 515 objective	72.17	N/A	N/A	N/A
Pang et al. [30]	Assign docs sentiments	SVM, NB, ME	Movie reviews	3-fold cross validation	700 (+) 700 (−)	77–82.9	N/A	N/A	N/A
Bo and Lee [1]	Assign docs sentiments	SVM, NB	Movie reviews	10-fold cross validation	1000 (+) 1000 (−)	86.4–87.2	N/A	N/A	N/A
Gamon and Michael [34]	Assign docs sentiments using 4-point scale	SVM	Customer feedback	10-fold cross validation	N/A	77.5	N/A	N/A	N/A
				10-fold cross validation	N/A	69.5	N/A	N/A	N/A
Yang et al. [35]	Assign topic sentiments	Class association rules	Consumer reviews	Macroaveraged Microaveraged (microaveraging means given a set of confusion tables, a new two-by-two contingency table is generated, each cell in the new table represents the sum of the number of documents from within the set of tables)	N/A	N/A	67.64	77.5	72.26
						N/A	67.47	74.4	72.26
		NB		Macroaveraged		N/A	73.96	66.79	70.19
				Microaveraged		N/A	68.3	66.11	67.19
Troussas et al. [36]	Assign docs sentiments	NB	Facebook 7000 status updates from 90 users	3-fold cross validation	1131 (+) 1131 (−)	N/A	77	68	72
Singh et al. [37]	Assign docs sentiments	NB	Product review, movie review	4-fold cross validation	N/A	45.6	83.1	N/A	81.2
		J-48				96.7	87.7	N/A	91.7
		BFTree				89.2	88.3	N/A	72.1
		OneR				90	91.3	N/A	97
Rathi [38]	Assign adjectives ±	SVM + KNN, hybrid	Tweet	N/A	267 (+) 264 (−) 168 (0)	76.17	N/A	N/A	N/A

Table 3: Related work of forecasting financial performance using textual information.

Authors	Material	Period	Methods	Key findings
Loughran and Mcdonald [42]	Annual report	1994–2008	Sentiment dictionary	Textual analysis can contribute to the ability to understand the impact of information on stock returns. Most importantly, researchers should be cautious when relying on word classification scheme derived outside the domain of business usage.
Schumaker et al. [43]	Financial news articles	2005	Machine learning	The financial news article prediction system can predict price increase or decrease in sentiment analysis with the accuracy of approximately 50%.
Hájek and Olej [44]	520 U.S. companies' annual report	2010	Machine learning	Sentiment information in annual reports can significantly improve the accuracy of the used classifiers.
Xin et al. [45]	U.S. manufacturing industry annual report	1997–2003	Machine learning	Using machine-learning approach employed with the results of text analysis can predict next year's earnings per share.
Hájek et al. [46]	448 U.S. companies' annual report	2008	Machine learning; neural networks	Sentiment information is an important determinant for forecasting financial performance.
Tong et al. [47]	Microblogs in China	2016.2–2016.9	Machine learning	There is a strong correlation between chat room postsentiment and stock price movement.

sentiment classification dictionary for the specific context, appraisal theory will be applied and the seed dictionary will be manually annotated for the domain of CEO letters.

2.4. Appraisal Theory. Although machine-learning methods show a better performance, the results of each experiment are not comparable because of various kinds of categorizations. A detailed theoretical framework is highly required to handle this difficulty. Researchers are asked to address more challenging tasks and attempt to perform more sophisticated sentiment analyses based on systematic and well-grounded sentiment theories and frameworks. These frameworks will become common theoretical platforms for comparing results across different studies.

To date, the existing sentiment techniques may not be sufficient for sentiment classification. Linguistic framework, therefore, has been employed as a theoretical platform for sentiment analysis. By far, computational linguists have projected several sentiment frameworks. Wiebe et al. [48] disclosed a sentiment framework based on private states with three types of expressions, that is, explicit mentions of private states, speech events expressing private states, and expressive subjective elements. The framework is designed to expand attitude types to include intention, warning, uncertainty, and evaluation. Asher, Benamara, and Mathieu [49] further denoted an annotation framework involving four categories, that is, reporting expressions, judgement expressions, advice expressions, and sentiment expressions to express various kinds of feelings and opinions. The most comprehensive linguistic theory of sentiment, however, is the appraisal framework [50], also known as interpersonal semantics, developed within the tradition of Systemic Functional Linguistics [51]. Appraisal theory is a framework employed in conveying the language of evaluation in text [52] and also an approach to linguistics that focuses on the semantics of text rather than its grammar [53].

The framework portrayed a taxonomy of the language to convey *attitude*, *engagement*, and *graduation* with respect to the evaluations of other people. The detailed taxonomy is depicted in Figure 3. Attitude considers how people conduct interpersonal interaction to express his or her opinions and emotions; engagement assesses the evaluation with respect to others' opinions; and graduation denotes how to strengthen or weaken the attitude through language functions.

Attitude, engagement, and graduation can be further divided into several subtypes, which are explained more in-depth in the following paragraphs.

Attitude can be further divided into three distinct subsystems: affect, appreciation, and judgement. Affect identifies the feelings and emotions of the author (happy and sad), which can be a behavioral process or an internal mental state. Appreciation represents the reaction that a person talks about (beautiful and ugly), such as impact, quality, composition, complexity, or valuation. Judgement considers the author's attitude towards the behavior of somebody (heroic and idiotic) and the evaluation may concern social sanctions or social esteem; social sanctions

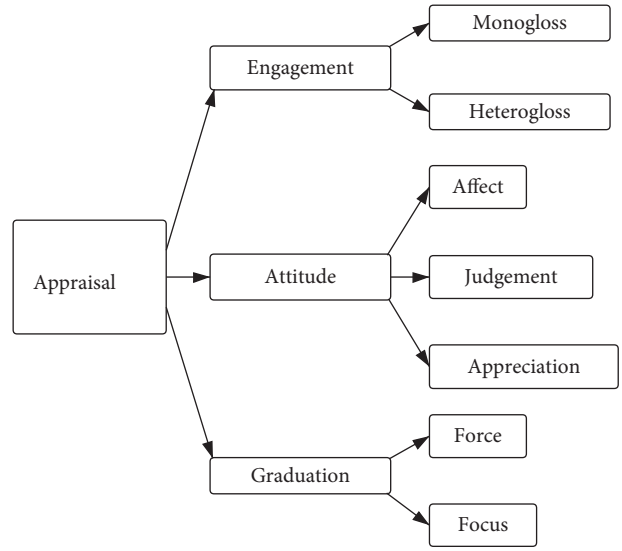


Figure 3: Outline of Martin and White's [52] appraisal framework (first two levels).

may involve veracity or propriety; and social esteem may involve the assessment of how normally someone behaves, how capable the person is, or how tenacious the person is.

Engagement contains two subcategories: monoglossic and heteroglossic. Monoglossic has no recognition of dialogistic alternatives; that is, the writer/speaker directly expressed the appraisal. Heteroglossic has the recognition of dialogistic alternative positions and voices, which means that the writer/speaker has either attributed to other methods or sources to make it credible.

Graduation also include two subclasses: force and focus. Force assesses the degree of intensity; focus covers to sharpen or soften the specification.

The main advantage of utilizing appraisal theory in sentiment identification is that it enables us to take a look deeper inside the thoughts of authors or publishers, revealing their real sensations by employing linguistic and psychological analysis of their texts. In our study, we want to extract the sentiment inside in the shareholders' letters and may possibly draw some statistical conclusions about the types of sentiment that companies of shareholders' letters express via the texts they write. At last, the analysis can help potential investors to make better decision-makings.

So far, scholars have made a great number of useful attempts at the level of sentiment analysis. Table 4 sums up the relevant sentiment analysis literature employing the appraisal framework. From the table, the research integrating the appraisal theory into the analysis of sentiment orientation is relatively scare. Korenek and Šimko [60] established the method of manually creating an emotional dictionary, utilizing the appraisal theory to evaluate emotional posts on Weibo posts; Taboada and Grieve [54] further made some improvements on how to aggregate the value of adjectives based on appraisal theory. In addition, Khoo et al. [50] extended the appraisal framework for analyzing long news reports, compared with microposts, and

Table 4: Related work of appraisal theory in sentiment analysis.

Authors	Models	Data source	Method	Attributes type	Accuracy (percent)
Taboada and Grieve [54]	Lexicon-based	Reviews	Taking into account text structure and adjectives frequency	Attitude	N/A
Whitelaw et al. [55]	Lexicon-based	Movie reviews	Analyzing appraisal adjectives and modifiers	Attitude, orientation, graduation, polarity	90.2
Read and Carroll [56]	Lexicon-based	Book reviews	Measuring interannotator agreement	Attitude, engagement, graduation, polarity	N/A
Argamon et al. [57]	NB, SVM	Documents	Automatic determining complex sentiment-related attributes	Attitude; orientation; force	N/A
Balahur et al. [58]	Robert Plutchik's wheel of emotion; Parrot's tree-structured list of emotions	ISEAR Corpus	EmotiNet (EmotiNet defines to store action chains and their corresponding emotional labels from several situations; in such a way, authors could be able to extract general patterns of appraisal)	Actor, action, object Emotion	N/A
Bloom [59]	Lexicon-based	MPQA 2.0 Corpus; UIC Review Corpus; Darmstadt Service Review Corpus; JDPA Sentiment Corpus; IIT Sentiment Corpus	Functional local appraisal grammar extractor	Attitude	44.6
Khoo et al. [50]	Lexicon-based	Political news article	Analyzing appraisal groups	Actor, attitude, engagement, polarity	N/A
Korenek and Šimko [60]	SVM	Microblog posts	Structuring sentiment analysis based on appraisal theory	Attitude, graduation, polarity, orientation, engagement	87.57
Cui and Shibamoto-Smith [61]	Lexicon-based	Self-constructed corpus of Chinese newspapers and websites	Discovering the lexical, syntactic, and semantic features of different types of sentiment parameters	Attitude Peripheral sentiment parameters (topic, source, field, process, degree)	N/A

assessed the framework's utility and possible problems, which is definitely a good attempt for our current study.

As indicated in Table 4, the aforementioned work has improved that appraisal framework is a highly useful tool for analyzing sentiment classification [50, 60]. This paper, thus, intends to augment the sentiment classification of president's letter by employing the appraisal theory. The appraisal theory proves to be useful in uncovering various aspects of sentiment that should be valuable to researchers and assists researchers to understand the feelings of shareholders' letters better. By following the doctor dissertation of Bloom [59], who laid the basis for sentiment classification by utilizing appraisal theory, a more systematic and well-grounded approach of sentiment analysis utilizing Martin and White's appraisal theory has been heuristically employed, which can successfully assist potential investors to understand corporations' attitudes accurately and make investment decisions more efficiently.

3. Research Methodology

As mentioned earlier, in order to achieve our goals in this study, a three-dimensional research framework has been established to enable us to systematically test different perspectives of text mining strategies with both quantitative and qualitative methods (see Figure 4).

The specific research questions that we address are the following:

RQ1: what are the prominent sentiment attributes in CEO's letters that can actively promote companies' images and avoid negative views?

RQ2: what are the main sentimental themes in CEO's letters? And which one can best encourage other companies' investment?

RQ3: how can the sentiment attributes of CEO's letters anticipate corporate financial performance that would be useful for stakeholders' decision-making?

The above questions in three dimensions evaluate CEO letters. In the microlevel of identifying sentiment attributes, appraisal theory and lexicon-based sentiment method have been applied to classify various sentiment attributes precisely. In the mesolevel analysis, we utilize NVivo qualitative data analysis software to do the thematic analysis of CEO letters. In the macrolevel, machine-learning approach was

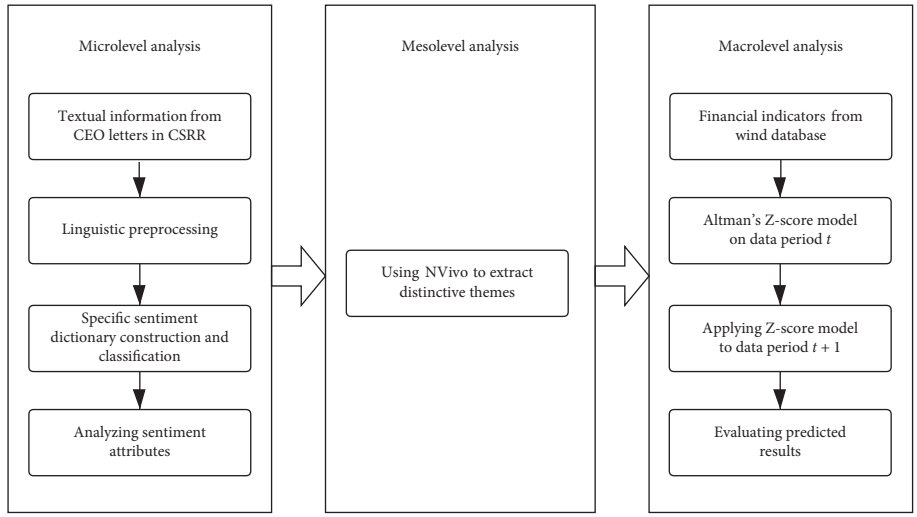


Figure 4: An overview of the workflow of the research framework.

selected to assess the power of CEO letters in predicting corporate financial performance in terms of the Z-score model.

3.1. Data Collection and Description. To date, various methods have been developed and introduced to measure sentiment. A suitable method to adopt for this study is to propose a specific sentiment dictionary based on the principles of appraisal theory. The advantage of utilizing the appraisal theory is that it allows researchers to categorize and compare sentiment attributes based on a systematic linguistic theory and also enables them to explore authors' thoughts more thoroughly. The major difference from the previous works which use the appraisal theory is that the whole basic appraisal tree has been selected and shareholders' letter specifics have been assessed.

In terms of the variety of genres, such as political news, movie reviews, and product reviews, the analysis results will significantly differ depending on distinct syntactic structure and lexical choices [56]. Therefore, in order to maintain consistency, the same genre has to been examined based on the appraisal framework. In this study, shareholders' letters are good candidates for this study because they are likely to contain similar languages in the same genre of writing; also, examples of appraisal's attributes can be easily detected. Most importantly, identifying valuable information from shareholders' letters is a sufficient way to help companies to improve the quality of released information and facilitate stakeholders' to make investment decisions.

All CSR reports with English version were downloaded from GRI's Sustainability Disclosure Database (https://database.globalreporting.org). This database can access to all types of sustainability reports from various industries relating to the reporting organizations. For the sake of preventing problems with both industry-specific attributes and different financial performance evaluation, the financial industries were excluded. Furthermore, in order to ensure accurate comparable information, all selected reports must have official English version and should be listed in the Hong Kong publicly regulated markets.

Finally, 41 Chinese companies' English version CSR reports from year 2016 had been collected. Then, 41 shareholders' letters were retrieved from the CSR reports in China in 2016. The corpus contains 35,670 tokens in total, named SLC. In addition, all the statistical data for calculating Z-score are gathered from the Wind financial database.

The year 2016 was selected because the United Nations General Assembly unanimously adopted the Resolution 70/1, Transforming our World: the 2030 Agenda for Sustainable Development. This historic document lays out 17 sustainable development goals, which aim to integrate and balance the three dimensions of sustainable development: economic, social, and environmental. The new goals and targets will come into effect on 1 January 2016 and will guide for the next fifteen years.

3.2. Data Preprocessing. Most CEO letters were released in PDF format. Firstly, letters were manually transformed.pdf documents to the essential.txt groups and texts were sorted to have only one sentence in each line. Then, the text documents were linguistically preprocessed using tokenization, part of speech tagging, and lemmatization:

Tokenisation—splitting texts into sentences and words

Part of speech tagging—adding a POS tag to each word in a sentence

Lemmatisation—converting a word into its basic lemma form

Deletion of stop words and company name

To control for bias in the analysis, in particular, shareholders' letters have a specific content that differs from other texts, which has to be coped with. For instance, the abbreviation of company's name is Best Buy, including the positive word of "Best", which may significantly convert the

polarity of the text. As a result, the pretreatments were a combination of tokenization, lemmatization, part-of-speech-tagging, and deletion of stop words and named entities, which gave the best result.

3.3. Specific Dictionary Construction and Classification. The main issue with the sentiment analysis of textual documents is the right choice of positive terms. Obviously, the categorization of words is not always unambiguous and requires context knowledge. This is due to the various meanings of words and domain specific tone of the words, respectively. Therefore, the appraisal theory has been employed in this study.

To evaluate the appraisal theory, a dictionary tagged with attributes from Martin and White's categorization (2005) has to be constructed. In terms of the work of Korenek and Šimko [60], they built a dictionary based on appraisal theory specializing in microblogs. By following their work, we accumulated approximately five hundred and fifty words from Martin and White's classification. In order to broaden the dictionary, WordNet database, Collins Thesaurus, and Merriam-Webster Thesaurus have been utilized to find synonyms to words identified in the previous step, and this formed the basic sentiment lexicons. The next step is to be aimed at extracting candidate sentiment word lists from the self-constructed SLC corpus, employing the statistical approach based on the amount of pointwise mutual information (PMI) ratio, which can be used to compare candidate words with the existing basic sentiment lexicons. PMI calculation has been defined as follows:

$$\mathrm{PMI}\left(\mathrm{word}_1, \mathrm{word}_2\right) = \log \frac{p\left(\mathrm{word}_1, \mathrm{word}_2\right)}{p\left(\mathrm{word}_1\right) p\left(\mathrm{word}_2\right)}. \quad (1)$$

In light of the PMI ratio, whether the word should be identified as a target or not must be decided. On completion of selecting targets, the target candidate words merged with the basic sentiment lexicons to construct a new dictionary specializing in shareholders' letter content, with approximately six hundred words in total.

Due to the specifics of shareholders' letters, the traditional lexicon-based approach cannot identify the complex context. So the statistical method of PMI ratio may not be accurate all the time. Manual screening is highly needed in this step.

To further categorize the new dictionary, each word is manually classified according to the attributes based on appraisal theory and in line with the research of [60], who assigned a 10-point Likert scale, from −5 to +5. The same scaling scope has been considered in this research. For the purpose of manual coding with participants' subjective judgement, eight human annotators, *a* to *h*, were invited to assign polarity independently. The annotators were all well-versed in the appraisal framework. They were asked to specify the type of attitude, engagement, or graduation present and assign a scaling and polarity to candidate words. In order to address the concern of inconsistent understanding regarding some ambiguous words, the annotation was executed over two rounds, punctuated by an intermediary analysis of agreement and disagreement among all annotators until a consensus has been achieved.

A partial example of entries is presented in Table 5. Each entry contains a word, a part-of-speech tag, a category, and subcategories according to the appraisal theory and appraisal value.

Based on the newly established sentiment dictionary for a specific corpus, we can further precisely analyze the sentiment characteristics of CEO letters from micro-, meso-, and macrolevel analysis.

3.4. Data Analysis

3.4.1. Microlevel Analysis. Regarding RQ1, we categorized CEO letters based on the specific established sentiment classification dictionary considering the following eleven categories of terms:

A. Positive attitude affect (e.g., happy, convinced, and satisfied)

B. Negative attitude affect (e.g., sorry and sad)

C. Positive attitude judgement (e.g., lucky, fortunate, and famous)

D. Negative attitude judgement (e.g., imperfect, unknown, and severe)

E. Positive attitude appreciation (e.g., exciting, dramatic, and excellent)

F. Negative attitude appreciation (e.g., imbalanced, disharmonious, and conflicting)

G. Positive engagement (e.g., certainly, obviously, naturally, and evidently)

H. Negative engagement (e.g., falsely and compellingly)

I. Positive graduation force (e.g., greatly, slightly, and somewhat)

J. Negative graduation force (e.g., small and remote)

K. Positive graduation focus (e.g., true and genuinely)

We assumed that well-performing corporations are being more positive, optimistic tones. Conversely, we expected a more active language in the case of poorly performing companies that need to take positive actions to improve their image and attract more investors.

3.4.2. Mesolevel Analysis. With the popularity of computer technology, a range of software packages emerged to assist with the analysis of qualitative data. It is generally accepted that computer-assisted qualitative data analysis software (CAQDAS) can enhance the data handling/analysis process if used appropriately and resolve analysts from complicated data analysis.

Table 5: A partial example of appraisal dictionary entries.

Word	Part of speech	Main category	Second level category	Other level categories	Appraisal value
Good	Adj.	Attitude	Judgement	Praise, propriety	3
Harmonious	Adj.	Attitude	Appreciation	Composition, balance	2
Advanced	Adj.	Attitude	Appreciation	Valuation	5
Clear	Adj.	Attitude	Appreciation	Composition, complexity	1
Never	Adv.	Engagement	N/A	N/A	−5
Incredible	Adj.	Attitude	Judgement	Admiration, normality	3
Largest	Adj.	Graduation	Force	Maximization	1

The software package NVivo (now update to version 12) is one of the most distinguished CAQDAS, which can help the analysis to work more efficiently and rigorously back up findings with grounded data [62].

In this study, NVivo has been prompted to do a thematic analysis for identifying the broad CSR topics existing in the shareholders' letters. Thematic analysis is a way of categorizing data from qualitative research through analyzing similar themes and interpreting the research findings [63]. In this study, thematic analysis can be employed to detect the main ideas existing in the CEO letters and through NVivo software to label or code different types of CSR.

NVivo allows nodes to have more than one dimension (tree branch). Therefore, we were able to identify where concepts may have more than one dimension or group them within a more general concept. This is a revolution in finding connections because it prompts the analyst to think about their concepts in more detail, facilitating conceptual clarity, and early discourse analysis [64]. Figure 5 reveals a sample of the tree node structure of CSR.

Coding stripes function is a useful function for researchers to annotate various segments in the whole documents, which may facilitate the comparison of categories. Figure 6 depicts an example of data coded at the ethical responsibility node.

Generally speaking, NVivo offers us a valuable tool to explore the complexities of potential relationships without forcing the data to fit specific categories. In this way, when we identified a possible relationship with CSR, we defined this in NVivo using the free code to represent it first and later created tree node to identify the internal relationship.

3.4.3. Macrolevel Analysis. The Altman's model of financial health (Altman's Z-score) was selected for the quantitative evaluation of the assessed companies. The reason of the selection was that this model was created for assessing company financial health in industrial branch with shares tradable on the Hong Kong publicly regulated markets and exactly such companies were selected for the study.

The scope of this so-called bankruptcy model is to predict the probability of survival or bankruptcy of the analyzed company. The nearer to the bankruptcy a company is, the better Altman's index works as a predictor of financial health. It predicts bankruptcy reliably about one to two years in advance. Altman's Z-score model uses the following relation to define the value of a company in industrial branch with shares publicly tradable on the stock market:

$$Z_i = 1.2X_{1,i} + 1.4X_{2,i} + 3.3X_{3,i} + 0.6X_{4,i} + 1.0X_{5,i}, \quad (2)$$

where i denotes the i-th company, X_1 is the working capital/total assets, X_2 is the retained earnings/total assets, X_3 is the earnings before interest and tax/total assets, X_4 is the market value of equity/total liabilities, and X_5 is the sales/total assets. Detailed variable definition can be found in Table 6.

The Z_i value is in range −4 to +8. The higher the value, the higher the financial health of a company. It holds true that if (1) $Z_i > 2.99$, the company is in the "safe zone" (a company with high probability to survive—financially strong company); (2) $1.80 \le Z_i \le 2.99$ "grey zone" (the future of the company cannot be determined clearly—a company with certain financial difficulties); and (3) $Z_i < 1.80$ "distress zone" (the company has serious financial problems—the company is endangered by bankruptcy).

Due to the special historical conditions of China's stock market formation, the types of stocks formed in China are different from those in foreign countries. In the stocks of listed companies, they can generally be divided into two categories: tradable shares and nontradable shares. In view of the fact that there is no market price for nontradable shares in the Hong Kong stock market, the model has made a slight change: X_4 = (share price ∗ tradable shares + net asset value per share ∗ nontradable shares)/total liabilities, and X_5 is the prime operating revenue/total assets.

The outputs of the forecasting models were represented by the classes of financial performance obtained using the Z-score bankruptcy model, namely, classes "safe zone," "grey zone," and "distress zone." In addition, we obtained additional output classes (increase, no change, and decrease). Given the fact that the classes were imbalanced in the dataset, we use the Synthetic Minority Oversampling Technique (SMOTE) algorithm [65] to modify the training dataset. The algorithm oversamples the minority classes so that all classes are presented equally in the training dataset.

The set of eleven sentiment attributes presented in the previous section was drawn from shareholders' letter. Following previous studies [46, 66], the input attributes were collected for Chinese companies in the year 2016, while the output financial performance (Z-score) was evaluated for the year 2017, and the change in the financial performance was measured as Z-score in 2017 related to its value in 2016. For the sake of preventing problems with both industry-specific attributes and different financial performance evaluation, the financial industry was excluded. As a result, among 41 Chinese companies, 9 companies were classified into "grey zone" and 32 companies were classified into "distress zone;"

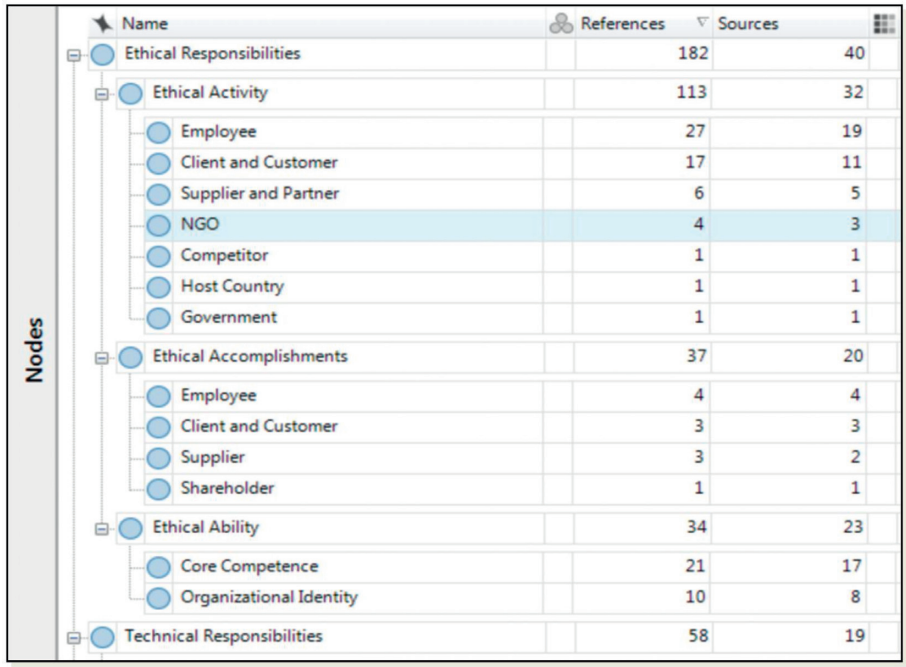


Figure 5: Tree node structure of CSR in China.

after making a comparison of financial situation between 2016 and 2017, only 2 companies were improved and the remaining 39 companies were unchanged.

In sum, the Z-score of Chinese companies in 2016 and 2017 could be spotted in Table 7.

Next, a various number of forecasting machine-learning methods have been explored, i.e., logistic regression, support vector machine, and naïve Bayes. Apart from logistic regression, the other two methods can process nonlinear data.

The logistic regression (LR) model has been used with a ridge estimator defined by Cessie and Houwelingen [67]. The classification performance of the logistic regression depends on the number of iterations and ridge factor, respectively.

Support vector machines (SVMs) are a set of related supervised learning methods, which are popular for performing classification and regression analysis using data analysis and pattern recognition.

Naïve Bayes (NB) is a simple multiclass classification algorithm with the assumption of independence between every pair of features. Naive Bayes can be trained very efficiently. Within a single pass on the training data, it computes the conditional probability distribution of each feature given label, and then it applies Bayes' theorem to compute the conditional probability distribution of the label given observation and use it for prediction.

In sum, logistic regression, support vector machine, naïve Bayes are three methods designed to forecast the accuracy of financial performance.

4. Experimental Results and Discussion

In the data filtering, we select 41 CEO letters in Chinese companies CSR report and try to do in-depth analysis at microlevel, mesolevel, and macrolevel, respectively.

4.1. Sentiment Attributes. Regarding the eleven sentiment attributes, the detailed statistical data can be found in Table 8. Among all the categories, positive attitude, judgement, appreciation, and positive graduation force are the top three most frequent sentiment attributes.

From the previous data collection part, we know that among 41 Chinese companies, 9 companies were classified into "grey zone" and 32 companies were classified into "distress zone." None of the companies were classified into the safe zone. Combing the eleven categorizations with our 2016 financial performance, interestingly, we found that poorly performing companies are expected to use a more active language to describe and evaluate their CSR. This phenomenon can be explained further by the impression management effect. Impression management refers to the process by which people try to manage and control the impression others make about themselves [68]. For corporations, the correct impression management can help companies to communicate with stakeholders smoothly. Companies try to use impression management to actively promote their images and avoid negative views. This is probably the reason why companies may encounter with a gloomy economic situation but still concentrate on shaping positive and optimistic images.

4.2. Sentimental Themes. The next segment is about to identify the sentimental themes in CEO letters through NVivo software. NVivo's coding stripes functions enable us to examine all the relevant text and identify the sentences which contributed to that relationship and also to find out which node the sentences belong to. After coding all the relevant text, we gathered comprehensive sentimental themes existing in CEO letters (see Table 9).

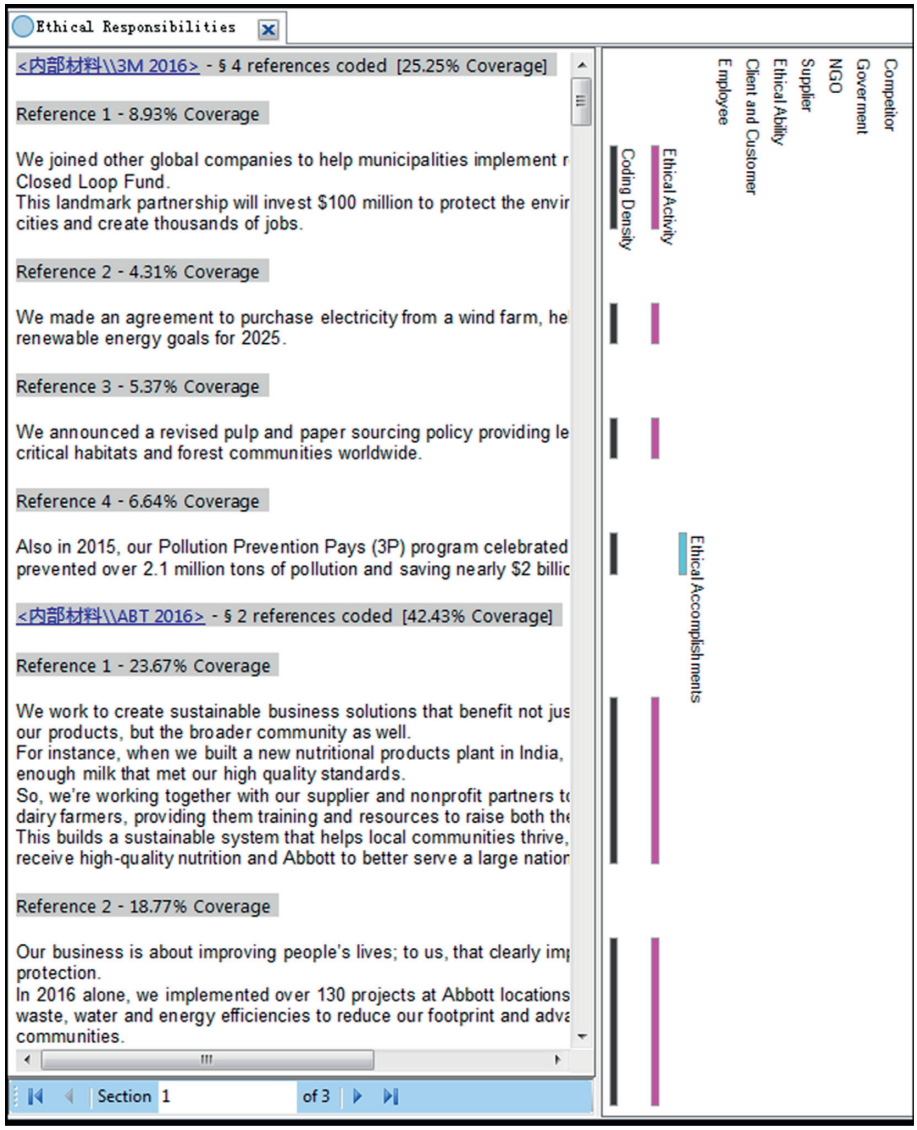


Figure 6: Coding stripes on the "ethical responsibilities" node.

Table 6: Financial variable definition.

Variable name	Definition
Working capital	Current assets minus current liabilities
Total assets	The final amount of all gross investments, cash and equivalents, receivables, and other assets as they are presented on the balance sheet
Retained earnings	The amount of net income left over for the business after it has paid out dividends to its shareholders
Earnings before interest and tax (EBIT)	Revenue minus expenses, excluding tax and interest
Market value of equity	The total dollar value of a company's equity
Total liabilities	The combined debts and obligations that an individual or company owes to outside parties
Sales	Total dollar amount collected for goods and services provided

According to Table 9, the most popular sentimental theme in CEO letters is the ethical responsibility. Business ethical responsibility for companies means a system of moral and ethical beliefs that guide companies' behaviors, values, and decisions and minimizing unjustified harm, suffering, waste, or destruction to people and the environment [69]. This result reveals that a large amount of companies (32 among 41 companies) have put great efforts into business ethics. The concept of business ethics began in the 1960s as corporations became more aware of a rising consumer-based society that showed concerns regarding the environment, social causes, and corporate responsibility. In fact,

Table 7: Z-score of Chinese companies in 2016 and 2017.

Stock code	Country	2016 Z-score	2016	2017 Z-score	2017
1958.hk	CN	1.260515439	Distress zone	1.482271118	Distress zone
0606.hk	CN	1.838183342	Grey zone	2.216679937	Grey zone
1800.hk	CN	0.867078825	Distress zone	0.864201511	Distress zone
1829.hk	CN	1.357941592	Distress zone	1.427298576	Distress zone
0941.hk	CN	2.936618457	Grey zone	2.789570355	Grey zone
3323.hk	CN	0.337950565	Distress zone	0.497012524	Distress zone
0390.hk	CN	1.192247784	Distress zone	1.15640714	Distress zone
3320.hk	CN	2.201481901	Grey zone	2.007730795	Grey zone
1055.hk	CN	0.551047016	Distress zone	0.659758361	Distress zone
0728.hk	CN	1.062436514	Distress zone	1.190768229	Distress zone
0688.hk	CN	1.961825278	Grey zone	1.972625832	Grey zone
2196.hk	CN	1.215563686	Distress zone	1.074862508	Distress zone
1618.hk	CN	0.89432166	Distress zone	0.879602037	Distress zone
0857.hk	CN	1.181842765	Distress zone	1.329173357	Distress zone
3377.hk	CN	1.083233034	Distress zone	1.142850174	Distress zone
0347.hk	CN	0.71319999	Distress zone	1.340165363	Distress zone
0992.hk	CN	1.775439241	Distress zone	1.686162129	Distress zone
0753.hk	CN	0.740145558	Distress zone	0.812868908	Distress zone
0883.hk	CN	1.927734935	Grey zone	2.336603263	Grey zone
0232.hk	CN	0.975494042	Distress zone	1.838159662	Grey zone
1186.hk	CN	1.251163858	Distress zone	1.239974657	Distress zone
2607.hk	CN	2.345340814	Grey zone	2.234627565	Grey zone
0763.hk	CN	1.059868156	Distress zone	1.363358965	Distress zone
0670.hk	CN	0.482355283	Distress zone	0.455072698	Distress zone
2202.hk	CN	0.807062775	Distress zone	0.687163013	Distress zone
0489.hk	CN	2.294304149	Grey zone	2.115933926	Grey zone
0836.hk	CN	0.99429478	Distress zone	0.843126034	Distress zone
0392.hk	CN	1.105575201	Distress zone	1.11398028	Distress zone
0604.hk	CN	1.226738962	Distress zone	0.889164039	Distress zone
2380.hk	CN	0.53081341	Distress zone	0.310240584	Distress zone
0257.hk	CN	1.719616264	Distress zone	1.540477349	Distress zone
0103.hk	CN	0.669137242	Distress zone	0.689465751	Distress zone
1211.hk	CN	1.261617869	Distress zone	1.124410212	Distress zone
0916.hk	CN	0.406631641	Distress zone	0.557081816	Distress zone
0175.hk	CN	2.405647254	Grey zone	4.48691365	Safe zone
1088.hk	CN	1.243041938	Distress zone	1.543226046	Distress zone
0123.hk	CN	0.919420683	Distress zone	1.015226322	Distress zone
2128.hk	CN	2.456387186	Grey zone	2.308100735	Grey zone
2866.hk	CN	0.125023125	Distress zone	0.230025232	Distress zone
3360.hk	CN	0.346809818	Distress zone	0.355677289	Distress zone
6166.hk	CN	1.678887209	Distress zone	1.73679922	Distress zone

Table 8: The frequency of eleven sentiment attributes.

Positive graduation focus	15
Negative graduation force	15
Positive graduation force	274
Negative engagement	7
Positive engagement	223
Negative attitude appreciation	6
Positive attitude appreciation	345
Negative attitude judgement	11
Positive attitude judgement	378
Negative attitude affect	2
Positive attitude affect	87

the importance of business ethics reaches far beyond the strength of a management tea bond or employee loyalty. As with all business initiatives, the ethical operation of a company is directly related to companies' short-term or long-term profitability. The reputation of a business in the surrounding community, other businesses, and individual investors is paramount in determining whether a company is a worthwhile investment. If a company is perceived to be unethical, investors are less inclined to support its operation.

In addition, in this study, we have divided the ethical responsibility into three parts: ethical accomplishment, ethical activity, and ethical ability. Ethical accomplishment means some awards that have been achieved by companies. Ethical ability expresses companies' core competence and organizational identity. Ethical activity makes investors to know about the activities and functions taking place in the company. Detailed samples are scheduled below:

(a) We implemented new energy efficiency projects worldwide, including the implementation of the ISO

Table 9: Main sentimental themes in CEO letters.

Responsibility type	Detailed categories	References	Sources
Ethical responsibility	Ethical ability	34	23
	Ethical activity	113	32
	Ethical accomplishments	37	20
Technical responsibility	Technical ability	8	6
	Technical activity	33	13
	Technical accomplishments	18	13
Economic responsibility	Economic ability	5	5
	Economic activity	12	9
	Economic accomplishments	39	23
Philanthropic responsibility	Philanthropic ability	2	2
	Philanthropic activity	47	26
	Philanthropic accomplishments	4	3
Administrative responsibility	Administrative ability	3	3
	Administrative activity	20	15
	Administrative accomplishments	2	2
Political responsibility	Political ability	12	10
	Political activity	11	6
	Political accomplishments	0	0
Legal responsibility	Legal ability	1	1
	Legal activity	2	1
	Legal accomplishments	0	0

50001 Energy Management System in all our European Union locations (Ethical activity, Lenovo 2016).

(b) In 2016, we have achieved safe flight hours of 2.38 million, transported 115 million passengers, eliminated incidents by human errors, continued to maintain the best safety record in China's civil aviation (Ethical accomplishments, China Southern Airlines 2016).

(c) Consequently, China Telecom is among the first batch of the national demonstration bases for entrepreneurship and innovation (Ethical ability, China Telecom 2016).

From the coding results, ethical activity occupies the largest proportion, which means that firms would prefer to demonstrate their actions and practices to the society. Corporations have more incentives to be ethical as the area of socially responsible and ethical investing keeps growing. An increasing number of investors are seeking ethically operating companies to invest, which drives more firms to take this issue seriously. With the consistent ethical behavior, an increasingly positive public image can be established, and to retain a positive image, companies must be committed to operating on an ethical foundation as it relates to the treatment of employees, respecting the surrounding environment and fair market practices in terms of price and consumer treatment.

4.3. Machine-Learning Approach Forecasting Financial Performance. We tested many machine-learning approaches and selected only the best outcomes. The measures of classification performance were, in addition to Acc, represented by the averages of standard statistics applied in classification tasks [70]: true-positive rate (TP), false-positive rate (FP), precision (Pre), recall (Re), F-measure (F-m), and ROC curve. F-m is the weighted harmonic mean of precision and recall. ROC is a plot of the true-positive rate against the false-positive rate for the different cutpoints of a diagnostic test. In order to validate the accuracy, the experiments were realized using 10-fold cross validation. The best results in terms of the accuracy of correctly classified instances Acc (%) are presented in Table 10 (for the financial performance classes).

Table 10: Average accuracy of the analyzed methods and weighted average TP, FP, Pre, Re, F-m, and ROC for classification of financial performance (safe zone, grey zone, and distress zone).

Model	Acc (%)	TP	FP	Precision	Recall	F-m	ROC
NB	0.6925	0.1187	0.3333	0.3097	0.1187	0.0451	0.3358
SVM	0.7022	0.1183	0.3333	0.3130	0.1183	0.0442	0.3298
LR	0.7046	0.1183	0.3333	0.3138	0.1183	0.0442	0.3324

The results obtained by modeling point out that the logistic regression is more suitable for forecasting financial performance, reaching the highest accuracy of 70.46%. This evidence suggests that there exists a linear relationship between the sentiment and financial performance.

5. Conclusions

CEO letter contains information about corporate social responsibility performance in CSRR, which is designated for stakeholders to make their investment decisions. In this study, sentiment analysis has been applied to the evaluation of CEO letters from three perspectives: sentiment dictionary (microlevel), sentimental themes (mesolevel), and machine learning (macrolevel). In the microlevel analysis, a designated sentiment dictionary has been constructed for classifying sentiment attributes. The results denote that no

matter the companies are in an active or passive economic situation, they are focusing on using a great proportion of positive words to establish a company image and attract investors. In the mesolevel analysis, a comprehensive tree node structure was identified to discover the sentimental topics relevant to CSR. In terms of outcome, the CEO letters contain a large quantity of ethical-related information, especially concentrating on the ethical activities that firms have organized. In the macrolevel, the logistic regression approach achieves the best result in forecasting future financial performance, which proves that there is a linear relationship between the sentiment and economic performance. In other words, sentiment information in CEO letters can be regarded as a vital determinant for forecasting financial performance.

The distinct contribution of this paper is threefold. Firstly, an advanced technique of sentiment analysis utilizing appraisal theory has been conducted, which is potentially useful for detecting the "concealed" information in letters. Secondly, a sentiment dictionary has been constructed successfully and specifically for shareholders' letters, which can significantly upgrade the accuracy of sentiment classification. Lastly, this research can guide companies to further enhance the technique of releasing nonfinancial information and display fundamental causes for diverse texts.

The current research has its own limitations. Utilizing Z-score, the quantitative assessment to define companies' economic situation may not be adequate. In the Z-score model, the stock market value is only a static value at some point in time, which cannot reflect a dynamic fluctuation. In fact, the need to interpret the firms' released information is to predict its future performance, it is insignificant which economic model was selected, and the critical point is the influence of sentiment on the perception of the company by its stakeholders.

In the future research, it is possible to apply other economic models for predicting financial performance. Especially with the popularity and high development of sentiment analysis, a great number of new approaches would be probed by text mining to detect more concealed information for investors' decision-making and to be applied to other languages as well.

Data Availability

All data, models, and code generated or used during the study appear in the submitted article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the 2019 Project of the National Social Science Foundation of China: On the Overseas CSR Driving Forces and Influencing Mechanism for Chinese Enterprises (Grant no. 19BGL116).

References

[1] P. Bo and L. Lee, "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts," in *Proceedings of the Meeting on Association for Computational Linguistics*, Philadelphia, PA, USA, June 2004.

[2] E. Abrahamson and E. Amir, "The information content of the president's letter to shareholders," *Journal of Business Finance & Accounting*, vol. 23, no. 8, pp. 1157–1182, 1996.

[3] P. Hájek, V. Olej, and R. Myšková, "Forecasting stock prices using sentiment information in annual reports - a neural network and support vector regression approach," *Wseas Transactions on Business & Economics*, vol. 10, no. 4, pp. 293–305, 2013.

[4] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.

[5] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis," *Computational Linguistics*, vol. 35, no. 3, pp. 399–433, 2009.

[6] C. J. Anderson and G. Imperia, "The corporate annual report: a photo analysis of male and female portrayals," *Journal of Business Communication*, vol. 29, no. 2, pp. 113–128, 1992.

[7] G. F. Kohut and A. H. Segars, "The president's letter to stockholders: an examination of corporate communication strategy," *Journal of Business Communication*, vol. 29, no. 1, pp. 7–21, 1992.

[8] L. Patelli and M. Pedrini, "Is the optimism in CEO's letters to shareholders sincere? Impression management versus communicative action during the economic crisis," *Journal of Business Ethics*, vol. 124, no. 1, pp. 19–34, 2014.

[9] P. Bo and L. Lee, *Opinion Mining and Sentiment Analysis*, Now Publishers Inc., Boston, MA, USA, 2008.

[10] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: opinion extraction and semantic classification of product reviews," in *Proceedings of the International Conference on World Wide Web*, Banff, Canada, May 2003.

[11] A. Kennedy and D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters," *Computational Intelligence*, vol. 22, no. 2, pp. 110–125, 2006.

[12] K. S. Srujan, S. S. Nikhil, H. R. Rao, K. Karthik, B. S. Harish, and H. M. K. Kumar, *Classification of Amazon Book Reviews Based on Sentiment Analysis*, Springer, Singapore, 2018.

[13] J. Feng, C. Gong, X. Li, and R. Y. K. Lau, "Automatic approach of sentiment lexicon generation for mobile shopping reviews," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 9839432, 13 pages, 2018.

[14] B. Huang, G. Yu, and H. R. Karimi, "The finding and dynamic detection of opinion leaders in social network," *Mathematical Problems in Engineering*, vol. 2014, no. 1, 7 pages, 2014.

[15] R. Quratulain, G. Sayeed, Sajjad, and Haider, "Lexicon-based sentiment analysis of teachers' evaluation," *Applied Computational Intelligence and Soft Computing*, vol. 2016, Article ID 2385429, 12 pages, 2016.

[16] M. Stewart, "The language of praise and criticism in a student evaluation survey," *Studies in Educational Evaluation*, vol. 45, pp. 1–9, 2015.

[17] C. L. Chen, C. L. Liu, Y. C. Chang, and H. P. Tsai, "Opinion mining for relating subjective expressions and annual earnings in US financial statements," *Journal of Information Science & Engineering*, vol. 29, no. 4, pp. 743–764, 2012.

[18] J. L. Hobson, W. J. Mayew, and M. Venkatachalam, "Discussion of analyzing speech to detect financial misreporting,"

Journal of Accounting Research, vol. 50, no. 2, pp. 393–400, 2012.

[19] C. Huang, X. Yang, X. Yang, and H. Sheng, "An empirical study of the effect of investor sentiment on returns of different industries," *Mathematical Problems in Engineering*, vol. 2014, Article ID 545723, 11 pages, 2014.

[20] C. Xie and Y. Wang, "Does online investor sentiment affect the asset price movement? Evidence from the Chinese stock market," *Mathematical Problems in Engineering*, vol. 2017, Article ID 2407086, 11 pages, 2017.

[21] M. Taboada, "Sentiment analysis: an overview from linguistics," *Linguistics*, vol. 2, no. 1, 2016.

[22] C. J. Hutto and E. Gilbert, *VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text*, in *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, Ann Arbor, MI, USA, 2014.

[23] P. D. Turney, "Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews," in *Proceedings of Annual Meeting of the Association for Computational Linguistics*, pp. 417–424, Philadelphia, PA, USA, July 2002.

[24] V. Hatzivassiloglou and K. R. Mckeown, "Predicting the semantic orientation of adjectives," in *Proceedings of the ACL*, pp. 174–181, Autrans, France, 1997.

[25] P. D. Turney and M. L. Littman, "Measuring praise and criticism," *Acm Transactions on Information Systems*, vol. 21, no. 4, pp. 315–346, 2003.

[26] M. Taboada, C. Anthony, and K. Voll, "Methods for creating semantic orientation dictionaries," in *Proceedings of the Language Resources & Evaluation*, Genoa, Italy, May 2006.

[27] X. Ding, L. Bing, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in *Proceedings of the International Conference on Web Search & Data Mining*, Cambridge, UK, 2008.

[28] A. Dey, M. Jenamani, and J. J. Thakkar, "Senti-N-Gram: an n-gram lexicon for sentiment analysis," *Expert Systems with Applications*, vol. 103, Article ID S095741741830143X, 2018.

[29] J. Brooke, M. Tofiloski, and M. Taboada, "Cross-linguistic sentiment analysis: from English to Spanish," in *Proceedings of the International Conference on Recent Advances in Natural Language Processing*, Borovets, Bulgaria, September 2009.

[30] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in *Proceedings of the Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, vol. 10, Philadelphia, PA, USA, July 2002.

[31] E. Boiy and M.-F. Moens, "A machine learning approach to sentiment analysis in multilingual Web texts," *Information Retrieval*, vol. 12, no. 5, pp. 526–558, 2009.

[32] L. I. Jun and M. Sun, "Experimental study on sentiment classification of Chinese review using machine learning techniques," in *Proceedings of the IEEE International Conference on Natural Language Processing & Knowledge Engineering*, Beijing; China, 2007.

[33] R. F. Bruce and J. M. Wiebe, *Recognizing Subjectivity: A Case Study in Manual Tagging*, Cambridge University Press, Cambridge, UK, 1999.

[34] Gamon and Michael, "Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis," in *Proceedings of the International Conference on Computational Linguistics*, 2004.

[35] C. Yang, X. Tang, Y. C. Wong, and C. P. Wei, "Understanding online consumer review opinion with sentiment analysis using machine learning," *Pacific Asia Journal of the Association for Information Systems*, vol. 2, no. 3, 2010.

[36] C. Troussas, M. Virvou, K. J. Espinosa, K. Llaguno, and J. Caro, "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning," in *Proceedings of the Fourth International Conference on Information*, Mikrolimano, Greece, 2013.

[37] J. Singh, G. Singh, and R. Singh, "Optimization of sentiment analysis using machine learning classifiers," *Human-centric Computing and Information Sciences*, vol. 7, no. 1, p. 32, 2017.

[38] M. Rathi, A. Malik, D. Varshney, R. Sharma, and S. Mendiratta, "Sentiment analysis of tweets using machine learning approach," in *Proceedings of the 2018 Eleventh International Conference on Contemporary Computing (IC3)*, Noida, India, August 2018.

[39] S. Yuan, H. Wang, and M. Zhu, "Sustainable strategy for corporate governance based on the sentiment analysis of financial reports with CSR," *Financial Innovation*, vol. 4, no. 1, p. 2, 2018.

[40] A. Aue and M. Gamon, "Customizing sentiment classifiers to new domains: a case study," in *Proceedings of the International Conference on Recent Advances in Natural Language Processing*, Varna, Bulgaria, September 2005.

[41] S. González-Bailón and G. Paltoglou, "Signals of public opinion in online communication," *The ANNALS of the American Academy of Political and Social Science*, vol. 659, no. 1, pp. 95–107, 2015.

[42] T. Loughran and B. Mcdonald, "When is a liability not a liability? Textual analysis, dictionaries, and 10-ks," *The Journal of Finance*, vol. 66, no. 1, pp. 35–65, 2011.

[43] R. P. Schumaker, Y. Zhang, C.-N. Huang, and H. Chen, "Evaluating sentiment in financial news articles," *Decision Support Systems*, vol. 53, no. 3, 2012.

[44] P. Hájek and V. Olej, "Evaluating sentiment in annual reports for financial distress prediction using neural networks and support vector machines," *Engineering Applications of Neural Networks*, Springer, Berlin, Germany, 2013.

[45] Y. Q. Xin, P. Srinivasan, and H. Yong, "Supervised learning models to predict firm performance with annual reports: an empirical study," *Journal of the American Society for Information Science & Technology*, vol. 65, no. 2, pp. 400–413, 2014.

[46] P. Hajek, V. Olej, and R. Myskova, "Forecasting corporate financial performance using sentiment in annual reports for stakeholders' decision-making," *Technological and Economic Development of Economy*, vol. 20, no. 4, pp. 721–738, 2014.

[47] S. Tong, W. Jia, P. Zhang, C. Yu, and D. Wang, "Predicting stock price returns using microblog sentiment for Chinese stock market," in *Proceedings of the 2017 3rd International Conference on Big Data Computing and Communications (BIGCOM)*, Chengdu, China, August 2017.

[48] J. Wiebe, T. Wilson, and C. Cardie, "Annotating expressions of opinions and emotions in language," *Language Resources and Evaluation*, vol. 39, no. 2-3, pp. 165–210, 2005.

[49] N. Asher, F. Benamara, and Y. Y. Mathieu, "Appraisal of opinion expressions in discourse," *Linguisticae Investigationes Revue Internationale De Linguistique Française Et De Linguistique Générale*, vol. 32, no. 2, pp. 279–292, 2009.

[50] C. S. G. Khoo, A. Nourbakhsh, and J. C. Na, "Sentiment analysis of online news text: a case study of appraisal theory," *Online Information Review*, vol. 36, no. 6, pp. 680–686, 2012.

[51] M. A. K. Halliday, *An Introduction to Functional Grammar*, Oxford University Press Inc., London, UK, 1994.

[52] J. R. Martin and P. R. R. White, *Language of Evaluation: Appraisal in English*, Palgrave Macmillan, London, UK, 2005.

[53] S. Eggins, *An Introduction to Systemic Functional Linguistics*, Pinter, London, UK, 1994.
[54] M. Taboada and J. Grieve, "Analyzing appraisal automatically," in *Proceedings of the Aaai Spring Symposium on Exploring Attitude & Affect in Text Theories & Applications*, pp. 158–161, Stanford, CA, USA, January 2004.
[55] C. Whitelaw, N. Garg, and S. Argamon, "Using appraisal groups for sentiment analysis," in *Proceedings of the ACM International Conference on Information & Knowledge Management*, Bremen, Germany, October 2005.
[56] J. Read and J. Carroll, "Annotating expressions of appraisal in english," in *Proceedings of the Linguistic Annotation Workshop*, Prague, Czech Republic, June 2007.
[57] S. Argamon, K. Bloom, A. Esuli, and F. Sebastiani, "Automatically determining attitude type and force for sentiment analysis," in *Proceedings of the Human Language Technology Challenges of the Information Society*, Poznan, Poland, October 2009.
[58] A. Balahur, J. M. Hermida, A. Montoyo, and R. Muñoz, *EmotiNet: A Knowledge Base for Emotion Detection in Text Built on the Appraisal Theories*, Springer, Berlin, Heidelberg, 2011.
[59] K. Bloom, "Sentiment analysis based on appraisal theory and functional local grammars," Dissertation thesis, Illinois Institute of Technology, Chicago, IL, USA, 2011.
[60] P. Korenek and M. Šimko, "Sentiment analysis on microblog utilizing appraisal theory," *World Wide Web*, vol. 17, no. 4, pp. 847–867, 2014.
[61] X.-L. Cui and J. S. Shibamoto-Smith, "A corpus-based study on Chinese sentiment parameters of Chinese sentiment discourse," *Chinese Language and Discourse*, vol. 5, no. 2, pp. 185–210, 2014.
[62] A. Edwardsjones, *Qualitative Data Analysis with NVIVO*, Sage Publications, vol. 15, p. 868, Thousand Oaks, CA, USA, 2010.
[63] R. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*, Sage Publications, Thousand Oaks, CA, USA, 1998.
[64] P. Bazeley, *Qualitative Data Analysis with NVivo*, Sage, London, UK, 2007.
[65] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, no. 1, pp. 321–357, 2002.
[66] P. Hájek, "Credit rating analysis using adaptive fuzzy rule-based systems: an industry-specific approach," *Central European Journal of Operations Research*, vol. 20, no. 3, pp. 421–434, 2012.
[67] S. L. Cessie and J. C. V. Houwelingen, "Ridge estimators in logistic regression," *Applied Statistics*, vol. 41, no. 1, pp. 191–201, 1992.
[68] E. Goffman, *Presentation of Self in Every Day Life*, vol. 21, pp. 14-15, Doubleday, New York, NY, USA, 1959.
[69] A. B. Carroll, "The pyramid of corporate social responsibility: toward the moral management of organizational stakeholders," *Business Horizons*, vol. 34, no. 4, pp. 39–48, 1991.
[70] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 1, no. 2, pp. 37–63, 2011.

Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 1747315, 7 pages
https://doi.org/10.1155/2020/1747315

Research Article

XGBDeepFM for CTR Predictions in Mobile Advertising Benefits from Ad Context

Han An and Jifan Ren

School of Economics and Management, Harbin Institute of Technology, Shenzhen 518055, China

Correspondence should be addressed to Jifan Ren; renjifan@hit.edu.cn

Received 28 February 2020; Revised 2 April 2020; Accepted 11 April 2020; Published 27 April 2020

Guest Editor: Shianghau Wu

The problem of click-through rate (CTR) prediction in mobile advertising is one of the most informative metrics used in mobile business activities, such as profit evaluation and resource management. In mobile advertising, CTR prediction is essential but challenging due to data sparsity. Moreover, existing methods often have difficulty in capturing the different orders of feature interactions simultaneously. In this study, a method was developed to obtain accurate CTR prediction by incorporating contextual features and feature interactions. We initially use extreme gradient boosting (XGBoost) as a feature engineering phase to select highly significant features. The selected features are mobile contextual attributes including time contextual, geography contextual, and other contextual attributes (e.g., weather condition) in actual mobile advertising situations. Our model, XGBoost deep factorization machine- (FM-) supported neutral network (XGBDeepFM), combines the power of XGBoost for feature selection, FM for two-order cross feature interaction, and the deep neural network for high-order feature learning in a united architecture. In a mobile advertising condition, our methods lead to significantly accurate CTR prediction in "wide and deep" type of model. In comparison with existing models, many experiments on commercial datasets show that the XGBDeepFM model has better value of area under curve and improves the effectiveness and efficiency of CTR prediction for mobile advertising.

1. Introduction

The task of click-through rate (CTR) prediction is crucial in advertising and recommendation areas; its main goal is to maximize the clicks to improve advertising revenue or user satisfaction [1–3]. In advertising area, CTR is an important indicator for measuring the effectiveness of advertising displays [4]. Advertiser's revenue relies heavily on the capability of CTR prediction. In recommendation area, the recommended items returned to users can be ranked by the predicted CTR [5]. This predicted probability helps recommendation systems know the users' interest on specific items such as news [6, 7], movies [8], tags [9], or commercial items [10], which influence the subsequent decision-making [10]. Recommendation solutions can be classified in terms of collaborative, content-based, knowledge-based, demographic, and hybrid [11]. Each strategy can benefit from the task of CTR prediction [12].

One of the core problems that mobile advertising strives to solve is providing the right ads to the right people at the right time and in the right context. Users' attention time has been greatly reduced; thus, no one has the time to watch useless and intrusive advertisements. Nevertheless, the answer may be in the hands of marketers, especially in a dynamic mobile world. Mobile contextual advertising is not only about finding the right users in the right context, including time, geography, and weather, but is also about connecting the advertisement with the user in ad context and providing a pleasant experience. Accurate CTR prediction is vital to marketers based on contextual features. Another key challenge for CTR prediction is learning low- and high-order feature interactions behind user behavior in a certain context. Some feature interactions are easy to capture. Low-order feature interactions (less than two orders) can be designed by experts' prior experience. However, high-order feature interactions can be difficult to understand. These

deep-feature interactions can be learned by deep neural networks (DNNs).

At present, researchers have proposed different methods of CTR prediction. Some ideas attempt to solve the two-order feature interactions. For example, a logistic regression (LR) model has been used to predict the CTR on Google Ads [13]. Factorization machines (FMs) have been used to consider two-order feature interactions [14]. In recent years, DNNs have been popular because of their capability to learn high-order feature interactions. For example, Zhang et al. studied feature representations and proposed the FM-supported neutral network (FNN) [15]. Qu et al. proposed the product-based neutral network (PNN), which learns high-order feature interactions by introducing a product layer [16]. Cheng et al. combined "wide" and "deep" (W&D) components in a W&D model for low- and high-order feature interactions [2].

The remainder of this paper is organized as follows. In Section 2, the extreme gradient boosting (XGBoost) deep FM-supported neutral network (XGBDeepFM) is proposed, considering the problems in contemporary research and the characteristics in the mobile advertising dataset. In Section 3, the experiment is designed, a comparison experiment is performed for CTR prediction, and the effectiveness and efficiency of the XGBDeepFM model are analyzed. Finally, Section 4 concludes the study.

2. Materials and Methods

According to the core challenge of computational advertising proposed by Border [3], the best match between a given user in a given context and a suitable advertisement should be determined. $((a_i, u_i, c_i), y_i)$ denotes the instance i of the dataset, where a_i denotes the advertisement, u_i represents the user, c_i denotes the context, and $y_i \in \{0, 1\}$ is the label of the clicking label.

We propose a united approach, namely, the XGBDeepFM model, which benefits from prior information from context and high-order feature interactions, as shown in Figure 1.

Our approach consists of three components, namely, the XGBoost, FM, and deep components. By using these three components, the proposed XGBDeepFM model can realize the full interactive combination modeling of the bilateral features (i.e., ad and user features) and contextual features.

$$\widehat{y} = \text{sigmoid}\left(y_{\text{XGBoost}} + y_{\text{FM}} + y_{\text{DNN}}\right). \tag{1}$$

$\widehat{y}$ denotes the predicted CTR, and y_{XGBoost}, y_{FM}, and y_{DNN} are the outputs of the XGBoost, FM, and deep components, respectively.

2.1. XGBoost Component. The XGBoost component is a scalable machine learning system for tree boosting [2]. In XGBoost, feature selection and combination are automatically performed to generate new discrete feature vectors as the input of the LR model. The depth of a decision tree determines the dimension of the feature intersection. For example, if the depth of the decision tree is four, then the final number of the leaf node is the number of orders (three order) of feature interactions. We use XGBoost to capture three-order feature interaction and perform feature selection among features. The objective function of the XGBoost is as follows:

$$\text{Obj} = \sum_{i=1}^{N} l\left(\widehat{y}_i, y_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right). \tag{2}$$

XGBoost uses the following forward distribution algorithm:

$$\begin{aligned} \widehat{y}_i^{(t)} &= \sum_{k=1}^{t} f_k\left(x_i\right) = \widehat{y}_i^{(t-1)} + f_t\left(x_i\right), \\ \text{Obj}^{(t)} &= \sum_{i=1}^{N} l\left(y_i, \widehat{y}_i^{(t-1)} + f_t\left(x_i\right)\right) + \Omega\left(f_t\right) + C, \end{aligned} \tag{3}$$

where $\widehat{y}_i^{(t)}$ is the predicted value of the time t of the iteration, that is, the predicted result of sample x_i by t trees and $\widehat{y}_i^{(t-1)}$ is the predicted value of the current $(t-1)$th iteration. Thus, when the model is initialized, the model has no tree, and the predicted result is a constant. Each iteration adds a new tree to the model, and the loss function then changes correspondingly. In addition, the training of $(t-1)$th trees is completed when the (t)th tree is added.

2.2. FM Component. The FM component is used to learn feature interactions [1]. FM models can capture two-order feature interactions as the inner product of respective feature latent vectors.

$$\widehat{y}\left(x\right) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \langle v_i, v_j \rangle x_i x_j, \tag{4}$$

where v_i, v_j denotes the latent vector. Each cross-term parameter w_{ij} is expressed by the inner product $\langle v_i, v_j \rangle$ of the latent vector. The objective function of FM is as follows:

$$\text{Obj}^{(t)} = \sum_{i=1}^{N} l\left(y_i, \widehat{y}_i^{(t-1)} + f_t\left(x_i\right)\right) + \Omega\left(f_t\right) + C. \tag{5}$$

2.3. Deep Component. The deep component is used to learn high-order (more than three orders) feature interactions. The original features are initially embedded such that the features of different fields are mapped to the same dimension of the embedding space. Similarly, the dimension of the implicit vectors is k. Here, we set two layers for the deep component, and the entire DNN component is then computed as follows:

$$\begin{aligned} a^{(0)} &= \left[e_1, e_2, \ldots, e_m\right], \\ a^{(1)} &= \sigma\left(W^{(0)} a^{(0)} + b^{(0)}\right), \\ a^{(2)} &= \sigma\left(W^{(1)} a^{(1)} + b^{(1)}\right), \\ \widehat{y}_{\text{DNN}} &= W^{(3)} a^{(2)} + b^{(3)}, \end{aligned} \tag{6}$$

where e_m denotes the embedding of discrete features and $\widehat{y}_{\text{DNN}}$ is the prediction of DNN for the CTR of mobile

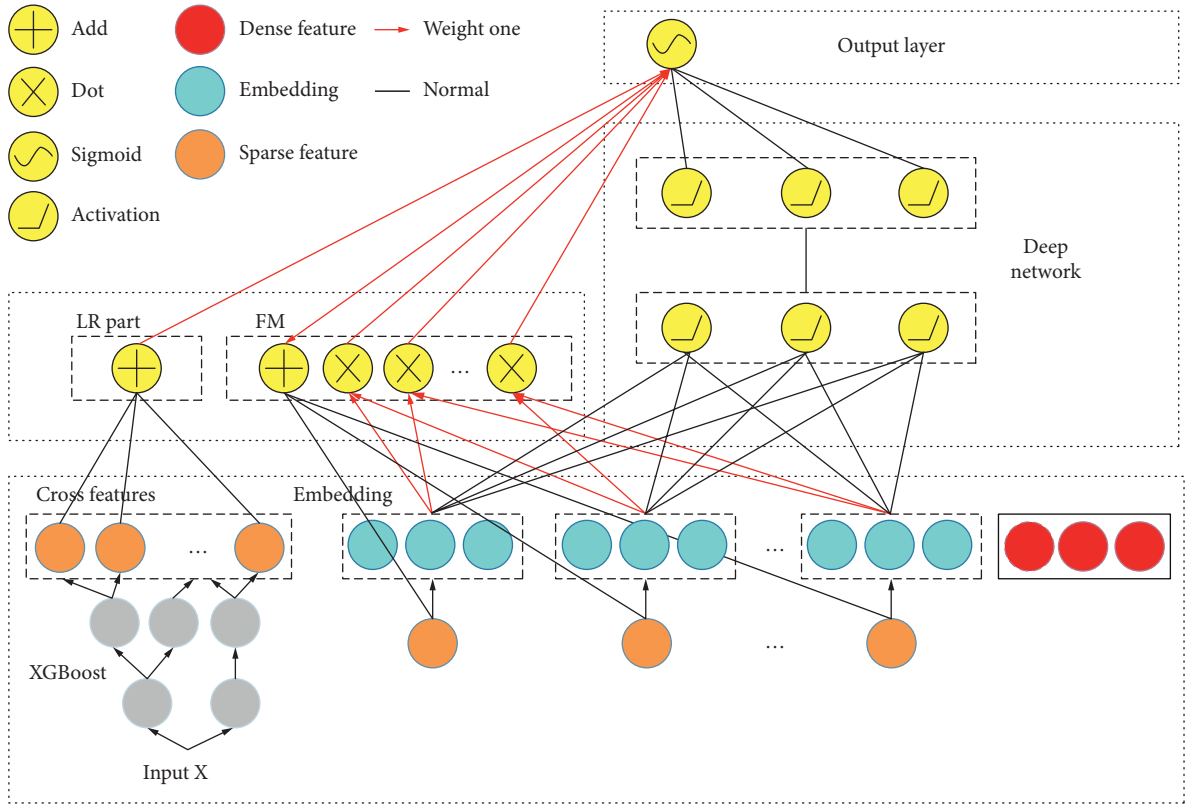


Figure 1: XGBDeepFM model.

advertising. After the feature selection of the XGBoost, the FM and deep components share the same feature embedding.

3. Experiment

3.1. Datasets. The dataset used in this study contains the O2O mobile ad data from a mobile Internet platform, which provides users with local life service information. The dataset covers offline scenes, such as catering, supermarkets, convenience stores, takeout, beauty salons, and cinemas. The platform not only provides rich ad information but also offers users' explicit and implicit behavior information. Such abundance of data brings great convenience in CTR prediction. The original experimental dataset contains attributes such as shop information, users' payment log, and users' browsing log in 2016 (see Tables 1 and 2).

Mobile contextual ad CTR has a significant relationship with weather. Thus, we also crawl weather information from a weather platform named http://WunderGround.com. The platform is a reliable source of historical weather forecast information on a global scale. In this study, 4,369,918 precise historical weather data of 122 cities on day and hour levels are crawled, as shown in Tables 3 and 4, respectively.

3.2. Evaluation Metrics. We use the area under the ROC curve (*AUC*) as our evaluation metric because it is not bias on the size of test or evaluation data. *AUC* measures the likelihood that given two random points, one from the positive and one from the negative class, the classifier will rank the point from the positive class higher than the one from the negative one. The larger the *AUC* is, the more accurate the CTR prediction of mobile advertising will be.

$$AUC = \sum_{i \in (P+N)} \frac{(TPR_i + TPR_{i-1})(FPR_i - FPR_{i-1})}{2}. \quad (7)$$

3.3. Feature Selection by XGBoost. A benefit of using XGBoost is that, after the boosted trees are constructed, importance scores that indicate how useful or valuable each feature is in the construction of the boosted decision trees within the model can be easily obtained. Thus, we choose XGBoost because this model is easily interpretable by human experts. Moreover, the depth of the decision tree can decide the order of feature interaction, which can make up for the FM and DNN components. We plot the feature importance calculated by the XGBoost model, as shown in Figure 2.

The ranking results of the feature importance show that contextual features are of high importance. For example, from the perspective of temporal features, the week ranks fourth in the importance of the model; from the perspective of geographical features, geographical location features rank ten; and from the perspective of temperature contextual features, pressure and body temperature, which are important features, rank second. This feature selection focuses on the integration of mobile ad bilateral factors (i.e., ad and user factors) and contextual factors, as shown in Table 5.

Table 1: Mobile ad shop information.

Field	Sample	Description
shop_id	000001	Shop id
city_name	Shenzhen	City name
location_id	001	Neighbour shops have the same location id
per_pay	3	Average pay
Score	1	Shop score
comment_cnt	2	Users' comment number
shop_level	1	Shop level
cate_1_name	Food	Category name
cate_2_name	Snack	Category name
cate_3_name	Other	Category name

Table 2: User payment behavior/browsing behavior.

Field	Sample	Description
user_id	0000000001	User id
shop_id	000001	Shop id
time_stamp	2016-10-10 11:00:00	Payment/browsing timestamp

Table 3: Weather data (day level).

Field	Sample	Description
city_name	Shenzhen	City name
Date	2016-01-01	Date
d_weather	Cloud	Weather
d_temp	11°C/4°C	Highest/lowest temperature
d_wind	SE	Wind direction

Table 4: Weather data (half-hour level).

Field	Sample	Description
city_name	Shenzhen	City name
Date	2016-01-01	Date
Time	12:30 AM	Time
h_temp	26.0	Temperature
h_bodytemp	24.0	Body temp
h_dew	24.0	Dew
h_humidity	89%	Humidity
h_pressure	1011	Pressure
h_visibility	10.0	Visibility
h_wind_dir	South	Wind direction
h_wind_speed	18.5	Wind speed
h_gust_speed	19	Gust speed
h_condition	Clouds	Weather condition

3.4. Model Comparison. We initially compare the performance of each component of the model (i.e., first-order linear, second-order FM, DNN, and XGB components) and its combination under the optimal settings. Then, we compare our proposed method with other models, namely, W&D, FNN, PNN, and XDeepFM.

Figure 3 presents the *AUC* results of in-model comparisons. We compare the predictive performance of different models and observe the following. First, the models with the FM component are better than those without it. The linear model, linear + DNN model, linear + XGB model, and linear + DNN + XGB model are improved by 0.1121, 0.0018, 0.0587, and 0.0003, respectively, after adding the FM component. Second, the models with the DNN component are better than those without it. The linear model, linear + FM model, linear + XGB model, and linear + FM + XGB model are improved by 0.2199, 0.1086, 0.1157, and 0.0573, respectively, after adding the DNN component. Third, the models with the XGB component are better than those without it. The linear model, linear + DNN model, linear + FM model, and linear + FM + DNN improved by 0.1049, 0.0007, 0.0515 and 0.0003, respectively, after adding the XGB component. The experimental results show that the FM part, DNN part, and XGB part have significant gains on the

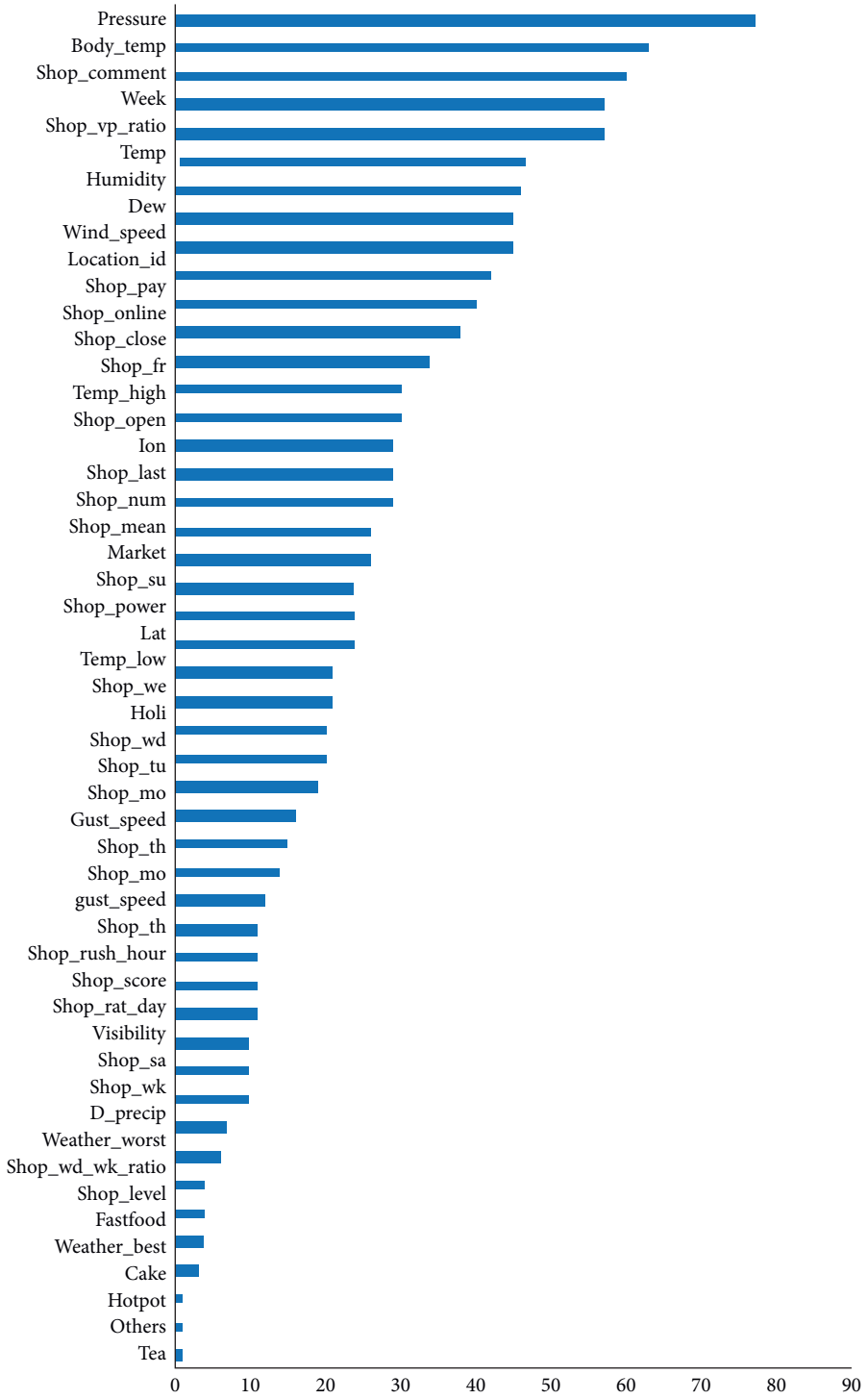


Figure 2: Feature importance ranking.

Table 5: Feature selection.

Category	Dimension	Description
Bilateral	Ad	Including price level, sales volume, ratings, categories, click-through rate
	User	Including user's category preference, price preference, location preference
Contextual	Time	Including seasonal characteristics, weekends, holidays, hour peak
	Geography	Including the user's active geographic location, location
	Weather	Including weather conditions, temperature, humidity, wind speed

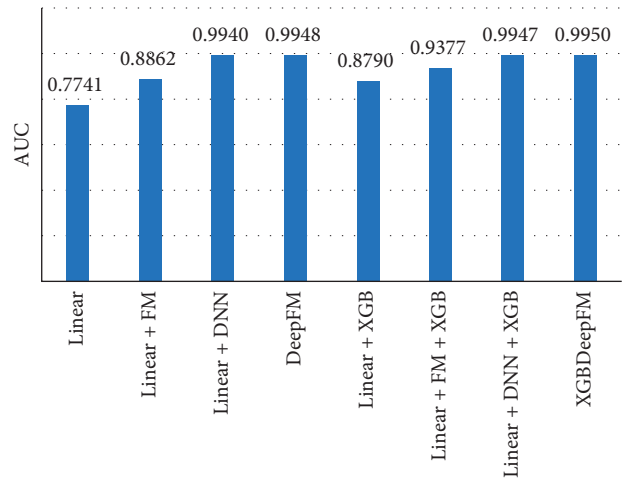


Figure 3: Prediction performance comparison of subcomponents.

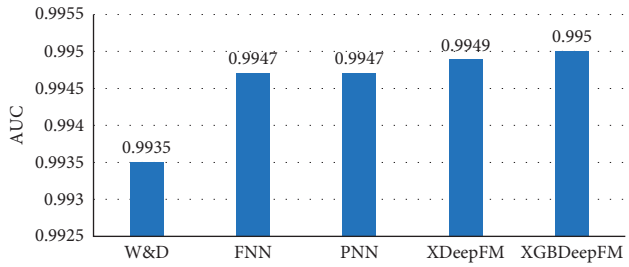


Figure 4: Prediction performance comparison.

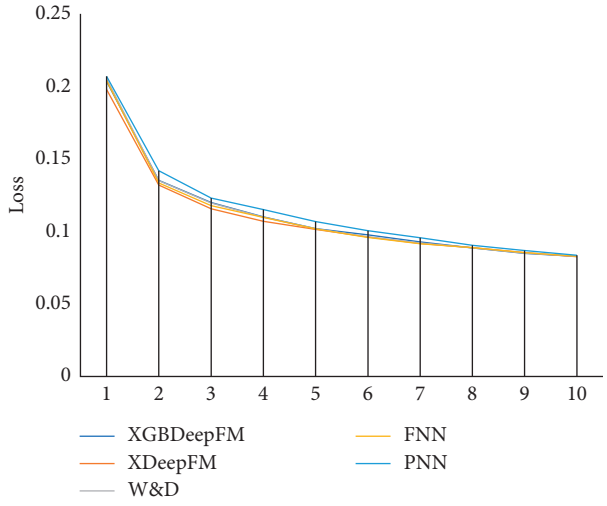


Figure 5: Convergence time comparison of XGBDeepFM with the deep learning model.

model, which are expected. The XGBDeepFM model has a strong information capacity for the CTR prediction of mobile advertising by integrating bilateral and contextual factors. Therefore, the prediction performance of the eight models is the best.

From the experimental results shown in Figure 4, XGBDeepFM is superior to all depth CTR models in terms of the *AUC* index. XGBDeepFM is 0.0015, 0.0003, 0.0003, and 0.0001 higher than W&D, FNN, PNN, and XDeepFM, respectively. Overall, the following results are obtained (Figures 3 and 4):

(1) Learning feature interactions instead of learning only linear features improves the performance of CTR prediction

(2) Learning low- and high-order feature interactions simultaneously contributes to CTR prediction

(3) Learning more important features based on the XGBoost model can improve the performance of a CTR prediction model

Figure 5 shows the comparison results of the convergence time of different models. The results show that the convergence speed of the XGBDeepFM algorithm is faster than that of PNN and FNN, only next to XDeepFM; the loss is the lowest after the 10th round of training.

4. Conclusions

In CTR prediction, the contextual features and interactions among ad, user, and contextual features are key factors that can affect the prediction performance. In this study, we propose the XGBDeepFM model. We initially include information on contextual features to improve the prediction accuracy from the perspective of time, geography, and weather. Then, a feature selection process is conducted to obtain important features. Low- and high-order features are obtained using the proposed XGBDeepFM model. We conduct extensive experiments to compare the effectiveness and efficiency of XGBDeepFM with other methods. Our experiment results demonstrate that (1) XGBDeepFM outperforms the state-of-art models in terms of *AUC*, and (2) the efficiency of XGBDeepFM outperforms most deep neural network models.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of China (grant nos. 71831005 and 71472056) and Shenzhen Key Research Base of Humanities and Social Sciences.

References

[1] H. Guo, R. Tang, Y. Ye et al., "DeepFM: a factorization-machine based neural network for CTR prediction," 2017, http://arxiv.org/abs/1703.04247.

[2] H. T. Cheng, L. Koc, J. Harmsen et al., "Wide & deep learning for recommender systems," in *Proceedings of the 1st Workshop*

on Deep Learning for Recommender Systems, pp. 7–10, Boston MA USA, September 2016.

[3] F. García-Sánchez, R. Colomo-Palacios, and R. Valencia-García, "A social-semantic recommender system for advertisements," *Information Processing & Management*, vol. 57, no. 2, p. 102153, 2020.

[4] Q. Wang, F. Liu, S. Xing, and X. Zhao, "Research on CTR prediction based on stacked autoencoder," *Applied Intelligence*, vol. 49, no. 8, pp. 2970–2981, 2019.

[5] J. Chen, B. Sun, H. Li et al., "Deep Ctr prediction in display advertising," in *Proceedings of the 24th ACM International Conference on Multimedia*, pp. 811–820, Amsterdam, The Netherlands, October 2016.

[6] M. Karimi, D. Jannach, and M. Jugovac, "News recommender systems-survey and roads ahead," *Information Processing & Management*, vol. 54, no. 6, pp. 1203–1227, 2018.

[7] P. Resnick, N. Iacovou, M. Suchak et al., "GroupLens: an open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, pp. 175–186, Chapel Hill, NC, USA, October 1994.

[8] W. Carrer-Neto, M. L. Hernández-Alcaraz, R. Valencia-García, and F. García-Sánchez, "Social knowledge-based recommender system. Application to the movies domain," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10990–11000, 2012.

[9] F. M. Belém, A. G. Heringer, J. M. Almeida, and M. A. Gonçalves, "Exploiting syntactic and neighbourhood attributes to address cold start in tag recommendation," *Information Processing & Management*, vol. 56, no. 3, pp. 771–790, 2019.

[10] L. Tang, Y. Jiang, L. Li et al., "Ensemble contextual bandits for personalized recommendation," in *Proceedings of the 8th ACM Conference on Recommender Systems*, pp. 73–80, Foster City, SV, USA, October 2014.

[11] C. C. Aggarwal, *An Introduction to Recommender systems*, pp. 1–28, Springer, Cham, Switzerland, 2016.

[12] T. Ragunathan, S. K. Battula, V. Jorika, C. Mounika, A. U. Sruthi, and M. D. Vani, "Advertisement posting based on consumer behaviour," *Procedia Computer Science*, vol. 50, pp. 329–334, 2015.

[13] H. B. McMahan, G. Holt, D. Sculley et al., "Ad click prediction: a view from the trenches," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1222–1230, Chicago, IL, USA, August 2013.

[14] S. Rendle and L. Schmidt-Thieme, "Pairwise interaction tensor factorization for personalized tag recommendation," in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pp. 81–90, New York NY, USA, February 2010.

[15] Y. Zhang, H. Dai, C. Xu et al., "Sequential Click Prediction for Sponsored Search with Recurrent Neural networks," in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, Québec City, Canada, July 2014.

[16] Y. Qu, H. Cai, K. Ren et al., "Product-based neural networks for user response prediction," in *Proceedings of the 2016 IEEE 16th International Conference on Data Mining*, IEEE, Barcelona, Spain, pp. 1149–1154, December 2016.