


# Emergency Drug Procurement Planning Based on Big-Data Driven Morbidity Prediction

Qin Song, Yu-Jun Zheng , Senior Member, IEEE, Yu-Jiao Huang, Zhi-Ge Xu, Wei-Guo Sheng , Member, IEEE, and Jun Yang 

**Abstract**—Due to the uncertainty of diseases, traditional approaches of drug procurement planning in hospitals often cause drug overstocking or understocking, which can have strong negative effects on healthcare services. This paper proposes a big-data driven approach, which uses a deep neural network to predict morbidities of acute gastrointestinal infections based on a huge amount of environmental data, and then constructs an optimization problem of drug procurement planning for maximizing the expected therapeutic effect on the predicted cases. The problem is solved by an efficient heuristic optimization algorithm. Computational experiments demonstrate the performance advantages of both the deep learning model and the heuristic algorithm over existing ones, and two real case studies in Central China show that the average prediction error of our approach is only 8% and the estimated recovery rate reaches 99%, much better than the currently used method. Our approach can also be extended for many other medical resource planning problems.

**Index Terms**—Big-data, drug procurement planning, deep learning, heuristic optimization, morbidity prediction, water wave optimization (WWO).

## I. INTRODUCTION

**D**RUG procurement is an important part of medical management in hospitals for keeping appropriate drug inventory levels to support medical activities. However, in existing ap-

Manuscript received May 31, 2018; revised September 7, 2018; accepted September 10, 2018. Date of publication September 17, 2018; date of current version December 3, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61473263, Grant 61573316, and Grant 61872123 and in part by the Natural Science Foundation of Zhejiang Province, China under Grant LY18F030023. Paper no. TII-18-1394. (Corresponding author: Yu-Jun Zheng.)

Q. Song is with the Scientific Research Institute, Hangzhou Normal University, Hangzhou 311121, China (e-mail: songqin@zjut.edu.cn).

Y.-J. Zheng is with the Hangzhou Institute of Service Engineering, Hangzhou Normal University, Hangzhou 311121, China, and also with the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China (e-mail: yujun.zheng@computer.org).

Y.-J. Huang and Z.-G. Xu are with the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China (e-mail: hjj0507@zjut.edu.cn; 3349634319@qq.com).

W.-G. Sheng is with the Hangzhou Institute of Service Engineering, Hangzhou Normal University, Hangzhou 311121, China (e-mail: w.sheng@ieee.org).

J. Yang is with the Medical College, Hangzhou Normal University, Hangzhou 311121, China (e-mail: gastate@zju.edu.cn).

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

Digital Object Identifier 10.1109/TII.2018.2870879

proaches of drug procurement planning, drug requirements are mainly estimated based on past experiences [1]–[5], which often leads to shortage of some drugs and overstocking of others—the latter can greatly increase the cost and cause waste of inventory and human resources, while the former can significantly decrease the ability of medical services.

In the literature, a few studies have been devoted to the use of mathematical or artificial intelligence methods, such as correlation analysis [6], [7], regression analysis [8], [9], pattern mining and matching [10], [11], and artificial neural networks (ANNs) [12], to predict drug consumption. However, these methods are mainly suitable for those diseases whose incidence rates are relatively stable or with obvious periodicity, and their prediction accuracies will significantly decrease when there are sudden changes of the external environment. For example, environmental pollution accidents may cause outbreaks of some diseases and surge of drug consumption, which will soon lead to stock-outs of some emergency drugs. In a traditional drug procurement planning, such changes are rarely taken into consideration.

The primary cause of the difficulty in drug requirement estimation and procurement planning is the uncertainty of diseases, that is, most diseases have many causal factors that are difficult to identify, measure, and evaluate, and thus it can be very difficult to predict their morbidities, especially in changing environments. In recent years, there has been an increased interest in the use of machine learning methods for morbidity prediction in order to improve healthcare services. Ture and Kurt [13] compare three ANNs and an autoregressive integrated moving average model to hepatitis A virus (HAV) forecasting, and their tests in forecasting time series from 1992 to 2004 in Turkey show that the basic multilayer perceptron network is more accurate than the others. Dmitriev and Kotin [14] use a simple three-layer ANN to predict short-term and long-term scarlet fever morbidities. Nevertheless, such time series predictions based on historical data do not consider etiologically influential factors and thus can also be very inaccurate in changing environments.

Bibi *et al.* [15] use an ANN to predict the effect of atmospheric changes on emergency department visits for respiratory symptoms, and employ a genetic algorithm (GA) to optimize the architecture and parameters of the network. Wu *et al.* [16] apply an ANN to predict the incidence of hemorrhagic fever with renal syndrome in Shenyang, China during 1984–2003, using input features, such as average temperature, relative humidity, precipitation and sunshine time, and epidemiologic information, such as rat density and viral carriage of rats. Wang *et al.* [17] study the

Granger causality between main air pollutants and respiratory mortality, and use an ANN to predict the mortality from 2005 to 2008 in Beijing. Junk *et al.* [18] conduct a similar study on the use of ANN for predicting respiratory morbidities associated with air pollution. Ma and Wang [19] predict meteorological morbidity from factors such as average humidity, temperature swing of 48 h, daily temperature range, and air pressure, and the experiments show that the prediction accuracy of ANN is close to 60%. However, these studies are aimed at a minority of diseases whose causal factors are relatively simple and easy to measure. For most diseases with complex and/or uncertain causal factors, their morbidity prediction is still challenging.

Recent development of big-data and deep-learning technologies provides great potentials for improving healthcare services in many aspects [20]–[25] and opens up a new way for predicting morbidities of complex diseases by automatically learning complex probabilistic relationship between the morbidities and a large number of (potentially) causal factors. Rodger [26] shows that using big data analytics on a Hadoop Hive data warehouse infrastructure can improve the prediction accuracy of survival, mortality, and morbidity of traumatic brain injury as well as other injury cases, and survival rates. Song *et al.* [27] use a deep neural network (DNN) to model the effect of food contamination on gastrointestinal infections, and show that its average prediction accuracy is close to 80% on the datasets from four counties in China from 2015 to 2016. Song *et al.* [28] extend the approach for predicting gastrointestinal infection morbidity based on 129 types of pollutants contained in soil and water, the results of which still have a high accuracy. However, the two models yield an overall morbidity for all types of gastrointestinal infections, i.e., they do not differentiate the morbidities of different types of gastrointestinal diseases. As different diseases require different types and amounts of drugs, the results of the models cannot be directly utilized for drug procurement planning.

In this paper, we extend the deep learning models of [27], and [28] for simultaneously predicting morbidities of ten types of acute gastrointestinal diseases based on historical morbidities, food contaminations, and predicted temperatures. As there are many alternative drugs that can be used for the same disease, we establish an optimization problem of drug procurement planning, which determines the types and amounts of drugs for each disease, such that the total expected therapeutic effect of all predicted disease cases is maximized under the procurement cost constraint. We then propose a heuristic algorithm based on water wave optimization (WVO) [29] for efficiently solving the problem. Fig. 1 presents the overall structure of our big-data driven approach for emergency drug procurement planning. Experimental results and real-world case studies demonstrate the feasibility and effectiveness of the proposed approach. The main contributions of this paper can be summarized as follows.

- 1) We propose a new evolutionary deep learning model for predicting morbidities based on a large number of potential influence factors.
- 2) We propose a new discrete WVO algorithm for drug procurement planning with the aim of maximizing the expected therapeutic effect on the predicted cases.

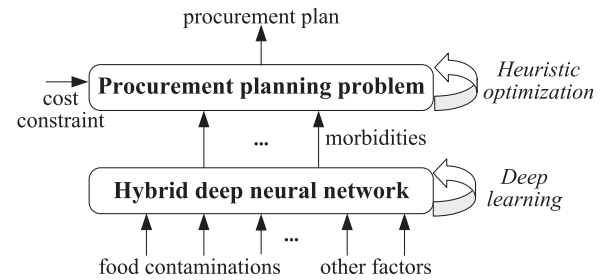


Fig. 1. Basic framework of the proposed big-data driven approach for emergency drug procurement planning.

- 3) The combination of the learning model and the optimization algorithm provides a novel and effective way to support medical preparation for disease control.

In the rest of this paper, Section II proposes our deep learning model for morbidity prediction, Section III presents the drug procurement planning problem, and Section IV proposes the WVO algorithm for the problem. Section V describes the experiments and case studies, and Section VI concludes.

## II. ACUTE GASTROINTESTINAL INFECTION MORBIDITY PREDICTION BASED ON DEEP LEARNING

### A. Model Inputs and Outputs

The model is to predict acute gastrointestinal morbidities based on a large number of contributing (or potentially contributing) factors. Ten types of common acute gastrointestinal diseases are considered, as shown in Table I. As the pathogenesis and dosages for children are often different from that for adults, for each disease we, respectively, calculate two morbidities, one for children and another for adults, and thereby the model has 20 output variables, denoted as  $o_1$ – $o_{20}$ . The prediction period is one week, i.e., the model predicts the morbidities of the diseases in the next week.

The inputs to the model can be divided into three parts as follows.

- 1) Food contaminant indices obtained in the current week (if a contaminant has been sampled and measured more than once in the week, the average value is used). This study aims at acute gastrointestinal morbidities in Central China, for which we select 123 types of food (summarized in Table II) and 264 types of contaminants (summarized in Table III). Different food has different types of contaminants, and the number of contaminant indices is 11 208. The corresponding input variables are denoted as  $v_1 \sim v_{11\,208}$ . The inputs can be modified according to different diseases and regions.
- 2) The predicted average, lowest, and highest temperatures in each day of the next week, denoted by  $t_1 \sim t_7$ ,  $t_1^L \sim t_7^L$ , and  $t_1^H \sim t_7^H$ , respectively.
- 3) The historical morbidities, including the morbidities in the previous three weeks, denoted by  $p_{i,1}$ ,  $p_{i,2}$ , and  $p_{i,3}$ , and morbidities in the corresponding week of the previous three years, denoted by  $q_{i,1}$ ,  $q_{i,2}$ , and  $q_{i,3}$  ( $1 \leq i \leq 20$ ).

**TABLE I**  
TYPES OF ACUTE GASTROINTESTINAL DISEASES CONSIDERED IN THE MODEL

1	2	3	4	5	6	7	8	9	10
acute gastritis	Crohn's disease	acute appendicitis	acute peptic ulcer (APU)	gastrointestinal bleeding	irritable bowel syndrome (IBS)	bacillary dysentery	typhoid fever	paratyphoid fever	food poisoning

**TABLE II**  
TYPES OF FOOD FOR ACUTE GASTROINTESTINAL MORBIDITY PREDICTION IN THIS STUDY

Class	Food
Cereals	rice, wheat, barley, corn, millet, black rice, sticky rice
Beans	soybean, mung soybean, red bean, black bean, broad bean, pea, cow pea, hyacinth bean, kidney bean, sword bean
Vegetables	Chinese cabbage, pakchoi cabbage, baby cabbage, celery cabbage, celery, lettuce, broccoli, Chinese broccoli, mustard leaf, leaf lettuce, okra, rape, spinach, water spinach, potherb mustard, amaranth, cauliflower, purslane, yam, carrot, celtuce, summer radish, tomato, cucumbers, lappa, radish, potato, sweet potato, pumpkin, bitter melon, white melon, loofah, chili pepper, bell pepper, green pepper, sweet pepper, pod pepper, pea sprout, soybean sprout, mung bean sprout, Chinese toon sprout, shiitake, button mushroom, oyster mushroom, needle mushroom, agaric, day lily, tremella, spring onion, Chinese onion, ginger, caraway, garlic, fragrant-flowered garlic, garlic sprouts
Fruits	apple, gala apple, bergamot pear, snow pear, mili pear, pineapple, orange, navel orange, pomelo, peach, nectarine, melon, watermelon, Hami melon, apricot, plum, cherry, bayberry, grape, longan, lychee, winter jujube, red jujube, sugarcane, pitaya
Meals & eggs	pork, beef, mutton, chicken, duck, egg, duck egg, quail egg
Aquatic	kelp, laver, carp, grass carp, yellow croaker, perch, crucian, pomfret, parabramis pekinensis, mandarin fish, prawn, crab, lobster, clam, abalone, holothurian, river prawn, river crab, river snail

**TABLE III**  
TYPES OF CONTAMINANTS USED FOR ACUTE GASTROINTESTINAL MORBIDITY PREDICTION IN THIS STUDY

Class	Subclass	Contaminants
Inorganic Contaminants	Heavy metals	Pb, Cd, Hg, Cu, Ni, As, Be, Bi, Sb, Tl, Cr, Mo, Zn, Au, Ag, Fe, V
	Others	cyanide, nitrate, nitrite, sulfate, carbonate
Organic Contaminants	Hydrocarbons	BTEX, PAHs, TPH
	Halogenated	HCFCs, chlorinated solvents, PCBs, dioxin
	Oxygenated	alcohols, phenols, ethers, esters, phthalate
	Dyes	Azo, QACs, benzidine, naphthylamine
	Pesticides	66 commonly used pesticides (see [33])
	Herbicides	18 commonly used herbicides (see [34])
	Fungicides	azoxystrobin, pyraclostrobin, prothioconazole, trifloxystrobin, copper fungicides, cyproconazole, mancozeb, epoxiconazole, tebuconazole, boscalid, metalaxyl, picoxystrobin, difenoconazole, propiconazole, chlorothalonil
	Endocrine disruptors	68 chemicals (see [35])
Pathogenic organisms	Others	PP, PPE, PS, PAEs, TCE, OCP
	Bacteria	salmonella, shigella, dysentery bacillus, plague bacillus, tubercle bacillus, typhoid bacillus, diphtheria bacillus, Francisella tularensis, Brucella, vibrio parahaemolyticus, vibrio cholerae, vibrio mimicus, vibrio fluvialis, clostridium tetani, clostridium botulinum, clostridium perfringens, staphylococcus aureus, Bacillus anthracis, Escherichia coli, Yersinia, helicobacter pylori, campylobacter jejuni, aeromonas hydrophila, roundworm eggs, hookworm eggs
	Fungi	candida albicans, aspergillus fumigatus, mucor racemosus
Antibiotics	Virus	rotavirus, norovirus, sapovirus, astrovirus
	$\beta$ -lactams	penicillin, amoxicillin, ceftriaxone, aztreonam, tienam
	Aminoglycoside	Amikacin, tobramycin, gentamycin, netilmicin, sisamicin, streptomycin
	Macrolide	erythromycin, roxithromycin, clarithromycin, azithromycin
	Polypeptide	polymyxin, vancomycin, Amiomycin
Others	tetracycline, quinolone, lincomycin, streptomycin, amphotericin, fosfomycin, nitrazole	

Note that the first two parts of the inputs are used for predicting all  $o_1 \sim o_{20}$ , while in the third part  $p_{i,1} \sim p_{i,3}$  and  $q_{i,1} \sim q_{i,3}$  are just used for predicting the corresponding  $o_i$  ( $1 \leq i \leq 20$ ). In comparison with others, the influences of the temperatures and historical morbidities on the future morbidities are relatively simple, while the influences of the contaminants can be very complex and uncertain. It is estimated that the volume of data to be collected is approximately 150~300 MB per week in a medium-sized county and 1~2 GB per week in a prefecture-level city. The total volume in a large area, accumulated over years, can be very huge. Moreover, there are a large variety of contaminant indices, which change constantly and contain

much noise and missing values. For such a highly complex big data analytics task, deep learning has inherent advantages of data abstraction and hierarchical learning. Therefore, we propose a DNN model that consists of two parts: The first is a deep denoising autoencoder (DDAE) [30] for learning the complex probabilistic distribution over the contaminants and the second is a Gaussian mixture model (GMM) [31], [32] for predicting the morbidities based on the outputs of the DDAE and the temperatures and historical morbidities. The overall structure of the hybrid DNN is illustrated in Fig. 2, and its two parts are, respectively, described in the following Sections II-B and II-C.

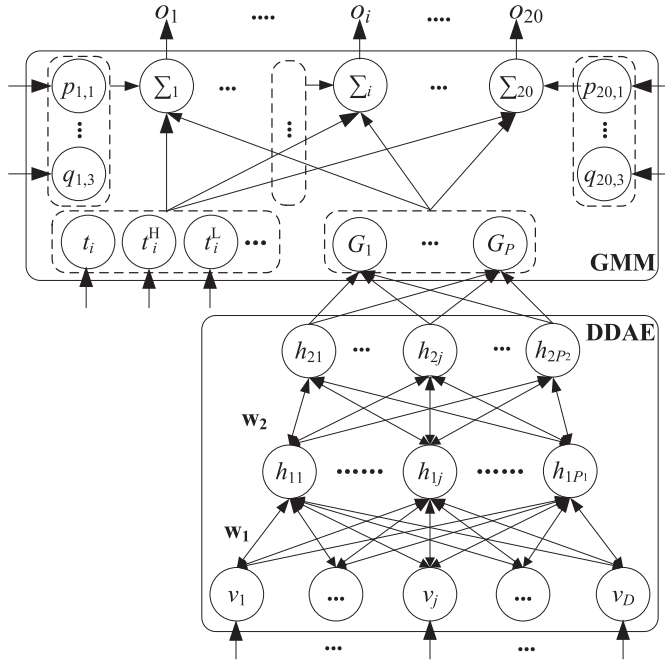


Fig. 2. Hybrid DNN for acute gastrointestinal morbidity prediction.

### B. Deep Denoising Autoencoder

The DDAE is used for modeling highly complex and uncertain influence effects of food contaminants on acute gastrointestinal infections. The basic building block of DDAE is autoencoder [36], which is an ANN with a single hidden layer that attempts to reconstruct an input  $\mathbf{v} \in [0, 1]^D$  by first transforming (encoding) it to a hidden representation  $\mathbf{z} \in [0, 1]^{D'}$  through an affine mapping as follows (where  $\theta = [\mathbf{w}, \mathbf{b}]$ ,  $\mathbf{w}$  is a  $D' \times D$  weight matrix and  $\mathbf{b}$  is a  $D'$ -dimensional bias vector):

$$f_{\theta}(\mathbf{v}) = s(\mathbf{w}\mathbf{v} + \mathbf{b}) \quad (1)$$

and then mapping (decoding)  $\mathbf{z}$  back to a reconstructed vector  $\mathbf{v}'$  in the input space with appropriately sized parameters  $\theta' = [\mathbf{w}', \mathbf{b}']$ :

$$g_{\theta'}(\mathbf{z}) = s(\mathbf{w}'\mathbf{z} + \mathbf{b}'). \quad (2)$$

The objective of autoencoder training is to minimize the average reconstruction error on the training dataset  $\mathcal{V}$ :

$$\min \frac{1}{|\mathcal{V}|} \sum_{\mathbf{v} \in \mathcal{V}} \mathcal{L}(\mathbf{v}, g_{\theta'}(f_{\theta}(\mathbf{v}))) \quad (3)$$

where  $\mathcal{L}$  is a loss function that can be the squared error

$$\mathcal{L}(\mathbf{v}, \mathbf{v}') = \|\mathbf{v} - \mathbf{v}'\|^2. \quad (4)$$

Denoising autoencoder (DAE) [37] is a variant of the basic autoencoder, which first corrupts an input  $\mathbf{v}$  into  $\tilde{\mathbf{v}}$  by means of a stochastic mapping  $q_{\mathcal{V}}(\tilde{\mathbf{v}}|\mathbf{v})$ , transforms the corrupted input  $\tilde{\mathbf{v}}$  to a hidden representation  $\mathbf{z} = f_{\theta}(\tilde{\mathbf{v}})$ , and then reconstructs a clean “repaired” input  $\mathbf{v}' = g_{\theta'}(\mathbf{z})$  as the basic autoencoder. The objective of DAE training is still to minimize the average reconstruction error, but  $\mathbf{v}'$  is now a deterministic function of  $\tilde{\mathbf{v}}$

and thus the result of a stochastic mapping of  $\mathbf{v}$

$$\min \frac{1}{|\mathcal{V}|} \sum_{\mathbf{v} \in \mathcal{V}} \mathcal{L}(\mathbf{v}, g_{\theta'}(f_{\theta}(\tilde{\mathbf{v}}))). \quad (5)$$

DDAE is an extension of DAE that has multiple hidden layers, where each layer uses denoising learning to capture complex correlations among the hidden features in the layer below. In this way, the whole DDAE gradually reduces a natural high-dimensional input to a low-dimensional manifold through a set of automatically discovered intermediate abstractions [30]. Note that input corruption is only used for the initial denoising learning of each individual layer. Once the mapping  $f_{\theta}$  has been learnt, it will henceforth be used on uncorrupted inputs. No corruption is applied to produce the representation that will serve as clean input for training the next layer.

### C. Gaussian Mixture Model

The GMM is built on the top of the DDAE for calculating each output morbidity  $o_i$  from the hidden feature vector  $\mathbf{z}$  of the top-most layer of the DDAE, the set of predicted temperatures  $\mathbf{t} = \{t_1, \dots, t_7, t_1^L, \dots, t_7^L, t_1^H, \dots, t_7^H\}$ , and the set of corresponding historical morbidities  $\mathbf{p}_i = \{p_{i,1}, p_{i,2}, p_{i,3}, q_{i,1}, q_{i,2}, q_{i,3}\}$  ( $1 \leq i \leq 20$ ):

$$o_i = \frac{1}{|\mathbf{z}| + |\mathbf{t}| + |\mathbf{p}_i|} \sum_{z_i \in \mathbf{z} \cup \mathbf{t} \cup \mathbf{p}_i} \log \left( \sum_{j=1}^{N_G} w_j \mathcal{N}(z_i; \mu_j, \sigma_j) \right) \quad (6)$$

where  $\mathcal{N}(z_i; \mu_j, \sigma_j)$  is a high-dimensional Gaussian function with mean  $\mu_j$  and diagonal covariance matrix  $\sigma_j$ ,  $N_G$  is the number of Gaussians, and  $w_j$  is the weight for Gaussian  $j$  subject to  $\sum_{j=1}^{N_G} w_j = 1$ .

### D. Model Training

The training of the hybrid DNN consists of two steps: pre-training of the DDAE and supervised training of the whole DNN. As the input dimension of the DDAE is very high (11 208 in our study), the classical greedy layer-wise algorithm [38] is easy to be trapped by local optima, and thus we employ the evolutionary learning algorithm proposed in [27] based on biogeography-based optimization [39] and its improved version [40]–[42] for both the steps. In DDAE pretraining, the algorithm is performed to minimize (3) layer by layer. After all the layers of the DDAE have been pretrained, the algorithm is performed on the whole hybrid DNN to minimize the average error between the predicted morbidities  $o_i(\mathbf{v})$  and actual morbidities  $\hat{o}_i(\mathbf{v})$  on the training dataset  $\mathcal{D}$

$$\min \frac{1}{20|\mathcal{V}|} \sum_{\mathbf{v} \in \mathcal{V}} \sum_{i=1}^{20} \mathcal{L}(o_i(\mathbf{v}), \hat{o}_i(\mathbf{v})). \quad (7)$$

As described earlier, the proposed DNN uses three parts of inputs for producing the output morbidities. Nevertheless, there can be many other factors, such as sanitary conditions and eating habits, that may have influences on the morbidities. Thus, it should be noted that we typically construct an instance of the DNN for predicting the morbidities in a given region, and use

samples from the region to train the model instance. In this way, other influence factors related with the region are implicitly built into the model [43]. In general, a model instance trained for one region is not suitable for predicting morbidities in another region.

### III. DRUG PROCUREMENT PLANNING PROBLEM

After estimating the morbidity  $o_i$  of each disease  $i$  in the next week, we can transform it to the number  $g_i$  of cases in the hospital based on the population size of the region as well as the portion of medical services carried by the hospital. If there is a significant increase in the cases of some diseases and the stock is insufficient to handle these cases, an emergency procurement is needed. After data abstraction and extraction by deep learning, the optimization problem of drug procurement planning built upon the predicted cases is much more manageable and much easier to solve than that built upon the raw big data.

Suppose that there are  $n$  types of drugs related with the acute gastrointestinal diseases, for each  $1 \leq i \leq 20$  and  $1 \leq j \leq n$  we have one of the following three possibilities:

- 1) drug  $j$  is not required in the treatment of disease  $i$ ;
- 2) drug  $j$  must be used in the treatment of disease  $i$ , and the amount required is denoted by  $r_{ij}$ ; and
- 3) the use of drug  $j$  is alternative in the treatment of disease  $i$ .

Considering the last possibility, we can identify  $K_i$  treatment items for each disease  $i$ , where each item  $k$  involves a set of alternative drugs. For example, the treatment items of acute gastritis (adults) include spasm and pain relief, diarrhea control, vomiting control, antiseptic treatment, and oral anti-inflammatory treatment. Candidate drugs for spasm and pain relief include atropine, propantheline, belladonna, anisodamine, etc. Let  $b_{ikj} = 1$  denote that drug  $j$  is a candidate for the  $k$ th treatment item of disease  $i$  and  $b_{ikj} = 0$  otherwise. We are also given a normalized dosage  $r_{ik}$  required by the  $k$ th item of disease  $i$ . The problem needs to decide  $x_{ikj}$ , the amount of drug  $j$  for the  $k$ th item of disease  $i$ , which must satisfy

$$\sum_{j=1}^n b_{ikj} x_{ikj} = g_i r_{ik}, \quad 1 \leq i \leq 20; 1 \leq k \leq K_i. \quad (8)$$

The problem objective is to maximize the total expected therapeutic effect of all disease cases subject to that the total procurement cost does not exceeds a predefined limit  $C$ . The consumption of drug  $j$  is

$$y_j = \sum_{i=1}^{20} \left( g_i r_{ij} + \sum_{k=1}^{K_i} x_{ikj} \right). \quad (9)$$

Let  $p_j$  be the unit price of drug  $j$ ,  $s_j$  be the current amount of the stock of drug  $j$ , and  $\hat{s}_j$  be the lower limit of  $s_j$  (that must be maintained for other diseases or unexpected cases). Thus, the amount of drug  $j$  to be purchased is

$$y'_j = \max(y_j - s_j + \hat{s}_j, 0) \quad (10)$$

which should not exceed the available supply  $a_j$  of drug  $j$  in the market

$$y'_j \leq a_j \quad (11)$$

and the cost constraint can be expressed as

$$c(\mathbf{x}) = \sum_{j=1}^n p_j y'_j \leq C. \quad (12)$$

Now, let us consider the therapeutic effects of alternative drugs. We use an index  $e_{ikj}$  to represent the therapeutic effect of drug  $j$  in the  $k$ th treatment item of disease  $i$ . The value is in the range of  $[0,1]$ : The higher the value, the better the effect. For example, based on a survey of gastroenterologists, the  $e_{ikj}$  values of atropine, propantheline, and belladonna for the ‘‘spasm and pain relief’’ item of acute gastritis (adults) are, respectively, set to 0.78, 0.65, and 0.8. We also use a weight  $w_{ik}$  to represent the importance of the  $k$ th item in disease  $i$  subject to  $\sum_{k=1}^{K_i} w_{ik} = 1$ . Therefore, for each disease case, if the drug for each  $k$ th item has been determined (and denoted as  $j_k$ ), the therapeutic effect of the case is evaluated as  $\sum_{k=1}^{K_i} w_{ik} e_{ikj_k}$ .

Here, we employ a first-come-first-served queuing model for determining drug  $j_k$  for each item of each disease case and thus evaluating its therapeutic effect. Given the number of cases  $g_1 \sim g_{20}$  in the next week, we divide the week into  $S$  time slots and assume that the cases are uniformly distributed in these time slots (the length of a slot typically varies from 15 to 60 min, depending on the number of cases and the reception ability of the hospital). Consequently, the number of cases in each slot is  $\bar{g}_i = g_i/S$  (if  $g_i < S$ , we can alternatively set the number to 1 and 0 in the slots). Let  $\Phi_{ik} = \{j | b_{ikj} = 1\}$ , i.e., the set of available alternative drugs that can be used for the  $k$ th item of disease  $i$ , and let the drugs in  $\Phi_{ik}$  be sorted in decreasing order of  $e_{ikj}$ . The queuing model assume that, whenever a new case arrives at the hospital, the doctor in charge always selects the most effective drug whose amount is larger than  $r_{ik}$  for each treatment item of the case. By simulating the arrival and treatment of all cases according to the procedure in Algorithm 1, we can evaluate the sum  $E_i(\mathbf{x})$  of the therapeutic effects of all cases of each diseases  $i$  ( $1 \leq i \leq 20$ ).

Thus, the objective function of the problem can be expressed as follows (where  $w_i$  is the importance weight of disease  $i$ ):

$$\max f(\mathbf{x}) = \sum_{i=1}^{20} w_i E_i(\mathbf{x}). \quad (13)$$

### IV. DISCRETE WWO ALGORITHM FOR THE PROBLEM

As we can see, the above-mentioned problem is an integer programming problem that is known to be *NP*-hard. Due to the large number of diseases and drugs, the solution space is huge, and the evaluation of its objective function involves the queuing procedure. So, it is impractical to use traditional exact algorithms to solve the problem. We have used some classical heuristic algorithms such as GA, but found the problem-solving performance is still unsatisfactory.

Here, we propose a heuristic algorithm based on the WWO metaheuristic [29] for the problem. WWO evolves a population

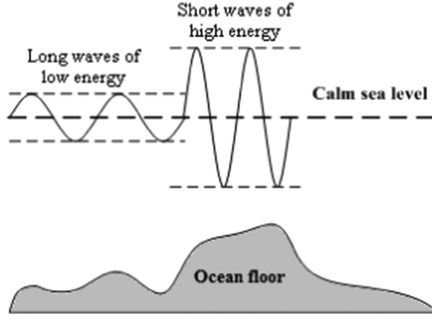


Fig. 3. Illustration of solution fitness (energy) and wavelengths in WWO.

**Alg. 1:** The procedure for drug selection and therapeutic effect evaluation based on the queuing model.

```

1 for  $i = 1$  to 20 do initialize  $E_i = 0$ ;
2 for  $j = 1$  to  $n$  do
3   let  $s'_j = y'_j + s_j - \hat{s}_j$  // available amount of drug  $j$ 
4 for  $s = 1$  to  $S$  do
5   for  $i = 1$  to 20 do
6     for  $g = 1$  to  $\bar{g}_i$  do
7       for  $k = 1$  to  $K_i$  do
8         let  $d = 1, j_k = \Phi_{ik}(d)$ ;
9         while  $s'_{j_k} < r_{ik}$  do
10           $d \leftarrow d + 1$ ;
11           $j_k \leftarrow \Phi_{ik}(d)$  // choose the most
           effective drug available
12           $s'_{j_k} \leftarrow s'_{j_k} - r_{ik}$  // consume the drug
13           $E_i \leftarrow E_i + w_{ik} e_{ikj_k}$ ;
14 return all  $E_i$ .
```

of solutions, each of which is analogous to a wave assigned with a wavelength proportional to its fitness. The fitter the solution, the smaller the wavelength, and the smaller the range it searches, as illustrated in Fig. 3. ; In this way, the algorithm can achieve a good dynamic balance between diversification and intensification.

For the considered problem, each solution vector consists of  $(\sum_{i=1}^{20} K_i)$  parts, the length of each part is  $|\Phi_{ik}|$ , and each component is an integer denoting the amount of drug assigned for the  $k$ th item of disease  $i$ . Initially, the algorithm initializes a population of  $N$  solutions, each of which is created by randomly distributing  $r_{ik}$  to  $|\Phi_{ik}|$  parts for all  $1 \leq i \leq 20$  and  $1 \leq k \leq K_i$ . The wavelength of each solution is initialized to 0.5. The algorithm then iteratively performs the following steps to evolve the solutions.

- 1) Calculate the objective function  $f(\mathbf{x})$  for each solution  $\mathbf{x}$ . Moreover, if the solution violates the constraint (12), a penalty item is added to the function as follows (where  $M$  is a large positive constant):

$$f(\mathbf{x}) = f(\mathbf{x}) + M \max(c(\mathbf{x}) - C, 0). \quad (14)$$

- 2) Update the wavelength of each  $\mathbf{x}$  according to its fitness as follows:

$$\lambda(\mathbf{x}) = \lambda(\mathbf{x}) \cdot \gamma^{-(f(\mathbf{x}) - f_{\min} + \epsilon) / (f_{\max} - f_{\min} + \epsilon)} \quad (15)$$

**Alg. 2:** The local improvement procedure on solution  $\mathbf{x}$  at the  $k$ -th item of disease  $i$ .

```

1 Let  $\hat{c} = C - c(\mathbf{x})$ ;
2 Let  $d = 1, d' = \max\{1 \leq d \leq |\Phi_{ik}| : x_{ik\Phi_{ik}(d)} > 0\}$ ;
3 Let  $j' = \Phi_{ik}(d')$ ;
4 while  $d < d' \wedge p_j - p_{j'} \leq \hat{c}$  do
5   Let  $j = \Phi_{ik}(d), \hat{a}_j = a_j - y'_j$ ;
6    $x_{ikj} \leftarrow x_{ikj} + 1$ ;
7    $\hat{a}_j \leftarrow \hat{a}_j - 1$ ;
8    $x_{ikj'} \leftarrow x_{ikj'} - 1$ ;
9    $\hat{c} \leftarrow \hat{c} - (p_j - p_{j'})$ ;
10  if  $\hat{a}_j = 0$  then  $d \leftarrow d + 1, j \leftarrow \Phi_{ik}(d), \hat{a}_j \leftarrow a_j - y'_j$ ;
11  if  $x_{ikj'} = 0$  then  $d' \leftarrow d' - 1, j' \leftarrow \Phi_{ik}(d')$ ;
```

where  $f_{\max}$  and  $f_{\min}$  are the maximum and minimum fitness values among the population,  $\epsilon$  is a very small number to avoid division-by-zero, and  $\gamma$  is the wavelength reduction coefficient set to 1.0026 according to [29], [44].

- 3) For each  $\mathbf{x}$ , update each of its components as follows:

$$x_{ikj} = \begin{cases} \text{rand}() \cdot 0.25^{\lambda(\mathbf{x})} \cdot x_{ikj} & \text{if } \text{rand}() \leq 0.5 \\ x_{ikj} & \text{else} \end{cases} \quad (16)$$

where  $\text{rand}$  produces a random number uniformly distributed in  $[0,1]$ , such that each dimension has a probability of 0.5 of being decreased: The higher the  $\lambda(\mathbf{x})$ , and more  $x_{ikj}$  is decreased. Afterwards, the decrease is added to other dimension(s) in the same part to maintain the constraint (8).

- 4) If the updated solution is better than the original  $\mathbf{x}$ , use it to replace the  $\mathbf{x}$  in the population.
- 5) If the updated solution is a new best known solution  $\mathbf{x}^*$ , perform a local search to produce a maximum of  $(\sum_{i=1}^{20} K_i)$  neighbor solutions, each of which being obtained by using Algorithm 2 to try to replace some drugs with more effective ones at a treatment item.
- 6) If a solution has been unchanged for  $\hat{g}$  generations (where  $\hat{g}$  is a control parameter typically set to 6), replace it with a new solution randomly generated.
- 7) Repeat the above-mentioned steps ; until the stop condition is met.

## V. EXPERIMENTS

The experiments of our approach consists of three parts: the test on morbidity prediction, the test on planning problem solving, and the test on the overall efficacy by case studies.

### A. Experiments on Morbidity Prediction

We first compare the morbidity prediction accuracy of our hybrid DNN with the following three models on the dataset of 200 tuples used in [27]:

- 1) a multiple linear regression (MLR) model [45];
- 2) a classical three-layer ANN with BP training; and
- 3) a DNN using the same structure as us but using the classical greedy layer-wise training [38], denoted as DNN-G.

**TABLE IV**  
THE PERCENTAGE ERRORS OF THE COMPARATIVE MODELS FOR MORBIDITY PREDICTION ON THE TEST SET

	MLR	ANN	DNN-G	DNN
Mean	69.26%	66.51%	35.85%	<b>20.09%</b>
Std	53.82%	36.20%	17.06%	11.89%
Max	716.36%	489.05%	270.20%	133.30%
Min	10.75%	3.59%	1.92%	0.86%
SR	10.74%	17.13%	46.18%	<b>72.38%</b>

The experimental environment is a computer with Intel i7-6500 2.5GH CPU, 8GB RAM, and a NVIDIA Quadro M500M card. The models are implemented with Python, using a set of basic computational functions taken from the Keras deep learning library (<https://github.com/fchollet/keras>). We use a 5-fold cross-validation for training and testing the models. Due to the page limit, here we only present the average percentage errors and the success rates (SRs) of the four comparative models over all 200 tuples and 20 diseases in **Table IV**. The best mean value and SR among the four models are shown in boldface. We consider a morbidity prediction (of a disease on a tuple) as “successful” if its error rate is less than 15% (otherwise the result can be very misleading for medical preparedness [27]). The results show that the traditional MLR has the highest error rate (nearly 70%) and the lowest SR (only about 10%), mainly due to that the relationship between the morbidities and the food contaminants is highly nonlinear and complicated and thus cannot be appropriately addressed by linear regression. Although the classical three-layer ANN is capable of modeling nonlinear systems, its error rate is still similar to MLR (but its SR is relatively higher), because its shallow structure is ineffective to learn from high-dimensional inputs. The deep learning models achieve much lower error rates and higher SRs than the MLR and the shallow ANN. Moreover, the DNN with evolutionary learning exhibits much better performance than the DNN with greedy layer-wise training, because the gradient-based algorithm is easy to be trapped in local optima. In summary, the proposed DNN model obtains the best prediction accuracy on the test set, and its error rate (about 20%) and SR (more than 70%) are generally acceptable in practice.

### B. Experiments on Drug Procurement Planning

Next, we compare the proposed WWO algorithm with three other metaheuristic algorithms, including standard GA, particle swarm optimization (PSO) [46], and differential evolution (DE) [47], for integer programming on eight test instances of the problem. The population size of WWO is set to 12. The control parameters of other algorithms are also fine-tuned on the test set. On each instance, each algorithm is run for 20 times, all with the same maximum running time of 8 h for a fair comparison. For the ease of comparison, we also run an exact branch-and-bound (BnB) algorithm to solve each instance with a maximum running time of 72 h, and use the best solution obtained as the benchmark solution. The algorithms are implemented with Visual C# 2015 based on our heuristic optimization algorithm library [48] (<http://www.compintell.cn/en/dataAndCode.html>).

**TABLE V**  
RESULTS OF THE COMPARATIVE ALGORITHMS ON THE DRUG PROCUREMENT PLANNING PROBLEM

#	$\sum g_i$	$\sum K_i$	GA	PSO	DE	WWO
1	326	336	$\dagger$ 0.997 (0.001)	<b>1.000</b> (0)	<b>1.000</b> (0)	<b>1.000</b> (0)
2	697	813	$\dagger$ 0.865 (0.026)	<b>1.000</b> (0)	$\dagger$ 0.992 (0.002)	<b>1.000</b> (0)
3	824	960	$\dagger$ 0.811 (0.039)	$\dagger$ 0.989 (0.005)	$\dagger$ 0.980 (0.008)	<b>1.000</b> (0)
4	2250	1021	$\dagger$ 0.773 (0.032)	0.973 (0.017)	$\dagger$ 0.954 (0.016)	<b>0.990</b> (0.002)
5	1566	1216	$\dagger$ 0.725 (0.036)	$\dagger$ 0.892 (0.016)	$\dagger$ 0.938 (0.012)	<b>0.972</b> (0.004)
6	3521	1453	$\dagger$ 0.686 (0.059)	$\dagger$ 0.837 (0.018)	$\dagger$ 0.902 (0.015)	<b>0.961</b> (0.005)
7	4946	1681	$\dagger$ 0.698 (0.063)	$\dagger$ 0.793 (0.023)	$\dagger$ 0.881 (0.024)	<b>0.955</b> (0.012)
8	7023	2625	$\dagger$ 0.725 (0.045)	$\dagger$ 0.817 (0.018)	$\dagger$ 0.910 (0.013)	<b>0.970</b> (0.009)

**Table V** presents the experimental results on the test instances. Columns 2 and 3, respectively, give the total number of cases and the total number of treatment items, which represent the sizes of the instances. Columns 4–7 show the results of the four algorithms in terms of  $f(\mathbf{x}^*)/f(\mathbf{x}^b)$  averaged over the 30 runs, where  $\mathbf{x}^*$  is the best solution found by the heuristic algorithm and  $\mathbf{x}^b$  is the benchmark solution. For each instance, the best mean value obtained is shown in boldface. The standard deviations are given in parenthesis. A superscript  $\dagger$  before an average value indicates that the result of the corresponding algorithm is significantly different from that of WWO, according to the non-parametric Wilcoxon rank-sum test with a confidence level of 95%. As we can see, the performance of GA is the worst among the four comparative algorithms, mainly because its standard crossover operator often produces many infeasible solutions. On small-size instances #1 and #2, PSO and WWO can always obtain the exact optimal solutions. With the increase of instance size, WWO exhibits more performance advantage over the other heuristics. PSO is efficient in solving small-size instances, but on large-size instances its performance degrades because of premature convergence. Comparatively, the performance of DE is slightly worse than PSO on small-size instances, but it shows more exploration ability and outperforms PSO on large-size instances. On each instance, our WWO algorithm always achieves the best performance among the four algorithms. According to statistical tests, the results of WWO are significantly better than GA on all eight instances, better than PSO on five instances, and better than DE on seven instances. WWO obtains the exact optimal solutions on instances #1–#3, and its ratios to the best-known objective values on other instances are also not less than 95%, which demonstrates it a very efficient algorithm for the considered problem.

In **Table V**, the results of the algorithms on the largest size instance #8 are even better than that on #6 and #7. This is because the exact BnB cannot terminate within 72 h for searching the large solution spaces of #7 and #8, i.e.,  $\mathbf{z}^b$  cannot guarantee

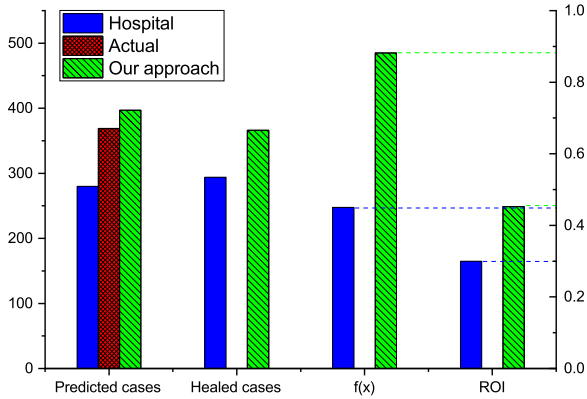


Fig. 4. Comparison of the hospital's method and our method in the first case study.

to be the exact optima, and the values of  $f(\mathbf{z}^*)/f(\mathbf{z}^b)$  increase.

### C. Case Studies

Furthermore, we use two case studies to validate the overall efficacy of our approach, which depends on both the morbidity prediction accuracy and the problem-solving performance. Both the case studies come from a county hospital in Jiangxi Province, Central China. The first occurred during the Spring Festival of 2016, where there was a rapid growth of acute gastrointestinal infections due to festive eating habits. In addition, there were several food poisoning accidents happened. Based on historical experiences, the hospital had anticipated the morbidity growth and estimated the total cases as 280, which was included as its previous drug procurement plan. Nevertheless, the actual number of cases was 369, which was much higher than anticipated and thus led to shortages of many drugs, and the average prediction error over the 20 diseases was 30.65%. Consequently, a portion of cases have to be treated with less effective alternative drugs. Even so, there were 22 cases that could not be handled by the hospital and had to be transferred to other hospitals. The actual recovery rate was 62.87% within 48 h and 79.67% within a week.

We collect the data of food contaminations, temperatures, and historical morbidities related with this case study, and use the hybrid DNN to predict the morbidities in this period. The predicted number of cases is 397, and the average prediction error over the 20 diseases is only 8.2%. We then use the WWO algorithm to produce the drug procurement plan for this instance, and evaluate the therapeutic effect of our plan as well as the actual plan adopted by the hospital based on (8)–(13) and the queuing model in Section III. Fig. 4 compares the predicted cases, the healed cases, the objective function values, and the returns on investment (ROIs, i.e., the ratio of  $f(\mathbf{x})$  to the procurement cost) of the two plans. As the number of cases predicted by our approach is much closer to the actual number, its therapeutic effect is much higher than that of the hospital's method: There is no case to be transferred out, and the recovery rate is 82.11% within 48 h and 99.19% within a week. Although its procurement

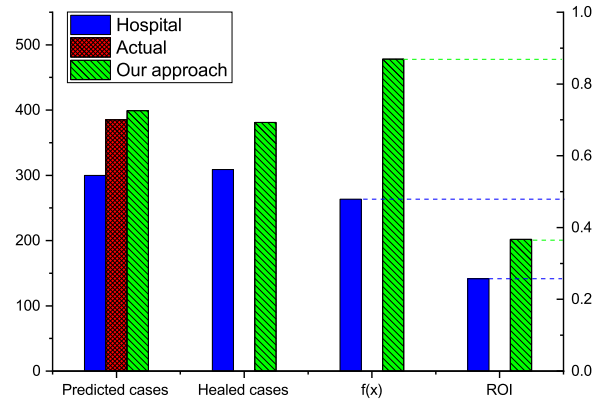


Fig. 5. Comparison of the hospital's method and our method in the second case study.

cost is higher due to overestimation, its ROI is still much higher than the hospital's method.

The second case study occurred in the summer of 2016, where the high morbidities of gastrointestinal infections were caused by bacterial contamination of food under high temperature. The hospital empirically estimated that there would be 300 cases of infections, and thus developed emergency procurement plan. However, the actual number of cases was 385, and the actual recovery rate was 71.17% within 48 h and 80.26% within a week. Using our approach, the hybrid DNN predicts that there will be 399 cases, and the estimated recovery rate is 80% within 48 h and 98.96% within a week. The results also indicate that our approach has much higher prediction accuracy, therapeutic effect, and ROI, as shown in Fig. 5.

From the two case studies, we can see that our approach integrating DNN-based morbidity prediction and heuristic optimization for drug procurement planning is quite efficient in preparing for rapid growth of acute gastrointestinal infections emergency in practice.

## VI. CONCLUSION

This paper proposes a big-data-driven approach that uses DNN to predict morbidities of gastrointestinal infection cases based on statistics of food contaminations and other related data, and then uses heuristic optimization to produce an emergency drug procurement plan to prepare for the upcoming cases. Computational experiments demonstrate the performance advantages of the proposed approach over existing approaches, and in two real case studies our approach exhibits an average prediction error of 8% and an estimated recovery rate of 99%, which is much better than the traditional approach adopted by the hospital.

With the development of big-data technologies, we can now identify and monitor a large number of direct and indirect influence factors of many other diseases, which makes our approach also applicable for medical preparation for such diseases. Our ongoing work includes three aspects. The first is to adapt our approach to other common epidemics such as HAV, influenza, and allergic dermatitis. The second is to develop a more complex but more accurate method for evaluating therapeutic effects by



considering different combinations of drugs. The third is to extend our approach for optimizing the configuration of more medical resources, such as physicians, nurses, facilities and equipments, in addition to drugs.

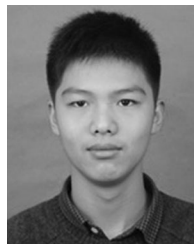
## REFERENCES

- [1] J. D. Quick, "Applying management science in developing countries: ABC analysis to plan public drug procurement," *Socio-Econ. Plan. Sci.*, vol. 16, no. 1, pp. 39–50, 1982.
- [2] M. Duggan and F. M. Scott Morton, "The distortionary effects of government procurement: Evidence from Medicaid prescription drug purchasing," *Quart. J. Econ.*, vol. 121, no. 1, pp. 1–30, 2006.
- [3] Y. Shu and R. Tong, "Common problems of hospital drugs procurement and its solutions," *China Pharm.*, vol. 22, no. 33, pp. 3114–3115, 2011.
- [4] P. V. Singh, A. Tatambhotla, and R. R. Kalvakuntla, "Replicating Tamil Nadu's drug procurement model," *Econ. Political Weekly*, vol. 47, no. 39, pp. 26–29, 2012.
- [5] A. L. Kjos, N. T. Binh, C. Robertson, and J. Rovers, "A drug procurement, storage and distribution model in public hospitals in a developing country," *Res. Social Administ. Pharmacy*, vol. 12, no. 3, pp. 371–383, 2016.
- [6] J. Federmann and V. Fikr, "Mathematical model for the prediction of drug consumption. I. correlative analysis of consumption factors," *Ceskoslovenske Zdravotnictvi*, vol. 18, no. 2, pp. 65–73, 1970.
- [7] E. Iosifidis *et al.*, "Differential correlation between rates of antimicrobial drug consumption and prevalence of antimicrobial resistance in a tertiary care hospital in Greece," *Infection Control Hospital Epidemiol.*, vol. 29, no. 7, pp. 615–622, 2008.
- [8] W. Wang, W. Zhang, J. Yin, and R. Xu, "A simple seasonal autoregressive model and its application in drug demand forecasting," *Chin. J. Health Statist.*, vol. 14, no. 2, pp. 46–47, 1997.
- [9] C.-h. Hu, "Research on the dynamic demand model of drugs," *Chin. J. Pharmaceutical Anal.*, vol. 30, no. 11, pp. 2229–2232, 2010.
- [10] M. S. Sohn, "Demand forecasting for developing drug inventory control model in a university hospital," *Korean J. Preventive Med.*, vol. 16, no. 1, pp. 113–120, 1983.
- [11] E.-I. M. A, D. S. M, and H. M. A, "Drug consumption prediction through temporal pattern matching," in *Proc. World Congr. Eng.*, vol. 3, 2013.
- [12] J. Gong and J. Li, "Drug demand combination prediction based on neural networks," *China Manag. Inform.*, vol. 17, no. 8, pp. 84–88, 2014.
- [13] M. Ture and I. Kurt, "Comparison of four different time series methods to forecast hepatitis A virus infection," *Expert Syst. Appl.*, vol. 31, no. 1, pp. 41–46, 2006.
- [14] A. Dmitriev and V. Kotin, "Time series prediction of morbidity using artificial neural networks," *Biomed. Eng.*, vol. 47, no. 1, pp. 43–45, 2013.
- [15] H. Bibi *et al.*, "Prediction of emergency department visits for respiratory symptoms using an artificial neural network," *Chest*, vol. 122, no. 5, pp. 1627–1632, 2002.
- [16] Z. Wu, W. Wu, P. Wang, and B. Zhou, "Prediction for incidence of hemorrhagic fever with renal syndrome with back propagation artificial neural network model," *Chin. J. Vector Bio Control*, vol. 17, no. 3, pp. 223–226, 2006.
- [17] Q. Wang, Y. Liu, and X. Pan, "Atmosphere pollutants and mortality rate of respiratory diseases in Beijing," *Sci. Total Environ.*, vol. 391, no. 1, pp. 143–148, 2008.
- [18] J. Junk, A. Krein, and A. Helbig, "Mortality rates and air pollution levels under different weather conditions: An example from Western Europe," *Int. J. Environ. Waste Manage.*, vol. 4, no. 1-2, pp. 197–212, 2009.
- [19] Y. X. Ma and S. G. Wang, "The application of artificial neural network in the forecasting on incidence of a disease," in *Proc. 3rd Int. Conf. Biomed. Eng. Informat.*, 2010, vol. 3, pp. 1269–1272.
- [20] L. Li, R. L. Ge, S. M. Zhou, and R. Valerdi, "Guest editorial integrated healthcare information systems," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 515–517, Jul. 2012.
- [21] B. Xu, L. D. Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in IoT-based information system for emergency medical services," *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1578–1586, May 2014.
- [22] R. Fang, S. Pouyanfar, Y. Yang, S.-C. Chen, and S. S. Iyengar, "Computational health informatics in the big data age: A survey," *ACM Comput. Surv.*, vol. 49, no. 1, pp. 12:1–12:36, 2016.
- [23] S.-M. Zhou *et al.*, "Defining disease phenotypes in primary care electronic health records by a machine learning approach: A case study in identifying rheumatoid arthritis," *PLOS ONE*, vol. 11, no. 5, pp. 1–14, 2016.
- [24] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [25] P. Li, Z. Chen, L. T. Yang, Q. Zhang, and M. J. Deen, "Deep convolutional computation model for feature learning on big data in Internet of things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 790–798, Feb. 2018.
- [26] J. A. Rodger, "Discovery of medical big data analytics: Improving the prediction of traumatic brain injury survival rates by data mining patient informatics processing software hybrid hadoop hive," *Informat. Med. Unlocked*, vol. 1, pp. 17–26, 2015.
- [27] Q. Song, Y.-J. Zheng, Y. Xue, W.-G. Sheng, and M.-R. Zhao, "An evolutionary deep neural network for predicting morbidity of gastrointestinal infections by food contamination," *Neurocomputing*, vol. 226, pp. 16–22, 2017.
- [28] Q. Song, M.-R. Zhao, X.-H. Zhou, Y. Xue, and Y.-J. Zheng, "Predicting gastrointestinal infection morbidity based on environmental pollutants: Deep learning versus traditional models," *Ecol. Indicators*, vol. 82, pp. 76–81, 2017.
- [29] Y.-J. Zheng, "Water wave optimization: A new nature-inspired metaheuristic," *Comput. Oper. Res.*, vol. 55, no. 1, pp. 1–11, 2015.
- [30] P. Vincent, H. Laroche, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, 2010.
- [31] J. L. Gauvain and C.-H. Lee, "Maximum a posteriori estimation for multivariate gaussian mixture observations of Markov chains," *IEEE Trans. Speech Audio Process.*, vol. 2, no. 2, pp. 291–298, Apr. 1994.
- [32] F. Cardinaux, C. Sanderson, and S. Marcel, "Comparison of MLP and GMM classifiers for face verification on XM2VTS," in *Audio- and Video-Based Biometric Person Authentication*, J. Kittler and M. S. Nixon, Eds. New York, NY, USA: Springer-Verlag, 2003, pp. 911–920.
- [33] H. Li and K. Huang, *Manual of Organic Pesticides and Intermediate Mass Spectrum*. Beijing, China: Chemical Industry Press, 2009.
- [34] M. Zhou *et al.*, *Heterocyclic Pesticides: Herbicides*. Beijing, China: Since Press, 2014.
- [35] L. H. Keith, "Environmental endocrine disruptors: an overview of the analytical challenge," in *Proc. 13th Ann. Waste Testing Quality Assurance Symp.*, Arlington, TX, USA, 1997.
- [36] G. E. Hinton, "Connectionist learning procedures," *Artif. Intell.*, vol. 40, no. 1, pp. 185–234, 1989.
- [37] P. Vincent, H. Laroche, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn.* New York, NY, USA, 2008, pp. 1096–1103.
- [38] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Advances in Neural Information Processing Systems*, vol. 19, J. P. Bernhard Schölkopf and T. Hoffman, Eds. Cambridge, MA, USA: MIT Press, 2007, pp. 153–160.
- [39] D. Simon, "Biogeography-based optimization," *IEEE Trans. Evol. Comput.*, vol. 12, no. 6, pp. 702–713, Dec. 2008.
- [40] Y.-J. Zheng, H.-F. Ling, and J.-Y. Xue, "Ecogeography-based optimization: Enhancing biogeography-based optimization with ecogeographic barriers and differentiations," *Comput. Oper. Res.*, vol. 50, pp. 115–127, 2014.
- [41] Y. J. Zheng, W. G. Sheng, X. M. Sun, and S. Y. Chen, "Airline passenger profiling based on fuzzy deep machine learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 12, pp. 2911–2923, Dec. 2017.
- [42] Y.-J. Zheng, S.-Y. Chen, Y. Xue, and J.-Y. Xue, "A Pythagorean-type fuzzy deep denoising autoencoder for industrial accident early warning," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 6, pp. 1561–1575, Dec. 2017.
- [43] Y.-J. Zheng, X.-H. Zhou, W.-G. Sheng, Y. Xue, and S.-Y. Chen, "Generative adversarial network based telecom fraud detection at the receiving bank," *Neural Netw.*, vol. 102, pp. 78–86, 2018.
- [44] Y.-J. Zheng and B. Zhang, "A simplified water wave optimization algorithm," in *Proc. IEEE Congr. Evol. Comput.*, 2015, pp. 807–813.
- [45] L. S. Aiken, S. G. West, and S. C. Pitts, *Multiple Linear Regression*. Hoboken, NJ, USA: Wiley, 2003.
- [46] E. C. Laskari, K. E. Parsopoulos, and M. N. Vrahatis, "Particle swarm optimization for integer programming," in *Proc. IEEE Congr. Evol. Comput.*, vol. 2, 2002, pp. 1582–1587.
- [47] M. Omran and A. Engelbrecht, "Differential evolution for integer programming problems," in *Proc. IEEE Congr. Evol. Comput.*, 2007, pp. 2237–2242.
- [48] Y. Zheng, H. Ling, J. Xue, and S. Chen, "Population classification in fire evacuation: A multiobjective particle swarm optimization approach," *IEEE Trans. Evol. Comput.*, vol. 18, no. 1, pp. 70–81, Feb. 2014.



**Qin Song** received the B.S. degree in biotechnology from Hainan Normal University, Haikou, China, in 2006, the M.S. degree in vegetable science from Beijing University of Agriculture, Beijing, China, in 2011, and the Ph.D. degree in environmental science and technology from Zhejiang University of Technology, Hangzhou, China, in 2018.

She is currently a Research Assistant with Hangzhou Normal University, Hangzhou, China. Her current research interests include environmental health and informatics.



**Zhi-Ge Xu** received the B.Sc. degree in software engineering from the Zhejiang University of Technology, Hangzhou, China, in 2017. He is currently working toward the M.Sc. degree in computer science and technology at Zhejiang University of Technology.

His research interests include neural networks and deep learning techniques.



**Yu-Jun Zheng** (M'06–SM'17) received the Ph.D. degree from the Institute of Software, Chinese Academy of Sciences, in 2010.

He is a Professor with Hangzhou Normal University, Hangzhou, China. He has authored more than 50 papers in famous journals including IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON FUZZY SYSTEMS, IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, etc. His research interests include bioinspired computation and its

applications.

In 2014, he received the runner-up of IFORS Prize for Development due to the work on intelligent scheduling of engineering rescuing tasks in the 2013 Dingxi earthquake. In 2018, he has been selected as a finalist for the Daniel H. Wagner Prize for Excellence in Operations Research Practice. He is an ACM member.



**Wei-Guo Sheng** (M'13) received the M.Sc. degree in information technology from the University of Nottingham, U.K., in 2002, and the Ph.D. degree in computer science from Brunel University, U.K., in 2005.

He was a Researcher with the University of Kent, U.K., and Royal Holloway, University of London, U.K. He is currently a Professor with Hangzhou Normal University, Hangzhou, China. His research interests include evolutionary computation, data mining/clustering, pattern recognition, and machine learning.



**Yu-Jiao Huang** received the B.S. degree in information and computer science, the M.S. degree in computational mathematics, and the Ph.D. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2008, 2010 and 2014, respectively.

She is currently a Lecturer with the Zhejiang University of Technology, Hangzhou, China. Her research interests include areas of artificial neural networks, stability theory, and dynamical systems.



**Jun Yang** received the B.S. degree in biochemistry and the M.S. degree in cell biology from Shandong University, Jinan, China, in 1991 and 1994, respectively, and the Ph.D. degree in biochemistry and molecular genetics from Georgia State University, Atlanta, USA, in 2000.

He is now a Professor with Hangzhou Normal University. His main research interests are public health management and environmental health.