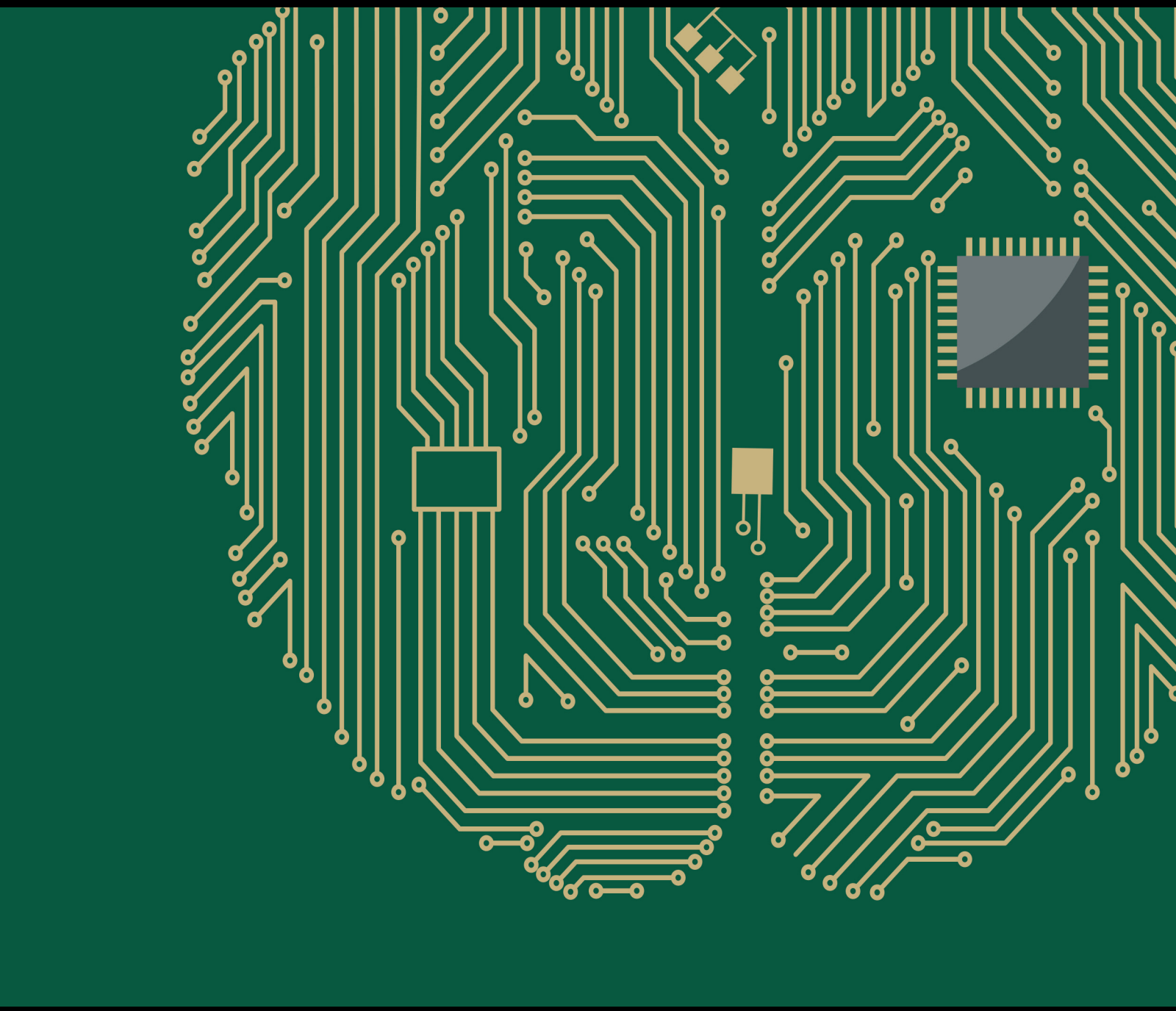
Computational Intelligence and Neuroscience

Recent Advances in Learning Theory

Guest Editors: Weihui Dai, Wlodzislaw Duch, Abdul Hanan Abdullah, Dongrong Xu, and Ye-Sho Chen



Recent Advances in Learning Theory

Computational Intelligence and Neuroscience

Recent Advances in Learning Theory

Guest Editors: Weihui Dai, Wlodzislaw Duch,
Abdul Hanan Abdullah, Dongrong Xu, and Ye-Sho Chen

Editorial Board

Contents

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 395948, 4 pages
http://dx.doi.org/10.1155/2015/395948

Editorial

Recent Advances in Learning Theory

Weihui Dai,[1] Wlodzislaw Duch,[2] Abdul Hanan Abdullah,[3] Dongrong Xu,[4] and Ye-Sho Chen[5]

[1] *School of Management, Fudan University, Shanghai 200433, China*
[2] *Department of Informatics, Nicolaus Copernicus University, Toruń, Poland*
[3] *Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia*
[4] *Psychiatry Department, New York State Psychiatric Institute, Columbia University, New York, NY 10032, USA*
[5] *E. J. Ourso College of Business, Louisiana State University, Baton Rouge, LA 70803, USA*

Correspondence should be addressed to Weihui Dai; whdai@fudan.edu.cn

Received 6 September 2015; Accepted 28 September 2015

1. Introduction

In the era facilitated by the Internet of Things, ubiquitous communications as well as cloud services, sensing means, and human-computer interfaces are becoming all-pervasive and online. This makes it more possible for us than ever before to study engineering problems, human activities, and social behaviors through machine learning analysis of the big data produced in the ubiquitous environment. Looking at the recent history and new trends, machine learning has made attractive progress in wide areas of applications, from natural language to nonverbal communication, from engineering application to humanities, arts, and social studies, and from the real world to cyber space.

In 1997, Dietterich summarized the development of machine learning in four directions: ensembles of classifiers, methods for scaling up supervised learning algorithm, reinforcement learning, and learning of complex stochastic models [1]. Latterly, Duch addressed machine learning as the foundations of computational intelligence comprehensively [2]. Since the beginning of the 21st century, research in machine learning has made progress in all of the four directions and has become focused on new challenges in learning from big data that cover a variety of application areas. One barrier that needs to be broken through is how to avoid the "curse of dimensionality" and ensure the generalization ability in the learning process. Owing to the efforts such as work of Koller and Friedman on Probabilistic Graphical Models [3] and the Compressive Sensing (CS) theory [4], the tentative path has been lightened. As the forethought by Wang [5], "predicting" the changes based on generalization ability and "describing" the knowledge discovered from huge data will be the two major tasks of machine learning in the future.

In today's society, machine learning has been in an extensive demand in the areas associated with human's psychology and behaviors, such as ubiquitous learning, e-commerce, online customer service, behavioral finance analysis, and government emergency management. We believe that machine learning could be the most promising, sometimes even the only, way to accomplish the complex computation on human psychology and behaviors in the ubiquitous environment.

However, in doing so, the machine learning research needs to pay attention to the following new aspects which may be beyond the ability of computer science and technology and requires more novel interdisciplinary ideas and methods: (1) a systematic model such as the social neuroscience mechanism [6] which can describe the neural activities and dominant process of human psychology and behaviors and helps the machine to understand its globally structural features and therefore reduce the computational cost by learning from the limited samples of a big data set, (2) the comprehensive context awareness in physical, cyber, and psychosocial spaces, as well as the information fusion processing and computing ability, which has been called Cyber Psychosocial and Physical (CPP) Computation by Dai [7], and (3) smart learning that enables the machine to cope

with both rational intelligence and emotional intelligence in the learning process.

The aim of this special issue is to bring together researchers working in different areas for exchanging and sharing with each other their progress in the new tendency of modern learning theory and applications. We were very pleased to see various new research ideas and many innovative contributions in the submitted manuscripts, which cover a wide range of recent advances in learning theory, techniques, and applications. What follows is a brief editorial review of the published papers in this special issue from four perspectives: Neuroscience in Learning Theory, Machine Learning in Psychological Computation and Behavior Analysis, Machine Learning in Public Management and Business Service, and Machine Learning in Knowledge Discovery and Human-Computer Engineering.

2. Neuroscience in Learning Theory

The paper titled "Neural Cognition and Affective Computing on Cyber Language" by S. Huang et al. analyzed the classification and cognitive characteristics of emotional symbols in cyber languages and put forward a mechanism model to show the dominant neural activities in that cognitive process. Through the comparative study of Chinese, English, and Spanish languages in their expressive patterns of emotions, an intelligent method of machine learning was proposed for affective computing on international cyber languages which can deal with the multisymbol information and mixed emotions in a cyber message and show their dynamic changes according to the characteristics of the neural cognition process.

The paper titled "Neural Basis of Intrinsic Motivation: Evidence from Event-Related Potentials" by J. Jin et al. employed event-related potentials (ERPs) to investigate the neural disparity between an interesting stop-watch (SW) task and a boring watch-stop (WS) task to understand the neural mechanism of intrinsic motivation. Research findings of this paper indicated that intrinsic motivation could be added as a candidate social factor in the construction of a machine learning model, and it provided the new possible indicators as well as the feature parameters for detecting and analyzing human intrinsic motivation based on machine learning technology in a wearable system.

The paper titled "A Neuroeconomics Analysis of Investment Process with Money Flow Information: The Error-Related Negativity" by C. Wang et al. studied the features of event-related potentials (ERPs) in the decision-making process of financial investment on stock market. Experimental results showed that ERN component was sensitive to the evaluation of the risk in the investment decision process and could be regarded as the early warning indicator to show a conflict perception or alert the brain to prepare for potential negative consequences. Findings in this paper have significant implications for exploring the neural cognitive effects and the basis of machine learning from ERP signals in the development of a financial intelligence system.

The paper titled "P300 and Decision Making under Risk and Ambiguity" by L. Wang et al. used ERPs to clarify and extend the current understanding of decision-making under risk and with ambiguity. Research findings pointed out that decision-making with ambiguity occupies larger amount of working memory and recalls more past experience, while decision-making under risk mainly mobilizes attention resources to calculate current information. This paper provided further understanding of brain mechanism in decision-making, which may be helpful in the design of humanoid robots based on machine learning theory.

The paper titled "Explore Awareness of Information Security: Insights from Cognitive Neuromechanism" by D. Han et al. took the online financial payment as a research example and conducted an experimental analysis of electrophysiological signals to study the awareness of information security. Its findings indicated that left hemisphere and beta rhythms of electroencephalogram (EEG) signal are sensitive to the cognitive degree of risks in the awareness of information security, which may be probably considered as the sign to assess people's cognition of potential risks in online financial payment. This paper contributed new knowledge to the understanding of EEG signals in the awareness of information security, which is of significance for the development of machine learning technology for the objective and technological assessment of information security awareness.

3. Machine Learning in Psychological Computation and Behavior Analysis

The paper titled "CyberPsychological Computation on Social Community of Ubiquitous Learning" by X. Zhou et al. studied the relationships between the ubiquitous learners' psychological reactions and their behavioral patterns in cyber space and summarized 15 common basic actions of the learners' habitual behaviors for the psychological assessment of their situations in the learning process. A CyberPsychological computation method based on BP-GA neural network was proposed by the authors, which can be used to estimate the learners' psychological states online according to their personalized behavioral patterns. Contributions of this paper provided new progresses in the field of machine learning on Cyber Psychosocial and Physical (CPP) Computation and have important significance for further research in wide application areas.

The paper titled "An Opinion Interactive Model Based on Individual Persuasiveness" by X. Zhou et al. discussed a common phenomenon in social interactive process which is associated with the propagation characteristics of public opinions and attitudes to a social event. Based on the Deffuant Model and its improvements, the authors considered the impacts of individual persuasiveness and conducted an experiment using multiagent simulation to show that the range of common opinion could be predicted when the initial distribution of opinions and persuasiveness are given. The interesting findings in this paper indicated that there would be some underlying rules for machine learning in the behavior analysis of a social interactive process.

In addition, studies of machine learning on affective computing were also reported by S. Huang et al. in the paper titled "Neural Cognition and Affective Computing on Cyber

Language" and by S. Gong et al. in the paper titled "Emotion Analysis of Telephone Complaints from Customer Based on Affective Computing."

4. Machine Learning in Public Management and Business Service

The paper titled "Information Dissemination of Public Health Emergency on Social Networks and Intelligent Computation" by H. Hu et al. emphasized that social networks had become the main information dissemination platform of public health emergency and caused high concerns in emergency management. The authors analyzed the complex characteristics of information dissemination in social networks in public health emergency and argued that the existing theoretical tools and modeling methods are not sufficient to accurately describe and predict the information dissemination in social networks. Based on the framework of TDF (Theory-Data-Feedback), a new intelligent computation method was constructed for the ACP (artificial societies, computational experiments, and parallel execution) simulation by this paper and reached highly precise results on the prediction of dynamic dissemination of emergency event's information on social networks.

The paper titled "The Large Scale Machine Learning in an Artificial Society: Prediction of the Ebola Outbreak in Beijing" by P. Zhang et al. proposed a new method for predicting the disease propagation based on artificial society. The authors established the virtual society system of "Artificial Beijing" which contained 19.6 millions of individuals and 8 millions of buildings in correspondence with the actual distributions of geodemographics, infrastructure, and social roles in Beijing. Through a large scale machine learning of individuals' behaviors, the propagation process of Ebola virus disease and its corresponding interventions for public health emergency management are well simulated and accurately predicted based on "Artificial Beijing."

The paper titled "Emotion Analysis of Telephone Complaints from Customer Based on Affective Computing" by S. Gong et al. studied the characteristics of telephone complaint speeches and proposed an analysis method based on affective computing technology, which can recognize the dynamic changes of customer emotions from the conversations between the service staff and the customer. Experimental results showed that this method is effective and could reach high recognition rates of happy and angry states. It has been successfully applied to the operation quality and service administration in telecom and Internet service company.

5. Machine Learning in Knowledge Discovery and Human-Computer Engineering

The paper titled "Exploiting Language Models to Classify Events from Twitter" by D.-T. Vo et al. studied the issue of event classification in Twitter based on the models of LDA-SP and ConceptNet. The authors explored a new method to compute tweets' similarity according to their features, including common term words and the relationships among their distinguishing term words, by the kNN classifier and experiments on Edinburgh Twitter Corpus. This paper gives a logical and comprehensive insight into the research issue, presenting an effective method which exhibits the valuable exploration by adopting the language model to improve the classification performance as shown in the experiments.

The paper titled "Design of Automatic Extraction Algorithm of Knowledge Points for MOOCs" by H. Chen et al. studied knowledge discovery and knowledge sharing in MOOCs (Massive Online Open Courses). The authors designed an automatic extracting course knowledge points (AECKP) algorithm for building the course ontology and employed the Vector Space Model (VSM) to calculate the similarity between knowledge points for the optimization of knowledge extraction. Experimental results showed that the proposed approach could achieve satisfactory results.

The paper titled "Intelligent Context-Aware and Adaptive Interface for Mobile LBS" by J. Feng and Y. Liu discussed an interesting issue about the intelligent interface for mobile LBS (Location Based Services). Through penetrating analysis of the requirements on the interface's design, the authors proposed a context-aware adaptive model for LBS interface along with the framework for its application and described the adaptive process of dynamic interaction between users and the interface. A standard usability evaluation test was conducted to show the performance improvements by the proposed model. Research work of this paper provided a valuable reference for the design of adaptive interface based on context awareness.

Finally, the paper titled "An Efficient Robust Eye Localization by Learning the Convolution Distribution Using Eye Template" by X. Li et al. presented a novel eye localization approach which explored only one-layer convolution map by eye template using a backpropagation (BP) network. Based on the extracted convolution distributed features and a multieye template set which considered the utilizations of both the global information and the local geometric features, the proposed method could obtain similar best results with greatly reduced training time and high prediction speed and provided a better comprehension compared with existing methods.

Acknowledgments

We are grateful to the editors for hosting this special issue and many thanks are due to the staff for their precious advice during the editorial process of the special issue.

Weihui Dai
Wlodzislaw Duch
Abdul Hanan Abdullah
Dongrong Xu
Ye-Sho Chen

References

[1] T. G. Dietterich, "Machine-learning research: four current directions," *AI Magazine*, vol. 18, no. 4, pp. 97–136, 1997.

[2] W. Duch, "Towards comprehensive foundations of computational intelligence," in *Challenges for Computational Intelligence*, W. Duch and J. Mandziuk, Eds., vol. 63 of *Springer Studies in Computational Intelligence*, pp. 261–316, Springer, 2007.

[3] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, MIT Press, Cambridge, Mass, USA, 2009.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] Y. Wang, "The puzzle of machine learning and historical revelation," in *Proceedings of the 9th Seminar of Machine Learning and Application*, 2011.

[6] J. T. Cacioppo and G. G. Berntson, "Social psychological contributions to the decade of the brain: doctrine of multilevel analysis," *American Psychologist*, vol. 47, no. 8, pp. 1019–1028, 1992.

[7] W. H. Dai, *Cyber Psychological and Physical (CPP) Computation Based on Social Neuromechanism*, Fudan University, Shanghai, China, 2015.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 506905, 9 pages
http://dx.doi.org/10.1155/2015/506905

Research Article

Emotion Analysis of Telephone Complaints from Customer Based on Affective Computing

Shuangping Gong,[1] Yonghui Dai,[2] Jun Ji,[3] Jinzhao Wang,[4] and Hai Sun[4]

[1] *Language and Culture Research Institute, National University of Defense Technology, Changsha 410074, China*
[2] *School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China*
[3] *Department of Operation Quality and Service Administration, China Unicom Co. Ltd., Shanghai Branch, Shanghai 200070, China*
[4] *School of Management, Fudan University, 220 Handan Road, Shanghai 200433, China*

Correspondence should be addressed to Jinzhao Wang; jzwang13@fudan.edu.cn and Hai Sun; sunhai@fudan.edu.cn

Received 19 February 2015; Revised 11 July 2015; Accepted 13 July 2015

Academic Editor: Ye-Sho Chen

Customer complaint has been the important feedback for modern enterprises to improve their product and service quality as well as the customer's loyalty. As one of the commonly used manners in customer complaint, telephone communication carries rich emotional information of speeches, which provides valuable resources for perceiving the customer's satisfaction and studying the complaint handling skills. This paper studies the characteristics of telephone complaint speeches and proposes an analysis method based on affective computing technology, which can recognize the dynamic changes of customer emotions from the conversations between the service staff and the customer. The recognition process includes speaker recognition, emotional feature parameter extraction, and dynamic emotion recognition. Experimental results show that this method is effective and can reach high recognition rates of happy and angry states. It has been successfully applied to the operation quality and service administration in telecom and Internet service company.

1. Introduction

Customer service has been playing an increasing important role in the competitive market in business administration in recent years. As one of the routines in customer service, the handling of telephone complaints from customers plays a significant role in showing the image of the enterprise, obtaining the feedback from the market, and improving the loyalty of the customers. Therefore, it has attracted high attention from enterprises and researchers. The existing researches on customer complaints mainly focus on the classification of complaints, record analysis, handling operation, information management, and so on [1–3]. However, less attention has been paid to the technical and intelligent analysis of the complaint speeches so far. Actually, the customer's speeches in telephone complaints usually carry rich emotions and provide valuable information for perceiving the customer's degree of satisfaction and studying the complaint handling skills.

Affective computing was proposed by Professor Picard in 1977, attempting to create a way of perception, recognition, and understanding of human emotion, which would make the computer system intelligent and sensitive so as to react friendly to human emotions [4]. It has become a burgeoning area of research in human-computer interaction during the past decades. Up to now, scholars have presented a lot of methods and models to deal with the affective computing issues from speech signals [5]. However, most researches are confined to isolated speeches and contain only one type of emotion [6–8]. The speeches in the telephone complaint of a real case are more complicated because they are embedded in the conversations between the service staff and the customer and are embodied with the dynamic changes of emotions ranging from excitement to calmness [5, 9]. Therefore, an

effective technology, which is independent of the speakers' emotional changes [10], should be first utilized to precisely distinguish the customer's speeches from those of the service staff's. On this basis, the conversations in telephone complaints can be segmented into separate speeches according to their different speakers. In the recognition of customer's emotions, the computing method ought to consider the dynamic changes of emotions in customer's continuous speech as well as the possible noise in telephone communication [5].

In this paper, we first study the characteristics of telephone complaint speeches and then conduct a cost-sensitive learning technology [10] to identify the different speakers and separate their speeches from the conversations. Thereafter, a robust method and signal process are proposed to recognize the customer's changing emotions. Furthermore, affective computing technology is explored by using support vector machine (SVM) to process the extracted MFCC parameters. The proposed method and technology have been successfully applied to the administration of telecom and Internet services.

This paper is organized as follows. In Section 2, theory and methodology are introduced as the basis of our research work; in Section 3, the characteristics of telephone complaint speeches are displayed and the speaker identification from conversations is discussed; in Section 4, the analysis method based on affective computing technology is proposed to recognize the dynamic changes of customer's emotions in telephone complaints; in Section 5, an experiment is illustrated to show the performance of the presented method; Section 6 is the conclusion and discussion of this paper.

Table 1: Classifications of basic emotions.

Scholars	Classifications of basic emotions
Arnold	Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, and sadness
Ekman, Friesen, and Ellsworth	Anger, disgust, fear, joy, sadness, and surprise
Frijda	Desire, happiness, interest, surprise, wonder, and sorrow
Gray	Rage and terror, anxiety, and joy
Izard	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, and surprise
James	Fear, grief, love, and rage
McDougall	Anger, disgust, elation, fear, subjection, tender-emotion, and wonder
Mowrer	Pain, pleasure
Oatley and Johnson-laird	Anger, disgust, anxiety, happiness, and sadness
Panksepp	Expectancy, fear, rage, and panic
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, and surprise
Tomkins	Anger, interest, contempt, disgust, distress, fear, joy, shame, and surprise
Watson	Fear, love, and rage
Weiner and Graham	Happiness, sadness

2. Theory and Methodology

2.1. Emotion Classification and Description. The classifications and the descriptions of emotions have been so diverse due to the different understanding of people's psychological experiences in a variety of applications. For a long time, scholars have not reached consensus on the classification of emotion. However, two types of emotion classifications have been widely accepted in psychological studies: one is the classification of basic emotions, and the other is the description of emotion in dimensions [11].

The basic emotion theory claimed that each type of emotions has its basic and disparate characteristics in human experiences, physiological arousal patterns, and the explicit modes. It suggested that all human complex emotions are the different combinations of basic emotions. And Ortony and Turner (1990) summarized the typical classifications of basic emotions proposed by the scholars in this field, which are shown in Table 1 [12].

Different from the basic emotion theory, dimension theory argued that the changes of human's emotions are continuous emotion. It suggests that emotions should be described in a dimensional space, and the similarities and differences between each emotion depend on the dimensions in the space distance. Russell and Peterson proposed the two-dimensional circumplex model for the sentiment classification, which included pleasant-unpleasant dimension and the strength dimension. They thought that affective states can be described by the above two dimensions [13–15]. In addition, Wundt put forth the three-dimensional theory of emotion [16], and he proposed that each emotion is one part of a continuum and different emotions are mapped to specific points in a space with three dimensions, among which, P (Pleasure-Displeasure) dimension reflects a positive or a negative evaluation such as comfortable or not comfortable and agreeable or disagreeable. A (Arousal-Nonarousal) dimension reflects the degree of physiological stimulation and takes some action preparations, which might be active or passive. D (Dominance-Submissiveness) dimension can reflect the strength and the desire for the control of a speaker; it accounts for dominant or submissive position. The continuous form of emotions in different dimensions is shown in Figure 1 [9].

2.2. Acoustic Parameters Related to Emotions. The commonly used acoustic parameters related to emotions can be divided into three categories [17], prosody parameters, spectral parameters, and sound quality parameters. In the above categories, prosody parameters such as the duration, pitch, and energy of a speech signal are the basic parameters for emotion recognition. Mel Frequency Cepstrum Coefficient (MFCC) parameters usually perform better than the other spectral parameters and are widely applied to speech recognition [18]. Sound quality parameters such as format

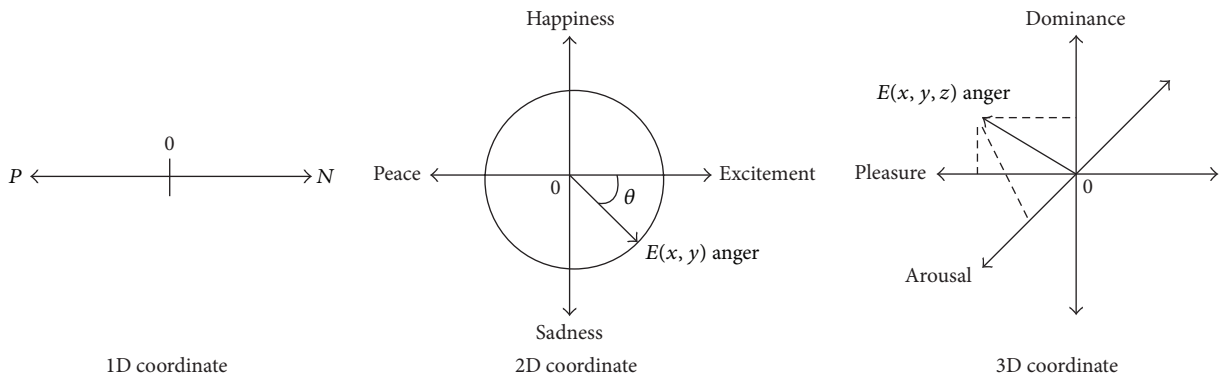


Figure 1: Continuous form of emotions in different dimensions.

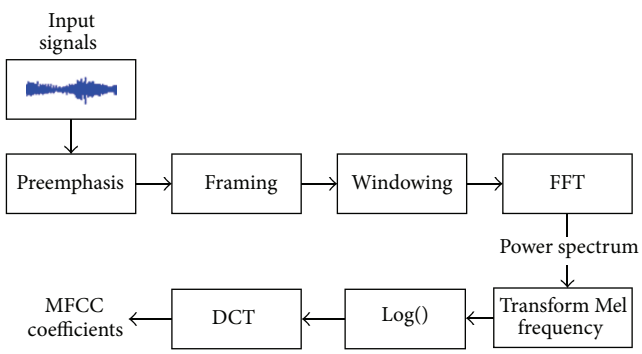


Figure 2: Extraction process of MFCC coefficients.

frequency and bandwidth are effective in differentiating the emotions associated with attitudes and intentions [19].

Recent experiments have shown that the combined parameters from different categories can acquire a more ideal performance [9, 19]. For example, the combination of short-time energy, pitch, short-time zero crossing rate, first formant, second formant, voice speed, number of voice breaks, and 12-order MFCC parameters was successfully applied to the dynamically affective computing on vocal social media [9]. In order to get the dynamic recognition of customer's emotions, the choice of acoustic parameters should take into account both affective features and voice features. Based on the previous research results, we adopt the short-time average energy, short-term zero crossing rate, pitch, formant, and 12-order MFCC coefficients as the feature parameters for speaker identification and emotion recognition in the conversations of telephone complaints.

Short-time average energy refers to the average energy of the speech signal. It is mainly used for acoustic boundary and the ligatures boundary. Short-time average energy can be expressed as in the following formula:

$$E_n = \sum_{i=0}^{N-1} x_n^2(i). \quad (1)$$

Short-time zero crossing rate refers to signal through the zero frequency in a frame. And it can be expressed as in the following formula:

$$Z_n = \sum_{i=0}^{N-1} \operatorname{sgn}[x_n(i)] - \operatorname{sgn}[x_n(i-1)]. \quad (2)$$

MFCC is derived from a type of cepstral representation of the audio clip. It reflects the characteristics of the short-time amplitude spectrum of speech. Extraction process of MFCC coefficients is shown in Figure 2.

2.3. Speech Emotion Recognition Algorithm. The main method of speech recognition includes k-nearest neighbor method (k-NN), artificial neural network (ANN), hidden Markov model (HMM), Gaussian mixture model (GMM), and support vector machine (SVM). Telephone complaints are continuous speeches in the conversations between the customer and the service staff with background noise. Among the proposed methods, ANN can simulate the complicated relationship between the input and output variables and utilize the hidden knowledge very well by sufficient training samples, while SVM has the advantages of superior stability, good generalization ability, and high efficiency [10, 20]. We will compare the performances of ANN with those of the SVM methods in our study.

2.3.1. Artificial Neural Network. A neural network is composed of a number of nodes, or units, connected by links. Each link has a numeric weight associated with it. Weights are the primary means of long-term storage in neural networks, and learning usually takes place by updating the weights. Some of the units are connected to the external environment and can be designated as input or output units.

Back-Propagation Neural Network (BPNN) is one of the most widely used artificial neural networks, and it adopts a kind of error back-propagation algorithm for training multilayer feed forward neural network. Therefore, it can learn and store the input-output mapping relationships. BPNN is based on gradient descent method which minimizes the total of the squared errors between the actual and the desired

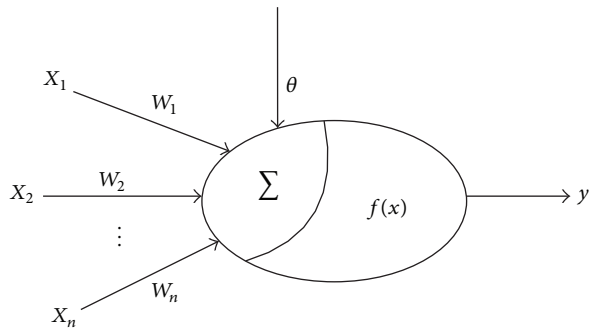


Figure 3: The simplified structure of neurons.

output values. In BPNN, the basic units of neural network are artificial neuron, which are simulating biological neurons with the simplified structure as shown in Figure 3.

The simplified structure neuron is a nonlinear element of a multiple input and a single output, and its relationship can be described as

$$I = \sum_{i=1}^{n} w_i x_i - \theta, \qquad y = f(I), \tag{3}$$

where x_i is the input value, θ is a threshold, and w_i means the strength of weight. $f(x)$ is the excitation function. Typically, a BPNN topology structure includes input layer, hidden layer, and output layer.

2.3.2. Support Vector Machine. Support vector machine is based on the statistical learning theory and VC dimension theory which is the structural risk minimization principle. Through a nonlinear mapping, SVM method takes the sample space to map a high dimensional feature space and makes nonlinear separable problem in the original sample space changed as a linear separable problem in the feature space.

When we use the SVM method for emotion recognition, the selection of kernel function is very important. Four kinds of kernel function are often used in SVM, which can be described as follows.

(i) Linear kernel function: $K(x, y) = x^T \cdot y$.

(ii) Polynomial kernel function: $K(x, y) = [(x \cdot y) + 1]d$.

(iii) RBF kernel function: $K(x, y) = \exp(-\|x - y\|^2/2\sigma^2)$.

(iv) Sigmoid kernel function: $K(x, y) = \tanh(a(x \cdot y) + b)$.

3. Telephone Complaints and Speaker Identification

3.1. Characteristics of Telephone Complaints. The speeches of telephone complaints occur in the conversations between the service staff and the customer. Therefore, the speakers should be identified so that we could find out the customer's speeches and deal with the emotion recognition of his speeches. In order to do well in spell recognition, the speeches in conversations will be first cut into a series of segmentations according to their continuous sound waves. Usually, the speech signals contain possible noises in telephone communication and the speaker's surroundings, so a bandwidth limited filter and the 50 Hz circuit noise elimination had to be adopted in the preprocessing procedure [5]. Through the careful analysis of a large number of complaints, we found that calmness, discontent, and anger are three typical emotional states which can satisfy the requirements of service management. Therefore, we will mainly discuss the recognition of these three typical emotions.

Table 2: Someone's speech characteristics.

Speech characteristics	Someone's emotional states		
	Calmness	Discontent	Anger
Mean-intensity (μv)	43.82	60.59	76.82
Maximum pitch (Hz)	315.59	408.13	532.11
Min pitch (Hz)	148.61	122.05	180.69
Mean pitch (Hz)	207.82	257.33	267.91
Pitch range (Hz)	107.77	286.08	321.42

Our previous study has indicated that intensity, pitch frequency, and spectrum parameter could be used as the prominent feature parameters for distinguishing those emotional states [5]. For example, someone's speech characteristics in three emotional states (calmness, discontent, and anger) are shown in Table 2.

From Table 2, we can find that calmness has the lowest mean-intensity with pitch range below 170 Hz. Discontent and anger have stronger mean-intensities, and their pitch ranges expand to more than 280 Hz. In particular, the pitch range of anger reaches more than 300 Hz.

3.2. Speaker Identification from Conversations. In order to precisely distinguish the customers' speech from the conversations, we adopted a robust speaker identification algorithm. The algorithm introduced a cost-sensitive learning technology to reweight the probability of the tested affective utterances in the pitch envelope level, which can effectively enhance the robustness in emotion-dependent speaker recognition as shown in Figure 4 [10].

4. Framework of Recognition

Based on previous analysis, we propose a recognition method of customer's emotions in telephone complaints as shown in Figure 5.

The recognition process includes preprocessing, feature parameter extraction, and emotion recognition.

4.1. Preprocessing of Speech Signals. In order to improve the quality of speech signals, the preprocessing aims to provide successor analysis services for feature extraction and speech emotion recognition, which may include speech unit segmentation, preemphasis, framing and windowing, and detecting endpoint [21–23].

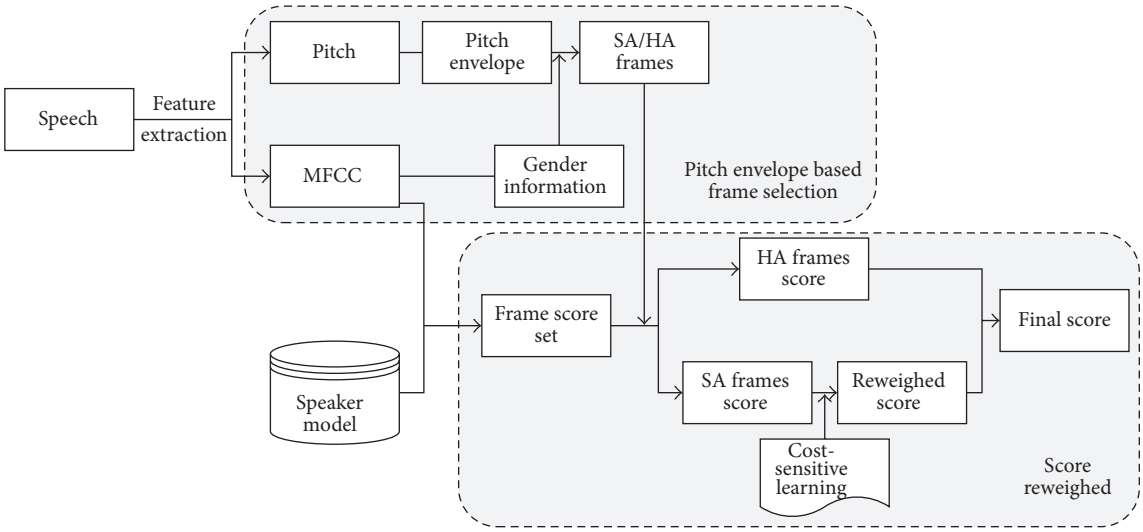


Figure 4: Speaker identification algorithm based on cost-sensitive learning technology.

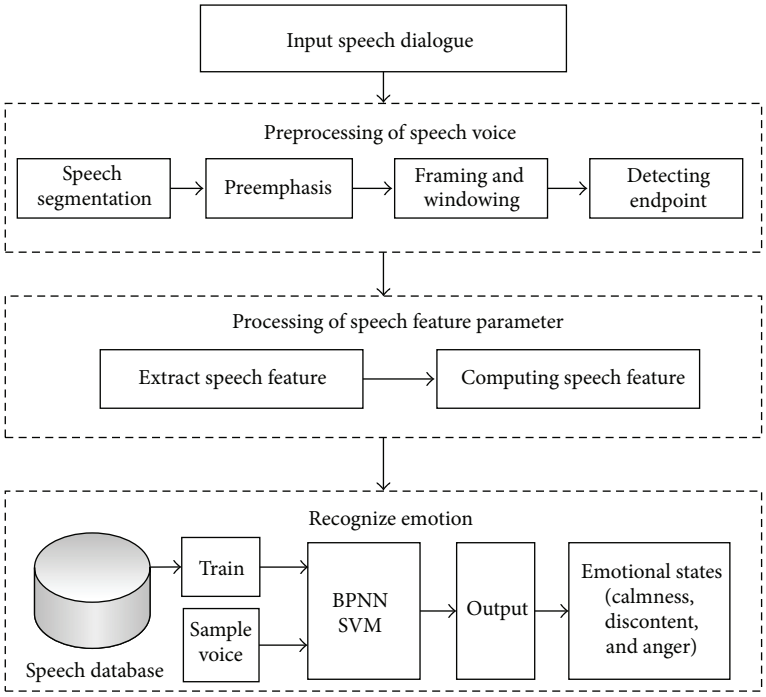


Figure 5: The process of speech emotion recognition.

4.1.1. Characteristics of Telephone Complaints. Because the high frequency part of the spectrum is relatively small in telephone voice signals, the preemphasis processing is usually utilized to enhance the high frequency part of the signals' amplitudes. This is frequently dealt with by the first order high pass filter as shown in the following:

$$H(z) = 1 - az^{-1}. \tag{4}$$

4.1.2. Framing and Windowing. Previous studies have found that the characteristics and physical characteristic parameters of telephone speeches can remain stable in the period of 10 ms–30 ms and can keep the short-time stationary [22, 24]. Therefore, it is necessary to split the speech signals into the time periods so as to analyze them based on the smallest units. This is processed by a frame with the length of 10 ms–30 ms.

Windowing is to select the window function after framing, and there are two factors to be considered, namely, the shape and the length. Generally speaking, we use three windows: rectangle window, Hamming window, and Hanning window.

The rectangular window is shown as in the following formula:

$$w(n) = \begin{cases} 1, & 0 \le n \le N-1 \\ 0, & \text{others.} \end{cases} \quad (5)$$

Hamming window (Hamming) is shown as in the following formula:

$$w(n) = \begin{cases} 0.54 - 0.46\cos\left[\dfrac{2\pi n}{N-1}\right], & 0 \le n \le N-1 \\ 0, & \text{others.} \end{cases} \quad (6)$$

Hanning is shown as in the following formula:

$$w(n) = \begin{cases} 0.5\left[1 - \cos\left(\dfrac{2\pi n}{N-1}\right)\right], & 0 \le n \le N-1 \\ 0, & \text{others.} \end{cases} \quad (7)$$

4.1.3. Endpoint Detecting. The main purpose of endpoint detecting is to use computer technology and digital processing to find the start point and the end point of emotional information contained in a section of speech signal. The basic parameters of the endpoint detecting are short-time energy, short-time average zero crossing rate, and short-time correlation function, and so forth. After detecting endpoints, we will employ the speaker identification algorithm to identify the customer's speeches and put them into the next step for feature parameter extraction.

4.2. Emotional Feature Parameter Extraction. Research findings in the fields of psychology and metrics have pointed out that prosody and voice quality in speeches are the most intuitive indicators to reflect the changes of a speaker's emotions. Statistical analysis shows that if someone is happy, he usually speaks very fast and the volume is high. However, if he is in time of sadness, he tends to speak slowly with relative small intensity. The emotional features which are commonly used in the researches of speech signals include short-time average energy, short-term zero crossing rate, pitch frequency, formant parameters, and Mel Frequency Cepstrum Coefficient (MFCC), as well as a variety of their derived variant forms, such as maximum, minimum, mean, range, and change of the covariance rate [25, 26]. In our study, we adopt MFCC parameters as the emotional features because they have very good ear perception features and contain the comprehensive characteristics of speech emotions. Besides, Mel scale has the advantages of simple calculation and easy distinguishing.

The calculation of Mel frequency is shown in the following formula:

$$\text{Mel}(f) = 2595\lg\left(1 + \frac{f}{700}\right). \quad (8)$$

4.3. Emotion Recognition. As we discussed above, the telephone complaints of a real case are included in the conversations between the service staff and the customer, so we should firstly identify the speakers. And it is performed by the algorithm based on cost-sensitive learning technology.

After detecting the customer's speeches, we adopted the affective computing technology based on BPNN and SVM methods and recognized the dynamic changes of customer's emotions from the extracted feature parameters of his complaints.

In the management of customers' telephone complaints, the typical emotions frequently concerned by the service staff are anger, discontent, and calmness. Anger and calmness are not likely to occur simultaneously; then, discontent can be regarded as their intermediate state. Therefore, anger, discontent, and calmness can reflect the customer's possible emotion changing period in a conversation and can be used to evaluate the service effects in telephone complaint management. In speech recognition algorithm, we used BPNN, SVM methods and cost-sensitive learning technology [10] to optimize recognition rate.

Table 3: Part of the ".wav" files of telephone complaints.

Recorders of telephone complaints		
(1) Anger 01-broadband fault.wav	(4) Discontent 01-broadband fault.wav	(7) Calmness 01-broadband fault.wav
(2) Anger 02-improper charges.wav	(5) Discontent 02-improper charges.wav	(8) Calmness 02-improper charges.wav
(3) Anger 03-harassing messages.wav	(6) Discontent 03-harassing messages.wav	(9) Calmness 03-harassing messages.wav

5. Experiment and Results

5.1. Experiment Data. We took CASIA Chinese speech emotional database and the records from the customer complaint service center of a telecom and Internet service company as experiment data. CASIA was developed by the National Laboratory of Pattern Recognition and Human-Computer Interaction research group at Chinese Academy of Sciences Institute of Automation [8]. It has been widely used as the standard corpus for Chinese language test. In this corpus, each speech with the same semantic texts is spoken by 2 men and 2 women in six different emotional tones: happy, sad, angry, surprise, fear, and neutral. Therefore, it can be used to evaluate the reliability and validity which are only related to the emotions [9].

The records from the customer complaint service center are 252 real speech samples saved as wav files at the sampling rate of 16 kHz. Each sample contains a whole conversation process between the staff and the customer with 6–12 sentences, which includes the dynamic changes of different emotional states. Table 3 shows the part of ".wav" files that we will mention in the following discussion.

Figure 6 shows the characteristic value (mean-intensity and pitch range) of some ".wav" file from Table 3. It can be seen that the values of anger state are much higher than those of calmness state.

Table 4: 12-order MFCC parameters.

	12-order MFCC coefficients				
1	11.9380	13.8424	14.5935	12.0397	16.7550
2	−3.6532	−1.2277	−0.5424	−0.9713	5.7762
3	−1.2243	−0.3909	1.0603	0.4253	3.1929
4	0.0357	1.8983	0.4281	−0.6828	7.2475
5	0.3830	0.7901	0.2184	−0.4252	0.3533
6	−0.6169	−0.2753	−0.0506	−0.3335	... −0.3407
7	1.3436	0.6748	−1.2993	−2.5699	1.1169
8	1.6690	3.4504	4.5318	4.7105	1.6278
9	−1.6734	1.3769	3.5090	4.7951	−1.5150
10	−0.3022	−0.6043	−0.8377	−0.3965	−0.4561
11	−0.9944	−1.0738	0.5673	1.1786	2.0980
12	0.2397	0.3336	0.0626	−0.1852	0.4053

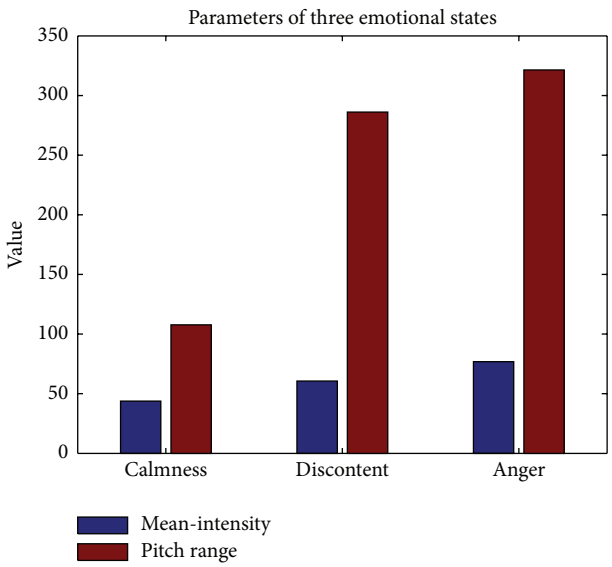


Figure 6: The characteristic value of three emotional states.

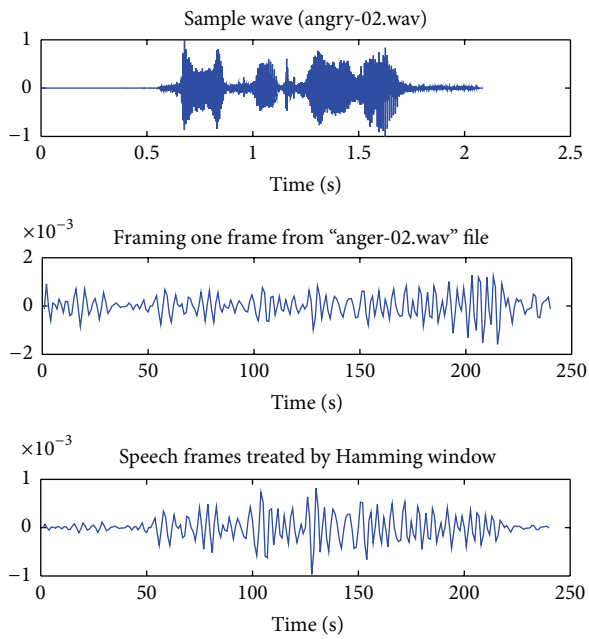


Figure 7: Framing and Hamming windowing of speech file.

In order to suppress the noises and highlight the main features of the speeches, framing and windowing are used to preprocess the samples and transfer the speech signals into frames. For example, the framing and Hamming windowing waveform of file "Anger-02.wav" whose frame length is 240 and the shift is 80 is as shown in Figure 7.

5.2. Feature Extraction. The intensity represents the strength of voice by the amplitude of speech signals. We considered the typical complaint situations such as the complaints of broadband fault and fee deduction of rubbish short messages. For example, "My home broadband has gone wrong. When would you repair it?" Getting fundamental frequency from this voice file is as shown in Figure 8.

The 12-order MFCC parameters extracted from the above sample speech are shown in Table 4 and Figure 9.

5.3. Recognition Results. After passing the test by CASIA database, we apply our method to the 252 real speech samples [5]. Each sample includes a dynamic conversation between the service staff and the customer. The three recognition methods results were shown in Table 5.

From Table 5, we can find that SVM with combined 12-order MFCC and short energy has the highest average recognition rate. It can reach 88.60%, 61.83%, and 89.80% in the dynamic recognition of calmness, discontent, and anger, respectively. The results also show that calmness is not a prominent emotional state and may be the neutral description of customer's psychological states. The proposed method has been successfully applied to the operation quality and service administration in telecom and Internet service company.

6. Conclusion and Discussion

This paper studied the characteristics of customers' telephone complaint speeches and discussed the factors of speaker identification from the conversations as well as the dynamic

Table 5: Average recognition rate of emotions.

Recognition methods	Recognition rate			
	Calmness	Discontent	Angry	Average
BPNN	81.22%	60.46%	80.92%	74.20%
SVM (12-order MFCC)	84.50%	61.40%	83.27%	76.39%
SVM (combined 12-order MFCC and short energy)	88.60%	61.83%	89.80%	80.08%

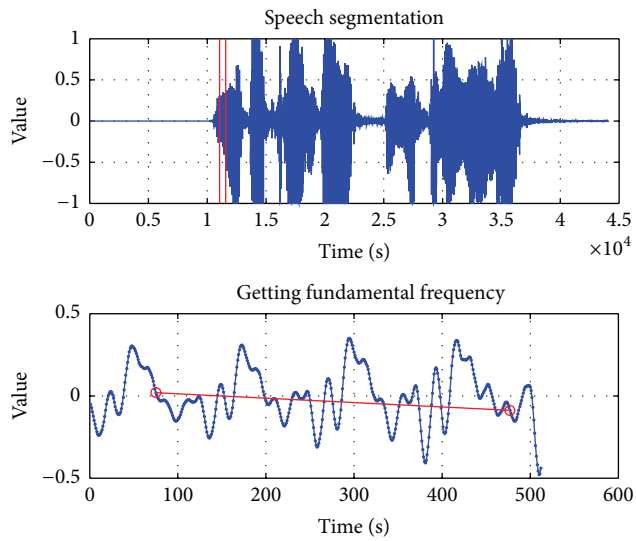


Figure 8: Fundamental frequency of the sample voice.

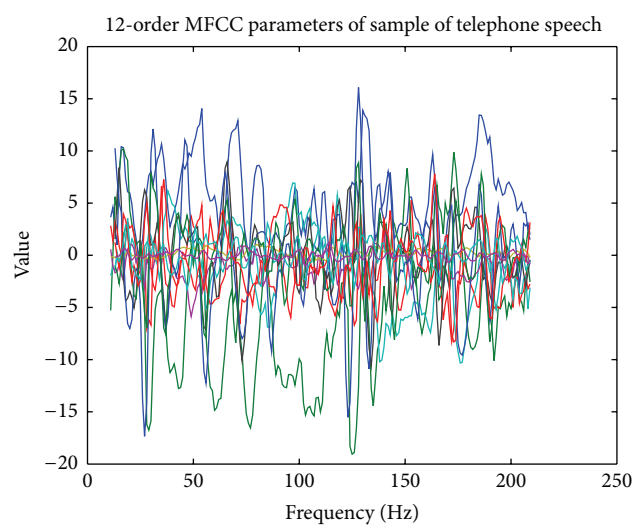


Figure 9: 12-order MFCC parameters extracted from a sample of telephone speech.

emotion recognition of customer speeches. In order to recognize the dynamic emotions in customers' complaints, we first employed the cost-sensitive learning technology to identify the speakers and then utilized the BPNN and SVM methods to realize the recognition of customers' dynamic emotions based on the affective computing from extracted feature parameters. Experimental results show that SVM with combined 12-order MFCC and short energy has the highest average recognition rate of 80.08%.

In the recognition of speech emotions, six typical emotions, namely, anger, distaste, fear, joy, sadness, and surprise, have been well researched based on the isolated speeches. Nwe et al. obtained the accuracy rate of 78% by HMM method [27]. Bhatti et al. developed a modular ANN and reported the correct rate of 83% [28]. The hybrid SVM method with combining acoustic features and linguistic information presented by Schuller et al. can achieve higher accuracy rates than HNN and ANN methods [29]. Emotion recognition from continuous conversations has been a new research issue in recent years [9, 30]. However, there are few studies on telephone complaints. Generally speaking, the reliable recognition rate may be 70%–80% [9].

Due to its specificity and complexity, the dynamic emotion analysis of customers' telephone complaints in the real application is expected for further researches. This paper provides a valuable reference on this issue. In the further improvement, the feature parameters which reflect the calm state may be explored and the quantitative evaluation should be made to show the strength and its dynamic changes of emotions. Besides, a knowledge base considering the differences of the customer's gender, age, and other attributes may be introduced to enhance the recognition rate.

Disclosure

Shuangping Gong and Yonghui Dai are the joint first authors of this paper. Jinzhao Wang and Hai Sun are the joint corresponding authors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors appreciate the anonymous reviewers for their helpful and constructive comments on the earlier draft of this paper. This research was supported in part by Philosophy and Social Sciences Plan, Hunan, China (no. 14YBA041), National Natural Science Foundation of China (no. 91324010), The 2nd Regular Meeting Project of Science and Technology Cooperation between China and Serbia (no. 2-9), Shanghai Philosophy and Social Sciences Plan, China (no. 2014BGL022), and Fudan's Undergraduate Research Opportunities Program, China (no. 15623).

References

[1] N. J. Chen, *Design and Implementation of Telecom Users Complaint Classification and Management System*, Beijing University of Posts and Telecommunications, Beijing, China, 2012.

[2] Y. Dai and M. Lu, "On establishment of the enterprise customer complaint handling system," *Economic Forum*, vol. 4, pp. 85–86, 2007.

[3] J. H. Li and B. Chen, "Telecom customer complaints and suggestions oriented intelligent analysis model," *Modern Telecommunication Technology*, vol. 43, no. 5, pp. 4–7, 2013.

[4] R. W. Picard, *Affective Computing*, MIT Press, Cambridge, Mass, USA, 1997.

[5] J. Ji, *Research on Emotion Analysis Technology of Customer Complaints from Telephone Speech*, Fudan University, Shanghai, China, 2014.

[6] L. Xu, M. X. Xu, and D. L. Yang, "Chinese emotional speech database for the emotional change detection," *Journal of Tsinghua University: Natural Science Edition*, vol. 49, no. 1, pp. 1413–1418, 2009.

[7] M. El Ayadi, M. S. Kamel, and F. Karray, "Survey on speech emotion recognition: features, classification schemes, and databases," *Pattern Recognition*, vol. 44, no. 3, pp. 572–587, 2011.

[8] W. J. Han and H. F. Li, "A brief review on emotional speech databases," *Intelligent Computer and Applications*, vol. 3, no. 1, pp. 5–7, 2013.

[9] W. H. Dai, D. M. Han, Y. H. Dai, and D. R. Xu, "Emotion recognition and affective computing on vocal social media," *Information & Management*, 2015.

[10] D. D. Li, Y. C. Yang, and W. H. Dai, "Cost-sensitive learning for emotion robust speaker recognition," *The Scientific World Journal*, vol. 2014, Article ID 628516, 9 pages, 2014.

[11] Y. J. Luo and J. H. Wu, "Emotional and mental control and cognitive strategy research," *Journal of Southwestern Normal University (Humanities and Social Sciences Edition)*, vol. 31, no. 2, pp. 26–29, 2005.

[12] A. Ortony and T. J. Turner, "What's basic about basic emotions?" *Psychological Review*, vol. 97, no. 3, pp. 315–331, 1990.

[13] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[14] J. A. Russell, "Core affect and the psychological construction of emotion," *Psychological Review*, vol. 110, no. 1, pp. 145–172, 2003.

[15] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and Psychopathology*, vol. 17, no. 3, pp. 715–734, 2005.

[16] K. Danziger, "Wundt's psychological experiment in the light of his philosophy of science," *Psychological Research*, vol. 42, no. 1-2, pp. 109–122, 1980.

[17] W.-J. Han, H.-F. Li, H.-B. Ruan, and L. Ma, "Review on speech emotion recognition," *Journal of Software*, vol. 25, no. 1, pp. 37–50, 2014.

[18] W. Guo, R. H. Wang, and L. R. Dai, "Mel-Cepstrum integrated with pitch and information of voiced/unvoiced," *Journal of Data Acquisition and Processing*, vol. 22, no. 2, pp. 229–233, 2007.

[19] S. Q. Zhang, Z. J. Zhao, B. C. Lei, and G. Y. Yang, "Speech emotion recognition by combining voice quality and prosody features," *Journal of Circuits and Systems*, vol. 14, no. 4, pp. 120–123, 2009.

[20] B. Schuller, G. Rigol, and M. Lang, "Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine—belief network architecture," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, pp. I577–I580, May 2004.

[21] B. X. Wang, D. Qu, and X. Peng, *Practical Fundamentals of Speech Recognition*, National Defence Industry Press, Beijing, China, 2005.

[22] L. Zhao, *Speech Signal Processing*, Machinery Industry Press, Beijing, China, 2nd edition, 2011.

[23] Y. L. Lin, G. Wei, and K. C. Yang, "Research progress of speech emotion recognition," *Journal of Circuits and Systems*, vol. 12, no. 1, pp. 90–98, 2007.

[24] Y. X. Pan, P. P. Shen, and L. P. Shen, "Speech emotion recognition using support vector machine," *International Journal of Smart Home*, vol. 6, no. 2, pp. 101–108, 2012.

[25] L. J. Xie, H. Wen, and N. F. Xiao, "Research on the calculation model of home service robot emotion," *Computer Engineering and Design*, vol. 33, no. 1, pp. 322–327, 2012.

[26] S. L. Li, R. Liu, L. Q. Zhang, and H. Liu, "Speech emotion recognition algorithm based on modified SVM," *Journal of Computer Applications*, vol. 33, no. 7, pp. 1938–1941, 2013.

[27] T. L. Nwe, S. W. Foo, and L. C. De Silva, "Speech emotion recognition using hidden Markov models," *Speech Communication*, vol. 41, no. 4, pp. 603–623, 2003.

[28] M. W. Bhatti, Y. Wang, and L. Guan, "A neural network approach for human emotion recognition in speech," in *Proceedings of the IEEE International Symposium on Cirquits and Systems (ISCAS '04)*, pp. II81–II84, May 2004.

[29] B. Schuller, G. Rigoll, and M. Lang, "Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 1, pp. 577–580, IEEE, May 2004.

[30] C. M. Lee and S. S. Narayanan, "Toward detecting emotions in spoken dialogs," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 2, pp. 293–303, 2005.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 181038, 10 pages
http://dx.doi.org/10.1155/2015/181038

Research Article

Information Dissemination of Public Health Emergency on Social Networks and Intelligent Computation

Hongzhi Hu,[1] Huajuan Mao,[2] Xiaohua Hu,[3] Feng Hu,[4] Xuemin Sun,[5] Zaiping Jing,[2] and Yunsuo Duan[6]

[1] *School of Management, Fudan University, Shanghai 200433, China*
[2] *Department of Vascular Surgery, Changhai Hospital, Second Military Medical University, Shanghai 200433, China*
[3] *Department of Information, Changhai Hospital, Second Military Medical University, Shanghai 200433, China*
[4] *Department of Respiratory Medicine, Shanghai Tongren Hospital, Shanghai Jiao Tong University, Shanghai 200336, China*
[5] *Department of General Surgery, Tongji Hospital Affiliated to Tongji University, Shanghai 200065, China*
[6] *Psychiatry Department, Columbia University/New York State Psychiatric Institute, New York City, NY 10032, USA*

Correspondence should be addressed to Feng Hu; hf3021@shtrhospital.com and Zaiping Jing; 1760051610@qq.com

Received 19 February 2015; Accepted 13 July 2015

Academic Editor: Abdul Hanan Abdullah

Due to the extensive social influence, public health emergency has attracted great attention in today's society. The booming social network is becoming a main information dissemination platform of those events and caused high concerns in emergency management, among which a good prediction of information dissemination in social networks is necessary for estimating the event's social impacts and making a proper strategy. However, information dissemination is largely affected by complex interactive activities and group behaviors in social network; the existing methods and models are limited to achieve a satisfactory prediction result due to the open changeable social connections and uncertain information processing behaviors. ACP (artificial societies, computational experiments, and parallel execution) provides an effective way to simulate the real situation. In order to obtain better information dissemination prediction in social networks, this paper proposes an intelligent computation method under the framework of TDF (Theory-Data-Feedback) based on ACP simulation system which was successfully applied to the analysis of A (H1N1) Flu emergency.

1. Introduction

Public health emergency refers to a sudden event which may cause serious damage to the social public health such as major infectious disease, mass unknown illnesses, and major food and occupational poisoning. In today's society, public health emergencies, such SARS in 2003, Bird Flu in 2006, A (H1N1) Flu in 2009, and EBHF in 2014, have triggered extensive social influence by the dissemination of related information.

In order to get a comprehensive understanding of unconventional emergencies which possibly bring serious impacts on the country and society and improve the government's coping capacity, a major research plan was launched by the NSFC (Natural Science Foundation of China) in 2009 and organized a large interdisciplinary research community with more than 300 scientists, technicians, and engineers. As an important part of the above research plan, our work is focused on the information dissemination and computation associated with public health emergency.

Social networks have become an open complex giant system that connected more than two-fifths of the world populations and acted as the main platform of information dissemination in public health emergency, which can easily cause the spread of false facts or rumors and bring about a serious panic without efficient management [1, 2]. Therefore, the prediction of information dissemination in social networks is necessary for estimating the social impacts and making a proper strategy in the management of public health emergency.

In the past decades, numerous research works have been devoted to social networks [3–7]. Scholars conducted the technologies of content analysis and topic detection to study

the information spread path as well as its key nodes [4, 5] and furthermore probed into the topological structure and organizational mechanism of the spread network [6–8]. It has been realized that information dissemination in social networks is affected not only by the variable network's structure and dynamically changing spread paths, but also by the complex interactive activities and group behaviors in social communities [7, 8]. Therefore, statistical models such as time regression or estimators based on the traditional machine learning technologies are difficult to achieve good prediction results [6, 9].

In recent years, in the view of complex adaptive systems, researchers put forward many models to describe the information dissemination characteristics in social networks, such as "the small world" model, information diffusion model of dynamics, and infectious disease diffusion model like SIS (Susceptible, Infected, and Susceptible) [10, 11]. Those research findings described the behavioral mechanism and theoretical characteristics in the process of information dissemination. However, there are some barriers to be applied in the prediction of real situation [9].

On the other hand, some researches were focused on data mining and established the predictive model based on the historical data and existing cases. There are also some limitations due to the dynamic changing social connections in cyber space and the uncertain processing behaviors which are largely affected by the cognition and reactions of members in social networks [6]. Considering these complex characteristics and factors, an effective computation method aiming at achieving a better prediction result of information dissemination needs to be explored.

2. Complex Characteristics of Information Dissemination in Social Networks

2.1. Network Structure of Information Dissemination. The online forum, Internet micro-blog, instant messaging platforms, mobile network, and various communication terminals consist of a highly complex system [12, 13] for people to share information with others and express their opinions to the public. The increasing scale of information dissemination makes existing mathematical model difficult to describe the real situations and simulate the dynamic process adequately.

Once an emergency event takes place, the information dissemination will usually experience a life cycle of five stages: incubation stage, outbreak stage, diffusion stage, decaying stage, and aftermath stage [8]. According to the existing research findings [6], the information dissemination in social networks appears in the forms of two kinds of patterns: the public pattern such as BBS and micro-blog, and the small world pattern such as Wechat and QQ. The network structure of information dissemination in public pattern is an open and unstable structure with a lot of uncertainty. However, the structure in the small world pattern is relatively stable but has higher efficiency and easily leads to group behaviors because it is usually based on the familiar members and a vocal media system [14].

In public health emergency, information dissemination in social network involves three roles of the participants [15]: information senders, information disseminators, and information receivers, as shown in Figure 1.

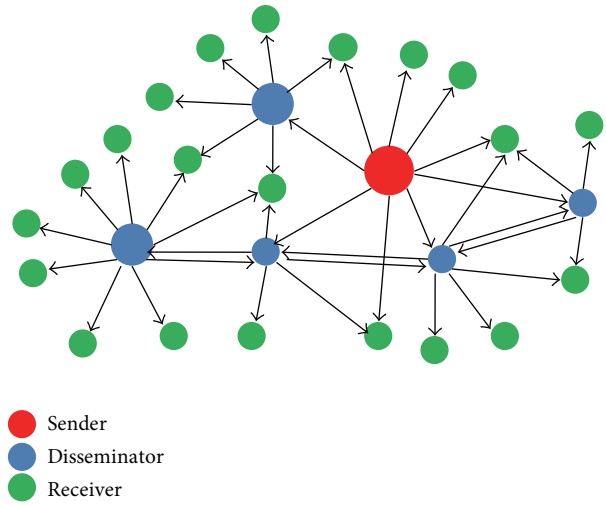


Figure 1: Participants of information dissemination in social networks.

Figure 2: The leading nodes and paths of information dissemination in social network.

However, the roles of participants are often changing by many factors during the process of dissemination with different kinds of events, such as interest and attention, the level of activity, and the time to catch relevant information. This will lead to dynamic changes of spread paths and network structure under different events. For example, Figure 2 shows its main leading nodes and paths of information dissemination in a social network which are expressed in different size and colors. The larger red nodes in the figure are the leading nodes, and the thicker links of each node are the main paths.

Different from the inflexible structure in electronic circuit, the connections in social networks are actually flexible and random. It means that the features of network structure are dependent on the dynamic strength among different nodes. For convenience, we define the expressions as follows.

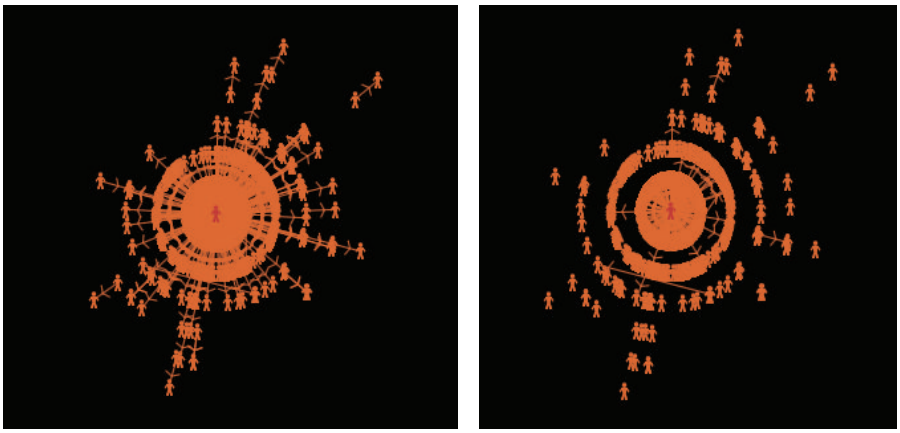


Figure 3: Dynamic structure of a micro-blog network when the spreading threshold varies from 0.3 to 0.5.

The connections in social networks can be regarded as a graph G, and

$$G = \{V, E\}, \tag{1}$$

where V is the collection of nodes and E is the collection of connections among the nodes:

$$\begin{aligned} V &= \{v_i\}, \quad i = 1, 2, 3, \ldots, M, \\ E &= \{e_j\}, \quad j = 1, 2, 3, \ldots, N. \end{aligned} \tag{2}$$

In the above, k_i is the number of active communication on connection e_j in a certain period t; then e_j is an effective connection if $k > 0$, or an ineffective connection if $k = 0$.

Figure 3 shows the dynamic structure of a micro-blog network when the spreading threshold varies from 0.3 to 0.5 [15]. It can be observed that the network structure is changing with time by different connection strengths; some connections may disappear but those strong connections still exist.

However, in the prediction of information dissemination in social networks, we should consider the complicated characteristics of its based network structure [6]. We will discuss this in the following section.

2.2. Cognitive Psychology and Group Behaviors. It has been verified by many real cases and empirical studies that information dissemination in social networks has close relationship with the participants' psychological reactions and their behaviors in particular situational context [6–8, 16, 17].

The process of information dissemination may be affected by many factors, among which the emotions and behaviors of participants have direct impacts on the development of public opinions. In social network, users can release and disseminate all types of information with real names or anonymous way, while the validities or authenticity of the information may be difficult to determine at first glance [18]. The information dissemination is thus beyond the technological control and subject to individual attitude, emotion, and behavior and largely depends on the trust of network relations and human cognitive process under specific situations. Furthermore, the individual cognitive nature of attitude, emotion, and behavior will easily result in the group behaviors in social networks.

Cognitive theory of emotion considers that emotion is affected by three factors: environmental events, physical condition, and cognitive process, among which cognitive process is the decisive factor [19]; the information processing especially elaborates on the content of the received information. According to limited cognitive resources, individuals usually identify and process information in terms of selective attention in cognitive processing [20].

In addition to the information content, people also consider the factors of information itself, for example, source credibility and information length. Some research shows that the social and cultural backgrounds and emotions of information disseminators will affect their cognitive process significantly, which is obvious in the Internet public opinions. For instance, some information will be largely reproduced, spread, and reposted in a very short time and then lead to large-scale riots and even affect daily life. In the above process, cognitive results and emotional reactions exhibited by groups are the core elements that lead directly to their coping strategies and behaviors; these mechanisms have been systematically studied by Bagozzi and Dholakia [21], Wheeless and Grotz [22], and Barnes and Olson [23] from the aspects of self-disclosure, opinion leaders, and opinion followers on social networks.

Recent researches of cognitive psychology and group behaviors in social networks have developed into a new interdisciplinary field which was called Cyber Psychology and defined as "understanding people how to react and behave within cyberspace" by Dr. Suler in his hypertext book *The Psychology of Cyberspace* [24]. The progress in social neuroscience [25] has provided new aspects for achieving a better study in this field. Remarkably, modern technology of functional magnetic resonance imaging (abbreviate to f-MRI) can directly track reactions involved with all types of situational information stimuli in cognitive process [26, 27], and ERPs (short for Event-related Potentials) can get high resolution in dynamic process which has been used to study sensitive reactions [28]. These will help to explore the brain mechanism of cognition in information dissemination.

Based on new theory and advanced experimental technologies, researchers have found complicated factors related to cognition and group behaviors which may affect the information dissemination of public health emergency in social networks, such as attention and interest on the events, psychological characteristics of people with different regions, cultural background, and specific situations context. However, the existing findings are mostly obtained by questionnaire surveys or experimental observations which may easily lead to the incomplete results due to the limitations of investigated samples or particular experimental cases. So, find a systemic way that is needed in applying in real environment [6].

According to our previous research [6], we found that the dissemination behaviors for an information receiver are related to the factors as in the following formula:

$$B_{t+1} = f\left(C, A_t, E_t, B_t\right), \tag{3}$$

where B_{t+1} is the dissemination behavior of an information receiver at time $t + 1$. C denotes the event's content; A_t, E_t, and B_t are the states of the information receiver's attitude, emotion, and his current behavior, respectively, at time t.

However, the quantitative relationship in formula (3) cannot be calculated accurately only by a simple mathematical formula. Some knowledge such as rules and behavioral characteristics of the information receiver are required [6]. In order to get an approximate estimation in the next time period, Kalman Filtering is a better choice shown in the following formula:

$$B_{t+1}\left(t \mid t-1\right) = LB_t\left(t-1 \mid t-1\right) + MU\left(t\right). \tag{4}$$

Information dissemination in social networks also pertains to the time period of occurrence, spread channels, type and quantity, and other dynamic circumstances. Moreover, the way of its dissemination is complicated, and the relations between them are complex [17]. Each participant may be regarded as a dissemination node, and it is not only the receiver but also a sender. The information dissemination through the Internet concerning the event is a complex mode and is featured to be divergent and fast. Daily interpersonal communications and telecommunication communications have been fused into information dissemination in whole social networks and constitute complex patterns of dissemination.

3. Intelligent Computation Method Based on TDF Framework

3.1. Limitations of Existing Methods. Unlike deep and wide research that was carried on epidemic spreading, quantitative research on the social networks' information dissemination has been relatively limited, and, furthermore, the existing methods do not sufficiently take into account the effects of social cognition, group behaviors, multichannel dissemination, and other characteristics or just make some macroscopic and statistical descriptive models. For the above reasons, we recognize that in practical applications, theoretical tools and modeling methods are difficult to adequately describe and accurately predict the information dissemination in social network. The main reason leads to the abovementioned problems being that the dissemination is affected by too many complex factors in dynamic changing environment due to cognitive psychology, emotional reactions, and behavioral intention of the network users which may result in uncontrollable processing behaviors and will continually produce new stimulus to other receivers. Under these circumstances, we find that it is not enough to solve this problem if we only rely on various theoretical models to describe basic features or artificial simulation systems derived from the historical data mining in such an open and uncertain environment.

On the whole, the existing methods have the following shortcomings: (1) It is difficult to apply a statistical model to predict actual situations in open and dynamic environment with macroscopic description based on sample data. (2) Some theoretical models can reflect the mechanism and some basic rules under an ideal condition, but there are huge variations as to actual situation in the real world. (3) Technologies of big data mining can provide valuable analysis and parameters for those models, but the future uncertainties may not accord with the changing tendency obtained by historical data. So, it needs further exploration to find a more effective method to deal with such complicated problem.

3.2. ACP Simulation System. Information dissemination in social networks involves an interdisciplinary study which covers psychology, sociology, mathematics, management science, and computer engineering. Such work is difficult to be clearly expressed just by mathematical computation formulas but should utilize the unstructured knowledge such as a variety of rules, empirical data, and cases. It is expected to exploit new ideas and methods from the integrated perspective of a real application environment.

For the purpose of exploring how to offer comprehensive solution to the scientific problems of complex social and economic system, Wang from Institute of Automation, Chinese Academy of Sciences, firstly proposed a parallel computational theory to solve management and control issues in complex system and put forth social computing method that combined the artificial society with computational experiments and parallel execution, which is called ACP (artificial societies, computational experiments, and the parallel execution) [29]. This method can deal with the difficult problems such as Cyber Psychological computation and make it computable; Figure 4 shows the ACP simulation system which has been successfully applied to the dynamic analysis in the national emergency management and responses of China [6, 30].

ACP simulation system provides an effective way to solve the above problem and new clue to further experimental research. By using this, complicated problems which are not easy to be accurately predicted, difficult to precisely model, and unable to repeat experiments in real social system all can be addressed. In this computing environment, the real-time system continuously exchanges data and synchronously modifies the artificial social context in the light of the feedback data. Through iterative interaction and persistent approximation, it can help to realize the parallel computation

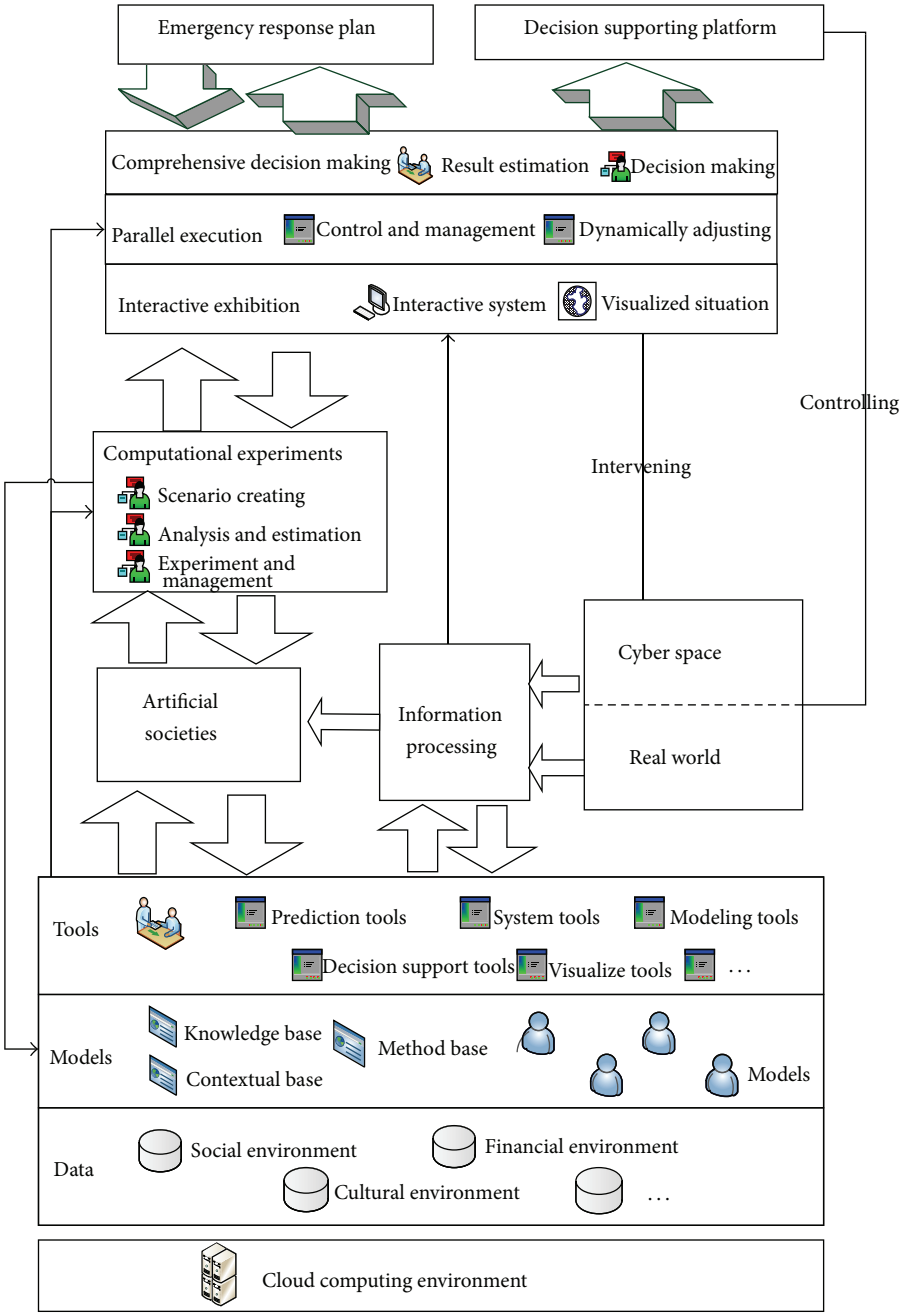


Figure 4: ACP simulation.

between artificial society and real society and finally achieve the goal of dynamic optimization management and control. The ACP method can also serve as a bridge between social problem and computing technology and break the dilemma of interdisciplinary study between social and computation science. The method also has important significance in virtual community information dissemination from the perspectives of fundamental theory, experimental methods, key technology, and practical applications.

Information dissemination in social networks has very close relationship with psychological and emotional behavior in particular situational context, as well as its temporal-evolution-characteristics and complex spreading mechanism. In the face of all uncertainties of future context, in order to gain relevant experiential knowledge and important data parameters, further study should focus on cognitive mechanism of network information, trend predictions, and reaction rules. Only through deep experimental observations, instant social psychological research in addition to large exact cases will be able to make more accurate analysis and evaluation on the information dissemination rules, evolution characteristics, and developing trends.

3.3. Intelligent Computation Based on TDF Framework. In order to get a better prediction of information dissemination in social networks, Professor Dai proposed a new TDF (Theory-Data-Feedback) framework to deal with the modeling problem [9]. His framework absorbed the merits of the mechanism model, data model, and social psychological feedback model interacted with instant online survey data and formed a new systematic analysis and modeling method. We extended that framework to be applied in ACP simulation system for analyzing the information dissemination of public health emergency as follows.

The mechanism model includes the basic laws of things change, existing research results, and related prior knowledge on similar problems, such as network information spreading mechanism and evolution characteristics. It lays foundation for normal operation of the whole model. In this part, information dissemination mechanism needs comprehensive analysis to determine the range of data acquisition in "input" part of the model, which is of great importance in accuracy of predicted results.

The data model stores historical data and reflects current state. This part accumulates relative data prepared for the construction of reaction rules on future uncertainties. In this part, by using methods like functional brain mapping analysis technology (i.e., f-MRI) or cognitive neurology analysis can obtain statistical characteristics of users' emotional and behavioral responses to situational context.

The feedback model is used to reflect response rules on uncertainties of future; this part will simulate the actual environment and predicts the following behavior if possible, so as to seek for effective solution or provide a basis for improvement to current solution in real-life situations. In order to reduce differences, the feedback system will begin a new round of optimization and evaluation procedure and generate error signal to further revise the assessment methods or parameters of artificial system through the observation of related state changes under the actual changes automatically.

By applying interactive iteration procedure and parallel computation analysis in both artificial and real society can obtain much better approximation descriptions of real environment and, in part, can predict the trend which in turn effectively manage and control changing problem in real social system. In this way, the model will be of great help to decision making for coping with dynamically changing situation. Figure 5 shows the TDF framework for intelligent computation on social networks' information dissemination of public health emergency.

The application and realizing process of TDF framework can be divided into the following three steps.

(1) Theory: Mechanism Model. The famous disease spread model firstly presented by Grassberger in 1983 [31] has been widely used to describe the essential features of information dissemination in social networks. In this model, each node in a social network has two states S and I. Here, S represents the initial state, when each node receives a message; it can be transformed into I with certain probability. After spreading a message, node in state I will return to state S which has the possibility to be infected or delete the message that may forget

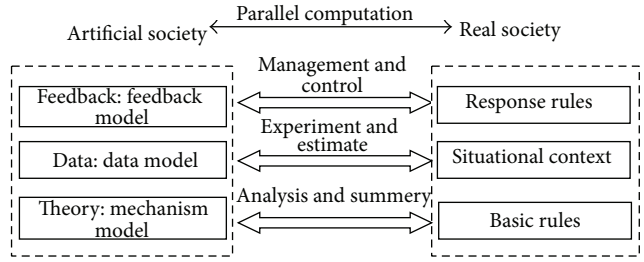


Figure 5: TDF framework.

Table 1: Attributes of agents in artificial society.

Attributes	Parameters
Serial number	$n = \{1, 2, 3, \ldots, N\}$
Social relationship	$\mathrm{Sr}(n) = [0, 1]$
Attention to this topic	$\mathrm{At}(T) = [0, 1]$
Attitude	$\mathrm{Ad}(T) = [-5, 5]$
Emotion	$E(t) = [-5, 5]$
Information dissemination role	(1) Opinion leader (2) Follower (3) Controller (4) The rest
Information dissemination state	$\mathrm{DS} = \{S(t), I(t), p(t)\}$

or not be interested with the topic with certain probability. Consider new comers and inactive ID users; the dynamic transmission model can be expressed as follows [32]:

$$\begin{aligned} \frac{dS_k(t)}{dt} &= b\left[1 - S_k(t) - I_k(t)\right] - \lambda k S_k(t)\,\Theta(t) \\ &\quad + \mu I_k(t) - d S_k(t), \\ \frac{dI_k(t)}{dt} &= \lambda k S_k(t)\,\Theta(t) - \mu I_k(t) - \varepsilon I_k(t), \end{aligned} \tag{5}$$

where $\Theta(t) = \sum(ip(i)I_i(t)/\sum kp(k))$.

It represents the probability that a given side connects with an information received node. Here k represents the degree of that node, λ represents the transmission threshold of a social network, t is a unit time, and $p(k)$ is a distribution function of k.

In order to endow the prior knowledge and basic rules of mechanism model into the agents' simulation system in artificial society, the key attributes of agents can be designed as in Table 1.

Hereafter, the basic rules of mechanism model can be designed as follows.

When receiving a message, the changes in attributes of an agent are decided by

$$\begin{aligned} &\mathrm{Agent}(t + \Delta t) \\ &\quad = I\left(Sd(a), \mathrm{At}(T), \mathrm{Ad}(T), E(t), p_r(t)\right). \end{aligned} \tag{6}$$

Here, $Sd(a)$ is the attributes of the agent who sent that message, $\mathrm{At}(T)$ is the attention to this topic by the

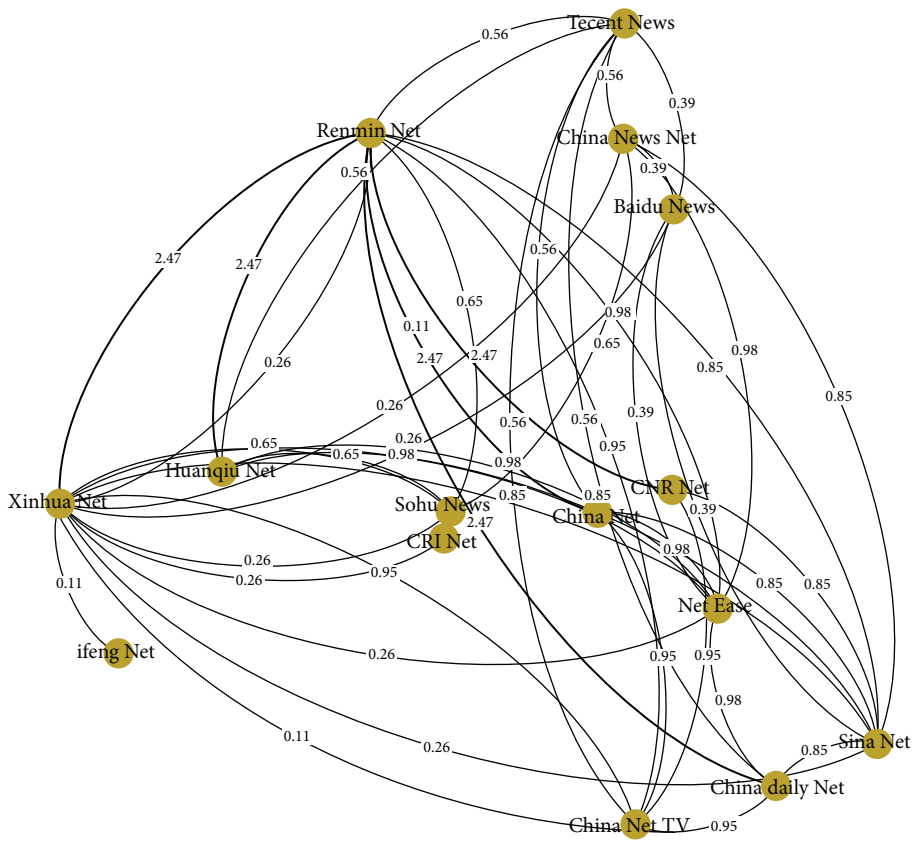


Figure 6: Relationship diagram of interconnections and strengths among websites.

information received agent, Ad(T) denotes the attitude to the event reflected in the information which is initially fixed by interest, $E(t)$ is emotion at time t, and $p_r(t)$ is the probability from state S to state I.

After sending a message, the changes in attributes of that agent are decided by

$$\text{Agent}\,(t+\Delta t) = O\left(E\left(t\right), p_s\left(t\right)\right). \tag{7}$$

Here, we mainly consider the changes of emotions which may be decayed as $E(t+\Delta t) = E(t) * e^{-k/\Delta t}$ according to the dynamic characteristics of emotions [6]. $p_s(t)$ is the probability returning from state I to state S.

The dissemination directions and scopes are decided by the social relationship Sr(n) of the information sending agent. Detailed relationships and parameters of formulas (6) and (7) will be achieved from data model by machine learning [32].

(2) Data: Data Model. Data model should be built from the historical data under specific situational context in the real society. We selected mainstream media websites covering the total number ranked top 15 that were released by Chinese Internet data platform. These sites represent the vast majority of Internet users' access channel to news in China. We use these 15 web sites as nodes to generate web site correlation model by analyzing the link number of each site that pointed to other sites by weighted graph. With this model, we may track information about Internet users' personal browsing behavior affected from the strength of interconnection network sites, which will further affect netizens' emotions and cognition. Considering the information dissemination of public health emergency, we can obtain the relationship diagram as shown in Figure 6.

In Figure 6, the size of the node indicates the number of Internet users, and the weight reveals the strength between each node. The attributes of agents contained in each node can be endowed by machine learning from historical data. Technologies such as computations of attentions, attitudes, emotions, and information dissemination roles have been well developed in the existing researches [6, 14, 32, 33]. If only considering the macro information dissemination among websites, we can regard each website as an agent and endow their statistical attributes by machine learning [32, 33].

(3) Feedback: Feedback Model. As discussed in this paper before, the aim of ACP simulation system is to establish an artificial society which can "follow" the changes in the real society and predict its trends in the future by the parallel computation and interactive iteration procedure. Therefore, the changed information in the real society should be returned into the artificial society so as to adjust the attribute parameters of intelligent agents.

We designed a feedback system which can acquire and track the new information disseminations in social networks in the real society. The Cyber Psychological computation method [34, 35] is employed to evaluate the real changes

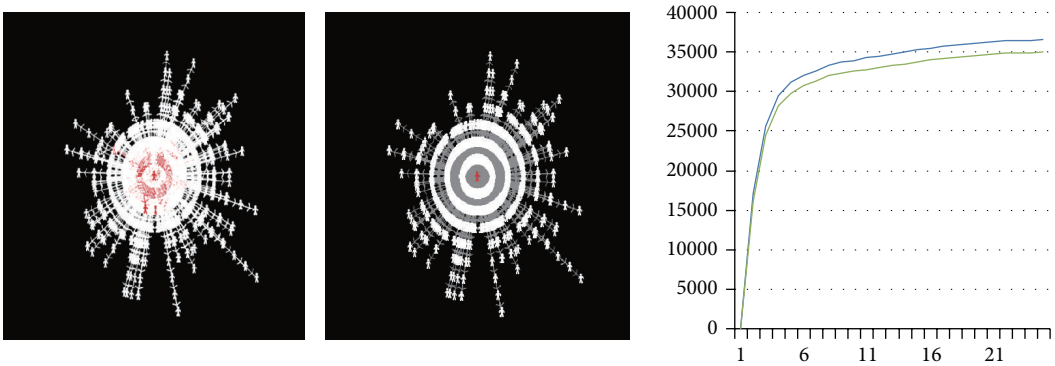


Figure 7: Information dissemination and prediction.

in attentions, attitudes, and emotions of the member in a social network. Corrections of the agents' attributes are usually made every one hour. The predictions of their changes in the artificial society are based on the Kalman Filtering model as shown in formula (4). Besides, an online survey is also utilized to update the new response rules of people's possible behaviors under the upcoming situations, which can be combined into the rules of agents' activities in the artificial society. Generally, the analyses and predictions of information disseminations in social networks are based on the dynamic changes of agents' attributes as the results of their interactive activities in the artificial society, which is different from the traditional methods based on the computation of a mathematical prediction model.

4. Experiment and Results

Compared with the traditional methods, ACP simulation and TDF parallel computation have the superiorities of the feedback adjustment ability which can follow the changes in the real society and the complex description ability which considers the inherent psychological and behavioral mechanisms of the participants in information disseminations. Therefore, the proposed method can achieve an ideal prediction result and has been successfully applied in the analysis of complex public health emergency such as Bird Flu, A (H1N1) Flu, and Ebola Outbreak [6, 32, 34, 36].

Table 2 shows the dynamic information dissemination in social networks, which was recorded from 16008 nodes when an A (H1N1) Flu emergency took place in China in 2009.

Figure 7 illustrates the information dissemination in social networks and the prediction of total number of disseminators. The left images in Figure 7 are the topology network of dissemination in social network within 24 hours. We point out the leading nodes (red nodes) and segment spreading groups into different strengths in this process. The right half in Figure 7 is the total number of disseminators. The blue line is the data of predication, and the green line is the actual number; it shows that the data of predication is very close to the real number. This indicated that the computation based on TDF framework can reach a very high precise prediction for the dynamic information dissemination in social networks. This method has been successfully applied to the national simulation platform for emergency management in China [32, 36].

Table 2: Information dissemination on social networks.

Time (hours)	Number of information dissemination
1	17107
2	8530
3	3896
4	1624
5	937
6	585
7	698
8	337
9	255
10	314
11	264
12	250
13	261
14	272
15	204
16	200
17	181
18	166
19	132
20	98
21	98
22	70
23	63
24	80

5. Conclusion and Discussion

This paper analyzed the complex characteristics of information dissemination in public health emergency of social network as well as its network structure, cognitive psychology, and group behaviors and argued that the existing theoretical tools and modeling methods are not sufficient to accurately describe and predict the information dissemination in social networks. Therefore, a new intelligent computation method

based on the framework of TDF (Theory-Data-Feedback) was constructed for the ACP simulation and prediction on the dynamic dissemination of emergency event's information and reached a high precise result.

Disclosure

Hongzhi Hu and Huajuan Mao are the joint first authors of this paper. Zaiping Jing and Feng Hu are the joint corresponding authors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by National High-tech R&D Program (863 Program) of China (no. 2012AA02A612), Natural Science Foundation of China (no. 91324010), and Shanghai Philosophy and Social Sciences Plan, China (no. 2014BGL022).

References

[1] D. S. Mileti and L. Peek, "The social psychology of public response to warnings of a nuclear power plant accident," *Journal of Hazardous Materials*, vol. 75, no. 2-3, pp. 181–194, 2000.

[2] F. Chierichetti, S. Lattanzi, and A. Panconesi, "Rumor spreading in social networks," *Theoretical Computer Science*, vol. 412, no. 24, pp. 2602–2610, 2011.

[3] M. Kilduff and W. Tsai, *Social Networks and Organizations*, Sage Publications, London, UK, 2003.

[4] X. Liu, *Research and Implementation of Topic Discovery and Tracking Technology Based on the Internet Data*, Fudan University, Shanghai, China, 2010.

[5] W. H. Dai, H. Z. Hu, T. Wu, and Y. H. Dai, "Information spread of emergency events: path searching on social networks," *The Scientific World Journal*, vol. 2014, Article ID 179620, 7 pages, 2014.

[6] W. H. Dai, L. H. Huang, D. R. Xu et al., *Information Classification, Dissemination and Cognitive Model in Unconventional Emergency*, National Natural Science Foundation of China, 2015.

[7] H. Z. Hu, D. Wang, W. H. Dai, and L. H. Huang, "Psychology and behavior mechanism of micro-blog information spreading," *African Journal of Business Management*, vol. 6, no. 35, pp. 9797–9807, 2012.

[8] W. H. Dai, X. Q. Wan, and X. Y. Liu, "Emergency event: internet spread, psychological impacts and emergency management," *Journal of Computers*, vol. 6, no. 8, pp. 1748–1755, 2011.

[9] W. H. Dai, "Public cognition morale mechanism of unexpected incidents in cities and the responding strategies," *Urban Management*, vol. 2014, no. 1, pp. 34–37, 2014.

[10] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[11] M. Draief, "Epidemic processes on complex networks—the effect of topology on the spread of epidemics," *Physica A*, vol. 363, no. 1, pp. 120–131, 2006.

[12] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[13] M. E. J. Newman, "The structure of scientific collaboration networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 2, pp. 404–409, 2001.

[14] W. H. Dai, D. M. Han, Y. H. Dai, and D. R. Xu, "Emotion recognition and affective computing on vocal social media," *Information and Management*, 2015.

[15] J. H. Ju, *Research on Dynamic Simulation of the Spread of Micro-Blog Information in Emergency Event*, Fudan University, Shanghai, China, 2014.

[16] Z. R. Liu, "From irrational network public opinion to study netizen group psychology," *Modern Communication*, vol. 3, article 167, 2007.

[17] Y. X. Zhang, *Research on Communication Mechanism of Sina Micro-Blog-Take Yaochen's Sina Micro-Blog for Example*, Southwest University, Nanjing, China, 2011.

[18] H.-P. Lu and K.-L. Hsiao, "Understanding intention to continuously share information on weblogs," *Internet Research*, vol. 17, no. 4, pp. 345–361, 2007.

[19] X. Y. Guo and Z. L. Yang, *Basic Experimental Psychology*, Academic Press, 2005.

[20] F. Liu and B. Zuo, "Intergroup emotions theory and the researches," *Advances in Psychological Science*, vol. 18, no. 6, pp. 940–947, 2010.

[21] R. P. Bagozzi and U. M. Dholakia, "Intentional social action in virtual communities," *Journal of Interactive Marketing*, vol. 16, no. 2, pp. 2–21, 2002.

[22] L. R. Wheeless and J. Grotz, "Conceptualization and measurement of reported self-disclosure," *Human Communication Research*, vol. 2, no. 4, pp. 338–346, 1976.

[23] H. L. Barnes and D. H. Olson, "Parent-adolescent communication and the circumplex model," *Child Development*, vol. 56, no. 2, pp. 438–447, 1985.

[24] J. Suler, *The Psychology of Cyberspace*, 2015, http://users.rider.edu/~suler/psycyber/psycyber.html.

[25] J. T. Cacioppo and G. G. Berntson, "Social psychological contributions to the decade of the brain: doctrine of multilevel analysis," *American Psychologist*, vol. 47, no. 8, pp. 1019–1028, 1992.

[26] D. Cohen, "Magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer," *Science*, vol. 175, no. 4022, pp. 664–666, 1972.

[27] Y. Ma, G. Zheng, J. Zhou, Z. Zhang, Y. Zhong, and G. Lu, "Review of methods for functional brain integration detection using fMRI," *Acta Biophysica Sinica*, vol. 27, no. 1, pp. 18–27, 2011.

[28] Y. D. Feng, *Research on the information framing effects under paroxysmal events based on brain potential analysis [Ph.D. thesis]*, Zhejiang University, Hangzhou, China, 2013.

[29] F. Y. Wang, "Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social-economic systems," *Complex Systems and Complexity Science*, vol. 1, no. 4, pp. 26–34, 2004.

[30] F. Y. Wang, X. G. Qiu, D. J. Zeng, Z. D. Cao, and Z. C. Fan, "A computational experimental platform for emergency response based on parallel systems," *Complex Systems and Complexity Science*, vol. 7, no. 4, pp. 1–10, 2010.

[31] P. Grassberger, "On the critical behavior of the general epidemic process and dynamical percolation," *Mathematical Biosciences*, vol. 63, no. 2, pp. 157–172, 1983.

[32] H. Z. Hu, *Network information dissemination and cyber psychological computation on emergency events [Ph.D. thesis]*, Fudan University, Shanghai, China, 2015.

[33] X. D. Zhao, *The Simulation Research on Information Emotional Cognition and Dissemination in Public Safety Incidents*, Fudan University, Shanghai, China, 2014.

[34] W. H. Dai, *Cyber psychological and physical (CPP) computation based on social neuromechanism [Ph.D. thesis]*, Fudan University, Shanghai, China, 2015.

[35] X. Zhou, G. H. Dai, S. Huang et al., "CyberPsychological computation on social community of ubiquitous learning," *Computational Intelligence and Neuroscience*. In press.

[36] P. Zhang, B. Chen, L. Ma et al., "The large scale machine learning in an artificial society: prediction of the Ebola outbreak in Beijing," *Computational Intelligence and Neuroscience*. In press.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 762403, 8 pages
http://dx.doi.org/10.1155/2015/762403

Research Article

Explore Awareness of Information Security: Insights from Cognitive Neuromechanism

Dongmei Han,[1,2] Yonghui Dai,[1] Tianlin Han,[1,3] and Xingyun Dai[4]

[1] *School of Information Management and Engineering, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, China*
[2] *Shanghai Financial Information Technology Key Research Laboratory, 777 Guoding Road, Shanghai 200433, China*
[3] *Shanghai Foreign Language Education Press, Shanghai International Studies University, 550 West Dalian Road, Shanghai 200083, China*
[4] *School of Management, Fudan University, 220 Handan Road, Shanghai 200433, China*

Correspondence should be addressed to Yonghui Dai; dyh822@163.com and Xingyun Dai; 082025041@fudan.edu.cn

Received 9 February 2015; Revised 6 May 2015; Accepted 15 June 2015

Academic Editor: Ye-Sho Chen

With the rapid development of the internet and information technology, the increasingly diversified portable mobile terminals, online shops, and social media have facilitated information exchange, social communication, and financial payment for people more and more than ever before. In the meantime, information security and privacy protection have been meeting with new severe challenges. Although we have taken a variety of information security measures in both management and technology, the actual effectiveness depends firstly on people's awareness of information security and the cognition of potential risks. In order to explore the new technology for the objective assessment of people's awareness and cognition on information security, this paper takes the online financial payment as example and conducts an experimental study based on the analysis of electrophysiological signals. Results indicate that left hemisphere and beta rhythms of electroencephalogram (EEG) signal are sensitive to the cognitive degree of risks in the awareness of information security, which may be probably considered as the sign to assess people's cognition of potential risks in online financial payment.

1. Introduction

Today's society is an information society. More and more people use information technologies in daily life and work. They are facilitated by increasingly diversified portable mobile terminals, online shopping, and social media in information exchange, social communication, and e-business. However, when people are enjoying the convenience from information technology, it is also facing the new severe challenges of information security, such as internet intrusion, sensitive information leak, and online payment fraud.

It is well known that information security is a complicated and systematic problem associated with technology, management, economy, and behavioral culture. Up to now, there are a lot of researches on this issue. Cavusoglu et al. studied risks related to information security; they pointed out that risks may have dire consequences, including corporate liability, monetary damage, and loss of credibility [1]. Ensuring information security has become one of the top managerial priorities in many organizations [2–4]. Kuner et al. took the PRISM project as an example which showed that both the offline and online activities had been reported to be related with extensive privacy; they argued that both privacy and security should be protected with individuals' confidence in the rule of law [5]. Numerous studies have shown that the biggest hidden danger of enterprise information security is the internal staff, rather than software vulnerabilities, and employees are often the weakest link in information security [6, 7].

In fact, many information security incidents are not all caused by technology, which happened often due to management oversights or people's weak awareness of information

security. For example, behavior of weak password, neglecting the operating system patch, and free use of unsafe mobile devices are related to the lack of recognition of the potential risks on information security. Since the awareness of information security depends on brain cognition of potential risk, it is very important to study brain cognition. A lot of scholars have made great achievements in cognitive research based on cognitive neuromechanism. Qin and Han assessed the neurocognitive processes involved in environmental risk identification by using event-related potential (ERP) and functional magnetic resonance imaging (fMRI); their findings show that an early detection in the ventral anterior cingulate cortex and a late retrieval of emotional experiences in posterior cingulate cortex can help identify dreadful environmental risks [8]. Wang et al. designed and evaluated the vocal emotion of humanoid robots based on brain mechanism; they found that stimulation from audio is related to some brain regional [9]. Dai studied the mechanism of public cognitive emotions when emergencies burst; he pointed out that it needs to consider the public psychology and cognitive ability and that it is easy to accept the way when the city emergency incident bursts out [10]. In addition, some scholars have done the research of brain cognition on investment behavior, framing effect, and microblog information spreading [11–13].

In our study, in order to explore the new technology for the objective assessment of people's awareness and cognition on information security, this paper takes the online financial payment as example and conducts an experimental study based on the analysis of electrophysiological signals.

This paper is organized as follows. In Section 2, the theory and method of cognitive model and EEG are presented. Then, trial is introduced in Section 3. Analysis and results are shown in Section 4. Finally, we provide a summary and discussion about our work in Section 5.

2. Theory and Methodology

Awareness is the human mind to reflect the objective material world, and it is the comprehension of feeling, thinking, and other psychological processes. In other words, awareness is a response to a stimulus of human brain. In order to study the information security awareness, cognitive psychology and EEG were used as the research theory and methods.

2.1. Cognitive Mechanism of Information Security

2.1.1. Cognitive Psychology. Cognition refers to all processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used [14]. Cognitive psychology usually takes human cognitive process as its major subject. It studies the cognitive activities from the viewpoint of information processing, including how humans learn, percept, imagine, memorize, and think of problems. So cognitive psychology is also called information processing psychology. Gagne is a famous scholar in the information processing theory, well-known for his outstanding contribution to information processing model of learning theory. In Gagne's theory, the learning processing was divided into eight stages, and each stage requires different information processing. Firstly, environmental stimuli affected learners; then these stimuli were encoded and were stored as image in the sensor register. These memory images can only store hundredths of a second. Then information entered short-term memory and was encoded again. It can maintain 2.5~3 seconds in here. However, short-term memory is limited to about seven "chunks" of information for most people. Once it exceeds this number, new information will replace the original information. In order to keep the original information, you can repeat it continuously. In this way, information in short-term memory can keep for a long time, but not more than one minute. Finally, the information entered long-term memory and it was encoded again. The majority of people believe that the long-term memory can be stored for a long time. Once you need to use this information, you can retrieve it from long-term memory. In here, information can directly enter response generator, or it can go back to the short-term memory. Meanwhile, expectation and executive control also affected this learning model [15]. After Gagne proposed information processing model, Model Human Processor (MHP) was presented and was used in cognitive modeling. Due to the fact that MHP can calculate the processing time after performing a certain task, it is especially suitable for our study. The processing of MHP is shown in Figure 1 [16]. It can be seen that MHP includes three subsystems, and each subsystem has its own processors and memories.

2.1.2. Cognitive Framework for Information Security Awareness. We know information cognition can be viewed as a process of information processing from the previous section. Previous research shows that visual stimuli can produce perceptual awareness [17–19]. Then, visual stimulation of information security was used in our study. And cognitive framework for information security awareness is shown in Figure 2.

From Figure 2, it can be seen that brain cognitive mechanism is closely related to selective attention. For example, when a person feels stimulation from field of information security, such that someone is surfing the internet with the public WiFi or somebody's computer does not install firewall, in the above scene, his brain starts to extract object features of the scene, and the selective attention mechanism begins running, which includes feeling, imagination, perception, and memory. Meanwhile, awareness is also accompanied by brain cognition mechanism which starts running.

2.2. EEG Signals Analysis

2.2.1. EEG Waves. The living human brain will continue to discharge, known as electroencephalogram (EEG) [20]. Brain and changes of electricity are the real time performance of brain activity. Generally, the level of volatility reflects brain excitability, and latency reflects the mental activities and processing speed and time evaluation. Human's brain waves frequency range is 0.1~100 Hz, and the frequency and amplitude of four basic brain waves are shown in Table 1 [21].

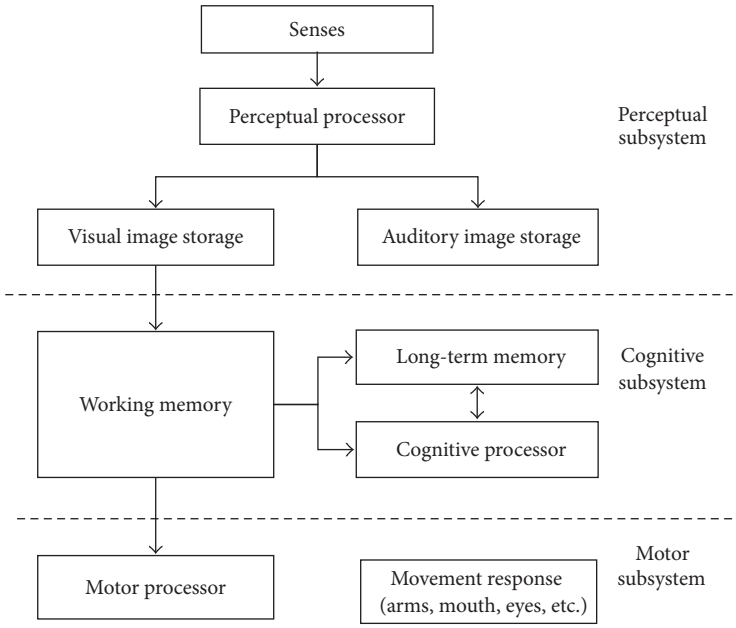


Figure 1: Model Human Processor.

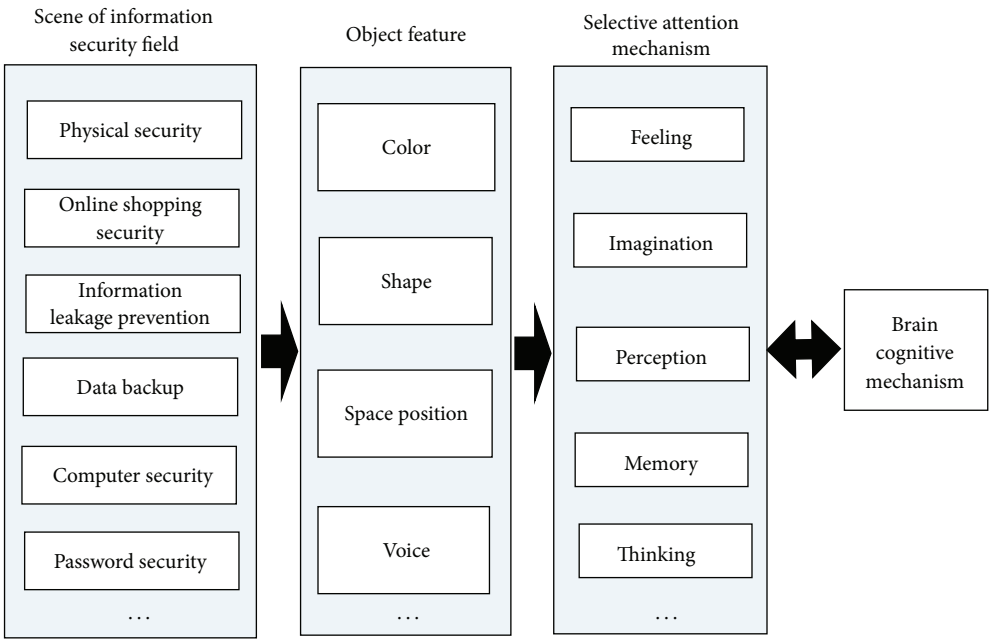


Figure 2: Cognitive framework for information security awareness.

EEG is closely related to human consciousness, and amplitude of EEG rhythm will increase or decrease when the brain activity increases. Previous research has suggested that α rhythms will appear in a relaxed state, β rhythms will appear in excited state, θ rhythms will appear in drowsy state, and δ rhythms usually appear in deep state [21].

2.2.2. EEG Signal Process. EEG signal process mainly includes data cleaning, denoising signal, feature extraction, and classification process. Among them, denoising signal and feature extraction algorithms include power spectrum density estimation, wavelet transform (WT), public space model, multidimensional statistical analysis, and model descriptor. Classification methods include Fisher's linear discriminant, Bayesian method, back-propagation neural network [22], and support vector machine. In our study, WT was used.

WT is a multifunctional multiscale analysis and filter based on combination of time-frequency analysis tool. It has the characteristic of multiresolution and can observe different detail by choosing different basic wavelet, which makes

Table 1: The frequency and amplitude of basic band.

Frequency band	Frequency (Hz)	Amplitude (μV)
δ	0.5~3.5	20~200
θ	4~7	100~150
α	8~13	20~200
β	14~30	5~20

the wavelet transform have the ability to characterize the local features of the signal in the time domain and frequency domain at the same time. Wavelet transform includes Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). CWT can be defined as follows:

$$\begin{aligned}\psi_{a,b}(t) &= \frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty} f(t)\,\psi^*\left(\frac{t-b}{a}\right)d(t)\\ &= \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right),\end{aligned} \tag{1}$$

where $\psi(t) \in L^2(R)$, $a, b \in R$, $a \neq 0$, then $\psi(t)$ is called basic wavelet, and a means expansion factor and b means translation factor.

For the discrete case, DWT can be defined as follows:

$$\psi_f(m,n) = \int_{-\infty}^{+\infty} f(t)\,\psi_{m,n}^*(t)\,d(t), \tag{2}$$

where $\psi(t) \in L^2(R)$, $m, n \in Z$.

In order to get high quality EEG signals for analysis, we adopt Discrete Wavelet Transform method and Mallat algorithm to renoise initial EEG signals. Mallat decomposition algorithm is shown as follows:

$$\begin{aligned}f_0 &= f_1 + d_1 = f_2 + d_2 + d_1 = \cdots\\ &= f_N + d_N + d_{N-1} + \cdots + d_2 + d_1,\end{aligned} \tag{3}$$

where f_0 means initial signal, f_N is the result of the approximation signal after decomposition (low frequency components), and d_N is the result of the error signal after decomposition (high frequency components).

3. Experiment

The formation process of EEG in our trial is shown in Figure 3.

From Figure 3, we can see experimenter watching specific scene and EEG device collecting EEG signals from experimenter. Once collecting signals finishes, the signal process begins to work, and EEG would be shown finally. EEG signal acquisition settings are as follows:

(i) sampling frequency: 128 Hz;

(ii) amplitude-frequency characteristic: 0.53 Hz–60 Hz;

(iii) electrode placement criteria: electrodes were placed according to the international 10–20 system [23], which is shown in Figure 4;

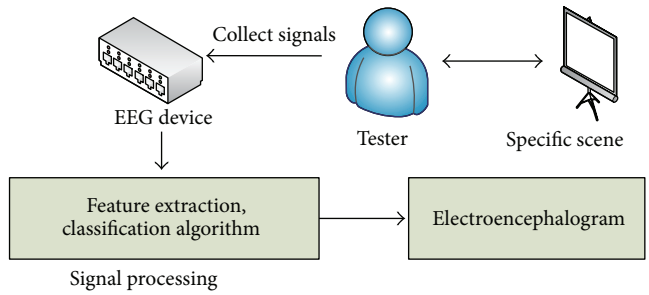


Figure 3: The formation process of EEG.

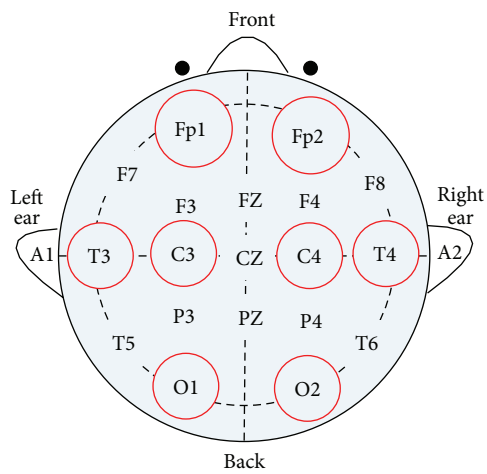


Figure 4: EEG electrodes location of international 10–20 system.

(iv) electrode channel selection: we choose eight positions of electrode as follows: frontal region (Fp1, Fp2), parietal region (T3, T4, C3, C4), and occipital region (O1, O2) [24];

(v) using a single-stage lead.

3.1. Experimental Overview. Our research involved human subjects, and we recruited 12 healthy adults to participate in our trail; among them, four had received information security awareness training, and eight had not received training. All of their education degrees are bachelor degree or above, with no history of mental illness. They were right-handed with an average age of 27.1 years and they represented 5.69 of the variance. The testing process was told to them before the experiment, and the agreement was signed.

3.2. Experimental Design. In order to research the human awareness of information security, nine experiment scenes were designed in our trial. Testers would make a choice when they take note of information security related pictures or hear fraud words. Tester may encounter fraud information in instant messaging, or access fishing website, or receive fraud text message in his mobile phone, or receive fraud message while using the online payment, and so forth. All of the above scenarios can be used as experimental scene, and sample pictures of trail are shown in Figure 5.

Table 2: Sample of experimental records.

Tester	Event	Minutes	Seconds	Channel 1	Channel 2	Channel 3	Channel 4	Channel 5	Channel 6	Channel 7	Channel 8	Baseline
Number 01	1	3.511	210.633	33.925	10.482	36.855	−1.352	5.072	7.101	47.675	−135.587	−738.123
Number 01	1	3.511	210.641	25.359	5.410	41.251	11.158	12.172	17.921	74.387	−129.163	−738.123
Number 01	1	3.511	210.648	35.954	31.784	50.719	32.798	27.388	29.417	64.920	−93.322	−738.123
Number 01	1	3.511	210.656	6.875	41.251	32.798	43.618	26.712	36.179	81.826	−66.610	−738.123
Number 01	1	3.511	210.664	53.085	70.330	78.445	78.445	66.948	59.848	120.710	−34.827	−738.123
Number 01	1	3.511	210.672	63.342	68.639	74.049	66.948	65.934	61.877	113.271	−3.381	−738.123
Number 01	1	3.511	210.680	84.869	104.818	130.178	125.782	113.271	113.948	159.594	29.079	−738.123
Number 01	1	3.511	210.688	168.385	151.479	197.802	160.609	167.709	138.292	180.220	51.395	−738.123
Number 01	1	3.512	210.695	194.759	149.789	201.860	160.270	192.392	152.494	207.608	104.480	−738.123
Number 01	1	3.512	210.703	257.875	200.507	253.593	218.090	234.658	196.788	234.996	143.026	−738.123
Number 01	1	3.512	210.711	238.377	187.658	207.270	184.277	180.220	170.752	183.263	123.415	−738.123
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Number 12	9	23.071	1384.227	−28.966	−75.063	−11.158	−52.747	−1.014	−29.417	26.374	−7.439	−738.123

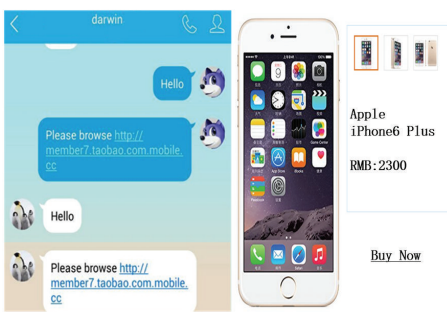


Figure 5: Sample pictures of trail.

The above website has two suspicions. Left graph uses this link http://www.shbillow.cn/index.mobile.cc.htm, to which the suffix "mobile.cc" was added, and it may be a fishing site. Right graph attracts customers with low price, and the price is too low for the normal price. Tester's information safety awareness may be arousing when he/she notices these scence.

Our experimental procedures are as follows:

(i) Tester wears electrode cap and puts electrode well. 8 channel recordings are used for electrode cap; 10–20 electrodes are put on standard position according to the International Institute of EEG. Tester seated in the most comfortable, as far as possible, position to ensure the comfort of the viewing test.

(ii) Tester connects to the computer and opens the EEG signal processing software and then checks whether the software works correctly. If there is no problem, then the experiment begins.

(iii) Tester closes his/her eyes, sits and rests, and calms him/herself, when the brain waves are smooth and then begins to record his/her brain waves signal.

(iv) Picture will be shown on the screen. Tester watches picture and listens to the sound with distance of 1 meter, and he or she responds to the prompt. After testing the current scene, another stimulus will appear at random intervals between 1000 ms and 2000 ms. During the interval, the screen background color is black, and the middle of the screen shows the symbol "+" with white color.

3.3. Experimental Records. In our experiment, records include tester number, event number, duration, eight-electrode value, and baseline electrode value. Sample of experimental records is shown in Table 2.

4. EEG Signal Process and Analysis

4.1. EEG Signal Process. Due to the fact that initial EEG signals include a lot of noise, they need to be processed. The process usually includes denoising and characteristics analysis [25]. In order to remove noise signals from the collected EEG signals, we adopted two processes. Firstly, baseline electrode voltage was replaced by the average electrode voltage, and it was recalculated for every electrode voltage. Some noise will be removed after the above steps. Contrast of initial EEG signal and denoising EEG signal is shown in Figure 6.

Secondly, wavelet transformation method was used for these EEG signals. Because the EEG signal below 30 Hz is worth studying, then we use wavelet filtering to filter above 30 Hz EEG signals. We select the db5 as wavelet packet and decompose EEG signals into four layers. In the process of wavelet decomposition, the best wavelet decomposition tree is shown in Figure 7.

According to sampling frequency which is $f_s = 128$ Hz, we can calculate frequency width of four layers of each subband as 4 Hz ($\Delta f = (1/2^4) \times (f_s/2) = 4$), and the four layers include 16-subband wavelet packet $S(4, i)$, where $i = 1, 2, 3, \ldots, 16$. Therefore, four kinds of rhythm waves (δ, θ, α,

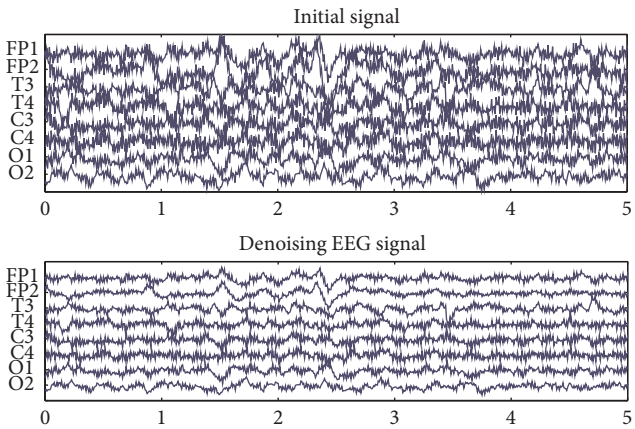


Figure 6: Contrast of initial EEG signal and denoising EEG signal.

and β) can be extracted by reconstruction. For example, δ, θ, α, and β rhythms can be extracted as follows:

$$\begin{aligned} \delta\,[0.5 \sim 3.5\,\text{Hz}]: &\ \{S(4,0)\}, \\ \theta\,[4 \sim 7\,\text{Hz}]: &\ \{S(4,1)\}, \\ \alpha\,[8 \sim 13\,\text{Hz}]: &\ \{S(4,2)\}, \\ \beta\,[14 \sim 30\,\text{Hz}]: &\ \{S(4,3), S(4,4), \ldots, S(4,7)\}. \end{aligned} \tag{4}$$

4.2. Characteristic Analysis. In order to analyze the correlation of EEG signal and safety awareness, four types of rhythm signal are extracted from wavelet transformation, which are shown in Figure 8.

In the selection of characteristic parameter, the rhythm energy and energy ratio of four types of rhythm were calculated, and both of them were used for characteristic analysis. Sample of rhythm energy and energy ratio of two test tasks (online payment and online chat) is shown in Table 3.

It can be seen from Table 3 that the alpha rhythm energy and energy ratio are relatively low in two test tasks, which is consistent with previous studies. Previous biomedical research results show that the alpha rhythm became inhibited or disappeared when people are feeling the external stimuli [26]. Our experiment proved that the beta rhythm is consistent with the distribution characteristics of the scalp. It also suggests that beta rhythms are easy to appear when the brain is thinking or exciting. Since information security awareness related to people's focus of attention who remain alert to stimuli for a prolonged period of time, and the beta rhythm is more active, then it can be used to research different brain cognition.

In addition, in order to do a comparative analysis, we choose energy ratio of beta rhythm of two test tasks as comparison; the results are shown in Figure 9. From Figure 9, we can clearly see that energy ratio of beta rhythm of left hemisphere (FP1, T3, C3, O1) is higher than that of the right hemisphere, which shows that the left hemisphere is more involved in reading related tasks.

From Figure 9, we also found that energy ratio of beta rhythm of test task 1 (online payment) is higher than that of test task 2 (online chat). The reasonable explanation is that the tester needs more attention and feels nervous in the online payment than those of the online chat. That is to say, visual stimuli are more likely to arouse the awareness of information security than aural stimuli. Furthermore, the energy ratio of parietal region (T3, C3) is higher than other regions, which showed that the parietal region was involved in awareness of information security related tasks.

In our experimental results, another finding showed that the EEG signals of tester who has been trained on information security were more active than those of untrained tester.

Table 3: Rhythm energy and energy ratio.

		Rhythm energy		Energy ratio	
		Task 1	Task 2	Task 1	Task 2
FP1	δ	0.3266	0.4593	0.4801	0.5872
	θ	0.1386	0.1528	0.2037	0.1923
	α	0.0555	0.0396	0.0817	0.0517
	β	0.1595	0.1305	0.2345	0.1688
FP2	δ	0.3436	0.4734	0.5021	0.5903
	θ	0.1525	0.1609	0.2229	0.2007
	α	0.0444	0.0364	0.0649	0.0454
	β	0.1437	0.1312	0.2101	0.1636
T3	δ	0.3304	0.4834	0.4740	0.5988
	θ	0.1388	0.1490	0.1991	0.1845
	α	0.0559	0.0388	0.0802	0.0472
	β	0.1720	0.1361	0.2467	0.1695
T4	δ	0.3278	0.4956	0.4927	0.6129
	θ	0.1385	0.1523	0.2081	0.1884
	α	0.0490	0.0337	0.0737	0.0417
	β	0.1501	0.1270	0.2255	0.1570
C3	δ	0.3174	0.4906	0.4572	0.6088
	θ	0.1402	0.1465	0.2020	0.1818
	α	0.0569	0.0390	0.0820	0.0384
	β	0.1797	0.1297	0.2588	0.1710
C4	δ	0.3370	0.4873	0.5072	0.6102
	θ	0.1307	0.1536	0.1968	0.1923
	α	0.0476	0.0330	0.0717	0.0414
	β	0.1490	0.1247	0.2243	0.1561
O1	δ	0.3126	0.4733	0.4833	0.6001
	θ	0.1350	0.1472	0.2001	0.1866
	α	0.0573	0.0378	0.0849	0.0479
	β	0.1698	0.1304	0.2317	0.1654
O2	δ	0.3720	0.5257	0.5014	0.6224
	θ	0.1482	0.1534	0.1998	0.1816
	α	0.0592	0.0377	0.0799	0.0447
	β	0.1624	0.1278	0.2189	0.1513

5. Discussion

Promotion of people's awareness of information security is the foundation and the precondition of information security of organization. In order to explore the new technology for

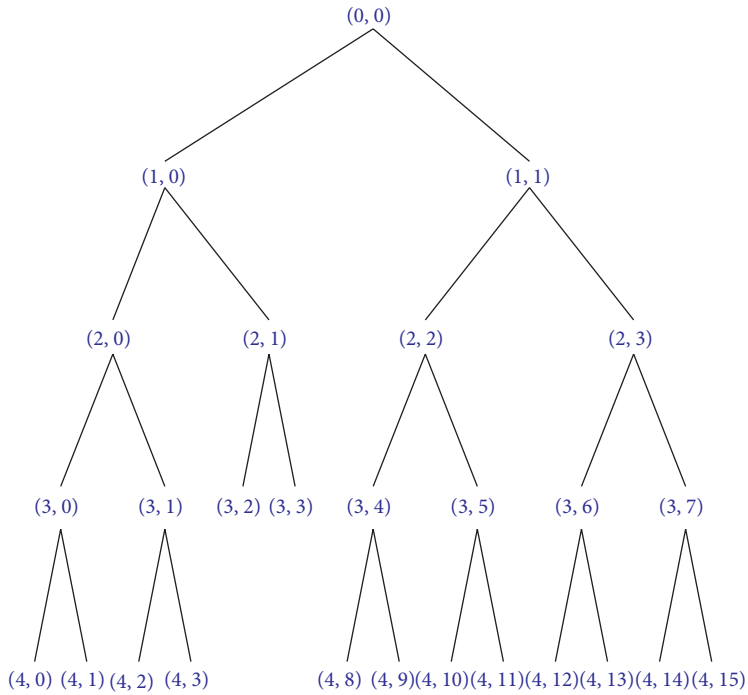


Figure 7: The best wavelet decomposition tree.

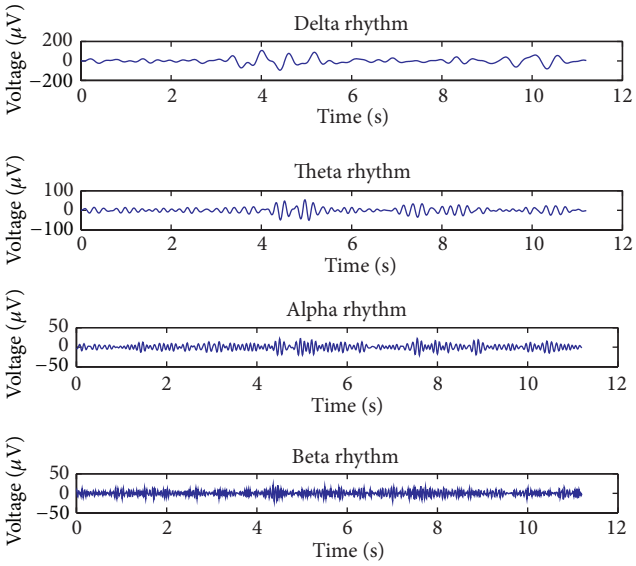


Figure 8: Four types of rhythm signals extracted from wavelet transformation.

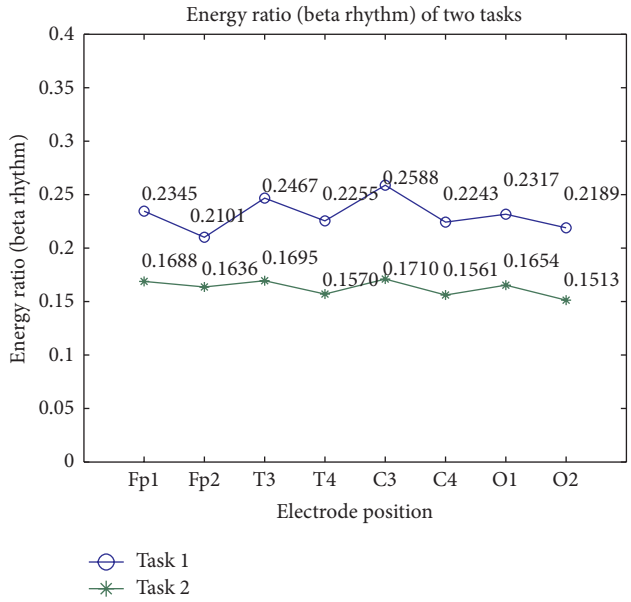


Figure 9: Energy ratio of beta rhythm of two test tasks.

the objective assessment of people's awareness of information security, this paper conducted cognitive study of information security awareness based on the analysis of EEG signals. We firstly discussed the theory and methodology of EEG signals on cognitive study and then presented a framework for the description of awareness and cognition of information security according to the brain mechanism. On this basis, an experiment was designed to test the reaction of EEG signals to the awareness of hidden problems in information security. This finding showed that the EEG signals could provide a good method for the objective assessment of people's awareness of information security.

In the future studies, we suggest that it can be combined with fMRI (functional magnetic resonance imaging) [27], PET (Positron Emission Tomography), and other measuring equipment to research cognition of individual information security.

Dislosure

Yonghui Dai and Xingyun Dai are the joint corresponding authors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors appreciate the anonymous reviewers for their helpful and constructive comments on the earlier draft. This work was supported partly by National Natural Science Foundation of China (no. 41174007 and no. 91324010), Shanghai Philosophy and Social Sciences Plan (no. 2014BGL022), Shanghai Science and Technology Innovation Project (no. 13511505200), and Humanity and Social Science Foundation of Ministry of Education of China (no. 13YJA630033). Many thanks are due to Chengfeng Lin and Jinzhao Wang for their assistance; both of them are undergraduate students of Fudan University, China.

References

[1] H. Cavusoglu, H. Cavusoglu, and S. Raghunathan, "Economics of IT security management: four improvements to current security practices," *Communications of the Association for Information Systems*, vol. 14, pp. 65–75, 2004.

[2] J. C. Brancheau, B. D. Janz, and J. C. Wetherbe, "Key issues in information systems management: 1994-95 SIM delphi results," *MIS Quarterly*, vol. 20, no. 2, pp. 225–242, 1996.

[3] D. F. Lohmeyer, J. McCrory, and S. Pogreb, "Managing information security," *Asia-Pacific Financial Markets*, vol. 2, pp. 12–16, 2002.

[4] S. Ransbotham and S. Mitra, "Choice and chance: a conceptual model of paths to information security compromise," *Information Systems Research*, vol. 20, no. 1, pp. 121–139, 2009.

[5] C. Kuner, F. H. Cate, C. Millard, and D. J. Svantesson, "PRISM and privacy: will this change everything?" *International Data Privacy Law*, vol. 3, no. 4, pp. 217–219, 2013.

[6] K. D. Mitnick and W. L. Simon, *The Art of Deception: Controlling the Human Element of Security*, John Wiley & Sons, New York, NY, USA, 2002.

[7] M. Warkentin and R. Willison, "Behavioral and policy issues in information systems security: the insider threat," *European Journal of Information Systems*, vol. 18, no. 2, pp. 101–105, 2009.

[8] J. G. Qin and S. H. Han, "Neurocognitive mechanisms underlying identification of environmental risks," *Neuropsychologia*, vol. 47, no. 2, pp. 397–405, 2009.

[9] Y. Wang, X. Hu, W. Dai, J. Zhou, and T. Kuo, "Vocal emotion of humanoid robots: a study from brain mechanism," *The Scientific World Journal*, vol. 2014, Article ID 216341, 7 pages, 2014.

[10] W. H. Dai, "Public cognition morale mechanism of unexpected incidents in cities and the responding strategies," *Shanghai Urban Management*, vol. 23, no. 1, pp. 34–37, 2014.

[11] S. K. Sahi, "Neurofinance and investment behaviour," *Studies in Economics and Finance*, vol. 29, no. 4, pp. 246–267, 2012.

[12] Q. G. Ma, Y. D. Feng, Q. Xu, J. Bian, and H. X. Tang, "Brain potentials associated with the outcome processing in framing effects," *Neuroscience Letters*, vol. 528, no. 2, pp. 110–113, 2012.

[13] H. Z. Hu, D. Wang, W. H. Dai, and L. H. Huang, "Psychology and behavior mechanism of micro-blog information spreading," *African Journal of Business Management*, vol. 6, no. 35, pp. 9797–9807, 2012.

[14] U. Neisser, *Cognitive Psychology*, Appleton-Century-Crofts, New York, NY, USA, 1967.

[15] R. M. Gagne, *Essentials of Learning for Instruction*, The Dryden Press, Hinsdale, Ill, USA, 1974.

[16] S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, 1986.

[17] N. K. Logothetis and J. D. Schall, "Neuronal correlates of subjective visual perception," *Science*, vol. 245, no. 4919, pp. 761–763, 1989.

[18] D. L. Sheinberg and N. K. Logothetis, "The role of temporal cortical areas in perceptual organization," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 7, pp. 3408–3413, 1997.

[19] D. A. Leopold and N. K. Logothetis, "Multistable phenomena: changing views in perception," *Trends in Cognitive Sciences*, vol. 3, no. 7, pp. 254–264, 1999.

[20] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*, Oxford University Press, Oxford, UK, 1995.

[21] Y. J. Luo, Y. Jiang, and K. Cheng, *Congnitive Neuroscience*, Peking University Press, Beijing, China, 2006.

[22] Y. Dai, D. Han, and W. Dai, "Modeling and computing of stock index forecasting based on neural network and Markov chain," *The Scientific World Journal*, vol. 2014, Article ID 124523, 9 pages, 2014.

[23] R. Thiruchselvam, J. Blechert, G. Sheppes, A. Rydstrom, and J. J. Gross, "The temporal dynamics of emotion regulation: an EEG study of distraction and reappraisal," *Biological Psychology*, vol. 87, no. 1, pp. 84–92, 2011.

[24] Q. G. Ma, J. Jin, and L. Wang, "The neural process of hazard perception and evaluation for warning signal words: evidence from event-related potentials," *Neuroscience Letters*, vol. 483, no. 3, pp. 206–210, 2010.

[25] K. E. Mathewson, C. Prudhomme, M. Fabiani, D. M. Beck, A. Lleras, and G. Gratton, "Making waves in the stream of consciousness: entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation," *Journal of Cognitive Neuroscience*, vol. 24, no. 12, pp. 2321–2333, 2012.

[26] G. Pfurtscheller and A. Aranibar, "Event-related cortical desynchronization detected by power measurements of scalp EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 42, no. 6, pp. 817–826, 1977.

[27] A. M. Goldfine, J. D. Victor, M. M. Conte, J. C. Bardin, and N. D. Schiff, "Determination of awareness in patients with severe brain injury using EEG power spectral analysis," *Clinical Neurophysiology*, vol. 122, no. 11, pp. 2157–2168, 2011.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 812650, 7 pages
http://dx.doi.org/10.1155/2015/812650

Research Article

CyberPsychological Computation on Social Community of Ubiquitous Learning

Xuan Zhou,[1] Genghui Dai,[2] Shuang Huang,[3] Xuemin Sun,[4] Feng Hu,[5] Hongzhi Hu,[6] and Mirjana Ivanović[7]

[1] *School of Humanities and Social Science, Sichuan Conservatory of Music, Chengdu 610021, China*
[2] *School of Marine Sciences, Sun Yat-Sen University, Guangzhou 510275, China*
[3] *Overseas Training Center, Shanghai International Studies University, Shanghai 200083, China*
[4] *Department of General Surgery, Tongji Hospital, Tongji University, Shanghai 200065, China*
[5] *Department of Respiratory Medicine, Shanghai Tongren Hospital, Shanghai Jiao Tong University, Shanghai 200336, China*
[6] *School of Management, Fudan University, Shanghai 200433, China*
[7] *Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, 21000 Novi Sad, Serbia*

Correspondence should be addressed to Genghui Dai; daigengh@mail2.sysu.edu.cn and Xuemin Sun; 1375734648@qq.com

Received 14 February 2015; Accepted 28 April 2015

Academic Editor: Ye-Sho Chen

Under the modern network environment, ubiquitous learning has been a popular way for people to study knowledge, exchange ideas, and share skills in the cyberspace. Existing research findings indicate that the learners' initiative and community cohesion play vital roles in the social communities of ubiquitous learning, and therefore how to stimulate the learners' interest and participation willingness so as to improve their enjoyable experiences in the learning process should be the primary consideration on this issue. This paper aims to explore an effective method to monitor the learners' psychological reactions based on their behavioral features in cyberspace and therefore provide useful references for adjusting the strategies in the learning process. In doing so, this paper firstly analyzes the psychological assessment of the learners' situations as well as their typical behavioral patterns and then discusses the relationship between the learners' psychological reactions and their observable features in cyberspace. Finally, this paper puts forward a CyberPsychological computation method to estimate the learners' psychological states online. Considering the diversity of learners' habitual behaviors in the reactions to their psychological changes, a BP-GA neural network is proposed for the computation based on their personalized behavioral patterns.

1. Introduction

In today's society, ubiquitous learning (U-learning) has been bringing about magical changes to traditional education [1–3]. Under the new learning patterns such as Webquest, ThinkQuest, Microlearning, and MOOCs (Massive Open Online Courses) [4–7], the learners can access online resources at any time, in any place, and through any multimedia terminal, and they can join an open community to study knowledge, exchange ideas, and share skills in cyberspace [4, 5].

In that circumstance, the participants of ubiquitous learning have formed an ecological social community which displays many new features and is sustained by the common motivation and goal, favorable atmosphere, and emotional communication, as well as enjoyable experiences [8, 9]. Many researchers have found that the effectiveness and sustainability of a social community of ubiquitous learning depend largely on the learners' initiative and the community cohesion, and therefore the most important question to be considered is how to stimulate the learners' interest and participation willingness, as well as improving their enjoyable experiences [8–10].

In the past century, learning theory has made great progress. Constructivism, situated cognition, and informal learning theories offer the guidance for ubiquitous learning by focusing on the individual's self-construction of knowledge and aim to provide new online intelligent, virtual

interactive, and seamless learning patterns for learners on the basis of such situational characteristics as location, time, and environment [3, 8]. Research findings in neuroscience lay the basis of neuroeducation, which provides the neural theory and the brain mechanism for the study of ubiquitous learning [11]. In particular, recent work in this field has been well supported by the advanced technologies such as fMRI (functional magnetic resonance imaging), ERPs (event-related potentials), and DTI (diffusion tensor imaging). In particular, the blood oxygenation level dependent functional magnetic resonance imaging (Bold-fMRI) has been successfully applied to the neural activity studies of cognition and emotions in education [10, 12–16].

However, the environment and patterns of ubiquitous learning in a social community differ from those in classroom completely. Different social media and interactive activities between the participants have all major impacts on the cognition, emotion, and attitude in the learning process [8, 16–18]. In recent years, vocal social media such as Wechat, QQ (China), ICQ, WhatsApp (USA), and Line (Japan) and various tools of instant voice messaging have been the popular means in the communication of ubiquitous learning. While facilitating conveying semantic information, vocal social media can also transmit abundant emotional information [19]. This variation has resulted in significant influence on not only improving the participants' experiences and senses of belonging to particular social groups and therefore enhancing their continuance intentions to these groups [20], but also strengthening the interpersonal relationships between the members within these groups as well as the community's cohesion and cognitive consistence in this community [19, 21]. A new study shows that the interactive activities in a social community can cause the propagation effects on the five interactional layers, information, emotion, attitude, behavior, and culture, and easily lead to some groups of "the small world" with close relationships in the community [16, 19].

Therefore, considering the new features and special environment in cyberspace, a further study of ubiquitous learning theory is required to provide comprehensive and systemic guidance for improving the organizational mode, learning behaviors, and teaching skills, which involves the interdisciplinary areas of pedagogy, psychology, sociology, organizational behavior, neuroscience, and information technology. Scholars have paid attention to the above issue in the 1990s. CyberPsychology, coined by Dr. John Suler in his hypertext book "The Psychology of Cyberspace" with the first version appearing in January of 1996, launched the original conceptual framework for understanding how people react to and behave within cyberspace [22]. The progress in CyberPsychology and social neuroscience [23] has provided the fresh theory and new method for the development of computational intelligence in cyberspace.

By affective computing [24] on the learners' reactions and their behavioral data, the ubiquitous learning system may have emotional intelligence [25] and make the learning more "smart" which was defined by Professor Dai in 2012 as that the machine can perceive and respond to human emotional needs and provide the full humanized services combining both rational and emotional intelligence [26]. This concept has been applied in wide areas such as smart education, smart city, smart healthcare, and smart service [27, 28].

This paper aims to explore an effective method to monitor and analyze the learners' psychological reactions based on their observable behaviors in cyberspace and therefore help to conduct the smart education in ubiquitous learning. This paper is organized as follows. Section 1 is an introduction to the research background and motivation. Section 2 discusses the research model. The CyberPsychological computation method is put forward in Section 3 with its experiment and result in Section 4. Section 5 is the conclusion and discussion.

2. Psychological Assessment and Behavioral Patterns

2.1. Psychological Assessment of Learners' Situations. Social community of ubiquitous learning is affected by a lot of factors associated with social, psychological, organizational, managerial, and technological aspects. Since the 1990s, a lot of scholars such as Webster and Hackley [29], Hill and Hannafin [30], and Hannafin et al. [31] have studied the influence factors and cognitive characteristics related to ubiquitous learning environment. In the late research, Chinese scholar concluded the influence factors as a LICE (Learner, Instructor, Curriculum, and Environment) model [32].

Recent experimental observations showed that a desirable learning atmosphere, good visual effects, pleasant voices, suitable topics and materials, and positive evaluation feedbacks were the most important factors that aroused the learners' interest in ubiquitous learning and contributed to the pleasant emotional experiences [10]. In order to adjust the strategies dynamically, the learners' psychological reactions such as attention, interest, emotion, and satisfaction are usually applied as the monitoring variables in the learning process [17]. Although the above four variables are actually not independent, the learners can get clear understanding of them and make accurate subjective assessment on each variable. Therefore, the dynamically psychological assessment of learners' situations can be expressed as formula

$$\mathrm{PA}(t_i) = \{A(t_i), I(t_i), E(t_i), S(t_i)\}, \quad i = 1, 2, 3, \ldots, N, \tag{1}$$

where $\mathrm{PA}(t_i)$ represents the vector of psychological assessment at different times t_i ($i = 1, 2, 3, \ldots, N$) given by the learners and $A(t_i)$, $I(t_i)$, $E(t_i)$, and $S(t_i)$ represent the scores of attention, interest, emotion, and satisfaction, respectively, at time t_i.

The calibration for scoring records is shown as in Table 1. Here, attention and interest are scaled from 0 to 10 with the intensity varying from low to strong, but emotion and satisfaction are scaled from −5 to +5 with the polarity and strength varying from the most negative to the extremely positive.

Table 1: Calibration for scoring records in psychological assessment.

Variables	Scores										
Attention	0	1	2	3	4	5	6	7	8	9	10
Interest	0	1	2	3	4	5	6	7	8	9	10
Emotion	−5	−4	−3	−2	−1	0	1	2	3	4	5
Satisfaction	−5	−4	−3	−2	−1	0	1	2	3	4	5

2.2. Behavioral Patterns in Learning Process. In ubiquitous learning, the learner's behaviors can be divided into three categories from simple to complex: operational behaviors, information exchanging behaviors, and problem-solving behaviors [33]. However, psychological reactions of the learners are mostly reflected in their habitual behaviors which are composed of a series of basic actions in cyberspace [8].

The pattern of each learner's habitual behaviors can be expressed as formula

$$\mathrm{BP}(p_a) = \{A(j), P(j)\}, \quad j = 1, 2, 3, \ldots, M, \tag{2}$$

where $\mathrm{BP}(p_a)$ represents the behavioral pattern related to psychological reaction p_a, $A(j)$ $(j = 1, 2, 3, \ldots, M)$ represents a series of basic actions in a certain period T under the psychological reaction p_a, and $P(j)$ $(j = 1, 2, 3, \ldots, M)$ is the parameters of $A(j)$.

The action's data are mainly acquired from the server's log files or by some online tracking tools. The switching frequency and retention time of webpages, the locations and movements of the mouse, and the keyboard operations are all the important parameters as the features to reflect the learners' psychological reactions. For example, if the learner has strong interest in something and is in a fairly good mood, he/she tends to stay on the interesting webpage longer, use his/her mouse and keyboard with higher frequency, answer the asked questions more quickly, and be more willing to make positive comments. On the contrary, if he/she is anxious and fretful, he/she will switch from one webpage to another frequently, move the mouse in a wide range quickly, and be more likely to give negative feedback [10]. Professor Dai proposed a CPP (CyberPsychological and Physical) computation method based on social neuromechanism and concluded 15 commonly basic actions of the user's habitual behaviors in cyberspace as in Table 2 [34].

2.3. Behavioral Features in Cyberspace. The psychological reactions of the learners are the results of a series of neural activities dominated by the brain mechanism [8], which will not only generate activated responses in their brains, but also result in the corresponding variation of physiological signals (e.g., EEG, ECG, EDR, respiration, and skin temperature) as well as external performances (e.g., speeches, facial expressions, gestures, and movements) and the possible subsequent behaviors [16, 23, 35, 36]. Therefore, the psychological computation in a real environment can be conducted by analyzing the expression patterns on the reactions of physiological signals, external performances, and subsequent behaviors based on the fusion features corresponding to the certain reactions in brain areas [10, 13, 26].

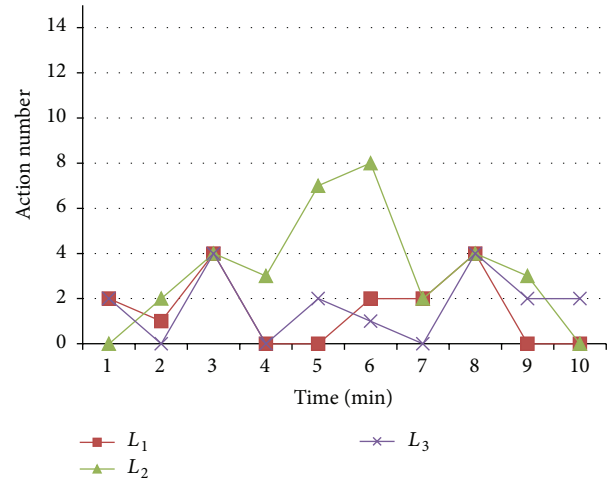


Figure 1: Actions of three learners to the same psychological reaction PA{9.3, 9.2, 2.5, 9.1} in a period of 10 minutes.

However, in the ubiquitous learning environment, the possible information we can obtain from the learners is their observable actions as well as the voice and video signals produced by their online activities under some circumstances. Considering the protection of the learners' privacy, CyberPsychological computation cannot be based on the content analysis and semantic detection of the above information. Fortunately, technologies of psychological computation from voice and video signals have developed quickly in recent years. For example, progress has been made on the voice signal based on its acoustic features parameters such as speech speed, voice intensity, pitch frequency, LPCC (Linear Prediction Cepstrum Coefficient), and MFCC (Mel Frequency Cepstrum Coefficient) [19, 37, 38].

In our research, we mainly consider the learners' behavioral features from their observable actions in cyberspace. The relationship between the learners' psychological reactions and their observable features is related to the habitual patterns of the learners' behaviors, which is dependent on a statistical study in the real cases. Figure 1 shows the actions of three learners (L_1, L_2, and L_3) to almost the same psychological reaction PA{9.3, 9.2, 2.5, 9.1} (within the relative errors of 10%) in a period of 10 minutes.

From Figure 1, we can find most of the learners' actions are Action number 2 and Action number 4, but there is diversity in the action's orders and frequencies for different learners, which indicates that the CyberPsychological computation should be based on the learners' personalized behavioral patterns.

3. CyberPsychological Computation Method and Technology

3.1. Computation Method. In order to monitor the learners' psychological reactions in the learning process, we put forward a CyberPsychological computation method as shown in Figure 2. The course information, learner's ID, and webpage information as well as the actions of mouse and keyboard are

Table 2: Basic actions and parameters of the learner's habitual behaviors in cyberspace.

Number	Actions	Parameters
0	No action	The object in screen center
1	Mouse: click	On a button or link, on another place
2	Mouse: scroll	Speed, the object in screen center when stopping scrolling
3	Mouse: move	Speed, radius
4	Mouse: open a new page	Null
5	Mouse: change a page	Null
6	Mouse: close a page	Null
7	Mouse: store a page	Null
8	Keyboard: input	Number of characters
9	Keyboard: delete	Number of characters
10	Mouse and keyboard: retrieve information	Number of keywords
11	Mouse and keyboard: post information on BBS	Number of characters
12	Mouse and keyboard: send information to other people	Number of characters, number of receivers
13	Mouse and keyboard: chat with other people	Number of characters, number of people chatted with
14	Streaming media: voice communication	Acoustic feature parameters
15	Streaming media: video communication	Visual feature parameters

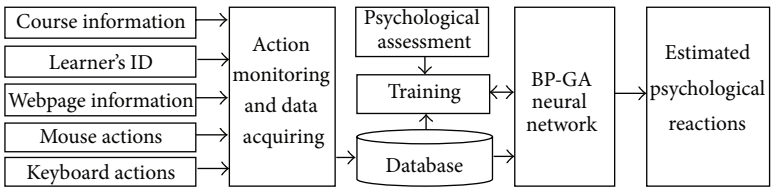


Figure 2: CyberPsychological computation method on social community of ubiquitous learning.

monitored and acquired by the computation system online. After preprocessing, those data are stored in a database for training or computation by a BP-GA neural network.

In the learning system, online activities such as answering questions, retrieving information, discussing problems, chatting with each other, and submitting assignments are all assigned in different functional areas on the webpage. So the layout structure and its related information of the webpage are extracted to assist in identifying the learners' action.

Every change of the elements on the webpage, such as moving, clicking, and scrolling of the mouse or the operation of the keyboard, will trigger the corresponding JavaScript function. In order to meet the requirements of real-time data collection, we adopted the PHP language to program the above JavaScript function and process the mouse and keyboard data acquisition, which can be realized by the intelligent multiagent technology [10].

As we discussed before in this paper, the learners' psychological reactions are exhibited in their diverse habitual behaviors and the CyberPsychological computation should be based on their personalized behavioral patterns. Therefore, we consider the learner's ID and his/her historical behavior patterns in our method. The computation on the learners' psychological reactions based on their behavioral features in the learning process can be regarded as a dynamic and nonlinear estimation problem by machine learning. So the subjective psychological assessment of learners' situations should be given by them in the sample training.

Nonlinear Regression, Kalman Filter, Artificial Neural Network (ANN), and SVR (Support Vector Regression) have been reported as successful technologies for solving the above estimating problem [10, 19]. However, the relationship between the learners' psychological reactions and their observable features in cyberspace is affected by many factors. BP neural network has the advantages of no specific requirements on the data distribution and no sensitivity to the influence of multicollinearity and outlier data and can be trained by increasing samples to reveal the implied relationship between input and output variables more comprehensively and obtain higher estimation precision, so it is appropriate to solve this problem. For the purpose of accelerating the convergence rate as well as improving the generalization ability of computation [39, 40], we adopt the technology of BP-GA neural network in our method.

3.2. BP-GA Neural Network. BP (back propagation) neural network is a nonlinear function to establish the uncertain and continuous relations between input and output variables based on trained samples by machine learning. It can continuously modify the network weights and thresholds through the error back propagation algorithm (BP algorithm) and reach the target of minimizing the mean square error.

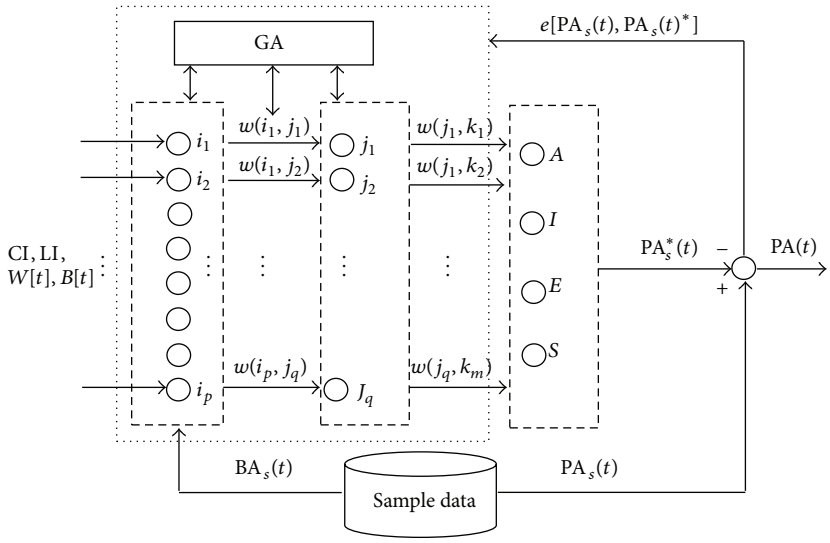


Figure 3: BP-GA neural network for CyberPsychological computation.

However, the traditional BP neural network easily causes the nonconvergence problem or falls into a local extremum. Those defects can be overcome by combining with a genetic algorithm (GA) [39, 40].

Figure 3 shows the proposed BP-GA neural network for CyberPsychological computation in our method. A3-layer BP neural network is mainly used to refine the generalized relationships between the input and output variables and produce the estimated psychological reactions of the learners based on their observable features in cyberspace. In this computation, GA contributes to the acceleration of convergence rate and prevents the network from going down to local minimum points by optimizing the weights and thresholds of the BP network.

The input layer has 12 nodes which are composed of the data of course information, learner's ID, mouse actions, and keyboard actions. The output layer has four nodes which correspond to the scores of the learner's psychological reactions in formula (1). The hidden layer is set with 22 nodes according to our best test result.

The BP-GA neural network operates firstly with the training by a group of sample data which are tested and recorded in the real cases. After it converges to a stable state, the BP-GA neural network can be applied to the computation in the learning process.

4. Experiment and Result

The experiment is based on a training course of life health and medical emergency rescue. In order to exclude the effects of the instructor, we edited this course into 20 lectures which are all taught by video tutorials and operated under a designed controlling program in the learning process. Every lecture runs for 50 minutes and is divided into 5 periods (from P1 to P5) with different activities arranged in each period. Figure 4 shows the design of learning process in each lecture of this course.

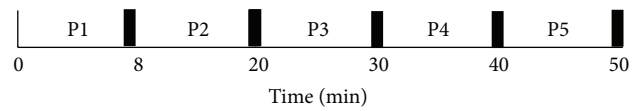


Figure 4: Learning process in each lecture of training course of life health and medical emergency rescue.

Table 3: Duration and activities in each period of a lecture.

Period	Duration	Activities
P1	8	Watching video tutorials
P2	12	Watching video tutorials
P3	10	Watching video tutorials
P4	10	Completing an online individual assignment
P5	10	Discussing with other learners

The learners are 16 social participants with more than 60 learning hours in a normal and stable environment of ubiquitous learning community. They are required to complete this course in 40 days. However, they can arrange the learning time freely in a ubiquitous learning environment. In this experiment, we assigned different activities in each period of a lecture and asked the learners to make a psychological assessment of their general situations in the final 1 minute of each period.

Table 3 shows the duration and activities in each period of a lecture.

In the end of our experiment, 11 participants successfully finished this course and provided completed subjective assessments of their psychological reactions in the learning process. We extracted randomly the records of 16 lectures as the training samples for BP-GA neural network and took the rest of the 4 lectures for the test of computation results. The above process was carried out 3 times. Table 4 is the test results of CyberPsychological computation.

Table 4: Test results of CyberPsychological computation.

Variable	Average score by assessment	Estimated result by computation	Relative error	Standard deviation
A	7.213	6.102	−15.4%	1.237
I	7.862	8.333	6.0%	0.6752
E	4.220	3.281	22.3%	1.401
S	8.581	7.114	17.1%	1.898

Table 4 shows that estimated relative errors of the learners' attention, interest, emotion, and satisfaction are −15.4%, 6.0%, 22.3%, and 17.1%, respectively. It indicates that our method can achieve an accuracy near 78% which can provide effective monitoring of the learners' psychological reactions in a real ubiquitous learning environment. From Table 4, we can find that the errors of attention, emotion, and satisfaction are larger than that of the interest. This is because the learner may have more actions to look for the further details of his/her interesting contents in this state. Also, the standard deviations of attention, emotion, and satisfaction are larger than the interest due to possibly less actions in the above states, which should cause more deviation errors in the computation because of lesser observation data. However, the computational accuracy can be improved markedly if an interactive activity is introduced. For example, the system may suggest a response to be made by the learner who has lesser actions by asking him/her a question or letting him/her to discuss with the other learners.

By analyzing learners' psychological reactions and their varying characteristics in the learning process, we can explore the statistical distribution and the changing rhythms of the learning community in different periods and scenarios and find the learners' ROI (Region of Interest). This has significant implications on the design of learning strategies such as the education scheme and teaching plans which are on-demand and more appealing and enjoyable to learners in order to provide state-of-the-art teaching models, skills, and technologies for smart education in a ubiquitous learning environment [10].

5. Conclusion and Discussion

With the development of modern information network and mobile communication technology, personalized learning in a ubiquitous learning environment has been made possible in web space accessible by a variety of multimedia terminals anytime and anywhere. It has brought about new changes to the theories, means, and patterns of traditional education. In the U-learning environment, learners' psychological reactions and learning experiences have significant impacts on stimulating their learning interest and improving the teaching efficiency.

This paper aims to explore an effective method to monitor the learners' psychological reactions based on their behavioral features in cyberspace and therefore provide useful references for adjusting the strategies in the learning process. As the comprehensive result of our research, a CyberPsychological computation method based on BP-GA neural network was proposed for estimating the learners' psychological states online. The experimental result shows that it can achieve accuracy near 78%.

Future researches should consider the dynamic psychological reactions of the learners through the studies of their physiological signals such as EEG, ECG, EDR, respiration, and skin temperature by a wearable device system and assist in obtaining a more precise psychological assessment of the learners' situations. Besides, the influence factors of different lecture's contents and the statistical distribution of the learning community as well as its varying characteristics in the learning process are all worthy of further explorations based on more cases and samples.

Disclosure

Xuan Zhou, Genghui Dai, and Shuang Huang are the joint first authors of this paper. Genghui Dai and Xuemin Sun are the joint corresponding authors.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by the 2nd Regular Meeting Project of Science and Technology Cooperation between China and Serbia (no. 2-9/2013), Shanghai Philosophy and Social Sciences Plan, China (no. 2014BGL022), Project of Scientific Research Fund of Sichuan Provincial Education Department (no. 15SB0219), and Project of Sichuan Conservatory of Music (CY2014173) and accomplished by an interdisciplinary team organized by Xuan Zhou, Genghui Dai, and Shuang Huang.

References

[1] X. F. Wei, Y. H. Zhang, and Z. H. Wei, "Transition from E-learning to U-learning: an interview with prof. Kinshuk," *Education Research*, vol. 2012, no. 2, pp. 4–8, 2012.

[2] Z. Q. Sa and D. D. Liu, "A research on college English teaching reform based on ubiquitous learning," *Overseas English*, vol. 2008, no. 8, pp. 32–33, 55, 2008.

[3] X. T. Yang, "Ubiquitous learning: theory, pattern, and resources," *Distance Education in China*, vol. 2011, no. 6, pp. 69–73, 2011.

[4] D. Gao, *Research and Practice in Network Teaching of Self-Study Pattern Based on Web Environment*, University of Electronic Science and Technology of China, Chengdu, China, 2007.

[5] Q. L. Zhang, *Online Learning and Teaching: New Concepts and New Patterns under the Background of the Internet*, Anhui Education Publishing House, Hefei, China, 2012.

[6] J. Wu, "Microlearning: a new pattern of students' self-study," *Chinese Journal of ICT in Education*, vol. 2012, no. 14, pp. 18–20, 2012.

[7] H. J. Chen, D. M. Han, Y. H. Dai, and L. N. Zhao, "Design of automatic extraction algorithm of knowledge points for

MOOCs," *Computational Intelligence and Neuroscience*. In press.

[8] W. H. Dai, S. Huang, X. Zhou, X. E. Yu, M. Ivanović, and D. R. Xu, "Emotional intelligence for ubiquitous learning: towards the smart foreign language education in the digital age," in *Proceedings of the 12th International Conference on IT Applications and Management (ITAM '14)*, pp. 87–93, 2014.

[9] M. Y. Xing, "The influences of affective factors on college English learning and teaching: theoretical and empirical research," *Foreign Languages and Their Teaching*, vol. 2003, no. 3, pp. 23–26, 2003.

[10] W. H. Dai, S. Huang, X. Zhou, X. E. Yu, M. Ivanović, and D. R. Xu, "Emotional intelligence system for ubiquitous smart foreign language education based on neural mechanism," *Journal of Information Technology Applications and Management*, vol. 21, no. 3, 2015.

[11] K. W. Fischer, A. M. Battro, and P. J. Lena, *The Educated Brain: Essays in Neuroeducation*, Cambridge University Press, Cambridge, UK, 2008.

[12] R. Horlings, *Emotion Recognition Using Brain Activity*, Delft University of Technology, Delft, The Netherlands, 2008.

[13] Y. H. Wang, X. H. Hu, W. H. Dai, J. Zhou, and T. Kuo, "Vocal emotion of humanoid robots: a study from brain mechanism," *The Scientific World Journal*, vol. 2014, Article ID 216341, 7 pages, 2014.

[14] Y. H. Wang, *An fMRI study of how a supra-modal decoding network exams interpretation [Ph.D. thesis]*, Tamkang University, Taipei, Taiwan, 2014.

[15] R. J. Davidson, D. C. Jackson, and N. H. Kalin, "Emotion, plasticity, context, and regulation: perspectives from affective neuroscience," *Psychological Bulletin*, vol. 126, no. 6, pp. 890–909, 2000.

[16] W. H. Dai, L. H. Huang, D. R. Xu et al., *Information Classification, Dissemination and Cognitive Model in Unconventional Emergency*, National Natural Science Foundation of China, 2015.

[17] X. Zhou, S. Huang, Y. S. Chen, and W. H. Dai, "Emotional intelligence system for smart education in ubiquitous learning," in *Proceedings of the International Conference on Information Technology and Applications (ITA '14)*, December 2014, in print.

[18] S. Huang, X. Zhou, W. H. Dai, and M. Ivanović, "Neural cognition and intelligent computing on the emotional symbols of cyber language," in *Proceedings of the International Conference on Intelligent Computing and Internet of Things (ICIT '15)*, 2015.

[19] W. H. Dai, D. M. Han, Y. H. Dai, and D. R. Xu, "Emotion recognition and affective computing on vocal social media," *Information & Management*, 2015.

[20] H. Lin, W. G. Fan, and P. Y. K. Chau, "Determinants of users' continuance of social networking sites: a self-regulation perspective," *Information & Management*, vol. 51, pp. 595–603, 2014.

[21] Y. Zhang, "Exploration of the new interpersonal communication pattern by Wechat in the mobile internet era," *News Research*, vol. 2014, no. 2, pp. 19–21, 2014.

[22] J. Suler, *The Psychology of Cyberspace*, 2014, http://users.rider.edu/~suler/psycyber/psycyber.html.

[23] J. T. Cacioppo and G. G. Berntson, "Social psychological contributions to the decade of the brain: doctrine of multilevel analysis," *American Psychologist*, vol. 47, no. 8, pp. 1019–1028, 1992.

[24] R. W. Picard, *Affective Computing*, MIT Press, Cambridge, Mass, USA, 1997.

[25] P. Salovery and J. D. Mayer, "Emotional intelligence," *Imagination, Cognition, and Personality*, vol. 9, no. 3, pp. 185–211, 1990.

[26] W. H. Dai, "Context awarensss and emotional intelligence: the gateway to smart city," *Urban Management*, no. 4, pp. 29–32, 2012.

[27] W. H. Dai, "The development pattern of smart city based on generalized ecological community," *Urban Management*, no. 6, pp. 33–37, 2013.

[28] Y. G. Zhou, *On-Demand Smart Services Based on Context Awareness*, Fudan University, Shanghai, China, 2012.

[29] J. Webster and P. Hackley, "Teaching effectiveness in technology-mediated distance learning," *Academy of Management Journal*, vol. 40, no. 6, pp. 1282–1309, 1997.

[30] J. R. Hill and M. J. Hannafin, "Cognitive strategies and learning from the World Wide Web," *Educational Technology Research & Development*, vol. 45, no. 4, pp. 37–64, 1997.

[31] M. Hannafin, J. R. Hill, K. Oliver, E. Glazer, and P. Sharma, "Cognitive and learning factors in web-based distance learning environments," in *Handbook of Distance Education*, M. G. Moore and W. G. Anderson, Eds., pp. 245–257, Lawrence Erlbaum Associates, London, UK, 2003.

[32] J. H. Zhang and J. P. Zhang, "The influence factors of e-learning and LICE model," *e-Education Research*, vol. 2009, no. 6, pp. 73–77, 2009.

[33] R. R. Lu, *The Research of Influence Factors of College Students' e-Learning Behavior*, Central China Normal University, Wuhan, China, 2012.

[34] W. H. Dai, "Cyber psychological and physical (CPP) computation method based on social neuromechanism," Working Paper, Fudan University, Shanghai, China, 2015.

[35] A. R. Damasio, *Descartes Error: Emotion, Reason and the Human Brain*, Gosset/Putnam Press, New York, NY, USA, 1994.

[36] D. Ververidis and C. Kotropoulos, "Emotional speech recognition: resources, features, and methods," *Speech Communication*, vol. 48, no. 9, pp. 1162–1181, 2006.

[37] D. D. Li, Y. C. Yang, and W. H. Dai, "Cost-sensitive learning for emotion robust speaker recognition," *The Scientific World Journal*, vol. 2014, Article ID 628516, 9 pages, 2014.

[38] D. R. Sanand and S. Umesh, "VTLN using analytically determined linear-transformation on conventional MFCC," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 5, pp. 1573–1584, 2012.

[39] M. H. Yao, "Application of improved genetic algorithm in optimizing BP neural networks weights," *Computer Engineering and Applications*, vol. 49, no. 24, pp. 49–54, 2013.

[40] H. R. Guo, Z. M. Li, X. Wang, and B. Xiong, "A method of improving generalization for BP network based on random GA," *Computer Technology and Development*, vol. 24, no. 1, pp. 105–108, 2014.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 701237, 12 pages
http://dx.doi.org/10.1155/2015/701237

Research Article

A Neuroeconomics Analysis of Investment Process with Money Flow Information: The Error-Related Negativity

Cuicui Wang,[1] João Paulo Vieito,[2] and Qingguo Ma[3,4]

[1] *School of Management, Hefei University of Technology, Hefei 230009, China*
[2] *School of Business Studies, Polytechnic Institute of Viana do Castelo, 4920311 Viana do Castelo, Portugal*
[3] *School of Management, Zhejiang University, Hangzhou 310058, China*
[4] *Neuromanagement Lab, Zhejiang University, Hangzhou 310058, China*

Correspondence should be addressed to Qingguo Ma; maqingguo3669@zju.edu.cn

Received 23 January 2015; Revised 13 April 2015; Accepted 25 April 2015

Academic Editor: Dongrong Xu

This investigation is among the first ones to analyze the neural basis of an investment process with money flow information of financial market, using a simplified task where volunteers had to choose to buy or not to buy stocks based on the display of positive or negative money flow information. After choosing "to buy" or "not to buy," participants were presented with feedback. At the same time, event-related potentials (ERPs) were used to record investor's brain activity and capture the event-related negativity (ERN) and feedback-related negativity (FRN) components. The results of ERN suggested that there might be a higher risk and more conflict when buying stocks with negative net money flow information than positive net money flow information, and the inverse was also true for the "not to buy" stocks option. The FRN component evoked by the bad outcome of a decision was more negative than that by the good outcome, which reflected the difference between the values of the actual and expected outcome. From the research, we could further understand how investors perceived money flow information of financial market and the neural cognitive effect in investment process.

1. Introduction

Most of the part of the literature that analyzes the relationship between investment behavior and stock market information has been done based on historical data, such as shares you own or stock gains; however, with this information alone, it is difficult to understand how an investor acts before and during the investment decision process. One may only understand what happened after the decision was made. Recently, various groups of researchers from the areas of economics, finance, and neuroscience worked together to attempt to understand how the human brain reacts when people make financial decisions [1, 2]. Moreover, with multidisciplinary knowledge and technique, we can further understand how investors perceive kinds of information in financial market. This paper is the first step to understand how money flow information is cognized and processed during investment with technique of neuroscience.

The Use of Neuroscience in Financial Market Investment. Neuroscience can use functional imaging of brain activity (such as functional magnetic resonance imaging (FMRI)) and other techniques (such as electroencephalogram (EEG)) to infer details about how the brain works during financial risk decision-making. Kuhnen and Knutson (2005) used FMRI to examine anticipatory neural activity in a financial decision-making task and found that the activation of the nucleus accumbens, which is a brain region, preceded risky choices as well as risk-seeking mistakes, while the anterior insula was activated prior to riskless choices as well as risk-aversion mistakes [3]. Moreover, the anterior cingulate cortex (ACC) was involved in risk assessment and in many other higher-order cognitive functions [4, 5]. Apart from various functional imaging techniques, EEG measurement is another technique which has the cost efficient advantage and excellent temporal (timing) resolution. Event-related potentials (ERPs)

are voltage fluctuations within EEG recordings which are time sequenced to specific events. Hewig et al. (2007), when investigating ERPs, found that the component of error-related negativity (ERN), one of the components which can reflect the cognitive process by recording brain evoked potentials and average superposition, was related to the degree of reward expectation, the amplitude of which was directly related to risk-taking and decision-making behavior [6]. Hajcak et al. (2006) also stated that the feedback-related negativity (FRN) component was related to the outcome of risk decision-making, which reflected the early appraisal of feedback based on a binary classification of good versus bad outcomes [7].

Although many investigations have been done in the area of decision-making under risk, few have analyzed the brain's electrical activity during realistic investment process. In this investigation, electroencephalogram technology was used to investigate behaviors and the electrical activity of investors' brain by displaying net money flow information of stocks. Money flow is an index measuring the strength of money going in and out of a stock or security in a period. Money flow index (MFI) is an important indicator for technical analysis which would help investors predict trend reversals and share price fluctuations [8]. Moreover, Gryc [9] chose MFI as one input for neural networks in predicting stock price fluctuations of companies. Net money flow (NMF) is calculated by the following formula:

$$\begin{aligned} \mathrm{NMF} &= \sum \left(\mathrm{MF}_{\mathrm{positive}} - \mathrm{MF}_{\mathrm{negative}}\right) \\ &= \sum \left(P_{\mathrm{high}} \times \mathrm{Volume}\right) - \sum \left(P_{\mathrm{low}} \times \mathrm{Volume}\right), \end{aligned} \tag{1}$$

where $\mathrm{MF}_{\mathrm{positive}}$ is the positive money flow, $\mathrm{MF}_{\mathrm{negative}}$ is the negative money flow, P_{high} and P_{low} are the prices of stocks when they are purchased at a higher price and lower price, respectively, and Volume is the number of shares purchased. Thus, there are two categories of net money flow: positive and negative. With this information, we investigated how a positive and/or negative net money flow was perceived and influenced the decision to buy or not to buy the stock, including the choices (the percentage of choosing "to buy" or "not to buy"), the time spent, and brain activities.

ERP Components Involved in Investment. One of the most commonly researched components related to risk decision-making is the ERN, which is a medial frontal negative component of an ERP, first described as a negative deflection response to errors in reaction time tasks by Falkenstein et al. (1991) and Gehring et al. (1993) [10, 11]. Larger ERNs were related to a decrease in error force, an increase in the likelihood of error correction, and slower responses in the following trial. They suggested that remedial action was one of the functions of the ERN process [11].

Holroyd et al. (2004) found that the relative outcome of the trial in relation to different possible outcomes mainly determined the amplitude of the ERN rather than the absolute outcome [12]. It corroborated the view that the relative or subjective relevance of the feedback moderated the amplitude of the ERN and a negative deviation from a reward expectation resulted in an ERN. In a gambling task where participants had to decide between larger and smaller amounts of money which turned out to be either wins or losses, ERN amplitude correlated with the differences in risk-taking behavior after monetary losses [13]. The above finding might suggest a role for the ERN in strategic adjustments of behavior and decision-making.

The literature suggests several theories about ERN, which has some controversies. One is an error detection theory, which assumes that ERN is a neural correlate of mismatch detected by comparing representations of the intended and the actual performed actions [10, 11, 14]. Another theory is a reinforcement learning theory, the basal ganglia of which continuously evaluates the outcome of ongoing behaviors against participants' expectations, and ERN plays a role in the strategic adjustment of behavior and decision-making [15, 16]. Furthermore, conflict monitoring theory assumes that ERN reflects the monitoring of response conflict during response selection [13, 16, 17]. The activation of the correct response gives rise to a transient period during which both the correct response tendency and the already executed incorrect response are activated [4].

Previous researches provided evidence for conflict monitoring theory of ERN by gambling tasks [4, 6]. Based on the blackjack gambling task, Hewig et al. (2007) found that, compared with the condition in which participants had cards with lower scores (<17), the hit decision evoked a more negative ERN for the condition in which the participants had cards with higher scores (>16) [6]. The hit decision when the cards had higher scores had more risk, and the conflict between the desire to win and the desire to be safe should be more severe. A recent study with a gambling task (where the participants who were required to choose to bet or not in each trial were presented with gain or loss feedback after each decision) observed that the ERN magnitudes were more negative for the "to bet" than for the "not to bet" choices, for both large and small stakes and choices involving large rather than small stakes were more negative [4]. It was supported that when the stakes were larger, the conflict to "to bet" was more severe. Since investment decision-making was similar to risk decision-making of a gamble, we speculated that if ERN was evoked in our experiment, conflict monitoring theory would give explanations for ERN component.

In addition to the ERN effect, there are two other components related with evaluating the outcome or feedback of behavioral performance. The first is feedback-related negativity (FRN), a negative ERP component occurring at 200–300 ms after the presentation of feedback, which is maximal over medial frontal scalp [18]. Another component is P300, which is a positive wave usually with peak occurring at about 300 ms poststimulus [19]. Yeung et al. used monetary gambling tasks and found that FRN was elicited by negative outcomes [20], while Hajcak et al. observed that larger P300 amplitudes were evoked by positive outcomes than negative outcomes [21].

This study aimed to use ERP and behavior data to provide further understanding of the cognitive processes involved in investment with money flow information. In this experiment, according to money flow information, participants were

required to decide whether to buy the stock or not. Stocks with negative money flows are normally perceived as being riskier and those with positive money flow less risky. Based on this assumption, one would expect different brain reaction times when volunteers make the decision "to buy" or "not to buy" in the different contexts. The literature citing the above suggests that ERN might reflect the level of riskiness of different choices. Accordingly, one would suggest that ERN could be found to reflect the investment decision process. Additionally, there was a feedback after each decision in this experiment, and FRN and P300 would be elicited by evaluating different kinds of feedbacks.

2. Materials and Methods

2.1. Participants. Twenty-nine major undergraduates (15 females, all right-handed) aged from 20 to 26 years (mean age, 23.5 years) from Zhejiang University participated in this experiment as paid volunteers. They were all familiar with stock investment and had some knowledge on money flow in stock markets but no experience of real investment. They were all native Chinese speakers and had normal or corrected to normal vision with no history of neurological or psychiatric abnormalities. There were only twenty-seven participants recording behavioral data, as some problems were found in the software codes and the behavioral data of the first two participants were not recorded. Before experiment, informed consent was obtained from all volunteers, and the research was approved by the Internal Review Board of Zhejiang University Neuromanagement Lab.

2.2. Electroencephalogram Recording. The experiment was performed in an electrically shielded and sound-attenuated cabin. Participants sat in a comfortable chair and a computer display was located 1 m away from their eyes. The electroencephalography (EEG) was recorded (band pass 0.05–100 Hz, sampling rate 500 Hz) with Neuroscan Synamp2 Amplifier (Scan 4.3.1, Neurosoft Labs Inc.), using Ag/AgCl electrodes placed at 64 scalp sites according to the extended international 10–20 system and referenced to left mastoid with a cephalic (forehead) location as ground. Vertical electrooculograms (EOG) were recorded with one pair of electrodes placed above and below the left eye, horizontal EOG with another pair 10 mm from the lateral canthi. Electrode impedances were maintained below 5 kΩ throughout the experiment. Before the formal experiment was presented, participants were given written and oral instructions and then practiced 30 trial runs.

2.3. Experimental Stimuli. The ERP was measured using a simplified investment task, in which participants were asked to make decisions on hundreds of stocks with net money flow information and then given the feedback. Net money flow as stimuli 1 (S1) contained two categories: positive net money flow information and negative net money flow information. Feedback as stimuli 2 (S2) included three categories: increase, decrease, and no significant change. In order to exclude other factors and keep ERP experiment simple, S1 and S2 only referred to the data of real financial market, but participants were not told the truth.

Table 1: The experimental stimuli of net money flow (S1) and the feedback (S2).

Positive net money flow (S1)	Negative net money flow (S1)	Increase (S2)	Decrease (S2)	No significant change (S2)
40 million	−40 million	+8.0%	−8.0%	+0.2%
42 million	−42 million	+8.2%	−8.2%	+0.4%
44 million	−44 million	+8.4%	−8.4%	+0.6%
46 million	−46 million	+8.6%	−8.6%	+0.8%
48 million	−48 million	+8.8%	−8.8%	+1.0%
50 million	−50 million	+9.0%	−9.0%	−0.2%
52 million	−52 million	+9.2%	−9.2%	−0.4%
54 million	−54 million	+9.4%	−9.4%	−0.6%
56 million	−56 million	+9.6%	−9.6%	−0.8%
58 million	−58 million	+9.8%	−9.8%	−1.0%

Information of S1 used referred to money flow information on stocks from June and July (2011), the prices of which were between 18 Chinese Yuan (CNY) (about \$2.89) and 22 CNY (about \$3.53), with an average of 20 CNY (about \$3.2). By ranking the stock according to the absolute value of the net money flow, it was found that the absolute value of the data which came in the first twenty varied greatly. The absolute value of 40 million to 60 million (CNY) was obtained from questionnaires which had also been administered to participants and this was large enough for nonexperienced participants. Therefore, ten figures as positive data from 40 million and 60 million (CNY) and ten figures as negative data from −60 million to −40 million (CNY) were chosen (see Table 1) as net money flow information for S1.

The S2 consisted of thirty pictures reflecting the increase/decrease of stock in percentage terms, divided into three categories: the percentage close to zero, the positive percentage between 8% and 10%, and the negative percentage between −10% and −8% (ten pictures per category, see Table 1). In order to distinguish positive and negative percentage data, positive feedbacks were highlighted in red and negative in green (in tune with real market). Each picture was digitized at 400 × 300 pixels with a gray background.

2.4. Experiment Procedure. The stimuli consisted of 600 pairs of stimuli (S1) – feedback stimuli (S2), that is, 10 figures (reflecting the net money flow of stocks) × 2 categories (negative or positive) × 10 feedback figures × 3 categories

(increase and decrease, no significant change). They were presented in the center of a screen with black words for the S1 and red or green words for the S2 on a grey background. E-prime software was employed to control the presentation of the stimuli in a random manner.

The experiment was composed of five blocks. For each trial of a block, firstly, there was a "+" presented for 500 millisecond (ms) at the beginning; then, S1 was presented for 1000 ms followed by a random interstimulus interval between S1 and S2, ranging from 300 ms to 700 ms (average mean was 500 ms). The feedback stimuli (S2) was presented for 500 ms, and the interval between the end of S2 and the onset of "+" was 1000 ms (see Figure 1).

Before the experiment, participants were asked to fill in a questionnaire, which was used to assess their familiarity with stock investment and net money flow. Participants were informed that we chose 600 different stocks from the real stock market, the prices of which were approximately 20 CNY in June or July 2011. The net money flow figures (S1) of all stocks were chosen after the morning trading one day in June or July 2011 from a real stock market. The feedback (S2) represented the real increased percentage of stocks following the closing on that day. Throughout each trial, they had to focus on the stimuli and buy one hundred shares or not, according to the net money flow (S1) stimuli. The objective was to obtain the best results from to buy or not to buy stock, as this would result in a monetary reward. Half of the participants were instructed "if buy, please press "1" with left hand; otherwise press "3" with the right," and the others were instructed to press "3" with right hand "if you buy"; otherwise press "1" with the left for the balance. They were told that all stocks were irrelevant and independent. In order to ensure that participants put their efforts in the experiment, there was an additional monetary reward paid depending on their performance, except for 15 CNY (about $2) as basic payment. After each block, we chose four trials randomly, and if participants chose to buy, the additional monetary reward took into account the share price in the end of the day (the data of feedback added 1 and multiplied by the basic price 20 CNY); otherwise the share price in the beginning (20 CNY) was taken into account. We averaged all share prices of stocks which were chosen from five blocks and gave participants as monetary reward of experiment. Additionally, participants were asked to make decision as soon as possible when the net money flow (S1) was presented.

2.5. Electroencephalogram Analysis. EEG recordings for ERN were extracted from −400 ms to 400 ms time-locked to the response onset, with the period from −400 ms to −200 ms prior to the response as baseline. EOG artifacts were corrected using the method proposed by [22]. Trials containing EOG activity or other artifacts (bursts of electromyographic activity or peak-to-peak deflection exceeding ±80 μV) were excluded from averaging. The remaining trials were corrected to baseline. ERPs were averaged for every participant in each of the four conditions according to categories of S1 and response categories for each electrode site, that is, 2 categories of S1 (the positive net money flow and the negative net money flow) × 2 response types (to buy or not). The averaged ERPs were digitally filtered with a low pass filter at 30 Hz (24 dB/Octave).

To quantify the response-locked ERN component, a 2 (net money flow category) × 2 (response category) × 5 (electrodes) within-subjects repeated measure ANOVA for ERN was conducted. The Greenhouse and Geisser (1959) correction was applied for the violation of sphericity assumption when necessary (uncorrected *df* was reported with the ε and corrected p values), and the Bonferroni correction was used for multiple paired comparisons. For the present report, five electrodes: Fz, FCz, Cz, CPz, and Pz (see the pink electrodes in Figure 2), were used for ERN statistical analysis. The restriction of the electrode set in the ERN statistical analyses was based on the hypothesis that a frontocentral midline maximum was expected for ERN and the ERN effects were the strongest of these [6].

For the feedback (S2), electroencephalogram recordings were extracted from −200 ms to 800 ms stimulus-locked to the onset of S2, with the period from −200 ms to 0 ms as baseline. Walsh and Anderson (2012) proved that FRN arose in the anterior cingulate cortex, which was similar with ERN [23]. Many researches chose Fz, FCz, Cz, CPz, and Pz electrodes for FRN analysis [23, 24]. Based on the above information, ERP amplitudes following feedback were also calculated for five midline electrodes (Fz, FCz, Cz, CPz, and Pz, see the pink electrodes in Figure 2).

If the participant decides to "to buy," normally they expect stock prices to rise in the future, while if the choice is "not to buy," their expectation is that the stock price will not rise in the future (it is better to drop). As such, the outcome of the decisions was classified into three categories: good feedback (the feedback was identical with expectation), bad feedback (the feedback was opposite to expectation), and constant feedback (the feedback had no significant change). Following that, a 3 (feedback categories) × 5 (electrodes) within-subjects repeated measure ANOVA for feedback stimuli was conducted.

Behavioral measures (choices in percentage and reaction times) were also considered via an analysis of variance (ANOVA) with response types ("to buy" or "not to buy") and the categories of net money flow (positive or negative) as two within-participant factors.

3. Results

3.1. Behavioral Data. The statistical results for the choices in percentage and for the reaction times (RTs) in four conditions were presented in Figure 3.

For percentage of choices, there was a significant difference in the response type factor ($F(1, 26) = 16.262$, $p = 0.000$), and the percentage of "to buy" (M = 0.288, SD = 0.009) was larger than that of "not to buy" (mean = 0.212, SD = 0.009). Therefore, this would suggest that the participants in our experiment were more risk taking. There was a significant interaction between the net money flow and response type factors ($F(1, 26) = 46.622$, $p = 0.000$), but with no difference between choices in the net money flow factor ($F(1, 26) = 0.151$,

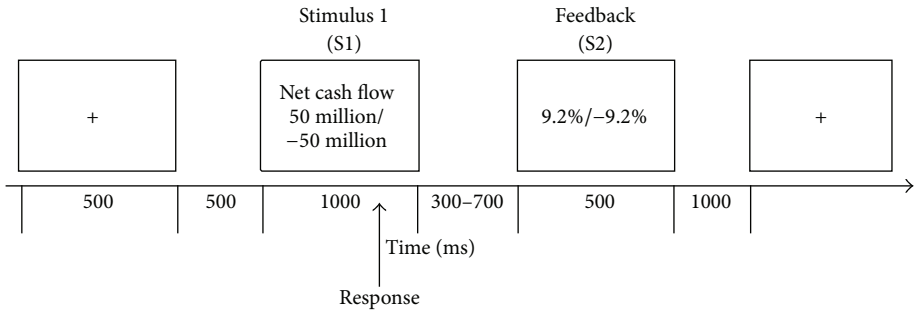


Figure 1: ERP paradigm.

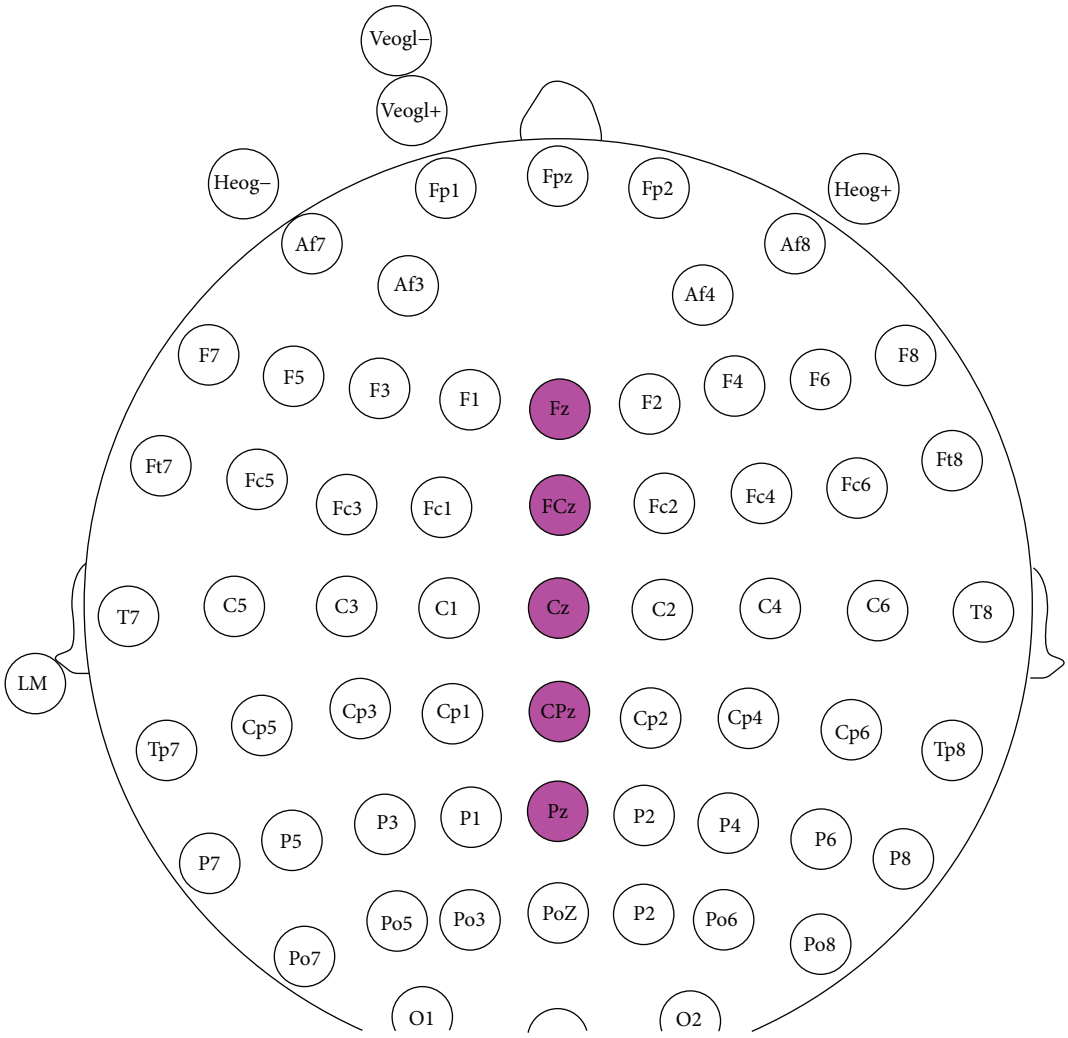


Figure 2: Description of brain EEG location of all the electrodes (the amplitudes of ERN and FRN were calculated for five midline pink electrodes (Fz, FCz, Cz, CPz, and Pz)).

$p = 0.701$). In the positive net money flow category, there was a significant difference in percentage between the "to buy" and "not to buy" choices ($F(1, 26) = 74.765$, $p = 0.000$) and the percentage of choices "to buy" (M = 0.338, SD = 0.010) was greater than that of "not to buy" (M = 0.162, SD = 0.010). However, no significant difference was found in the category of negative money flow.

For RTs, there was a significant difference in the net money flow factor ($F(1, 26) = 9.098$, $p = 0.006$), however, with no significant difference in the factor of response types ($F(1, 26) = 0.769$, $p = 0.389$). It suggested that participants responded faster in the condition of positive net money flow (M = 459.07 ms, SD = 118.97) than in the condition of negative net money flow (M = 469.35 ms, SD = 115.31). There was a

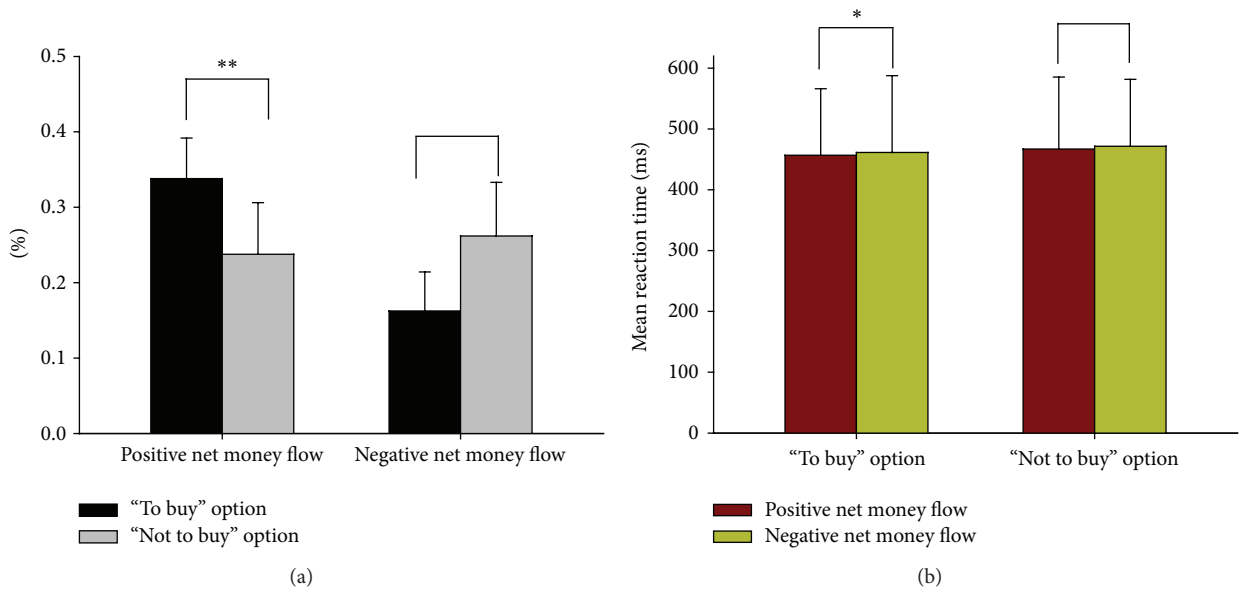


Figure 3: (a) Comparing "to buy" option and "not to buy" option in percentage in positive and negative net money flow category, respectively. Error bars indicate SD of choices in percentage. (b) Mean RTs of positive and negative net money flow sorted by response "to buy" and "not to buy," respectively. Error bars indicate SD of RT.

significant difference when the RTs of "to buy" in the two categories of the net money flow were compared ($F(1, 26) = 5.211$, $p = 0.031$). It was faster for the RTs of "to buy" in the positive net money flow than in the negative net money flow. However, no significant difference was found when comparing the RTs of "not to buy" in the two categories of the net money flow ($F(1, 26) = 4.480$, $p = 0.037$). The ANOVA revealed that the RTs had no significant interaction with the factors of net money flow and response types ($F(1, 26) = 0.000$, $p = 0.997$).

3.2. ERP Data

3.2.1. ERP Data of Stimuli (S1). The response-locked grand average ERPs of S1 for "to buy" response in the positive net money flow, "not to buy" response in positive net money flow, "to buy" response in the negative net money flow, and "not to buy" in the negative net money flow were displayed in Figure 4. Response onset was presented at 0 msec. From the figure, we can see the negative net money flow evoked significantly larger negative but less negative amplitude than positive net money flow for ERN in the responses "to buy" and "not to buy," respectively.

The ANOVA for the mean amplitude of ERN in the 0 to 40 ms time window revealed no significant effects for the net money flow ($F(1, 28) = 0.931$, $p = 0.343$) and response type factors ($F(1, 28) = 2.859$, $p = 0.102$). However, significant effects for the electrodes factor were detected ($F(4, 112) = 80.890$, $\varepsilon = 0.479$, $p = 0.000$). For the amplitude of electrodes, the mean negative voltages distributed from frontal to parietal regions showed a decline trend: frontal electrode Fz (M = −2.665 μV, SD = 0.627), frontal-central electrode FCz (M = −0.812 μV, SD = 0.738), central electrode Cz (M = 1.161 μV, SD = 0.763), central-partial electrode CPz (M = 2.722 μV, SD = 0.706), and partial electrode Pz (M = 2.958 μV, SD = 0.627). There were significant differences among each electrode except that there was no significant difference between the CPz and Pz electrodes.

The ANOVA revealed that the ERN had significant interaction with the net money flow and response type factors ($F(1, 28) = 10.810$, $p = 0.003$). However, there was no significant interaction with the net money flow and electrodes factors, or the response type and electrodes, or the net money flow × response type × electrodes. Furthermore, for ERN in the responses "to buy," the negative net money flow (M = 0.521, SD = 0.664) evoked significantly larger negative amplitude than positive net money flow (M = 1.512, SD = 0.704) ($F(1, 28) = 7.294$, $p = 0.012$), while, for ERN in the responses "not to buy," the positive net money flow (M = −0.390, SD = 0.848) evoked significantly larger negative amplitude than negative net money flow (M = 1.048, SD = 0.669) ($F(1, 28) = 8.406$, $p = 0.007$). Table 2 showed all the statistical results of within-subjects repeated measure ANOVA for ERN.

3.2.2. Feedback Stimuli (S2). For feedback stimuli, the grand average for the Fz, FCz, Cz, CPz, and Pz electrodes is shown in Figure 5. Feedback stimuli onset was presented at 0 msec. In stimulus-locked ERPs analysis, −200 ms to 0 ms was chosen as the baseline. From this figure, the amplitude of FRN and P300 evoked by "bad feedback" and "constant feedback" was more negative than "good feedback."

In the case of "bad feedback," it is possible to observe that the FRN appeared to peak at approximately 250 ms after the feedback stimulus. The ANOVA for the mean amplitude of FRN in the 200 to 300 ms time window was

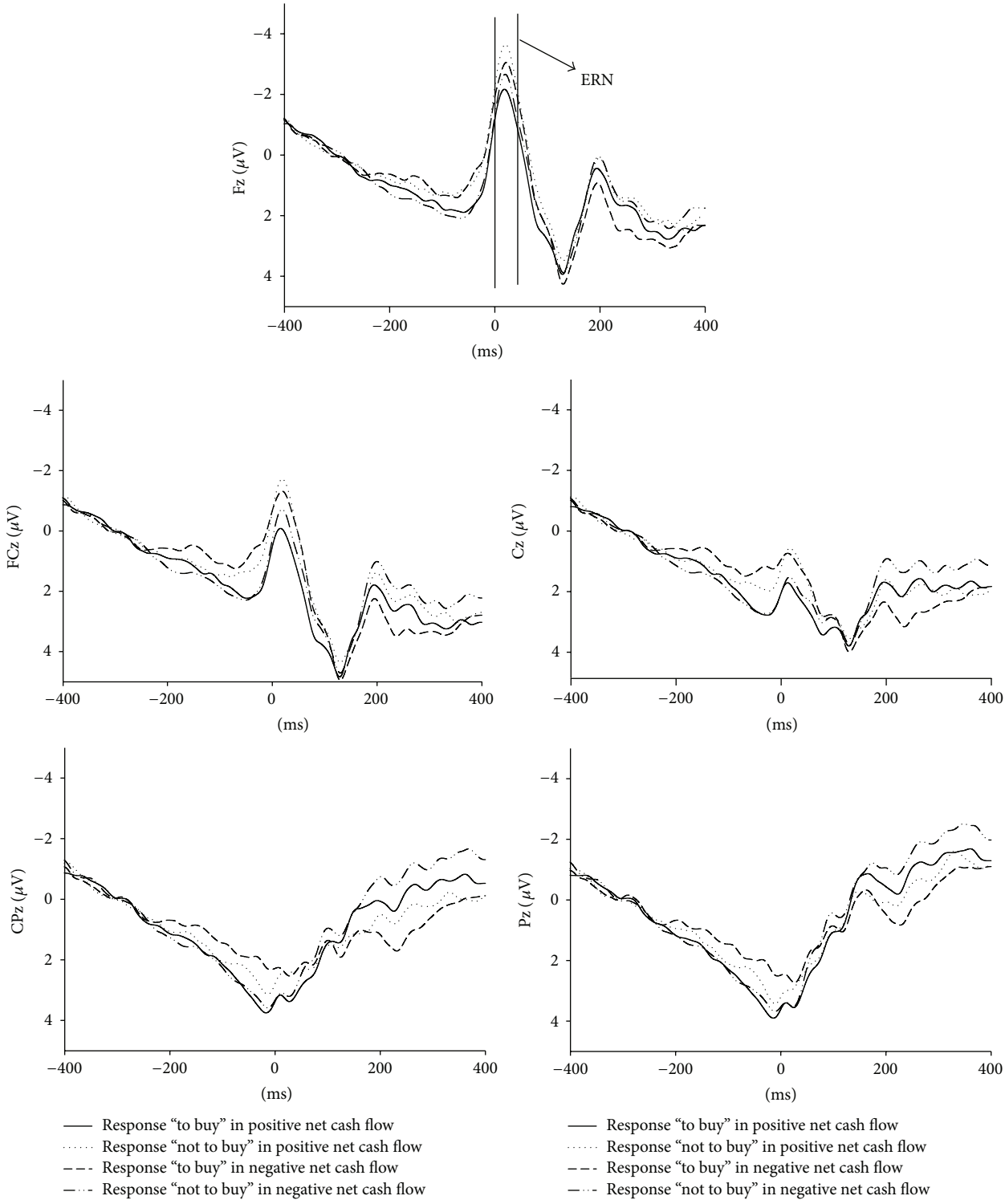


Figure 4: The grand averaged ERPs of S1 from channels Fz, FCz, Cz, CPz, and Pz are described separately for the "to buy" response in the positive net money flow, "not to buy" response in the positive net money flow, "to buy" response in the negative net money flow, and "not to buy" response in the negative net money flow.

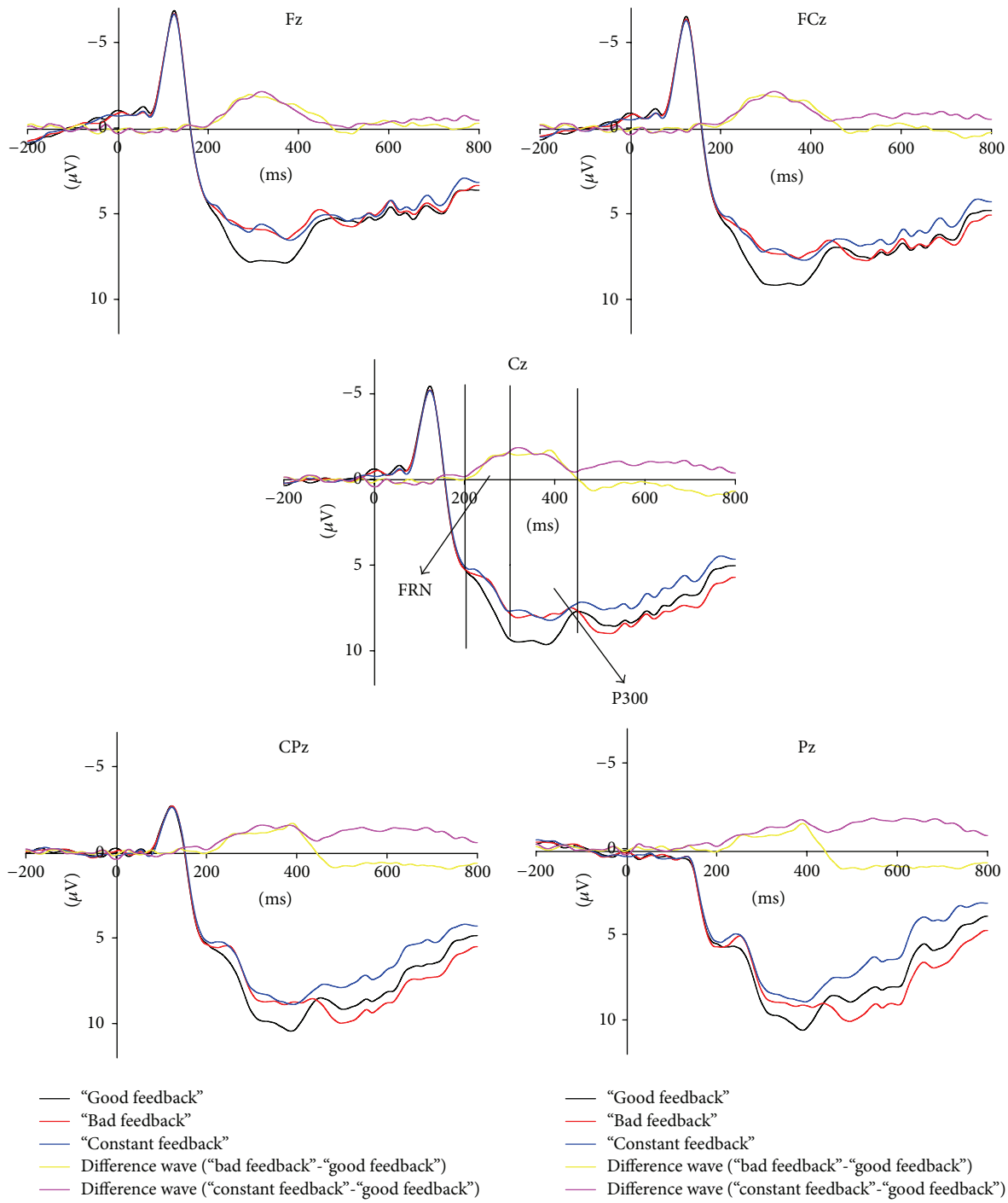


Figure 5: The grand averaged ERPs of S2 from channels Fz, FCz, Cz, CPz, and Pz separately for the "good feedback," "band feedback," and "constant feedback."

conducted, and the main effect of feedback categories (good feedback versus bad feedback versus constant feedback) was found ($F(2, 56) = 12.111$, $p = 0.000$) with no significant effect of electrode location. By pairwise comparisons, the amplitude evoked by "bad feedback" ($p = 0.005$) and "constant feedback" ($p = 0.000$) was more negative than that evoked by "good feedback." However, there was no significant difference between "bad feedback" and "constant feedback" ($p = 1$).

In the time window of 300 ms to 450 ms after the feedback presentation for P300, there were main significant effects in the feedback category ($F(2, 56) = 18.506$, $p = 0.000$) and electrode location ($F(4, 112) = 10.106$, $\varepsilon = 0.354$, $p = 0.000$). By pairwise comparisons, the amplitude evoked by "bad feedback" ($p = 0.000$) and "constant feedback" ($p = 0.000$) was more negative than that evoked by "good feedback." However, there was no significant difference between "bad feedback" and "constant feedback" ($p = 0.723$).

Table 2: The statistical results of within-subjects repeated measure ANOVA for ERN with factors Net money flow (positive/negative), response type ("to buy"/"not to buy"), and electrodes (Fz, FCz, Cz, CPz, and Pz).

Factor	*df*, error	Greenhouse and Geisser (ε)	F/t value	p value
Main effect				
Net money flow	1, 28	—	0.931	0.343
Response type	1, 28	—	2.859	0.102
Electrodes	4, 112	0.478	80.89	**0**
Interaction effect				
Interaction of net money flow and response type	1, 28	—	10.81	**0.003**
Interaction of net money flow and electrodes	4, 112	0.481	0.794	0.453
Interaction of response type and electrodes	4, 112	0.533	2.6	0.079
Interaction of net money response, response type, and electrodes	4, 112	0.415	1.166	0.33
Simple effect				
Net money flow in response of "to buy"	1, 28	—	7.294	**0.012**
Net money flow in response of "not to buy"	1, 28	—	8.406	**0.007**

4. Discussion

In this experiment, we used ERPs to examine the perception of money flow information and further understand the cognitive processes involved in investment. It was observed that ERN components were more negative for the "to buy" option but were less negative for the "not to buy" option in the negative net money flow than in the positive net money flow. We suggested that the factor of net money flow affected the risk level of stocks and the findings for the ERN effect in the investment decision process were sensitive to the risk evaluation of the choices. ERN may be used to act as an alerting system to prepare the brain for the potential negative consequences of actions. For the FRN component, actual results that were different from expectations evoked more negative amplitude than those that were in line with expectations, reflecting the different cognitions for different kinds of feedback results.

The behavioral data which demonstrated that participants were more likely to choose "to buy" than "not to buy" reflected risk taking, which was consistent with previous investigations [4, 6]. Compared with the choices in the negative net money flow condition, there was larger percentage of "to buy" option and less percentage of "not to buy" option in the positive net money flow condition. As a positive net money flow condition of the stock was regarded as having a larger possibility of good performance in the future which may be considered low risk, the conflict to choose "not to buy" was severe; however, the same was true for the negative net money flow condition. This was also supported by the RTs results. Compared to the RTs in the negative net money flow, it was easier and implied significantly shorter response time to choose the "to buy" option in the positive net money flow.

Clearly, the amplitude of ERN evoked by the decision "to buy" was significantly more negative in the negative net money flow than in the positive net money flow; however, ERN elicited by the decision "not to buy" had an inverse effect. The component of ERN reflected the neural basis in the investment decision process with money flow information. Decision-making on investment referred to the process by which one action (e.g., "to buy" or "not to buy") was chosen based on the assessment of the potential costs and benefits associated with it. In this experiment, participants were paid according to the outcome of their choices; therefore, they had to compare risk with reward, choose one action, and evaluate the outcome obtained for that particular action. In real stock markets, we only know that investors refer to money flow information in investment decision-making, with no profound understanding of its effect and the cognitive process. We speculated from ERN results that stocks with larger positive net money flow were often considered as low-risk stocks and had a stronger possibility to increase. However, stocks with large negative net money flow were often considered as high-risk stock and had a greater probability to decrease in price. For the purpose of earning greater returns, the conflict to choose "to buy" in positive net money flows was smaller than in negative net money flows; however, this decision was severer for the "not to buy" option. The findings provided further evidence that conflict monitoring theory was one important function of ERN component.

In this investigation, the amplitude of ERN was more negative in the frontal, frontal-central, and central areas than in the central partial and partial areas. This may be related to the source of the ERN. Many investigations by dipole source analysis of the ERN, which can speculate the source of the ERN components according to the EEG scalp distribution, suggested that this component was generated in the anterior cingulate cortex (ACC) [25–27]. The ACC is connected to the areas of the brain which are responsible for drive and arousal, for motor output and for cognition, and which deal with risk assessment [28–30]. According to an extended version of the conflict monitoring theory [31], the monitored ACC is activated not only during different stimulus-response mappings in the selection of the responses, but also when different

internal desires are aroused [4]. During the investment with different money flow information, the corresponding levels of desire to win and desire to be safe with action "to buy" and "not to buy" were not the same. And the conflict of taking a risk when choosing "to buy" in a negative net money flow or rejecting by choosing "not to buy" in a positive net money flow may well be detected by the ACC and the ERN effect was thus ensured.

FRN follows the display of negative feedback [32]. It can be used as a tool for studying reward valuation and decision-making, which reflects the difference between the values of actual and expected feedback or reward prediction errors [23, 33]. In this experiment, when the participant chose to buy the stock in question, their expectation was that the price of the stock would increase, and when the participant chose not to buy the stock, they hoped the price of the stock would not rise. If the feedback was not in line with their predictions, FRN was evoked. There was no difference between "bad feedback" and "constant feedback" in the amplitude of FRN, but there was a significant difference between "good feedback" and "bad feedback." The results may show that the effect derived from the stock for which prices did not change was the same as that for stock for which the prices dropped in the mind of the investors. However, in the later time window of 300 ms to 450 ms, the finding of P300 increased amplitudes for "good feedback" reflected different emotional state to feedback stimuli. Previous research found that P300 amplitude was larger for emotional stimuli than for neural stimuli [33]. Furthermore, the positive emotional stimuli induced larger P300 than the negative stimuli [34]. In this study, "good feedback" induced a positive emotional state, but "bad feedback" or "constant feedback" made investors have a bad mood.

There were some differences between ERN and FRN evoked in our experiment and traditional ERN and FRN. Traditionally, the ERN peaks 100 ms after response errors, and FRN emerges at 200 ms and peaks 300 ms after negative feedback onset [23]. However, in our study, the ERN peaked around 20 ms after response onset, and FRN also emerged at 200 ms but with peak of about 250 ms after negative feedback (bad feedback or constant feedback). It is because our paradigm was more complex than the usual paradigms used for ERN and FRN research. Typical ERN appears in speeded reaction tasks, in which errors are due to impulsive responding. But, in our experiment, it was a more complicated cognitive process under investment with money flow information. ERN was evoked not by error response but by conflict/risk perception. For FRN in our experiment, we speculated that as the feedback was closely connected with decision behavior of participants, the feedback stimuli may be processed into cognition earlier and the latency of FRN was less than traditional latency (300 ms).

One might ask why there was no significant difference between the choices of "to buy" and "not to buy" in a negative net money flow and why there was no significantly different effect in the ERN between response of "to buy" and "not to buy" in a negative net money flow. The main reason was that people preferred to take risk. In the gambling tasks, Hajcak et al. (2006) and Yu and Zhou (2009) found that people were more likely to choose "to buy" than "not to buy" in risk decision [4, 7]. This result was also supported by the data from the experiment which demonstrated that the percentage of participants choosing "to buy" was larger than that of those choosing "not to buy." However, in the negative net money flow, data suggested that participants experienced some conflict when deciding to buy with higher risks, and this to some extent inhibited participants' fondness for taking risk. Therefore, the number of choices "to buy" may decrease but there was no significant difference with the "not to buy" choices. This may explain why no significant difference was found in the ERN amplitude between the "to buy" and "not to buy" responses in a negative net money flow.

Findings reported in this paper were a good starting point to further understand the cognitive processes in investment behavior of financial market. However, real financial market is a complex environment in which there are so many factors (such as industry information, government policy, and other persons' behavior) affecting the investment of decision-making. The lab studies using EEG or FMRI method only focus on one or two factors and control other factors. As our study only aimed at the role of money flow information in investment, further research should investigate how other factors of financial market affect the cognitive process of investment. Moreover, our study only recruited college students with stock knowledge but no speculation experience, and in further research we can study whether the neural basis of investment could be modulated by individual investing experience. Although more refining and simplifying of the paradigm used in this paper was required, it is a small forward step to use basic neuroscience research methods and apply them to real world practice.

The paper used EEG technique to infer details about how the brain works during financial risk decision-making, which provided the new insights into the financial computation and financial intelligence analysis. In financial decision-making, many models posited distinct cognitive and emotional contributions to decision-making under uncertainty, and cognitive processes typically involved exact computations according to a cost-benefit calculus [35]. This paper could provide a new reference for the cognitive processes in the financial computation based on EEG signals.

5. Conclusion

Using a simplified investment task in which a participant had to choose to buy or not to buy one stock with net money flow information, it was demonstrated that the ERN was more negative for the "to buy" response but less negative for the "not to buy" response in the negative net money flow than in the positive net money flow. FRN was larger when evoked by "bad feedback" than when evoked by a "good feedback," which reflected the difference between the values of actual and expected outcomes. From this experiment, we could further understand the neural cognitive effect of money flow information in the process of investment. The ERN which was generated by ACC reflected the function of conflict monitoring in risk investment. This component may be used

as an early warning index to alert the brain to prepare for potential negative consequences when investing in high-risk stock or other risky actions.

Disclosure

Some contents of this paper as a working paper were exchanged and discussed in World Finance & Banking Symposium on 17th-18th December, 2012 in Shanghai, China.

Conflict of Interests

There is no conflict of interests.

Acknowledgments

This work was supported by "the Fundamental Research Funds for the Central Universities" (2015), Grants nos. 71371167 and 71201045 from the National Natural Science Foundation, and Innovation Program for Students of Zhejiang Province (2014). The authors would like to thank Chi Sheh as discussant of the paper from the University of the West, which gave many good suggestions. The authors checked and assured that this paper was not published in any other journal.

References

[1] M. Hubert, "Does neuroeconomics give new impetus to economic and consumer research?" *Journal of Economic Psychology*, vol. 31, no. 5, pp. 812–817, 2010.

[2] U. Witt and M. Binder, "Disentangling motivational and experiential aspects of 'utility'—a neuroeconomics perspective," *Journal of Economic Psychology*, vol. 36, pp. 27–40, 2013.

[3] C. M. Kuhnen and B. Knutson, "The neural basis of financial risk taking," *Neuron*, vol. 47, no. 5, pp. 763–770, 2005.

[4] R. Yu and X. Zhou, "To bet or not to bet? The error negativity or error-related negativity associated with risk-taking choices," *Journal of Cognitive Neuroscience*, vol. 21, no. 4, pp. 684–696, 2009.

[5] A. C. K. van Duijvenvoorde, H. M. Huizenga, L. H. Somerville et al., "Neural correlates of expected risks and returns in risky choice across development," *Journal of Neuroscience*, vol. 35, no. 4, pp. 1549–1560, 2015.

[6] J. Hewig, R. Trippe, H. Hecht, M. G. H. Coles, C. B. Holroyd, and W. H. R. Miltner, "Decision-making in blackjack: an electrophysiological analysis," *Cerebral Cortex*, vol. 17, no. 4, pp. 865–877, 2007.

[7] G. Hajcak, J. S. Moser, C. B. Holroyd, and R. F. Simons, "The feedback-related negativity reflects the binary evaluation of good versus bad outcomes," *Biological Psychology*, vol. 71, no. 2, pp. 148–154, 2006.

[8] W. Brock, J. Lakonishok, and B. LeBaron, "Simple technical trading rules and the stochastic properties of stock returns," *The Journal of Finance*, vol. 47, no. 5, pp. 1731–1764, 1992.

[9] W. Gryc, "Neural network predictions of stock price fluctuations," Tech. Rep., 2010, http://i2r.org/nnstocks.pdf.

[10] M. Falkenstein, J. Hohnsbein, J. Hoormann, and L. Blanke, "Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 6, pp. 447–455, 1991.

[11] W. J. Gehring, B. Goss, M. G. Coles, D. E. Meyer, and E. Donchin, "A neural system for error detection and compensation," *Psychological Science*, vol. 4, no. 6, pp. 385–390, 1993.

[12] C. B. Holroyd, J. T. Larsen, and J. D. Cohen, "Context dependence of the event-related brain potential associated with reward and punishment," *Psychophysiology*, vol. 41, no. 2, pp. 245–253, 2004.

[13] N. Yeung and A. G. Sanfey, "Independent coding of reward magnitude and valence in the human brain," *The Journal of Neuroscience*, vol. 24, no. 28, pp. 6258–6264, 2004.

[14] M. G. H. Coles, M. K. Scheffers, and C. B. Holroyd, "Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing," *Biological Psychology*, vol. 56, no. 3, pp. 173–189, 2001.

[15] C. B. Holroyd and M. G. H. Coles, "The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity," *Psychological Review*, vol. 109, no. 4, pp. 679–709, 2002.

[16] C. B. Holroyd, N. Yeung, M. G. H. Coles, and J. D. Cohen, "A mechanism for error detection in speeded response time tasks," *Journal of Experimental Psychology: General*, vol. 134, no. 2, pp. 163–191, 2005.

[17] M. M. Botvinick, C. S. Carter, T. S. Braver, D. M. Barch, and J. D. Cohen, "Conflict monitoring and cognitive control," *Psychological Review*, vol. 108, no. 3, pp. 624–652, 2001.

[18] R. Yu, Y. Luo, Z. Ye, and X. Zhou, "Does the FRN in brain potentials reflect motivational/affective consequence of outcome evaluation?" *Progress in Natural Science*, vol. 17, pp. 136–143, 2007.

[19] S. Sutton, M. Braren, J. Zubin, and E. R. John, "Evoked-potential correlates of stimulus uncertainty," *Science*, vol. 150, no. 3700, pp. 1187–1188, 1965.

[20] N. Yeung, C. B. Holroyd, and J. D. Cohen, "ERP correlates of feedback and reward processing in the presence and absence of response choice," *Cerebral Cortex*, vol. 15, no. 5, pp. 535–544, 2005.

[21] G. Hajcak, C. B. Holroyd, J. S. Moser, and R. F. Simons, "Brain potentials associated with expected and unexpected good and bad outcomes," *Psychophysiology*, vol. 42, no. 2, pp. 161–170, 2005.

[22] H. V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, "A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP," *Psychophysiology*, vol. 23, no. 6, pp. 695–703, 1986.

[23] M. M. Walsh and J. R. Anderson, "Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice," *Neuroscience and Biobehavioral Reviews*, vol. 36, no. 8, pp. 1870–1884, 2012.

[24] G. F. Potts, L. E. Martin, S.-M. Kamp, and E. Donchin, "Neural response to action and reward prediction errors: comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations," *Psychophysiology*, vol. 48, no. 2, pp. 218–228, 2011.

[25] W. J. Gehring, J. Himle, and L. G. Nisenson, "Action-monitoring dysfunction in obsessive-compulsive disorder," *Psychological Science*, vol. 11, no. 1, pp. 1–6, 2000.

[26] W. H. R. Miltner, J. Brauer, H. Hecht, R. Trippe, and M. G. H. Coles, "Parallel brain activity for self-generated and observed errors," in *Errors, Conflicts, and the Brain: Current Opinions on Performance Monitoring*, pp. 124–129, 2004.

[27] H. T. van Schie, R. B. Mars, M. G. H. Coles, and H. Bekkering, "Modulation of activity in medial frontal and motor cortices during error observation," *Nature Neuroscience*, vol. 7, no. 5, pp. 549–554, 2004.

[28] M. Ernst, E. E. Nelson, E. B. McClure et al., "Choice selection and reward anticipation: an fMRI study," *Neuropsychologia*, vol. 42, no. 12, pp. 1585–1597, 2004.

[29] A. N. McCoy and M. L. Platt, "Risk-sensitive neurons in macaque posterior cingulate cortex," *Nature Neuroscience*, vol. 8, no. 9, pp. 1220–1227, 2005.

[30] T. Paus, "Primate anterior cingulate cortex: where motor control, drive and cognition interface," *Nature Reviews Neuroscience*, vol. 2, no. 6, pp. 417–424, 2001.

[31] V. Van Veen, J. D. Cohen, M. M. Botvinick, V. A. Stenger, and C. S. Carter, "Anterior cingulate cortex, conflict monitoring, and levels of processing," *NeuroImage*, vol. 14, no. 6, pp. 1302–1308, 2001.

[32] W. H. R. Miltner, C. H. Braun, and M. G. H. Coles, "Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a 'generic' neural system for error detection," *Journal of Cognitive Neuroscience*, vol. 9, no. 6, pp. 788–798, 1997.

[33] Y. Bai, K. Katahira, and H. Ohira, "Valence-separated representation of reward prediction error in feedback-related negativity and positivity," *Neuroreport*, vol. 26, no. 3, pp. 157–162, 2015.

[34] V. S. Johnston, D. R. Miller, and M. H. Burleson, "Multiple P3s to emotional stimuli and their theoretical significance," *Psychophysiology*, vol. 23, no. 6, pp. 684–694, 1986.

[35] S. R. Quartz, "Reason, emotion and decision-making: risk and reward computation with feeling," *Trends in Cognitive Sciences*, vol. 13, no. 5, pp. 209–215, 2009.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 108417, 7 pages
http://dx.doi.org/10.1155/2015/108417

Research Article

P300 and Decision Making under Risk and Ambiguity

Lei Wang,[1,2] Jiehui Zheng,[1,2] Shenwei Huang,[1,2] and Haoye Sun[1,2]

[1]*Department of Management Science and Engineering, School of Management, Zhejiang University, Hangzhou 310058, China*
[2]*Neuromanagement Lab, Zhejiang University, Hangzhou 310027, China*

Correspondence should be addressed to Lei Wang; wang_lei@zju.edu.cn

Received 10 February 2015; Revised 16 April 2015; Accepted 16 April 2015

Academic Editor: Justin Dauwels

Our study aims to contrast the neural temporal features of early stage of decision making in the context of risk and ambiguity. In monetary gambles under ambiguous or risky conditions, 12 participants were asked to make a decision to bet or not, with the event-related potentials (ERPs) recorded meantime. The proportion of choosing to bet in ambiguous condition was significantly lower than that in risky condition. An ERP component identified as P300 was found. The P300 amplitude elicited in risky condition was significantly larger than that in ambiguous condition. The lower bet rate in ambiguous condition and the smaller P300 amplitude elicited by ambiguous stimuli revealed that people showed much more aversion in the ambiguous condition than in the risky condition. The ERP results may suggest that decision making under ambiguity occupies higher working memory and recalls more past experience while decision making under risk mainly mobilizes attentional resources to calculate current information. These findings extended the current understanding of underlying mechanism for early assessment stage of decision making and explored the difference between the decision making under risk and ambiguity.

1. Introduction

Risk and ambiguity are two conditions in which the likelihood of outcomes is uncertain [1]. But differences are here to stay; in the condition of risk, the probability distribution of possible outcomes is well defined, which can be used to calculate the expectancies of outcomes and compare between choices. The probability of outcomes determines the riskiness of risk condition that high probability brings lower risk and vice versa. However, under the ambiguous condition, participants are unknown of the probabilities of outcome [2–4]. The participants tend to subjectively add probability of the outcome in decision making, and it is difficult to be described by the theoretical models accurately. The distinction between the two uncertain conditions was first illustrated by the Ellsberg Paradox, which indicates the so-called phenomenon of ambiguity aversion [5] that means the peoples' preference to bet in risky conditions rather than ambiguous conditions.

Researchers have investigated the underlying mechanism of ambiguity aversion for a long time and put forward some explanations, such as the competence hypothesis [5] and the comparative ignorance [6]. Rode and colleagues put these hypotheses into two categories: cognitive approach and motivational approach [7]. Cognitive mechanism regarded ambiguity as a second-order probability distribution of option [8, 9]. Participants can obtain probability information from former experience. Motivational approach focused on effective factors that come from the lack of information. Frisch and Baron put forward that the risky prospect is more justifiable than the ambiguous one due to missing potentially available probabilistic information of the latter [10]. Fox and Weber found people feeling less confident for issues that they did not fully understand [11]. And a bad outcome in ambiguous conditions may be ascribed to the incompetence or thoughtless choice [12], while in risky conditions the poor judgment cannot be blamed for when the outcome is undesirable. Since all required information is provided, participants are more likely to attribute the bad outcome to bad luck [5].

In the latest ten years, numerous empirical studies about uncertainty decision making using functional magnetic resonance imaging (fMRI) tools have attempted to contrast these two types of uncertainty at neural level, and identified the neural mechanism related to decision making under ambiguity and risk [3, 13–15]. Hsu et al. suggested a general neural circuit responding to degrees of uncertainty [14]. In

their first treatment of Card-Deck that compared the pure risk (where probabilities were known with certainty) against pure ambiguity as baseline, the participants were asked to choose between betting on one of the two options and taking a fixed monetary reward in each trail. Finally, they compared the difference between ambiguity and certain conditions, and difference between risk and certain conditions, respectively. However, in our experiment, we asked the participants to choose to bet or not in ambiguous and risky conditions, in both of which they would get no fixed payouts if they gave up betting. This design made us compare the difference of decision making between risky and ambiguous conditions, which was not further studied in Hsu et al.'s work [14].

Compared with previous fMRI studies [3, 4], event-related brain potentials (ERPs) offer better temporal resolutions for researchers to study how a cognitive process is taking place in real-time. Decision making is a continuous process, which can be divided into several stages temporally, including assessment and formation of preferences among possible options, selection, and feedback or evaluation of an outcome. Until now, a lot of work has been conducted for the feedback stage of decision making while little attention is allocated to earlier stages, such as assessment of the options. Previous ERP studies about decision making under uncertainty mainly focused on the feedback stage and explored components such as FRN and P300 [16–19]. Zhang's studies indicated that the P300 component was sensitive to risky decision making [20]. Zhou et al. used a risky gambling game and found out that the P300 and FRN were quite different between the conditions of win and loss [21]. Xu and her colleagues applied event-related brain potential (ERP) to explore how an uncertain (risk and ambiguity) cue was processed. They designed a gambling task called "wheel of fortune" and found out that a larger P300 was elicited by the unexpected cue under uncertain condition [22]. Gu and his colleagues found that P3 was larger in the positive outcome condition than the other three conditions (negative, neutral, and ambiguous) by using a monetary gambling task [23]. These studies revealed that both the risky and ambiguous conditions would evoke the P300. Previous studies, including the above two studies, mainly focused on the feedback stage and viewed P300 as a typical indicator to reflect rewarding processing in decision-making [19, 24, 25]. However, until now few ERP studies have focused on the early stage of decision making before the feedback stage and explored the corresponding neural mechanism.

P300 is one of the most commonly studied components of ERPs for decision making [26], and it usually emerges in the late period after stimuli onset (300–600 ms). In task relevant paradigms, the amplitude of P300 is generally considered as a representation of memory load [27–29]. Task load can be divided into two dimensions: driving task load and working memory load [30]. P300 is related to memory processes in the evaluation of stimuli for the subsequent response [31, 32]. Task-related information is updated through learning and forgetting in working memory and P300 is elicited at the same time [33]. The decrease of P300 amplitude was observed in several tasks employing high memory load [34–36]. Since ambiguous tasks provide less definite information than risky ones, individuals need to not only mobilize some attentional resources to analyze current stimuli, but also recall a large amount of past practices and memory to get probability information to form a clear expectance of outcome and reduce cognitive strain in dealing with ambiguous tasks. So it is more effortful and difficult for participants to make decisions under the ambiguous condition, which may induce a higher working memory load [37, 38].

Compared with the previous studies, the present research focuses on the temporal electrophysiological changes and difference at the early stimuli assessment stage of decision making between the ambiguous and risky conditions, which we believe can help better understand the process of decision making. For exploring the cognitive process of these two decision types, we designed a monetary gambling game in which the probabilities of outcomes were known (risk) or unknown (ambiguity) while the outcomes were varied across trials but balanced in pairs. Considering that the P300 is a typical neural indicator of decision making and can be interpreted as a reflection of memory load and is inversely related to difficulty of decision making [37, 38], we speculated that it would elicit a smaller P300 amplitude under ambiguity than that under risk decision making.

2. Materials and Methods

2.1. Participants. Participants were recruited from the student population of the Zhejiang University. A total number of 12 right-handed participants took part in the experiment (5 females; average age: 22.58 years, standard deviation (SD) = 1.55 years, range 20–25 years). All participants had no history of neurological or psychiatric disorders.

2.2. Materials. Prior to the experiment, participants were informed that the purpose of the experiment was to investigate brain waves during gambling. An informed consent, approved by the Internal Review Board of Neuromanagement Lab, Zhejiang University, was obtained from each participant before formal experiment.

The game consists of two types of primes: risky stimuli and ambiguous stimuli. Risky stimuli presented a monetary value with 50% probability for gain and loss each. Ambiguous stimuli showed only a monetary value but concealed its corresponding probability on gain or loss. Risky stimuli and ambiguous stimuli appeared in the experiment with equal frequency. The monetary value was a random integer ranging from 11 to 190 for each trail in both conditions. Each stimulus was made into a picture and digitized at 200 × 150 pixels (Figure 1).

2.3. Procedure. The EEG participant was seated in a chair 1 meter in front of a Dell 22 in. CRT display (screen resolution: 1024 × 768; refresh rate: 120 Hz; color quality: highest 32 bit). Stimuli were presented sequentially in the center of a computer screen with a visual angle of 2.58°× 2.4°. In order to draw participants' attention, the screen presents a "+" for a fixed duration of 300 ms at the beginning of each trial. Then, a stimulus of risk or ambiguity was shown after a mean delay of 700 ms. After the presentation of the risky or ambiguous stimuli, if the participants chose to bet, a 500 ms blank was

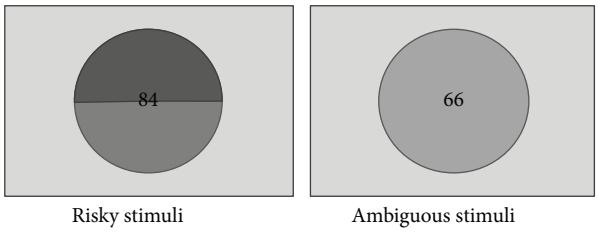


Figure 1: Stimulus: risk stimuli and ambiguous stimuli.

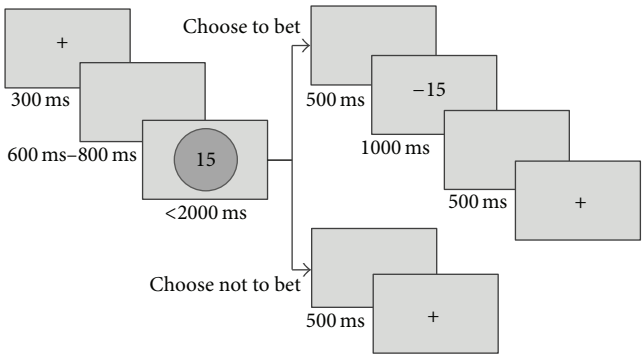


Figure 2: Task procedure.

presented followed by the outcome with 50% probability of lose or win for 1000 ms; otherwise it went to the next trial after a 500 ms blank (see Figure 2).

All subjects were asked to read the experiment instructions before the experiment. A practice block was administered before the formal test. The formal test had 6 blocks with 60 trials each. Each condition (risk versus ambiguity) had 180 trials distributed randomly in 6 blocks. Each gambling value was presented twice: one in risky condition and the other in ambiguous condition. In half of the trials in each condition, the feedback was positive (the participants would gain if they decided to bet); in the other half of trials, the feedback was negative. The total gains and total losses were counterbalanced in each condition and across the two conditions. However, the participants were blind to this design; they were only told to choose to bet or not and that they would be rewarded according to their performance. The presentation of stimuli and recording of the participant's responses were controlled by STIM 2 software (Stim2, Neurosoft Labs, Inc., Sterling, USA).

In summary, all the stimuli probabilities and the monetary magnitude presented in risk and ambiguity conditions were the same in our experiment design. The only difference between risk and ambiguity conditions was whether the participants knew the probability of outcome. Before the experiment, the participants were told that their final payment was decided by their performance in the experiment. That is, 35-Yuan basic payments for their participation and an additional gain or loss based on the mean of total investment outcomes, by which means we motivated the participants to make decisions carefully and effectively in all trails in order to achieve the maximum benefit.

2.4. Electroencephalography (EEG) Recording. The EEG and the EOG (electrooculograms) were recorded and preprocessed by Neuroscan Synamp2 Amplifier (Scan4.5, Neurosoft Labs, Inc., Sterling, USA), with the reference to the left mastoid. In order to keep the impedances of electrodes below 5 kOhm all electrode sites were cleaned with electrode jelly and gently abraded prior to electrode fixation. EEG and EOG were amplified with a 64 channel AC amplifier (input impedance: 10 MOhm). Vertical Electrooculogram (EOG) was recorded supra and infra-orbitally at the left eye. Horizontal electrooculograms were recorded from electrodes placed 1.5 cm lateral to the left and right external canthi. Band-pass was set to 0.05–100 Hz; the signals were digitized online at 500 Hz and stored for later analyses.

Electroencephalogram recordings were segmented for the epoch from 200 ms before appearance of the stimulus picture for decision type to 800 ms after the stimulus onset, with the first 200 ms pretargets as a baseline. Trails were contaminated by amplifier clipping, wherein bursts of electromyography activity and peak-to-peak deflection exceeding $\pm 80\,\mu$V were excluded. Finally, EEG waveforms were averaged separately for each participant, each experimental condition, and each electrode. In addition, SPSS statistical software (SPSS Inc., SPSS Inc., Chicago, Illinois, USA) was used for data statistical analyses.

3. Results

3.1. Behavioral Data. Behavior data are shown in Figure 3. A paired-sample t-test showed that there was no significant difference between the two decision types of ambiguity and risk about reaction time (RTs), $t = -0.800$, $P = 0.439 > 0.05$. In contrast to the RTs, the mean proportion of choosing to bet in ambiguous condition (57.24%, SD = 17.09%) was less than that in risky condition (68.49%, SD = 13.17%). A paired-sample t-test showed that the difference was significant ($t = -2.250$, $P = 0.044 < 0.05$). The results indicated that participants would rather bet in risky condition than bet in ambiguous condition, but it took them the same time to decide whether or not to bet in different conditions.

3.2. ERP Analyses. The component P300 was analyzed as well (Figure 4). The P300 amplitude, peaking at approximately 500 ms after stimulus onset, is mainly distributed in the center scalp areas [32, 39, 40]. Similar to the previous studies, this study selected 9 electrode sites (FC3, FCZ, FC4, C3, CZ, C4, CP3, CPZ, and CP4) for statistical analysis. Two-way repeated measure ANOVA testing across two levels of decision types and nine levels of electrodes were computed on P300 amplitude.

We measured the P300 average amplitudes in the shaded 450–550 ms time window for both risky and ambiguous conditions (in Figure 4). A 2 (decision type: risk and ambiguity) × 9 (electrode: FC3, FCZ, FC4, C3, CZ, C4, CP3, CPZ, and CP4) with subjects repeated measure ANOVA showed that there was a main effect for decision type [$F(1, 11) = 17.147$, $P = 0.002 < 0.05$]. The grand average amplitude of 9 electrodes of risky condition ($M = 5.595\,\mu$V) was significantly larger than that of ambiguous condition ($M = 4.513\,\mu$V).

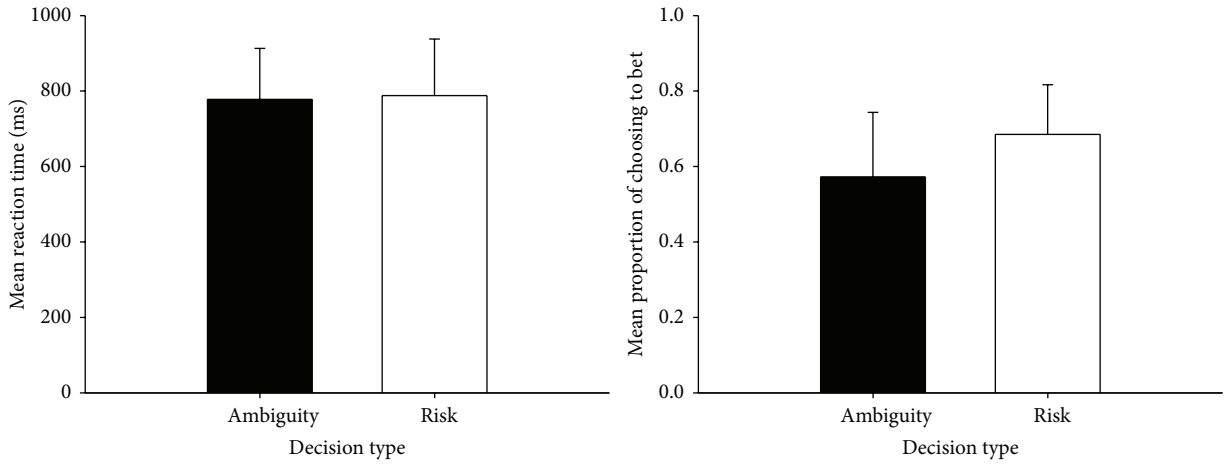


Figure 3: Behavioral data: reaction time and proportion of choosing to bet for decision under ambiguity and risk.

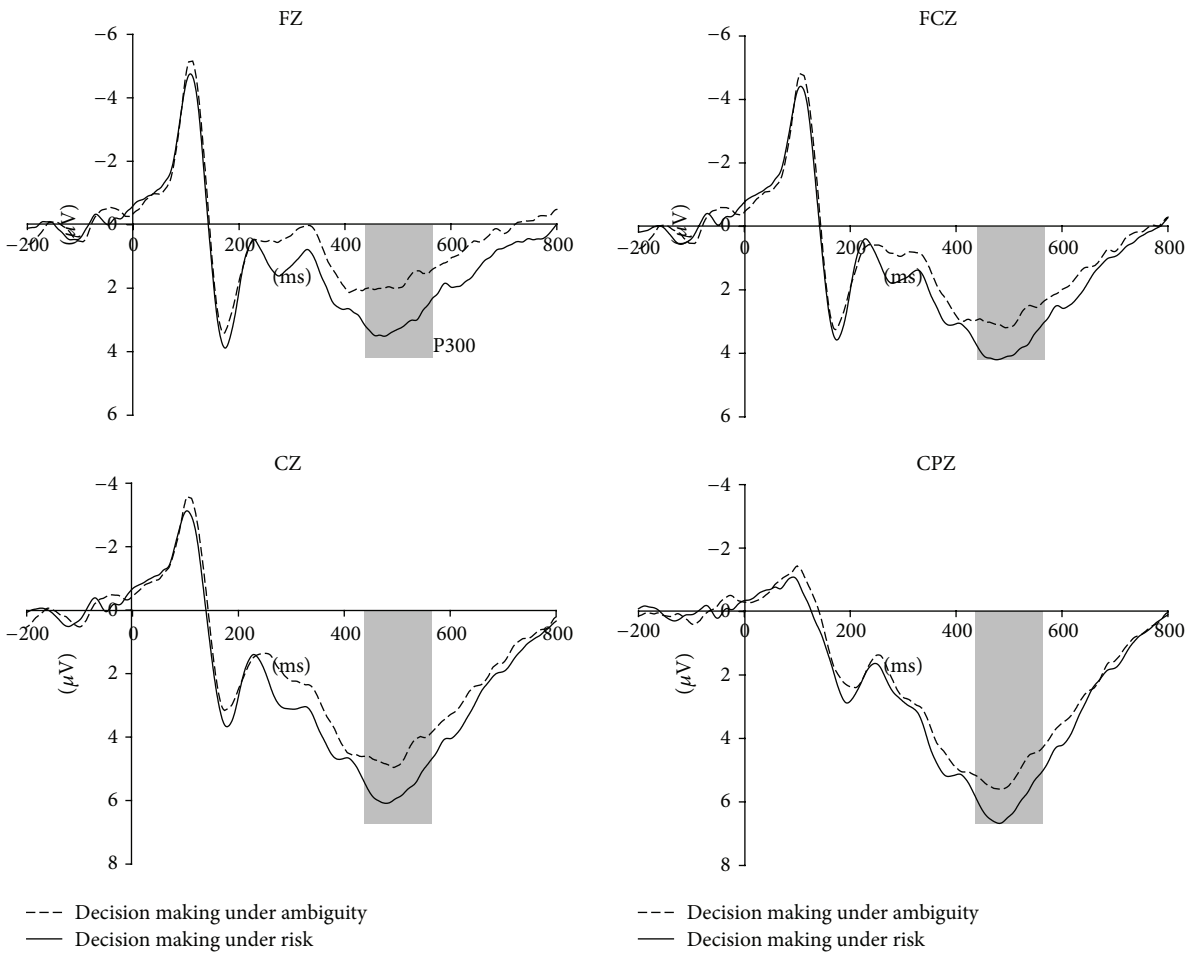


Figure 4: Grand averaged ERP waveforms for two stimulus conditions at electrode sites FZ, FCZ, CZ, and CPZ.

4. Discussion

Our behavioral data indicated ambiguity aversion with the results that the bet rate in risky condition was significantly higher than that in ambiguous condition. Besides, the ERP results showed that P300 amplitude elicited in ambiguous condition was significantly smaller than that evoked in risky condition, which revealed that higher working memory load was needed in the stage of assessment when making decision in ambiguous condition.

A large number of previous studies have demonstrated that people held different attitudes towards risk and ambiguity. According to Smith and colleagues [4], individual behavior may be affected by attitudes about payoffs (gains and losses) and beliefs about outcomes (risk and ambiguity). No matter the outcome is gain or loss, people are always ambiguity adverse. That is to say, people tend to be more adverse to ambiguity than to risk [5, 41, 42]. Thus our behavioral result showing that the proportion of subjects choosing to bet in ambiguous condition was significantly less than that in risky condition was consistent with the aforementioned studies [8, 9, 43].

P300 were elicited in both risky and ambiguous conditions, which was consistent with prior studies [22, 23]. Many studies related with decision making found that P300 was sensitive to many factors, such as the magnitude of reward [21, 44], the valence of reward [21], and interpersonal relationship in reward processing [18]. However, almost all of those studies focused on the feedback stage in which the features of reward matter. Compared with them, our experiment studied the P300 evoked in the stage of processing stimuli, which was earlier than feedback stage and did not involve the reward processing. Thus the factors influencing the modulation of P300 in our study were different from that in previous studies. Besides, we also controlled some other factors; that is, both the stimuli probabilities and the monetary magnitude presented in risky and ambiguous conditions were the same in our experiment design. The main difference between these two conditions was whether the probabilities were blind to the participants or not.

In gamble games in both risky and ambiguous conditions, people need to calculate the expectances of outcomes based on the probabilities of outcome and specific monetary value before they make a decision. In our experiment design, both of the two required pieces of information were provided in risky condition but only monetary value was given to participant in ambiguous condition. The lack of probabilities may induce more effort to recall a large amount of past practices and memory to get probability information to form a clear expectance of outcome and reduce cognitive strain in dealing with ambiguous tasks. As we know, learning from feedback plays a role in guiding decision making. Personal experience of similar situations has an effect on current decision. Positive or negative emotion induced from previous experience facilitates present information process [45]. In addition, working memory holds and manages information which exerts influence on subsequent behaviors in the short term [46]. Besides, such process of working memories appears universally in reality especially when dealing with decision making under incomplete information. According to previous studies, ambiguity can be regarded as a second-order probability distribution of option [8, 9] and people can obtain probability information from former experience. Thus, participants in ambiguous condition would learn the experience from former trails and infer the current probability, which induced a higher working memory load [38, 39]. But this process may not be expected in risky condition since the probability was definite and the outcomes were more likely to be attributed to the luck at present. Thus, we thought this difference resulted in the decrease of P300 amplitude in ambiguous condition.

Prior studies indicated that, in stimulus processing, the P300 could be considered as a representation of working memory [27–29] and it was widely demonstrated that the P300 amplitude is inversely proportional to working memory [34–36]. Thus, the lower P300 amplitude elicited in ambiguous condition revealed that participants employed higher memory load at assessment stage when they made decisions. Besides, as we know, the ambiguous decision making is relative to emotional process, and people would experience more negative emotion (such as being more worried and anxious) and be less confident, since they cannot exactly know the outcome [11]. As discussed above, people would need to trace back for evidences which required greater number of certainty cues in order to neutralize this negative feeling and make themselves confident under ambiguous conditions [47, 48]. They would not only pay attention to current stimuli (i.e., monetary value) as they did in risky condition, but also recall the past experience about the gamble to calculate a more clear expectance of outcome [49]. All in all, due to the integrality of information, it is much easier for decision maker to calculate expected value in risky condition than in ambiguous condition. Participants focus mainly on calculating explicit information for logical strategies in risky condition while they are more likely to mobilize past experience to figure out the possible probability [50] to reduce cognitive stress with insufficient information in ambiguous condition. This additional effort and difficulty resulted in a high working memory load with a lower P300 elicited. Therefore, our results of lower P300 amplitude in ambiguous condition than in risky one indicated the different cognitive mechanism between ambiguity and risk.

5. Conclusion

In this study, we used ERPs to clarify and extend the current understanding of decision making under risk and ambiguity. Particularly, our findings suggested that the P300 amplitude can be applied to reflect information processing at the early stage (the stage of assessment) of decision making. In our research, the P300 amplitude elicited by gambling in ambiguous condition was significantly smaller than that evoked in risky condition, showing that participants met with higher working memory under ambiguity to mobilize past experience to calculate the expected value and reduce cognitive strain than under risk. Furthermore, our behavioral result validated ambiguity aversion phenomenon with the finding of a lower bet rate in ambiguous condition.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the Nature Science Foundation of China under Contracts no. 71471163 and no. 71071135.

References

[1] F. H. Knight, *Risk, Uncertainly and Profit*, Library of Economics and Liberty, 1921.

[2] A. Bechara, "The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage," *Brain and Cognition*, vol. 55, no. 1, pp. 30–40, 2004.

[3] S. A. Huettel, C. J. Stowe, E. M. Gordon, B. T. Warner, and M. L. Platt, "Neural signatures of economic preferences for risk and ambiguity," *Neuron*, vol. 49, no. 5, pp. 765–775, 2006.

[4] K. Smith, J. Dickhaut, K. McCabe, and J. V. Pardo, "Neuronal substrates for choice under ambiguity, risk, gains, and losses," *Management Science*, vol. 48, no. 6, pp. 711–718, 2002.

[5] C. Heath and A. Tversky, "Preference and belief: ambiguity and competence in choice under uncertainty," *Journal of Risk and Uncertainty*, vol. 4, no. 1, pp. 5–28, 1991.

[6] C. R. Fox and A. Tversky, "A belief-based account of decision under uncertainty," *Management Science*, vol. 44, no. 7, pp. 879–895, 1998.

[7] C. Rode, L. Cosmides, W. Hell, and J. Tooby, "When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory," *Cognition*, vol. 72, no. 3, pp. 269–304, 1999.

[8] M. H. Freeston, J. Rhéaume, H. Letarte, M. J. Dugas, and R. Ladouceur, "Why do people worry?" *Personality and Individual Differences*, vol. 17, no. 6, pp. 791–802, 1994.

[9] H. J. Einhorn and R. M. Hogarth, *Decision Making under Ambiguity: A Note*, Springer, 1988.

[10] D. Frisch and J. Baron, "Ambiguity and rationality," *Journal of Behavioral Decision Making*, vol. 1, no. 3, pp. 149–157, 1988.

[11] C. R. Fox and M. Weber, "Ambiguity aversion, comparative ignorance, and decision context," *Organizational Behavior and Human Decision Processes*, vol. 88, no. 1, pp. 476–498, 2002.

[12] J. Baron and J. C. Hershey, "Outcome bias in decision evaluation," *Journal of Personality and Social Psychology*, vol. 54, no. 4, pp. 569–579, 1988.

[13] D. R. Bach, B. Seymour, and R. J. Dolan, "Neural activity associated with the passive prediction of ambiguity and risk for aversive events," *The Journal of Neuroscience*, vol. 29, no. 6, pp. 1648–1656, 2009.

[14] M. Hsu, M. Bhatt, R. Adolphs, D. Tranel, and C. F. Camerer, "Neural systems responding to degrees of uncertainty in human decision-making," *Science*, vol. 310, no. 5754, pp. 1680–1683, 2005.

[15] I. Levy, J. Snell, A. J. Nelson, A. Rustichini, and P. W. Glimcher, "Neural representation of subjective value under risk and ambiguity," *Journal of Neurophysiology*, vol. 103, no. 2, pp. 1036–1047, 2010.

[16] G. Hajcak, J. S. Moser, C. B. Holroyd, and R. F. Simons, "The feedback-related negativity reflects the binary evaluation of good versus bad outcomes," *Biological Psychology*, vol. 71, no. 2, pp. 148–154, 2006.

[17] A. Sato, A. Yasuda, H. Ohira et al., "Effects of value and reward magnitude on feedback negativity and P300," *NeuroReport*, vol. 16, no. 4, pp. 407–411, 2005.

[18] Y. Wu and X. Zhou, "The P300 and reward valence, magnitude, and expectancy in outcome evaluation," *Brain Research*, vol. 1286, pp. 114–122, 2009.

[19] N. Yeung and A. G. Sanfey, "Independent coding of reward magnitude and valence in the human brain," *Journal of Neuroscience*, vol. 24, no. 28, pp. 6258–6264, 2004.

[20] D. Zhang, R. Gu, L. S. Broster et al., "Linking brain electrical signals elicited by current outcomes with future risk decision-making," *Frontiers in Behavioral Neuroscience*, vol. 8, article 84, 2014.

[21] Z. Zhou, R. Yu, and X. Zhou, "To do or not to do? Action enlarges the FRN and P300 effects in outcome evaluation," *Neuropsychologia*, vol. 48, no. 12, pp. 3606–3613, 2010.

[22] Q. Xu, Q. Shen, P. Chen, Q. Ma, D. Sun, and Y. Pan, "How an uncertain cue modulates subsequent monetary outcome evaluation: an ERP study," *Neuroscience Letters*, vol. 505, no. 2, pp. 200–204, 2011.

[23] R. Gu, Y. Ge, Y. Jiang, and Y.-J. Luo, "Anxiety and outcome evaluation: the good, the bad and the ambiguous," *Biological Psychology*, vol. 85, no. 2, pp. 200–206, 2010.

[24] C. Bellebaum, D. Polezzi, and I. Daum, "It is less than you expected: the feedback-related negativity reflects violations of reward magnitude expectations," *Neuropsychologia*, vol. 48, no. 11, pp. 3343–3350, 2010.

[25] S. Chandra, S. C. Srivastava, and Y.-L. Theng, "Evaluating the role of trust in consumer adoption of mobile payment systems: an empirical analysis," *Communications of the Association for Information Systems*, vol. 27, no. 1, pp. 561–588, 2010.

[26] R. San Martín, "Event-related potential studies of outcome processing and feedback-guided learning," *Frontiers in Human Neuroscience*, vol. 6, p. 304, 2012.

[27] D. Karis, M. Fabiani, and E. Donchin, "'P300' and memory: individual differences in the von Restorff effect," *Cognitive Psychology*, vol. 16, no. 2, pp. 177–216, 1984.

[28] A. Azizian and J. Polich, "Evidence for attentional gradient in the serial position memory curve from event-related potentials," *Journal of Cognitive Neuroscience*, vol. 19, no. 12, pp. 2071–2081, 2007.

[29] J. Polich, "Updating P300: an integrative theory of P3a and P3b," *Clinical Neurophysiology*, vol. 118, no. 10, pp. 2128–2148, 2007.

[30] S. Lei and M. Roetting, "Influence of task combination on EEG spectrum modulation for driver workload estimation," *Human Factors*, vol. 53, no. 2, pp. 168–179, 2011.

[31] E. Donchin, A. Kramer, and C. Wickens, "Applications of brain event-related potentials to problems in engineering psychology," in *Psychophysiology: Systems, Processes, and Applications*, pp. 702–718, Guilford, 1986.

[32] B. Palmer, V. T. Nasman, and G. F. Wilson, "Task decision difficulty: effects on ERPs in a same-different letter classification task," *Biological Psychology*, vol. 38, no. 2-3, pp. 199–214, 1994.

[33] B. K. Bonala and B. H. Jansen, "A computational model for generation of the P300 evoked potential component," *Journal of Integrative Neuroscience*, vol. 11, no. 3, pp. 277–294, 2012.

[34] A. Kok, "Event-related-potential (ERP) reflections of mental resources: a review and synthesis," *Biological Psychology*, vol. 45, no. 1–3, pp. 19–56, 1997.

[35] A. Kramer, W. Schneider, A. Fisk, and E. Donchin, "The effects of practice and task structure on components of the event-related brain potential," *Psychophysiology*, vol. 23, no. 1, pp. 33–47, 1986.

[36] A. Mecklinger, A. F. Kramer, and D. L. Strayer, "Event related potentials and EEG components in a semantic memory search task," *Psychophysiology*, vol. 29, no. 1, pp. 104–119, 1992.

[37] A. Kok, "On the utility of P3 amplitude as a measure of processing capacity," *Psychophysiology*, vol. 38, no. 3, pp. 557–577, 2001.

[38] J. Polich, "Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli," *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 68, no. 4, pp. 311–320, 1987.

[39] T. R. H. Cutmore and T. D. Muckert, "Event-related potentials can reveal differences between two decision-making groups," *Biological Psychology*, vol. 47, no. 2, pp. 159–179, 1998.

[40] J. Yuan, Q. Zhang, A. Chen et al., "Are we sensitive to valence differences in emotionally negative stimuli? Electrophysiological evidence from an ERP study," *Neuropsychologia*, vol. 45, no. 12, pp. 2764–2771, 2007.

[41] K. R. MacCrimmon and S. Larsson, "Utility theory: axioms versus paradoxes," in *Expected Utility Hypotheses and the Allais Paradox*, vol. 21, pp. 333–409, Springer, 1979.

[42] S. P. Curley, J. F. Yates, and R. A. Abrams, "Psychological sources of ambiguity avoidance," *Organizational Behavior and Human Decision Processes*, vol. 38, no. 2, pp. 230–256, 1986.

[43] D. Ellsberg, "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics*, vol. 75, no. 4, pp. 643–669, 1961.

[44] H. Nittono, Y. Otsuka, and P. Ullsperger, "Asymmetrical effects of frequent gains and frequent losses in a gambling task," *Neuroreport*, vol. 19, no. 13, pp. 1345–1349, 2008.

[45] S. M. Carpenter, E. Peters, D. Västfjäll, and A. M. Isen, "Positive feelings facilitate working memory and complex decision making among older adults," *Cognition and Emotion*, vol. 27, no. 1, pp. 184–192, 2013.

[46] J. M. Fuster, "Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory," *Journal of Neurophysiology*, vol. 36, no. 1, pp. 61–78, 1973.

[47] M. J. Dugas, M. H. Freeston, and R. Ladouceur, "Intolerance of uncertainty and problem orientation in worry," *Cognitive Therapy and Research*, vol. 21, no. 6, pp. 593–606, 1997.

[48] A. L. Krain, S. Hefton, D. S. Pine et al., "An fMRI examination of developmental differences in the neural correlates of uncertainty and decision-making," *Journal of Child Psychology and Psychiatry*, vol. 47, no. 10, pp. 1023–1030, 2006.

[49] M. Brand, F. Grabenhorst, K. Starcke, M. M. P. Vandekerckhove, and H. J. Markowitsch, "Role of the amygdala in decisions under ambiguity and decisions under risk: evidence from patients with Urbach-Wiethe disease," *Neuropsychologia*, vol. 45, no. 6, pp. 1305–1317, 2007.

[50] A. Achtziger, C. Alós-Ferrer, S. Hügelschäfer, and M. Steinhauser, "The neural basis of belief updating and rational decision making," *Social Cognitive and Affective Neuroscience*, vol. 9, no. 1, Article ID nss099, pp. 55–62, 2014.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 345160, 10 pages
http://dx.doi.org/10.1155/2015/345160

Research Article

An Opinion Interactive Model Based on Individual Persuasiveness

Xin Zhou, Bin Chen, Liang Liu, Liang Ma, and Xiaogang Qiu

School of Information System and Management, National University of Defense Technology, Changsha 410073, China

Correspondence should be addressed to Xin Zhou; zhouxinnudt@163.com

Received 3 February 2015; Revised 13 April 2015; Accepted 15 April 2015

Academic Editor: Ye-Sho Chen

In order to study the formation process of group opinion in real life, we put forward a new opinion interactive model based on Deffuant model and its improved models in this paper because current models of opinion dynamics lack considering individual persuasiveness. Our model has following advantages: firstly persuasiveness is added to individual's attributes reflecting the importance of persuasiveness, which means that all the individuals are different from others; secondly probability is introduced in the course of interaction which simulates the uncertainty of interaction. In Monte Carlo simulation experiments, sensitivity analysis including the influence of randomness, initial persuasiveness distribution, and number of individuals is studied at first; what comes next is that the range of common opinion based on the initial persuasiveness distribution can be predicted. Simulation experiment results show that when the initial values of agents are fixed, no matter how many times independently replicated experiments, the common opinion will converge at a certain point; however the number of iterations will not always be the same; the range of common opinion can be predicted when initial distribution of opinion and persuasiveness are given. As a result, this model can reflect and interpret some phenomena of opinion interaction in realistic society.

1. Introduction

In the last decades, there are fruitful research achievements in learning behaviors of individual and complex group phenomena. There is little doubt that the behavior of individual and group phenomenon is inextricably linked. Group consists of lots of individuals, while common behaviors of plenty of individuals constitute many macroscopic emergences. Hence, it is an effective way to study the group phenomenon based on individual's behavior. Since the human society can also be considered as a complex multicomponent system consisting of individuals interacting with themselves and with their material environment, it is a challenge to develop a strategy allowing of a general quantitative modeling procedure for collective dynamic macro processes in the society [1]. In social dynamic, there is a premise that in social phenomena the basic constituents are not particles but humans and every individual interact with a limited number of peers, usually negligible compared to the total number of people in the system. So many macroscopic phenomena naturally call for a statistical physics approach to social behavior, which means the attempt to understand regularities at large scale as collective effects of the interaction among single individuals, considered as relatively simple entities [2].

In the research of cognitive learning behavior of individual, Brenner [3] sums up individual's cognitive behaviors as nonconscious learning [4, 5], routine-based learning [6], and belief learning [7, 8] according to the strength of the individual consciousness. However, in the research of complex emergence phenomenon, many researchers have put forward a number of models according to social dynamic, which derives from statistic physics, such as opinion evolution [9, 10], cultural dissemination [11], disease transmission [12, 13], and the spread of rumor [14, 15]. Among them, opinion dynamics has been a hot research field, which reflects and interprets a wide range of social phenomena ranging from collective decision making, finding and not finding of consensus, emergence of political parties, minority opinion survival, emergence of extremism, and so on [16]. In this paper, an opinion emergence phenomenon is studied based on opinion dynamics.

So far, models of opinion dynamics provided by scholars can depict the opinion interactive situation well in realistic society in some special situation. Based on the form of opinion, opinion dynamics can be divided into discrete model and continuous model. In discrete model, opinion is a discrete numerical variable which includes Voter Model [17], Sznajd Model [18], and Galam Model [19]. The idea of Voter Model is that an agent may be influenced by a neighbor so as to change its voting choice or opinion to the neighbor's and such local influences give rise to a global process of the collective voting results of the whole population. Sznajd Model states that one is easier to be persuaded by two or more people who sharing the same opinion than by a single person. Galam Model depicts that group consent is formed by the process of minority subordinating to the majority. In continuous model, opinion is a continuous variable which includes Deffuant model [20] and HK Model [21] mainly. Common places of these two models are that two individuals will not exchange their opinions until the difference of their opinions is below a threshold. However, in Deffuant model individual can only interact with one individual in each step while it can interact with many individuals in HK Model. Besides, there are a lot of expended models based on these two basic models.

In the early research, individual is homogeneous. With the deepening of researching, many scholars add the heterogeneous attribute to individuals, such as softhead individuals, amiable individuals, bigoted individuals, stubborn individuals, opinion leaders, and authoritative individuals [22]. In the past, the major networks of group are one-dimensional ring, grid, regular network, and fully connected network [22]. However, complex network, especially the discovery of small world network and scale-free network, injects new life into the research of opinion dynamics. The opinion dynamics in complex network [23, 24] and coevolution of opinion in self-adaptive network also become a hot topic gradually [25, 26].

This paper puts forward a new opinion interactive model based on Deffuant models and its improved models. This model emphasizes the importance of individual's persuasiveness and simulates the variation trend of group opinion. With this aim, the remainder of this paper is organized as follows. Previous work about Deffuant model is introduced in Section 2. A new model is then given in Section 3. After that, through simulation experiments, three problems in Section 4 are solved: the influence of randomness, initial persuasiveness distribution of group, and number of individuals on the common opinion, the influence of randomness, initial persuasiveness distribution of group, and number of individuals on the iteration of experiment, and the prediction of the interval of opinion common based on the initial distribution persuasiveness of group. Finally, in Section 5, we have a discussion about implication of our model.

2. Previous Works on Deffuant Model

The main idea of Deffuant model [20] is that, considering a population of N agents with continuous opinions O, at each time step any two agents are randomly chosen to meet. They readjust their opinion when their difference of opinion is smaller in magnitude than a threshold d. Suppose that the two agents have opinion O_i and O_j and that $|O_i - O_j| < d$, i and j represent the ith and the jth individual, respectively; opinions are then adjusted according to

$$\begin{aligned} O_i &= O_i + \mu \cdot \left(O_j - O_i\right), \\ O_j &= O_j + \mu \cdot \left(O_i - O_j\right), \end{aligned} \tag{1}$$

where μ is the convergent parameter taken between 0 and 0.5 during the simulations. The rationale for the threshold condition is that agents only interact when their opinion are already close enough; otherwise they do not even bother to discuss.

Honestly speaking, there are plenty of models after proposing of Deffuant model, especially some heterogeneous models. Lorenz [27] studies heterogeneous bounds of confidence, where two kinds of individuals, namely, open-minded and closed-minded individuals, are studied in the paper. The difference between open-minded and closed-minded individual is that they possess different d. In addition, extremism individual is studied by Weisbuch et al. [28]. The extremism model is based on two more assumptions: a few extremists with extreme opinions at the ends of the opinion spectrum and with very low threshold for interaction are introduced; whenever the threshold allows interaction, both opinions and threshold are readjusted according to similar expressions. That means in extremism model, the threshold d will change with the interaction; in other words, the more "tolerant" agent (with larger d) can be influenced by the less tolerant (with smaller d) while the less tolerant agent is not. What is more, truth seekers are discussed by Hegselmann and Krause [29] and Malarz [30]. In true seekers model, two parameters are added into Deffuant model: $T \in [0, 1]$ and α_i which represent the true opinion and the strength of the attraction to the truth for ith agent, respectively.

However, all the improved Deffuant models can only deal with one aspect problem. Most of them focus on the threshold d, while they seldom take the parameter μ into consideration. What is more, most of them divided group into several categories, like open-minded individual and close-minded individual. Nonetheless, it is true in realistic society that all the individuals are different from others. So, this paper analyzes a case that different individuals own different μ, namely, persuasiveness.

3. Model

Suppose a scene that a group of people discuss a topic. Everyone has its own attitude and interest to a certain topic. However people cannot only insist on their own opinion, because as an individual of society, it should take other people's opinion into account. After discussion, the group should reach an agreement. This scene often appears in reality society, such as the discussion of some topics in Congress and the discussion of entertainment place where to go in the weekend among office members.

This paper simulates the scene through modeling and experiments. Any two individuals chosen randomly in group can exchange their opinions. Opinions are published according to the round, and opinions of current round are only

affected by opinions of last round. In the process of interaction, the sequence of speech of individual is ignored which means that the speech of the whole group is parallel. Individual acquires the last round opinion of itself and the opposite individual's opinion and then calculates the current round opinion of itself in line with some behavior rules and after that publishes its new opinion in the next round. With the evolution of group opinion, it forms a series of polymerization of macroopinion cluster eventually.

In this paper, the model is introduced in the form of agent. Individuals are abstracted as agents while each group consists of a lot of agents. Every agent has the same attributes and behavior rules, while they may be different from concrete attribute values or concrete behaviors. Through the interaction of microscopic interactive process, some macroscopic emergence phenomena of group can be found.

3.1. Attributes of Agent. Attributes of agent are abstracted as index, opinion, and persuasiveness. Explanation to these attributes is as follows.

3.1.1. Index. To distinguish from other agents, every agent has a unique identity.

3.1.2. Opinion. To a certain topic in the discussion, every agent has its own opinion. Opinion is the position or attitude that one observes things. In the mathematic model, opinion is abstracted as a discrete variable or continuous variable. Although it is maybe too simple to model the complex human society, it has some advantages to some certain problems, such as "turning left or turning right," "approve or disapprove," and "going to classroom or going to library." Due to the variety of opinions to a certain topic and for the convenience of analysis, the opinion is modeled by continuous interval between 0 to 1 referring to Deffuant model. Among it, 0 and 1 represent the opposite opinion and 0.5 represents neutrality that opinion values have different meanings in different cases.

3.1.3. Persuasiveness. Persuasiveness is the ability to persuade the other individual. In order to protect its own interest, everyone wants to persuade others and make them accept its opinion in the discussion. However, the status is often unequal. Some people are of higher qualification, elder age, or an opinion leader in some fields. So their opinions have a guiding role to others and their persuasiveness is higher than others. Most people have similar knowledge to a certain topic, and they do not have a deep research in it. So their opinions are just reference to others and their persuasiveness is lower. Drawing lessons from the mathematic model of opinion, the persuasiveness is modeled by continuous interval ranging from 0 to 1. Among it, 0 represents the lowest persuasion, 1 represents the highest persuasion, and 0.5 represents that individual has their own opinions to a topic but does not have an insight into it.

3.2. The Opinion Updating Rule of Agent. The paper does not take topology of network into consideration in this model. What we focus on is the influence of individual's persuasiveness and interactive probability on the process of interaction. In real life, if the difference between two individuals is low, the probability of thorough interaction between them may be high. However, it cannot be guaranteed that two sides will interact, because there are many other factors influencing interacting such as personal character. If the difference between two sides is large, the probability of thorough interaction is small. Similar to the former, it cannot guarantee that two sides will not interact of which the probability of interaction is just low. As to persuasiveness, an individual with high persuasiveness can change other's opinion easier than that with low persuasiveness.

Suppose that there are N individuals in finite set $A = \{1, 2, \ldots, N\}$, forming $N \times 1$ opinion column vector $\{O_1(t), O_2(t), \ldots, O_N(t)\}$, where $O_i(t)$ is the ith individual's opinion in the t round, $O_i(t) \in [0, 1]$, $i \in A$, $t \geq 0$, where $t = 0$ is the initial opinion. Similarly, there are $N \times 1$ persuasiveness column vector $\{P_1, P_2, \ldots, P_N\}$, where P_i is the ith individual's persuasiveness, $P_i \in [0, 1]$, $i \in A$. In our model, the persuasiveness of each individual is fixed and will not change with the time. For the convenience of description, the agent whose index is i is used to be the object of study, which is denoted as Agent i. With the above ideas, the opinion updating rules of Agent i are described as follows:

(1) each time Agent i randomly chooses an agent as partner, denoting the agent as Agent j;

(2) if $(p < (1 - |O_i(t) - O_j(t)|))$, then enter into step (3); else exiting;

(3)

$$O_i(t+1) = O_i(t) + P_j \cdot \left(O_j(t) - O_i(t)\right). \quad (2)$$

Among formula (2), p is the random probability which means the probability of interaction, $p \in [0, 1]$, and obeys uniform distribution.

This model draws the lessons from Deffuant model and some other homogeneous Deffaunt models. Honestly speaking, if every individual possess the same and fixed parameters p and P_i, there are no difference between our model and Deffuant model. However, when these two parameters are difference from person to person, many new issues should be solved. Figure 1 shows the classical opinion evolution processes of Deffuant model.

From Figure 1 we can know that, with the increasing of the iteration, the group opinion gradually converges, forming the stable opinion cluster and reaching an agreement. Because the relationship network of the group is full connected graph, so the result of the model is influenced by randomness, initial opinion distribution, and persuasiveness of individual.

4. Simulation

This serial of experiments set a certain scale of agents, which does not take the topology of network into consideration. When the absolute difference of two opinions is less than 0.001, we assume that their opinions are the same.

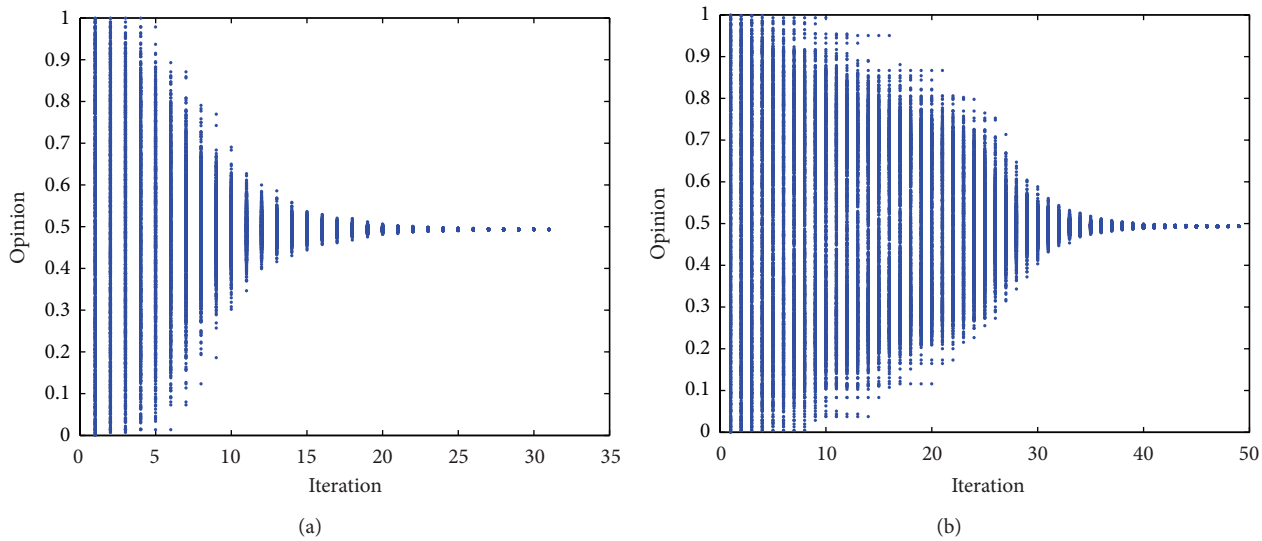


Figure 1: Classical opinion evolution processes (initial opinion distributions of group obey the uniform distribution ranging from 0 to 1); number of individuals is 1000. The horizontal ordinate represents the iteration number, and vertical ordinate represents individual opinion. (a) $p = 0.5$, $P_i = 0.3$; (b) $p = 0.3$, $P_i = 0.3$, $i \in A$.

The simulation will come to an end when all the agents have the same opinion. The experiment flow is as follows:

(1) assign initial values to all the agents;

(2) in each time step, every agent can randomly choose one agent to exchange opinion. If the total number of individuals is odd, an agent will miss a turn in this round of interaction that it will not exchange its opinion with others;

(3) all the agents change their opinions based on the opinion updating rule in section three except the one who misses a turn;

(4) repeating step (2) and step (3), and if the whole group reaches an agreement, the simulation experiment stops.

In addition, for the convenience of analysis of the evolution character in macroscopic aspect, two indexes are defined here. Number of iterations (NOI) is the total time steps when all the agents have the same opinion, which indicates the convergent speed of group opinion. Common opinion (CO) is the opinion of group when all the agents have the same opinion, which determines variation trend of group opinion.

In the opinion updating model, there are two factors influencing CO and NOI. One factor is initial distribution of persuasiveness and opinion and the corresponding relationship between them. The other factor is the probability in the interaction. It is liable that the group may obey different kinds of distributions. However, for the convenience of analysis, we assume three persuasiveness distributions, namely, normal distribution, exponent distribution, and uniform distribution. Three experiments have been done in this paper. Experiment 4.1 and Experiment 4.2 are based on the same initial condition. The difference between them is that Experiment 4.1 focuses on the analysis of factors (influence of randomness, initial distribution, and number of individuals) influencing CO, while Experiment 4.2 stresses factors (influence of randomness, initial distribution, and number of individuals) influencing the NOI. Experiment 4.3 takes two initial distributions (normal distribution and exponential distribution) into account to predict the range of CO.

4.1. Influence of Randomness, Initial Distribution, and Number of Individuals on CO. In this experiment, we discuss three group experiments and each group experiment is also divided into three experiments. These three group experiments have the same initial opinion distribution obeying uniform distribution. The difference among them is that initial persuasiveness distribution of the first group experiment follows normal distribution, which can be denoted as "Nor," initial persuasiveness distribution of the second group experiment follows uniform distribution which is denoted as "Unif," and the remained group experiment follows the exponent distribution which can be represented as "Exp." Each group experiment is also divided into three experiments, which are distinguished by the number of individuals. Each group experiment, respectively, has 100 agents, 1000 agents, and 10000 agents. So there are nine experiments in total; three experiments of "Nor" can be denoted as "Nor100," "Nor1000," and "Nor10000," respectively; three experiments of "Unif" are represented as "Unif100," "Unif1000," and "Unif10000," respectively; three experiments of "Exp" are denoted as "Exp100," "Exp1000," and "Exp10000," respectively. Each experiment is independently repeatedly calculated for 100 times.

Random number generator separately generates initial opinion distribution and initial persuasiveness distribution of group based on above requirements. Suppose that the initial opinion distribution of all the experiments obeys the uniform distribution ranging from 0 to 1. As for "Nor100," "Nor1000,"

Table 1: COs of nine experiments.

Group name	Nor100	Nor1000	Nor10000
Mean value	0.5818	0.5978	0.5838
Deviation	$7.7987 * 10^{-6}$	$1.5884 * 10^{-6}$	$2.7965 * 10^{-6}$
Group name	Unif100	Unif1000	Unif10000
Mean value	0.6353	0.6643	0.6656
Deviation	$1.1000 * 10^{-5}$	$2.5195 * 10^{-6}$	$4.9222 * 10^{-6}$
Group name	Exp100	Exp1000	Exp10000
Mean value	0.7272	0.7482	0.7500
Deviation	$1.1030 * 10^{-5}$	$3.0887 * 10^{-6}$	$5.6305 * 10^{-6}$

and "Nor10000," the initial persuasiveness of group follows normal distribution of which the mean value μ is 0.5 and standard deviation σ is 0.1667. With regard to "Exp100," "Exp1000," and "Exp10000," the initial persuasiveness of group obeys exponent distribution that the parameter λ is equal to 5. As to "Unif100," "Unif1000," and "Unif10000," the initial persuasiveness of group follows uniform distribution of which the low bound is 0 and high bound is 1. If the generated persuasiveness is bigger than 0.99 or smaller than 0.01, it will be generated again. Here an extreme situation is taken into consideration; that initial opinion distribution and initial persuasiveness distribution are one-to-one correspondence from small to large in all these nine experiments, which can be labeled as positive sequence correspondence. Reasons for considering the situation is that on one hand it reflects some phenomena in real life, such as one proposal satisfies the desire of high persuasive people while it is not in line with the interest of low persuasive people; on the other hand it can deduce general situations.

Mean values and standard deviations of COs of nine experiments are shown in Table 1.

Based on Table 1, three phenomena can be found. (1) From each experiment, all the independently repeated experiments of "Nor100" converge at 0.5818; all the independently repeated experiments of "Nor1000" converge at 0.5978; other experiments similar to "Nor100" and "Nor1000" converge at different fixed values. (2) Classifying experiments based on the distribution, no matter how many agents they are, COs of "Nor100," "Nor1000," and "Nor10000" are all around 0.58; COs of "Unif100," "Unif1000," and "Unif10000" are all around 0.66; COs of "Exp100," "Exp1000," and "Exp10000" are all around 0.74. (3) Classifying experiments according to number of individuals, the CO of "Unif" is about 0.8 larger than "Nor" and that of "Exp" is about 0.8 larger than "Unif" when number of individuals is the same.

Without formula derivation, some conclusions can be drawn through these experiment results. (1) CO will converge at a fixed point when all the initial values of agents are fixed. From the perspective of experiment results, all COs of one group experiment are the same. From the perspective of mathematic, the order of magnitude of standard deviation is almost near $1.0 * 10^{-5}$ in each group experiment in Table 1. The experiment precondition has supposed that two agent will reach an agreement when the difference of their opinions is less than $1.0 * 10^{-3}$. The deviation can only influence the opinion value at 10^{-4} based on 3σ principle, so it will not change the opinion value at 10^{-3}. In this way, the opinion will converge at a fixed point. (2) If initial persuasiveness of group obeys the same distribution and parameter, the number of individuals has little influence on CO. (3) The initial persuasiveness distribution has a big impact on CO, and different initial persuasiveness distributions leads to different COs. (4) Randomness has no effect on CO. According to the opinion updating rule, there are two random processes in the interaction, one process is two agents are randomly chosen to interact, and the other one is whether exchanging their opinions or not is based on probability. However, all the COs of 100 independently repeated experiments are the same in each group experiment. So it makes no difference to the CO based on experiment results.

4.2. Influence of Randomness, Initial Distribution, and Number of Individuals on NOI. The initial condition of this experiment is the same with Experiment 4.1. In this experiment, the influence of randomness, initial distribution of agent's attributes, and the number of individuals on NOI are discussed. Figure 2 shows the NOI of nine experiments. Among it, Figures 2(a), 2(b), and 2(c) show the difference of NOI when they have the same initial distribution but different numbers of individuals; Figures 2(d), 2(e), and 2(f) show the difference of NOI when they have the same number of individuals but different initial distributions. In these figures, horizontal coordinate represents different experiments, and vertical coordinate represents the mean value and deviation of NOI.

Many phenomena can be found from Figure 2. (1) From Figures 2(a), 2(b), and 2(c), the mean value of NOI of three distributions increases with the rising of number of individuals. Among them, the rising speed of mean value of "Exp" is the most obvious, where the mean value of "Exp10000" is about 15 larger than "Exp100"; the second is "Unif" where the mean value of "Unif10000" is about 4 larger than "Unif100"; and the least is "Nor" where the mean value of "Nor10000" is about 2 larger than "Nor100." (2) From Figures 2(d), 2(e), and 2(f), the mean value of NOI of "Exp" is largest, and that of "Nor" is smallest when in the same number of individuals.

Without formula derivation, three conclusions can be acquired based on simulation experiment results. (1) Randomness is the basic cause of differences of NOI. The NOI is the same without randomness in the condition of the same initial value. (2) The more the initial distribution of persuasiveness is concentrated around 0.5, the less the NOI will be. From formula (1) we can deduce that it just needs one interaction that two sides of the interaction reach an agreement when their persuasiveness is 0.5. If their persuasiveness is far away from 0.5, it needs more than once to come to an agreement. Experiment results demonstrate it. In "Exp100," "Exp1000," and "Exp10000," the initial persuasiveness distribution obeys exponent distribution of which the number of low persuasiveness individuals is large and the number of high persuasiveness individuals is small. In "Nor100," "Nor1000," and "Nor10000," initial persuasiveness distribution follows normal distribution of which mean value is 0.5 and deviation is 0.1667. The NOI of "Exp" is much higher

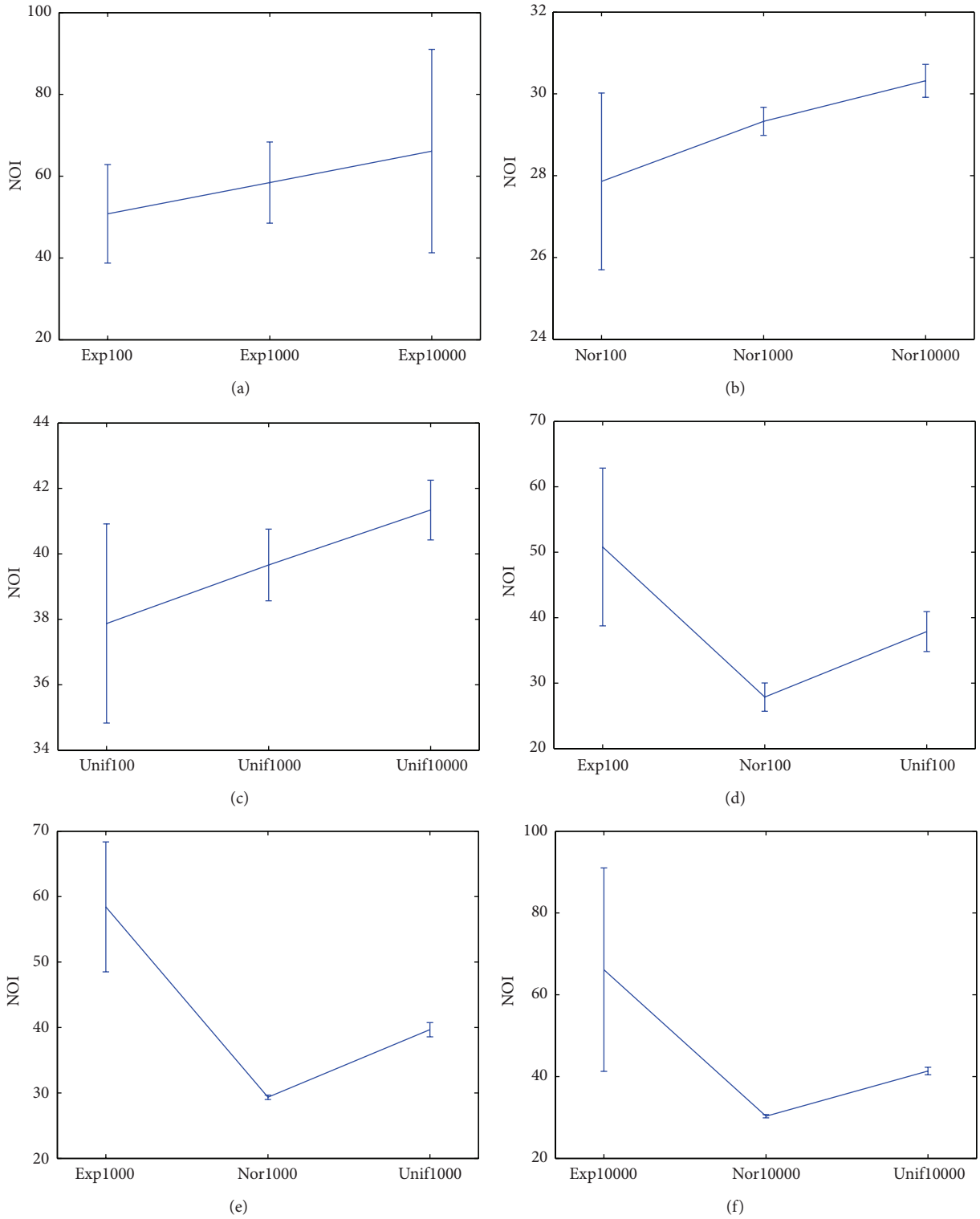


Figure 2: NOIs of nine experiments.

than that of "Nor" in the same number of agents. (3) With the increasing of number of agents, the NOI increases gradually. However, the speed of increasing depends on the initial persuasiveness distribution.

4.3. Two Predicted Experiments. In this experiment, a situation is considered that initial opinion distribution and initial persuasiveness distribution of group are known but we do not know the corresponding relationship between them. The CO

cannot be acquired because the value of opinion and persuasiveness of every agent cannot be acquired in advance. Calculating the boundary condition is a proper solution which can decrease the degree of error of predicting CO. There are two situations of discussion in real life: one scene is that to a topic most people have their own opinion while only a few people have in-depth knowledge or have no idea of it; the other scene is that most people do not have idea of it while only a few people have an insight into the topic. Normal distribution and exponent distribution can substitute for these two scenes. Through simulation experiment, the range of CO in the light of initial persuasiveness distribution and initial opinion distribution can be predicted.

4.3.1. Normal Distribution. The range of CO is based on two boundaries: one boundary is corresponding of positive sequence between initial persuasiveness distribution and initial opinion distribution, and the other boundary is corresponding of negative sequence between them. Opinion distribution and persuasiveness distribution are one-to-one correspondence from small to large in positive sequence corresponding while they are one-to-one correspondence that one distribution is from small to large and the other distribution is from large to small in negative sequence corresponding.

Positive sequence corresponding is taken into account at first. Suppose that the initial persuasiveness of group obeys normal distribution and initial opinion of group obeys the uniform distribution ranging from 0 to 1. The number of individuals is 10000. Because the range of persuasiveness is from 0 to 1, in order to confine the persuasiveness within the boundary it will generate persuasiveness again if persuasiveness is bigger than 0.99 or smaller than 0.01. The normal distribution has two parameters (deviation and mean value). A series of discrete values of standard deviation which are 0.02, 0.04, 0.06, 0.08, 0.1, 0.12, and 0.14 separately and discrete values of mean value which are 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, and 0.7 separately are picked up. Figure 3 shows experiment results. In Figure 3(a), horizontal ordinate represents the mean value, vertical ordinate represents the CO, and a series of different colors and shapes represent different standard deviations. In Figure 3(b), "miu" is the mean value, "sigma" is standard deviation, and CO is common opinion of group.

From Figure 3, the relationship among CO, μ, and parameter σ is estimated to obey a distribution which may be similar to two-dimensional normal distribution. Based on two-dimensional normal distribution, a proximate parameter equation is given as formula (3). The formula is one of the most suitable curves according to the simulation data and there are maybe other curves suiting the simulation data:

$$f(x, y) = a \cdot e^{(-b\cdot(x-u)^2 - c\cdot(y-v)^2 + d\cdot(x-u)\cdot(y-v))}. \quad (3)$$

Among it, parameters a, b, c, d, u, and v are real number ranging from negative infinity to positive negative. x represents mean value μ, y represents deviation σ, and $f(x, y)$ is CO. After fitting the curve, the relationship among CO, μ, and σ is as formula (4). Consider the following:

$$\mathrm{CO} = 0.5352 * e^{(0.2804(\mu-1.007)^2 - 0.5445(\sigma-0.1305)^2 - 2.008(\mu-1.007)(\sigma-0.1305))}. \quad (4)$$

Table 2: Indexes of fitted curve.

Index	Value
SSE	0.0005162
RMSE	0.003009
R-square	0.9896
Adjusted R-square	0.9887

Table 2 shows four indexes of fitted curve. Sum of squared error (SSE) and root mean squared error (RMSE) are near 0; coefficient of determination (R-square) and degree-of-freedom adjusted coefficient of determination (adjusted R-square) are near 1. It means that the effect of fitting is well and unknown data can be predicted successfully.

Figure 4 shows the result of fitted curve. Curves in Figures 4(a) and 4(b) are based on formula (4). Among them, the meaning of ordinates is the same with Figure 3.

In addition, the negative sequence corresponding is considered. All the conditions are the same with the above except the relationship of sequence corresponding between initial opinion distribution and initial persuasiveness distribution. However, it is not necessary to calculate again because the initial opinion distribution obeys uniform distribution ranging from 0 to 1 which has a symmetrical characteristic. CO of negative sequence corresponding is symmetrical with that of positive sequence corresponding and the axis of symmetry is $x = 0.5$, where x is the horizontal ordinate of initial opinion distribution.

Suppose that the initial persuasiveness distribution obeys normal distribution which mean that value is μ_0 and standard deviation is σ_0. Positive sequence corresponding is denoted as CO1 and the CO of negative sequence corresponding CO2. CO1 can be calculated from $f(\mu_0, \sigma_0)$, where $f(\cdot, \cdot)$ is formula (4) while CO2 = 1 − CO1. So the range of CO can be predicted approximately which is from CO1 to CO2.

4.3.2. Exponent Distribution. Positive sequence corresponding is taken into account at first. Suppose that the initial persuasiveness of group obeys exponent distribution and initial opinion of group obeys the uniform distribution ranging from 0 to 1. The number of individuals is 10000. Because the range of persuasiveness is from 0 to 1, in order to confirm persuasiveness within boundary, it will generate persuasiveness again if persuasiveness is larger than 0.99 or smaller than 0.01. The parameter in exponent distribution is denoted as λ. A series of discrete values of λ are set as 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, and 14 separately. Figure 5 shows the experiment results. Horizontal coordinate represents parameter λ, while vertical coordinate represents the CO of group.

Figure 5 shows that parameter λ is nonsignificant where all COs are around 0.73. Hence, a conclusion can be acquired that the CO is equal to 0.73 approximately based on

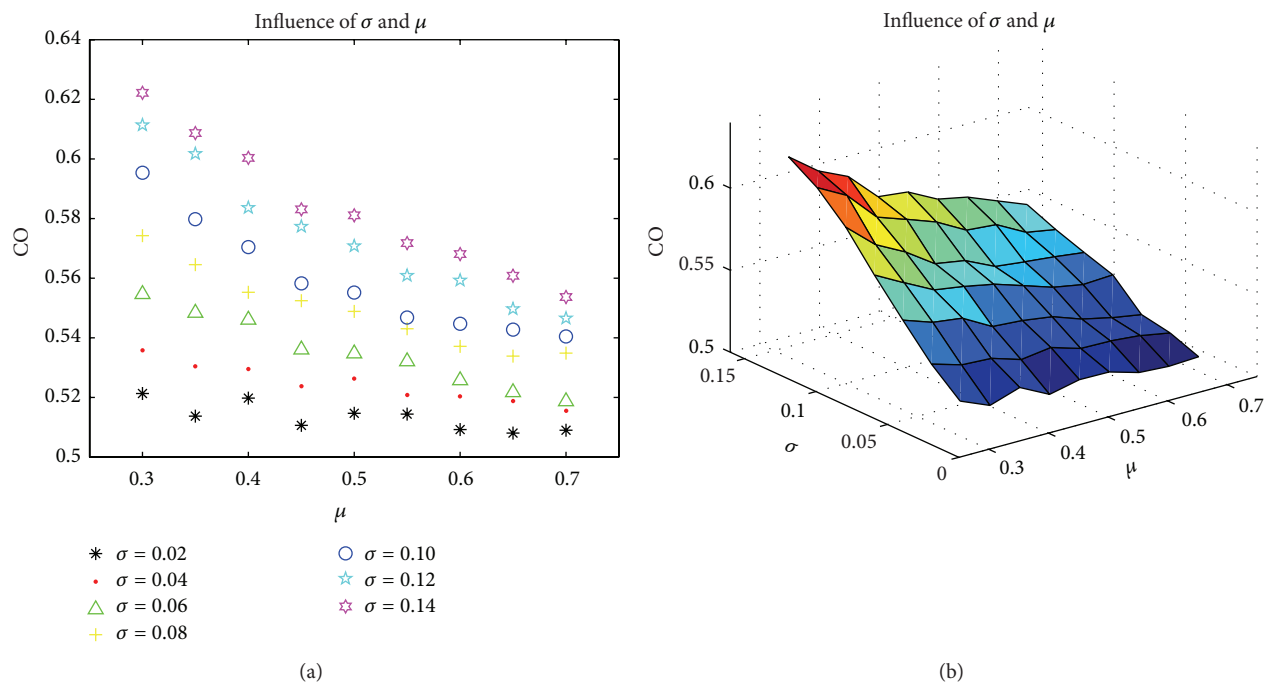


FIGURE 3: The influence of μ and σ on CO.

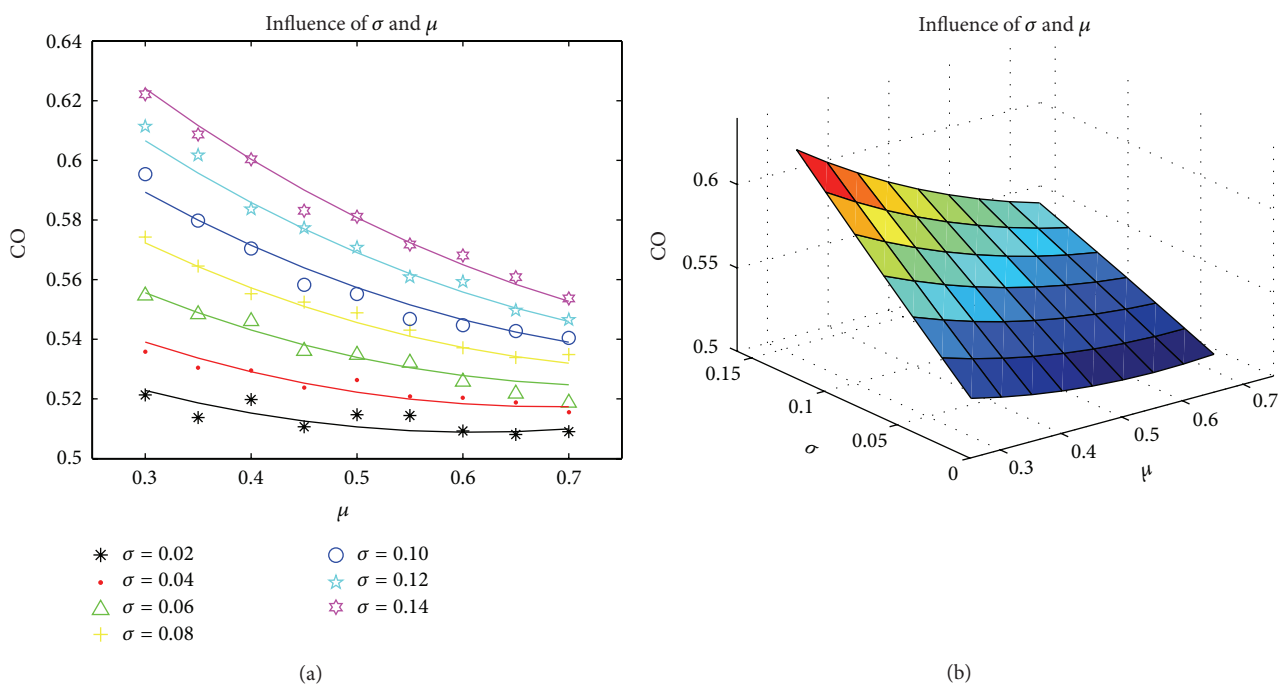


FIGURE 4: The fitted curve based on formula (4).

the premise that initial persuasiveness obeys exponent distribution and initial opinion obeys uniform distribution ranging from 0 to 1 which are positive sequence corresponding between them.

In addition, the negative sequence corresponding is considered. All the conditions are the same with the above except the relationship of sequence corresponding. However, it is not necessary to calculate again because the initial opinion distribution obeys uniform distribution ranging from 0 to 1 which has a symmetrical characteristic. The calculating process is the same with Experiment 4.3.1.

Suppose that the initial persuasiveness distribution obeys exponent distribution of which parameter is λ_0. The result of positive sequence corresponding is denoted as CO1 and

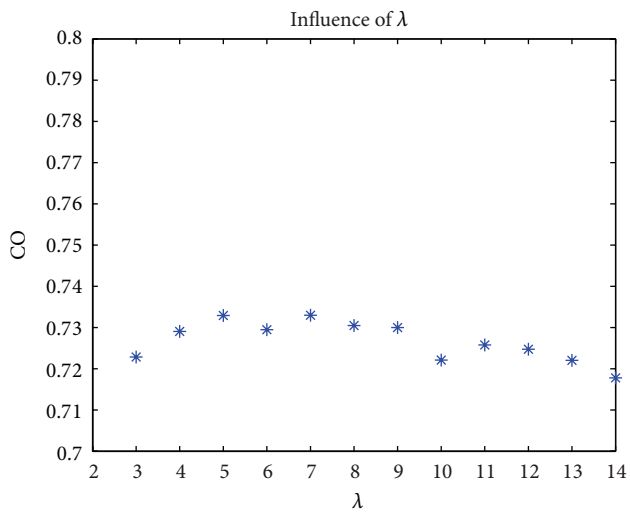


Figure 5: The influence of parameter λ on CO.

negative sequence corresponding as CO2. CO1 = 0.73 while CO2 = 1 − CO1 = 0.27. So range of CO can be predicted ranging from 0.27 to 0.73 approximately.

4.4. Conclusion. Through simulation experiments, some meaningful conclusions and predicted methods are summarized as follows.

(1) No matter how many independently replicate experiments we do, if initial values of all the individuals are fixed, the CO of group is fixed.

(2) Randomness, initial distribution, and number of agents have a great influence on the NOI of experiment. The more the initial persuasiveness distribution concentrates on 0.5, the less the NOI is. In the realistic society, the persuasiveness value of 0.5 represents that individual has their own opinions to a topic but does not have an insight into it.

(3) A method of predicting interval of CO is put forward in the condition that the distribution of initial opinion and initial persuasiveness are known in advance. When initial opinion distribution obeys uniform distribution ranging from 0 to 1 and initial persuasive distribution obeys normal distribution which mean that value is μ_0 and standard deviation is σ_0, the predicted interval of CO is from $f(\mu_0, \sigma_0)$ to $1 - f(\mu_0, \sigma_0)$ where function $f(\cdot, \cdot)$ is formula (4). When initial opinion distribution obeys uniform distribution ranging from 0 to 1 and initial persuasive distribution obeys exponent distribution, the predicted interval of CO is from 0.27 to 0.73.

5. Discussion

The paper focuses on the influence of individual persuasiveness on opinion interaction and builds a new opinion interactive model based on Deffuant models and its improved models. The most different thing between our model and Deffuant model is that all individuals in our model are different from others, namely, owning different persuasiveness. If such a mechanism of taking a decision by a community is correct, our model leads to the following conclusions.

(1) In a closed (isolated) community there are a variety of possible opinions ranging from person to person in original state. After a short and long time our model tend to be one of the "ordered" states, there is a possibility of reflecting the case that the group takes a common decision.

(2) As for a certain discussion topic of closed community, when the initial conditions of the whole group are known, the result may be predicted ahead of time. What may not be predicted easily is the convergent time because of many uncertain factors, like number of individuals, individual persuasiveness, and so forth.

(3) One of the most important factors is the distribution of persuasiveness. Taking Section 4.3.2, for example, the initial persuasiveness distribution of group follows the exponent distribution. Although the proportion of high persuasiveness individuals is much less than that of low persuasiveness individuals, the common opinion is always near the opinion of high persuasiveness individual. By what I mean is that the public (with low persuasiveness) may easily be influenced by opinion leader (with high persuasiveness) and change their opinions to approach the opinion leader which frequently occur in today's society; especially some people are unaware of the truth. So, in order to make the group approach one's opinion, the most crucial thing is to improve his persuasiveness, rather than the number of individuals.

To sum up, the proposed very simple rules trigger a rather complicated dynamics. However, one can doubt if these rules properly describe real mechanisms of taking a decision. There are of course other possibilities within the improved model. Because humans are exactly the opposite of such simple entities, such as atoms and molecules, indeed the detailed behavior of each of them is already the complex outcome of one's interest, benefits, and especially many physiological and psychological processes. However in the model, we just consider a certain case mentioned in Section 3. And these simulation results reflect and interpret some social phenomena of opinion interaction to an extent; for example, persuasiveness of individual is far more significant than the number of individuals in the discussion, which has an important guiding significance of predicting the common opinion of group.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

In this paper, the research was supported by National Nature and Science Foundation of China (nos. 91024030, 71303252, 61403402, 41201544, and 71373282).

References

[1] W. Weidlich, "Sociodynamics—a systematic approach to mathematical modelling in the social sciences," *Nonlinear Phenomena in Complex Systems*, vol. 5, no. 4, pp. 479–487, 2002.

[2] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics," *Reviews of Modern Physics*, vol. 81, no. 2, pp. 591–646, 2009.

[3] T. Brenner, "Agent learning representation: advice on modelling economic learning," in *Handbook of Computational Economics*, vol. 2, pp. 895–947, North Holland, 2006.

[4] W. B. Arthur, "Designing economic agents that act like human agents: a behavioral approach to bounded rationality," *The American Economic Review*, vol. 81, no. 2, pp. 353–359, 1991.

[5] A. E. Roth and I. Erev, "Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term," *Games and Economic Behavior*, vol. 8, no. 1, pp. 164–212, 1995.

[6] C. M. Anderson and C. F. Camerer, "Experience-weighted attraction learning in sender-receiver signaling games," *Economic Theory*, vol. 16, no. 3, pp. 689–718, 2000.

[7] J. S. Jordan, "Bayesian learning in normal form games," *Games and Economic Behavior*, vol. 3, no. 1, pp. 60–81, 1991.

[8] B. Edmonds, "Modelling bounded rationality in agent-based simulations using the evolution of mental models," in *Computational Techniques for Modelling Learning in Economics*, vol. 11 of *Advances in Computational Economics*, pp. 305–332, 1999.

[9] Y. Liu, F. Ding, and Z. Zhang, "A survey of models of consensus formation and evolution," *Journal of Beijing Jiaotong University*, vol. 34, no. 5, pp. 83–88, 2010.

[10] F. Jia, "Rational inference based opinion formation model," *Journal of Beijing Jiaotong University*, vol. 35, no. 2, pp. 60–65, 2011.

[11] X. Hao-Xiang, W. Hui-Li, and X. Zhao-Guo, "Opinion dynamics: a multidisciplinary review and perspective on future research," *International Journal of Knowledge and Systems Science*, vol. 2, pp. 72–91, 2011.

[12] Y.-L. Lu, G.-P. Jiang, and Y.-R. Song, "Epidemic spreading on a scale-free network with awareness," *Chinese Physics B*, vol. 21, no. 10, Article ID 100207, 2012.

[13] Y.-W. Gong, Y.-R. Song, and G.-P. Jiang, "Epidemic spreading in scale-free networks including the effect of individual vigilance," *Chinese Physics B*, vol. 21, no. 1, Article ID 010205, 2012.

[14] Z. Fang, S. Guang-Ya, and L. Pi, "A survey for rumor propagation models," *Complex System and Complexity Science*, vol. 6, no. 4, pp. 1–11, 2009.

[15] Y.-R. Gu and L.-L. Xia, "The propagation and inhibition of rumors in online social network," *Acta Physica Sinica*, vol. 61, no. 23, Article ID 238701, 2012.

[16] J. Lorenz, "Continuous opinion dynamics under bounded confidence: a survey," *International Journal of Modern Physics C*, vol. 18, no. 12, pp. 1819–1838, 2007.

[17] V. Sood and S. Redner, "Voter model on heterogeneous graphs," *Physical Review Letters*, vol. 94, no. 17, Article ID 178701, 2005.

[18] K. Sznajd-Weron and J. Sznajd, "Opinion evolution in closed community," *International Journal of Modern Physics C*, vol. 11, no. 6, pp. 1157–1165, 2000.

[19] S. Galam, "Minority opinion spreading in random geometry," *European Physical Journal B*, vol. 25, no. 4, pp. 403–406, 2002.

[20] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, pp. 87–98, 2000.

[21] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence: models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, 2002.

[22] S. Jiong-Ming, L. Bao-Hong, L. Qi, and M. Hong-Xu, "Trust, evolution, and consensus of opinions in a social group," *Acta Physica Sinica*, vol. 63, no. 5, Article ID 050501, 2014.

[23] P. P. Li, D. F. Zheng, and P. M. Hui, "Dynamics of opinion formation in a small-world network," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 5, Article ID 056128, 2006.

[24] C.-F. Feng, J.-Y. Guan, Z.-X. Wu, and Y.-H. Wang, "Effects of average degree of network on an order-disorder transition in opinion dynamics," *Chinese Physics B*, vol. 19, no. 6, Article ID 060203, 2010.

[25] M.-H. He, D.-M. Zhang, H.-Y. Wang, X.-G. Li, and P.-J. Fang, "Public opinion evolution model with the variable topology structure based on scale free network," *Acta Physica Sinica*, vol. 59, no. 8, pp. 5175–5181, 2010.

[26] B. Kozma and A. Barrat, "Consensus formation on adaptive networks," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 77, no. 1, Article ID 016102, 2008.

[27] J. Lorenz, "Heterogeneous bounds of confidence: meet, discuss and find consensus!," *Complexity*, vol. 15, no. 4, pp. 43–52, 2010.

[28] G. Weisbuch, G. Deffuant, and F. Amblard, "Persuasion dynamics," *Physica A*, vol. 353, no. 1–4, pp. 555–575, 2005.

[29] R. Hegselmann and U. Krause, "Truth and cognitive division of labour: first steps towards a computer aided social epistemology," *Journal of Artificial Societies and Social Simulation*, vol. 9, no. 3, article 10, 2006.

[30] K. Malarz, "Truth seekers in opinion dynamics models," *International Journal of Modern Physics C*, vol. 17, no. 10, pp. 1521–1524, 2006.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 709072, 10 pages
http://dx.doi.org/10.1155/2015/709072

Research Article

An Efficient Robust Eye Localization by Learning the Convolution Distribution Using Eye Template

Xuan Li,[1] Yong Dou,[1] Xin Niu,[1] Jiaqing Xu,[1] and Ruorong Xiao[2]

[1] *Science and Technology on Parallel and Distributed Processing Laboratory, School of Computer, National University of Defense Technology, Changsha 410073, China*
[2] *Informatization Office, National University of Defense Technology, Changsha 410073, China*

Correspondence should be addressed to Xuan Li; lixuan@nudt.edu.cn

Received 30 January 2015; Accepted 14 April 2015

Academic Editor: Ye-Sho Chen

Eye localization is a fundamental process in many facial analyses. In practical use, it is often challenged by illumination, head pose, facial expression, occlusion, and other factors. It remains great difficulty to achieve high accuracy with short prediction time and low training cost at the same time. This paper presents a novel eye localization approach which explores only one-layer convolution map by eye template using a BP network. Results showed that the proposed method is robust to handle many difficult situations. In experiments, accuracy of 98% and 96%, respectively, on the BioID and LFPW test sets could be achieved in 10 fps prediction rate with only 15-minute training cost. In comparison with other robust models, the proposed method could obtain similar best results with greatly reduced training time and high prediction speed.

1. Introduction

Eye localization is essential to many face analyses. In analysis of the human sentiment, eye focus, and head pose, the location of the eye is indispensable to extract the corresponding information there [1]. In face tracing, eye localization is often required in real time. In face recognition, many algorithms ask for the alignment of the face images based on eye location [2]. Inaccurate location may result in the failure of the recognition [3, 4].

However, real-world eye localization is filled with challenges. Face pictures are commonly taken by a projection from the 3D space to the 2D plane. Appearance of the face image could be influenced by the head pose, facial expression, and illumination. Texture around eyes is therefore full of change. Moreover, eyes may be occluded by stuffs like glasses and hair, as shown in Figure 1. To work in any unexpected cases, the algorithm should be robust to those impacts.

In the design of the eye localization algorithm in practical use, prediction accuracy, rate, and the training cost are the most concerned factors. A robust algorithm should keep high prediction accuracy for varying cases with diverse face poses, facial expressions in complex environment with occlusion, and illumination changes. For real time applications, high prediction rate is required. For some online learning systems like the one used for public security, short training time is also in demand to quickly adapt the algorithm to different working places. Low training cost is also of benefit for the tuning of the algorithm. To improve the accuracy in the difficult environment, complex model is often applied. However, the over complicated model will increase the training cost and the prediction time. How to select an approach with enough complexity to achieve high prediction accuracy, high prediction rate, and low training cost at the same time is still a challenge.

Eye localization approaches could be mainly divided into the texture based and the structure based. Texture based methods [5–8] learn the features from the image textures. For the methods exploring local textures [5, 6], high prediction rate could be achieved with simple training. However, they are usually not robust to the situation with occlusion and distortion due to the limited information from the local area. On the other hand, methods like [7, 8] study the global texture feature from entire face image by convolution networks. High

Figure 1: Eye localization result using our method.

prediction accuracy could be obtained by these approaches with high prediction rate. However, the training cost becomes considerable. A long training time is required due to a large number of the model parameters. Proper selection of the model parameters often needs repeated test as well. The structure based approaches [9–12] explore the predefined critical facial points. Eye locations could be detected mainly by the structure information. Although high accuracy could be achieved by a simple training, the prediction often involves an iterative optimization. And the iteration times and predication accuracy highly depend on the initialization. Therefore, the prediction accuracy and rate are usually not stable.

In this study, it was found that there is regular response distribution on the convolution map generated by eye template. This distribution reflects the spatial relationship among the major facial objects. By a nonlinear learning model, such distribution could be explored to predict the location of the eyes. Instead of using local response like conventional template methods, global information could be explored according to such distribution. In this way, eye locations could be accurately predicted and even occlusion or distortion occurs. Besides, high prediction rate could be expected with a noniterative prediction model. To this end, the convolutional networks [7, 8] could be explored. However, conventional convolutional networks learn from the raw image, which may need more layers to map the complex nonlinear relation between the face image and eye locations. On the other hand, convolution map produced by eye template contains concise distribution information, where different facial objects have stable response patterns. Since many unnecessary textures are ignored, distribution could be learned using relatively shallow networks and the training time can be much reduced.

Based on this principle, an efficient and robust eye location algorithm is proposed in this paper. The algorithm explores the face convolution map by the eye template using a BP network. To enhance the performance of the eye template, a novel template training method was proposed by noise suppression control. Besides, a FFT- (fast Fourier transform-) based convolution approach was designed to further improve the training and prediction speed. Eventually, the proposed algorithm could achieve accuracy of 98% and 96%, respectively, on the BioID [13] and LFPW [14] test sets with a prediction rate of 10 fps. However, the training time was only 15 minutes for 13,466 samples.

The rest of the paper is organized as follows. Section 2 reviews the related work on eye detection. Section 3 describes the finding of the convolution distribution. Section 4 gives a detailed description of the proposed approach. Section 5 discusses the experiments and results. Conclusion is given in Section 6.

2. Related Work

Eye localization techniques have been greatly developed in recent years. According to the properties of the explored features, methods could be mainly divided into two classes: the structure based and the texture based approaches.

The structure based approaches explore some critical points on a specific face structure model for the prediction. Typical structure based approaches include ASM [9] and AAM [11]. In these methods, eye locations are estimated by the structure information and the local texture around the points through an iterative process. The iteration time and the prediction accuracy are affected by the initial structure. Moreover, the optimization employs a least square method whose robustness is poor to the case with occlusion.

The texture based approaches extract the face texture to predict the locations of the eyes. Typical structure based approaches include [5, 7, 8, 15]. Instead of using limited critical points by the structure based approach, all the textures from the entire face image could be explored. Previous studies [7] have showed the robustness of the face textures to occlusion and distortion. The texture based approaches are usually noniterative. For the extraction of the texture features, approaches like convolution [5, 7], LBP [16], HOG [17], HAAR [18], and so forth are commonly used. Within them, best prediction accuracy has been reported by the the convolution based approach [7, 8]. Like other texture extraction, convolution operation is also computation intensive since the whole image needs to be scanned over by a texture window. However, it was also found that the convolution feature extraction can be well accelerated in frequency domain by FFT.

Currently, there are mainly two kinds of eye localization using convolution feature extraction. Methods like [5, 19, 20] tried to explore the peak response by a predefined template (as the convolution kernel) with an ideal eye pattern. Since the convolution response of the eye template on the other facial objects like cheek or lip was hard to explain, the remaining part of the response map was ignored, and eyes were only located by the peak response. These approaches have low training cost and high prediction rate. But they are not robust to the occlusion and changing face pose, since the response on the eye by such fixed eye template will greatly decrease in these situations. Prediction by the peak response may eventually fail in these cases. In contrast to exploring the peak response by the complex eye template, other methods try to code the face by the templates of basic image elements, most of which are line segments. Relation between the eye locations

and the code distribution of the face could be learned by nonlinear models. Due to the use of global information, the prediction accuracy could be improved and the effects of occlusion and distortion by pose changes could be relieved. However, such nonlinear relationship is usually complex. Models used in those approaches such as Deep CNN [7, 8] are commonly with a large number of parameters, which leads to a costly training phase and a low prediction rate. Moreover, search of the optimal setting of these complex models is also difficult.

In this paper, it was found that there is relative stable global distribution pattern in the convolution response generated by the eye template and the face image. Based on this phenomenon, a novel eye location approach is proposed. This approach explored the global face information produced by the eye template for location prediction to avoid arbitrary decision by only the peak response. Meanwhile, using specific eye template rather than the randomly learned kernels, the convolutional network could be greatly simplified with only one layer. Properties of the response distribution as the new finding will be described in Section 4. Section 5 gives the proposed algorithm based on these findings.

3. Regulation on the Convolution Response Map by Eye Template

The global information of the convolution response map by eye template has long been ignored, since the response distribution of the face is not explicit. Therefore, only the peak response was evaluated to indicate the location of the eyes, as the eye templates normally have the largest convolution response value at the place of the eyes. However, these approaches are usually with low prediction accuracy in practical use. The facial expression, occlusion, and changing pose will all affect the appearance of the eyes, which will lead to lower response there. Moreover, it is also difficult to distinguish the left eyes from the right.

However, it was found in the experiments that the convolution response map of the face image by eye template has regular global distribution, where the facial objects like nose and mouth have relatively stable patterns.

To demonstrate this phenomenon, an experiment which tests the similarity of the response samples from the same facial object has been conducted. KL distance [21] was used to quantitatively measure the similarities of those response patterns as follows:

$$\mathrm{KL}\left(p \parallel q\right) = \sum_{\nu \in \Omega} p\left(\nu\right) \ln \frac{p\left(\nu\right)}{q\left(\nu\right)}, \tag{1}$$

where p and q are the two test distributions. KL tends to be small when p and q are similar.

In the experiments, 100 front face images were randomly selected from the BioID test set. They were scaled to 128×128 and preprocessed to remove the shadow effects by Tan-Triggs [22]. 100 convolution response maps were further generated by a pretrained eye template as shown in Figure 2(a). Five comparison groups (rows (b), (c), (d), (e), and (f) in Figure 2) were defined by the cropped texture samples at five locations (with the color squares and center points) of the left eye (red), right eye (orange), nose (green), left mouth corner (blue), and right mouth corner (black). A texture sample was picked up from each group as the reference to compare with the other samples in the same group with KL distance. In KL measurement, p (in the first column of Figure 2) was used for the reference texture and q for the texture samples in the same group. p and q were all normalized to $[0, 1]$. Zero value is replaced by 10^{-9} to avoid the zero division. To demonstrate the similarity of the samples in each facial object group, a group (row (g) in Figure 2) with randomly cropped textures over the face images was also prepared for comparison. Additional group (row (h) in Figure 2) formed by the samples from the right eye with occlusions was prepared for study as well. Colors were used to enhance the visibility of the response maps with the increasing value from blue to red.

From Figure 2, stable responses of the eyes, noses, and mouths ((b)–(g)) can be observed. The average KL distances of these groups were about 2.23, 2.85, 4.35, 5.34, and 5.38, which were significantly smaller than those of the random group $T = 12.78$. This indicates the similarity of the response patterns of the samples from the same facial object.

From Figures 2(b) and 2(c), it can be observed that the distributions of the textures of the left and right eyes are stable unimodal. However similarity between the two eyes is also evident, which implies the difficulty to tell the difference from the left to the right using only the peak response. The distribution of nose responses in Figure 2(d) is multimodal with the peak values in the half part of the stable M pattern. The average KL distance in nose group is relatively bigger than that of eyes, which indicates more diversity in this area. The distributions of the left (Figure 2(e)) and right (Figure 2(f)) mouth corners also showed regular unimodal patterns with more variation than that of the mouth. As illustrated in Figure 2(h), the distribution of the samples from occluded eye has irregular pattern and is more similar to the randomly selected patches in group (g) with the average KL distance 8.39.

From the above experiments, it could be noted that a relative stable global distribution pattern could be achieved using convolution by the eye template. The convolution responses of the facial objects have distinctive and stable patterns. These properties could be explored to form a stable spatial relationship between the eye locations and the positions of other facial objects. Even when the information of part of the face is destroyed by the occlusion or distortion, the location of the eyes could also be predicted from the rest of the stable patterns.

4. Architecture

The architecture of the proposed eye location model mainly consists of 3 stages, that is, convolutional feature extraction, downsampling, and a BP neuronetwork as illustrated in Figure 3. In contrast to many other convolutional networks, only one convolution layer was employed in our work. Moreover, the pretrained eye templates were used as the convolution kernels. Reduced convolution layers could simplify the model and significantly short the training time, while

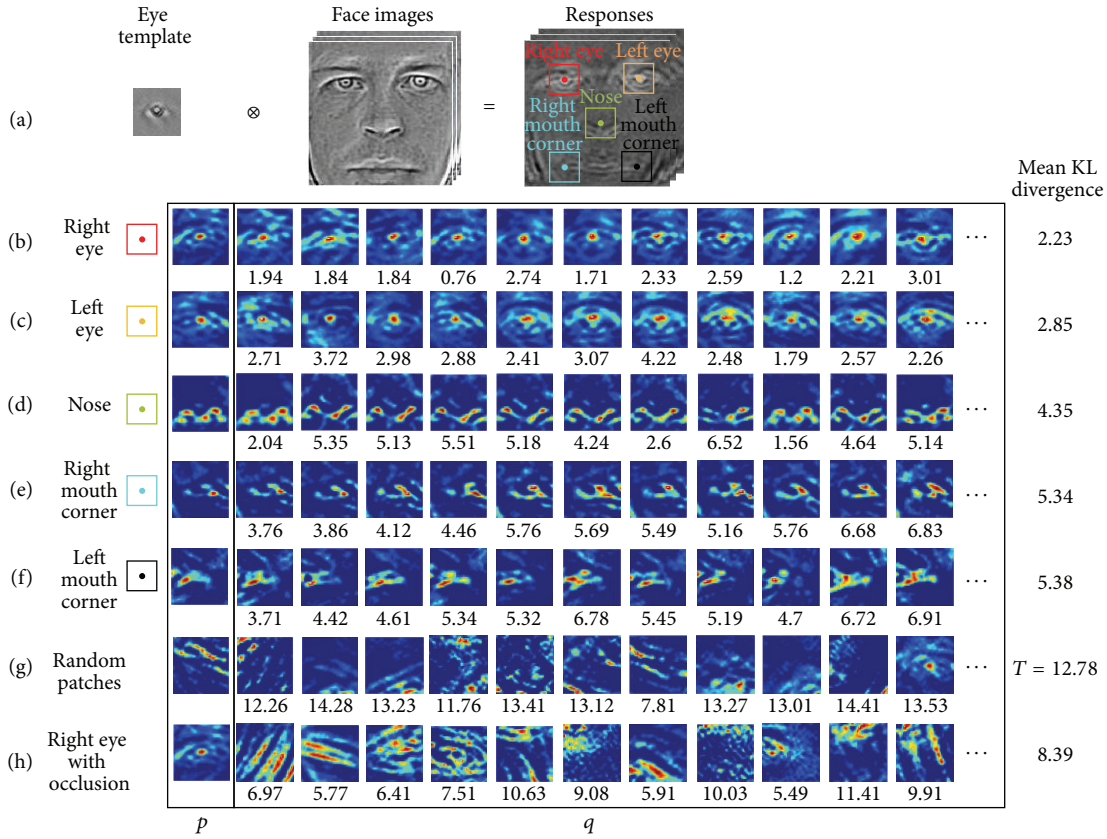


Figure 2: Convolution responses for different facial objects.

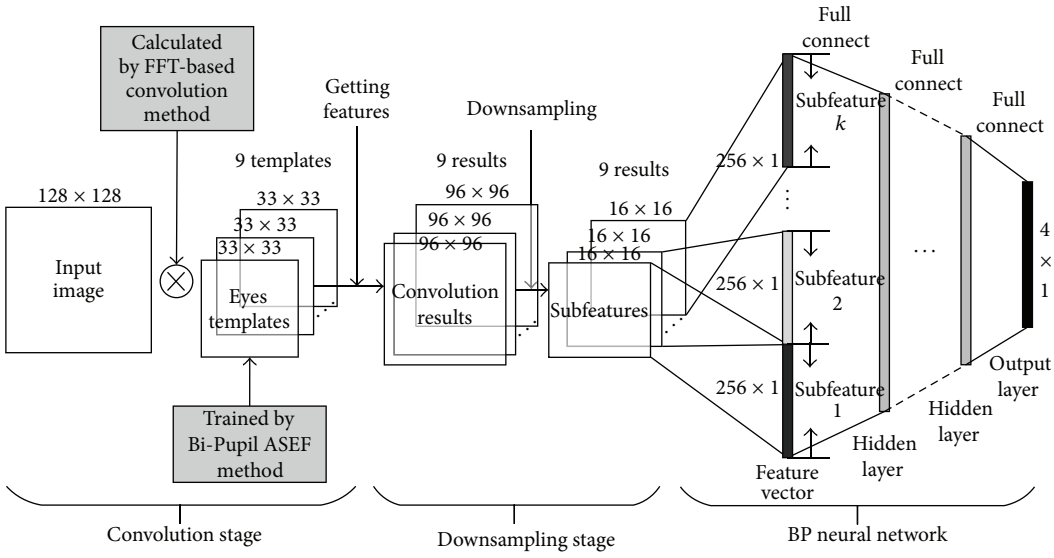


Figure 3: Architecture of the proposed eye localization model.

the feature extraction ability may be limited. However, it was found that by proper preprocessing and selection of the convolution kernels, the relationship between the eye locations, and the convolution response could be well learned by a BP network. For image preprocessing, the Tan-Triggs method was selected to reduce the illumination effects. The eye templates were trained by a proposed Bi-Pupil ASEF approach. Convolution operations were implemented by a FFT-based method to further reduce the computation. Details of these approaches are given in the following subsections.

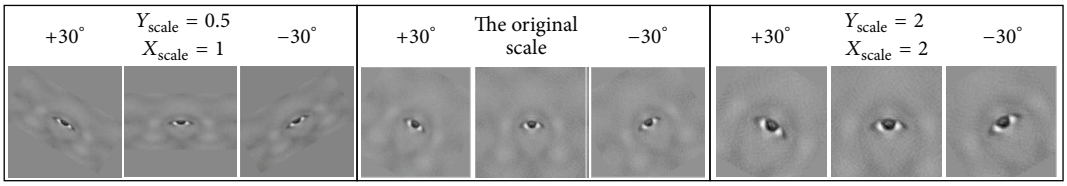


Figure 4: Multieye template set.

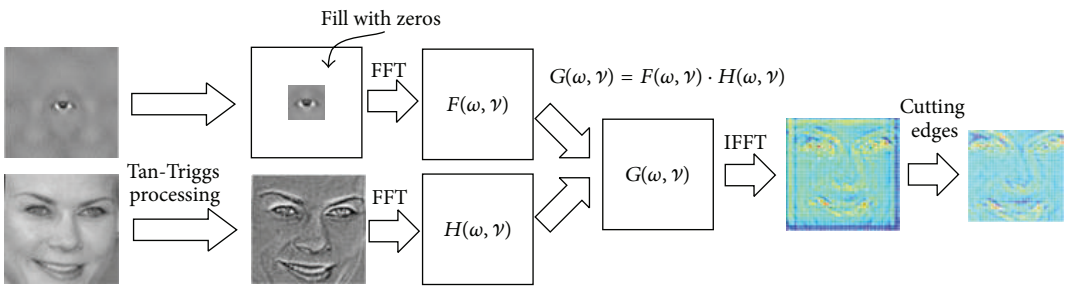


Figure 5: The process of the FFT-based convolution.

4.1. Eye Template Training by Bi-Pupil ASEF. The eye templates are trained by a proposed Bi-Pupil ASEF which is based on ASEF [5]. In ASEF, an eye template sample could be given by each eye image sample with a specific response function. The eye template could be obtained by averaging the template samples to stress the features in common. At the location of the eyes, peak response could be normally observed by the ASEF templates. To efficiently synthesize the characters of the left and right eyes for feature extraction, a modified response function was explored in the proposed Bi-Pupil ASEF. And a multieye template set was employed to cope with the change of scale and rotation.

4.1.1. Bi-Pupil Response Function. It was found the response maps generated by the left and right eyes have certain relations. To efficiently generate response maps and reduce the cost of the separate training of the two eyes, a Bi-Pupil response function has been proposed instead of the one used in ASEF. Consider

$$g(x, y) = e^{-((x-x_{le})^2+(y-y_{le})^2)/\sigma^2} + e^{-((x-x_{re})^2+(y-y_{re})^2)/\sigma^2}, \quad (2)$$

where (x_{le}, y_{le}), (x_{re}, y_{re}), respectively, give the positions of the left and right eyes. By this function, similar high responses could be achieved on both eyes. The convolution response map could be taken as the average of the ones by left and right eyes templates. However, the training task is reduced to only one template for two eyes. Moreover, by average of the two eyes, the coupling of the template to certain samples could be further decreased. As the conventional ASEF, the Bi-Pupil ASEF eye template training is noniterative. With the increased number of the training samples, the template eventually converges. Therefore, the well trained eye template could also be used to other data sets.

4.1.2. Change Invariant Multieye Template Set. To make the template invariant to the changes like eye closing, scaling, and rotation in the real-world use a multieye template set was generated with 3 rotation types in 3 scale levels as illustrated in Figure 4. Accordingly, nine convolution response maps were produced to code the face with the concerns of these situations.

4.2. FFT-Based Convolution Method. In the ASEF based feature extraction, the convolution of the face image and the eye template is implemented by dot product in the frequency domain through FFT. Although FFT could reduce the computation for convolution, it will also bring the frequency effects on the image edge, where the response may be wrongly computed with the periodic edge information from the other side. To mitigate such effects, a cosine window approach [5] has been proposed. However, it was found such approach may fail due to the greatly reduced eye texture when the eyes are not in the central zone of the image. To reduce the computation for convolution operation without introducing additional errors in practical use, a modified FFT-based approach was proposed as illustrated in Figure 5.

We use a zero filled window function to reduce the frequency effects. Comparing to the cosine window method, the zero filled window function is applied on the eyes templates other than the face images. This method can avoid the reduction of the eye texture which happens in the cosine window method. To keep the feature extraction ability and to reduce the frequency effects, the size of zero filled area on the eye templates should be chosen carefully. Based on our experience, this method, that is, only keeping the eyes texture and filling the other area with zeros on the templates, is a good choice. In our experiment, the size of the face images is 128×128, and the size of the eye region is approximately 33×33. According to the size of the eye region, only the 33×33 center area of eye templates is retained.

After the fringe of the eye template is zeroed out, the zero filled eye template and the preprocessed eye image are transformed by FFT. With an IFFT on the dot product of

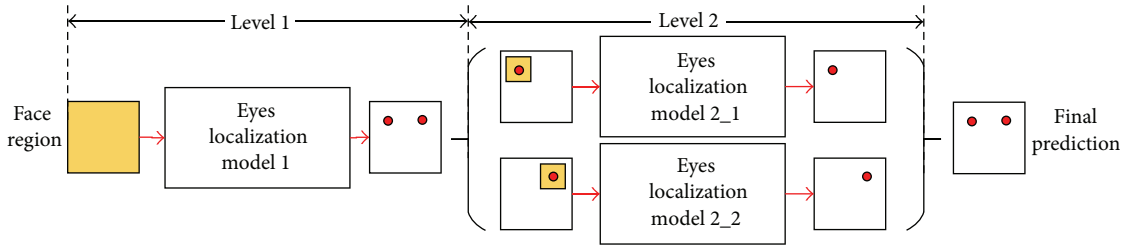


Figure 6: The structure of the two-level cascade enhancement.

the two frequency maps, an initial convolution map could be achieved. By cutting out the inaccurate edges with a width of 16 pixels, the final convolution result of 96×96 is obtained. Such result is the same to that by normal convolution.

To demonstrate the effect of the FFT-based approach for reducing the convolution computation, a computation complexity analysis is given as follows. The analysis is only based on the multiplication, since it is the most time-consuming operation. Given N the face image width and M the eye templates, the total number of multiplications in a normal convolution in the spatial domain is

$$C_{\text{spatial}} \approx O\left(M^2 (N - M + 1)^2\right). \tag{3}$$

With 3 FFT (two FFT and one IFFT) and the product of the complex matrices in frequency domain, the multiplication number of the FFT-based convolution could be

$$C_{\text{freq}} \approx O\left(3N^2 \log N + 4N^2\right). \tag{4}$$

As in this experiment, $N = 128$ and $M = 33$, for one convolution map, about 10,000,000 multiplications are required for the normal convolution and 40,0000 for the FFT-based approach. Considering 9 maps by the multieye templates, the reduction of the computation is significant. Decreasing of the training and prediction time could benefit from such reduction of convolution computation.

4.3. The BP Network and the Cascade Enhancement. Before the further processing, the convolution response maps were firstly normalized with the mean 0 and the standard deviation 1.0. It was found such normalization is significant to improve the prediction accuracy, since the unbalance between the distributions of the response maps could be well reduced. To improve the invariance of the extracted features, max pooling was employed to downsample the normalized response maps to 16×16. The nine downsampled matrices were further vectorized and catenated as a vector with the length of 2304 to input to the BP network. The fully connected BP network employed the sigmoid activation function as illustrated in Figure 3. As the output, the location coordinates were also normalized into $[0, 1]$. The BP network is trained by conjugate gradient descent with the least square error between the label and prediction as the error function. By comparison, a 4-layer network with the hidden neurons of [30-20-20-10] was selected considering the prediction accuracy and training time.

To further improve the prediction accuracy, a two-level cascade enhancement scheme has been proposed as illustrated in Figure 6, just as [7] does. The first level gives the initial positions of the eyes from the entire face image. Then, the location of each eye was revised in the second level within a square centered by the initial position. The width of the square is one-fourth of the face region.

The second level only learns the deviation Δx between the prediction by the first level and the labeled location. Since each model in the second level only learns one eye, the single eye response function in original ASEF was used to generate the eye template in the second level. The final prediction x is the sum of the output of the first level x_1 and the second level Δx.

5. Experiment

The data used for this experiment is the same as [7]. The data set consists of 13,466 training and 2,551 test images. The training samples were selected from the LFW [23] data set and internet. The test set was formed by the complete BioID and LFPW data sets. The BioID data was mainly composed of the regular front images, while the other data such as LFPW were collected from many complex environments including wide changes in head pose, illumination, scale, clarity, and occlusion. Therefore, a robust algorithm should have high accuracy on the LFPW data set. All the images were marked with face region, the locations of the eyes, and nose and two mouth corners.

For evaluation, a relative prediction error is defined as follows:

$$\text{err} = \frac{\sqrt{(x - x_t)^2 + (y - y_t)^2}}{l}, \tag{5}$$

where (x, y) and (x_t, y_t) are, respectively, the predicted position and the ground truth and l is the biocular distance. Based on this error, the accuracy of the algorithm could be defined as the ratio of the samples, which have the prediction error less than 0.1, to the whole test samples. The mean error is also used as another indicator to evaluate the prediction results in the following experiments.

The proposed algorithm was implemented in Matlab 2014a, and the experiments were conducted on a desktop computer with a 3.3 GHz CPU.

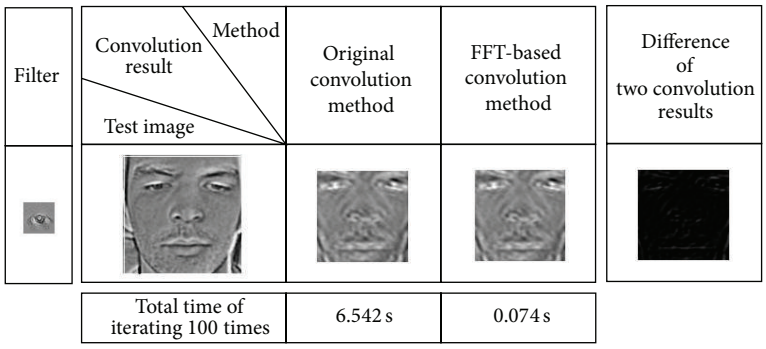


Figure 7: Comparison between the original convolution and the FFT-Based approach.

Table 1: Accuracies of the involved schemes.

Method	BioID		LFPW	
	Right eye	Left eye	Right eye	Left eye
M1	95.7%	96.6%	95.5%	95.4%
M2	96.8%	96.1%	94.7%	94.8%
M3	96.7%	96.8%	96.1%	95.8%
M4	94.2%	94.8%	93.2%	92.7%
Standard	**98.1%**	**98.2%**	**96.8%**	**96.9%**

5.1. Effects of the Improvement Schemes. In this section, the effects of the above-mentioned improvement schemes will be discussed.

5.1.1. Effects on Prediction Accuracy. In the following experiments, the schemes of the Tan-Triggs preprocessing, Bi-Pupil ASEF eye template training, normalization of the convolution response map, and the cascade enhancement schemes on the prediction accuracy were tested. The proposed model with all the schemes was noted as the standard configuration. As show in Table 1, four compared configurations were prepared: Tan-Triggs is replaced by $\log(x + 1)$ function (M1), where x is the pixel values. Bi-Pupil ASEF is replaced with the original ASEF (M2) which train templates using only right eye location, response map normalization is replaced by no action (M3), and the cascade enhancement is replaced by the only first-level raw model (M4).

As demonstrated in Table 1, all the schemes could effectively improve the prediction accuracy. Using Tan-Triggs rather than the Log function to reduce the illumination effects could increase the accuracy by about 1.5%–2%. The Bi-Pupil ASEF templates can raise the accuracy by about 2%–3% than the original ASEF templates. Response normalization was also of benefit to the accuracy increment about 0.5%–1.5%. In comparison with the raw model, the cascade enhancement could bring 3%-4% accuracy improvement.

5.1.2. Acceleration by FFT-Based Convolution. To demonstrate the performance of the FFT-based approach, a convolution of a face sample and an eye template has been repeated 100 times. As illustrated in Figure 7, there is slight difference between the results by the FFT-based convolution and the direct convolution. However, such difference hardly has any effect on the further analysis with the extracted feature. For the computation efficiency, with the reduced operations, the FFT-based approach could generate almost 80 speed-up ratio in this experiment. For the models with a large numbers of convolution, this approach could significantly decrease the training and prediction time.

5.2. Comparison with Other Approaches. To demonstrate the efficiency of the proposed method, comparisons with many state-of-the-art eye location approaches including the texture based and the structure based were conducted considering the prediction accuracy, prediction rate, and training time.

For the texture based approaches, ASEF, Template-SVR, and Deep CNN [7] have been compared. Deep CNN is currently one of the best approaches with highest accuracy in eye location. ASEF use the maximum pixel position of convolution result directly to predict eye's location. The Template-SVR approach is formed by replacing the BP network in the proposed model with the nu-SVR implemented by libSVM [24] to draw a fair comparison on the efficiencies of the 2 nonlinear mapping models. ASEF and Template-SVR were implemented in Matlab, while the results of Deep CNN were obtained from [7] with the same training and test data in this paper.

Besides, some leading structure based approaches like CBDS [10], BORMAN [12], and one included in a commercial software LUXAND [25] localizing were also compared. However, their prediction results were obtained from [7]. The test data sets of them were the same to that used in this paper, but the training data are unknown.

5.2.1. Prediction Accuracy. The prediction accuracies of the compared approaches were listed in Table 2. The top 2 results were marked in bold. It should be noted that the robust algorithm usually has high accuracy on LFPW, since samples in this test set were collected in varying situations.

As shown in Table 2, the best results were produced by Deep CNN. However, the proposed method could also achieve similar high accurate results close to Deep CNN. Overall, the proposed method could be the second best one. Averagely, the accuracy of the proposed method was about only 2% lower than that by Deep CNN, while the mean

Table 2: Comparison of accuracies among the state-of-the-art eye localization approaches. Note the robust algorithms with high accuracies on LFPW.

Method	Accuracy				Mean error			
	BioID		LFPW		BioID		LFPW	
	Right eye	Left eye	Right eye	Left eye	Right eye	Left eye	Right eye	Left eye
Our method	98.1%	98.2%	**96.8%**	**96.9%**	**2.7%**	**2.4%**	**3.4%**	**3.1%**
Deep CNN	**99.9%**	**100%**	**99.1%**	**99.4%**	**1.7%**	**1.5%**	**2.1%**	**2.0%**
ASEF	1.2%	0.2%	2.4%	0.6%	121.4%	88%	81.2%	99.2%
nu-SVR	96.1%	95.9%	92.8%	92.8%	4.2%	4.1%	4.9%	4.9%
BORMAN	79.1%	75.8%	78.2%	92.8%	7.1%	7.8%	7.8%	8.8%
CBDS	97.7%	**98.9%**	87.9%	91.9%	4.1%	3.9%	7.2%	7%
LUXAND	**98.9%**	98.66%	95.6%	96.8%	4.1%	3.7%	5.6%	4.5%

error was 1.1% higher. However, it should be noted that, to achieve such results, the Deep CNN approach should explore the information from 5 feature points with a more complex structure which had 3 cascade levels and several deep models with 3 convolution layers in each level. In comparison, the proposed method only used 2 feature points by a 2-level cascade structure with single convolution layer model in each level. By introducing more feature points for reference with corresponding model structure, there is still large improvement space.

For other texture based approaches, it could be found, without exploration of the global information, the ASEF can hardly work in the complex environment. Considering the supervised learning model, SVR could also well map the relationship between the response map and the eye locations. Nevertheless, the prediction accuracy by nu-SVR is about 4–6% lower than the proposed method by the 4-layer BP network.

In comparison with the structure based approaches, the accuracy of the proposed method is slightly lower than that of CBDS and LUXAND on the BioID by about 0.7%. That might be because of the fact that the samples in BioID are all the regular front face image, which is suited for such structure based approaches. While on the LFPW where pose changes and occlusion commonly exist, the proposed method was better than these methods. Moreover, the mean error of the proposed method was obviously lower than the structure based approach, which also demonstrates the robustness of the proposed method.

5.2.2. Prediction Rate. As mentioned before, the structure based approach employs an iterative prediction approach. The iterative time is not stable and highly related to the study case and initialization. For the situation with side face or occlusion, it may cost seconds to process one image.

Therefore, the experiment on prediction rate was focused on the texture based approach. To simplify the analysis, the face detection time, which is highly related to different situations, was not considered in this experiment. The rate is only calculated for the prediction from the detected face image with the same size. The prediction rate of the Matlab implemented ASEF, nu-SVR, and the proposed method were recorded. Result of the C++ coded Deep CNN was the referenced data.

Table 3: Prediction rates of the texture based approaches.

Method	ASEF	nu-SVR	Deep CNN	Our method
fps	66.7	0.82	8.3	10.5

As shown in Table 3, without nonlinear BP network, ASEF could achieve the highest prediction rate by about 66 fps. The prediction rate by nu-SVR is the lowest with 0.82 fps. In comparison, by the BP network, the proposed approach could achieve 10.5 fps. The well optimized C++ Deep CNN had reported a prediction rate of 8.3 fps which is close to our method.

5.2.3. Training Time. For the structure based approach, usually a well structure model can be obtained by limited samples without too much training time. However, most of the structure based approaches are not as robust as the high accuracy texture based approach. Texture based approach with simple model could be trained rapidly. In the experiment, training of the ASEF with more than 13,000 samples only costs about 200 seconds. However, it is hard to obtain notable improvement through extensive training for those simple models. Therefore, the comparison was focused on the high accuracy texture based approach.

To obtain the above-mentioned accuracy, the proposed method with 13,466 training samples was only trained in 15 minutes including the training of 3 eye templates and 3 BP network in the 2-level cascade model. The low training cost makes our model easy to be tuned and suitable for online learning task.

In comparison, training of other high accuracy models was very slow. The training cost of CNN with back-propagation will be dramatically increased with additional layer. With 3 cascade levels and several deep models with 3 convolution layers in each level, there are totally more than 160,000 parameters to be trained for eye localization. Training of the Deep CNN could cost hours and even days. Due to the long training time, selection of an optimal configuration (like the number of layers and kernels, kernel size, etc.) for the deep model will become difficult. Searching of an optimal configuration will also increase the cost of the training. The nu-SVR model is also time-consuming. For the training

of the Template-SVR approach, this costs several hours. Optimization of those models is therefore difficult.

6. Conclusion

The stable response distribution by convolution of the face image and eye templates has been discovered. This distribution pattern is beneficial to the eye location. A novel eye location approach has been proposed by learning the distribution of the convolution response maps. The proposed approach only used one convolution layer with a specific eye template and a BP network. In comparison with many state-of-the-art approaches, comparable best prediction accuracy could be achieved by the proposed method with high prediction rate and less training time. The proposed method, which is robust to the pose changes, distortion, and occlusion, can be well used in the complex environment. It has been demonstrated that with proper selection of the template as the convolution kernel, a shallow convolution model could produce similar accurate results to that by deep convolution models with high prediction rate and greatly reduced training time.

In the future work, localization of other critical facial points with the proposed approach will be studied. Efficiency of the kernel selection in the convolution networks will be further analyzed.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is mainly supported by the National Natural Science Foundation of China under Grants U1435219, 61125201, and 61402507.

References

[1] R. Valenti, N. Sebe, and T. Gevers, "Combining head pose and eye location information for gaze estimation," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 802–815, 2012.

[2] R. Gross, S. Baker, I. Matthews, and T. Kanade, *Handbook of Face Recognition*, 2005.

[3] H. K. Ekenel and R. Stiefelhagen, "Why is facial occlusion a challenging problem?" in *Advances in Biometrics*, pp. 299–308, Springer, 2009.

[4] B. Kroon, A. Hanjalic, and S. M. P. Maas, "Eye localization for face matching: is it always useful and under what conditions?" in *Proceedings of the International Conference on Image and Video Retrieval (CIVR '08)*, pp. 379–387, ACM, July 2008.

[5] D. S. Bolme, B. A. Draper, and J. R. Beveridge, "Average of synthetic exact filters," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 2105–2112, IEEE, June 2009.

[6] Z.-H. Zhou and X. Geng, "Projection functions for eye detection," *Pattern Recognition*, vol. 37, no. 5, pp. 1049–1056, 2004.

[7] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 3476–3483, IEEE, Portland, Ore, USA, June 2013.

[8] E. Zhou, H. Fan, Z. Cao, Y. Jiang, and Q. Yin, "Extensive facial landmark localization with coarse-to-fine convolutional network cascade," in *Proceedings of the 14th IEEE International Conference on Computer Vision Workshops (ICCVW '13)*, pp. 386–391, IEEE, December 2013.

[9] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[10] L. Liang, R. Xiao, F. Wen, and J. Sun, "Face alignment via component-based discriminative search," in *Computer Vision—ECCV 2008*, vol. 5303 of *Lecture Notes in Computer Science*, pp. 72–85, Springer, Berlin, Germany, 2008.

[11] T. F. Cooles, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.

[12] M. Valstar, B. Martinez, X. Binefa, and M. Pantic, "Facial point detection using boosted regression and graph models," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2729–2736, IEEE, June 2010.

[13] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the hausdorff distance," in *Audio-and Video-Based Biometric Person Authentication*, pp. 90–95, Springer, 2001.

[14] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar, "Localizing parts of faces using a consensus of exemplars," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 545–552, IEEE, June 2011.

[15] J. Song, Z. Chi, and J. Liu, "A robust eye detection method using combined binary edge and intensity information," *Pattern Recognition*, vol. 39, no. 6, pp. 1110–1125, 2006.

[16] B. Kroon, S. Maas, S. Boughorbel, and A. Hanjalic, "Eye localization in low and standard definition content with application to face matching," *Computer Vision and Image Understanding*, vol. 113, no. 8, pp. 921–933, 2009.

[17] D. Monzo, A. Albiol, and J. Sastre, "Precise eye localization using HOG descriptors," *Machine Vision and Applications*, vol. 22, no. 3, pp. 471–480, 2011.

[18] Y. Ma, X. Ding, Z. Wang, and N. Wang, "Robust precise eye location under probabilistic framework," in *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition (FGR '04)*, pp. 339–344, IEEE, May 2004.

[19] E. Vazquez-Fernandez, D. Gonzalez-Jimenez, and L. L. Yu, "Improved average of synthetic exact filters for precise eye localisation under realistic conditions," *IET Biometrics*, vol. 2, no. 1, pp. 10–20, 2013.

[20] V. Štruc, J. Ž. Gros, and N. Pavešić, "Principal directions of synthetic exact filters for robust real-time eye localization," in *Biometrics and ID Management*, vol. 6583 of *Lecture Notes in Computer Science*, pp. 180–192, Springer, 2011.

[21] S. Kullback, "The kullback-leibler distance," 1987.

[22] X. Tan and B. Triggs, "Preprocessing and feature sets for robust face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '07)*, vol. 7, pp. 1–8, 2007.

[23] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," Tech. Rep. 07-49, University of Massachusetts, Amherst, Mass, USA, 2007.

[24] C.-C. Chang and C.-J. Lin, "LIBSVM: a Library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.

[25] http://www.luxand.com/facesdk/.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 749326, 10 pages
http://dx.doi.org/10.1155/2015/749326

Research Article

Neural Cognition and Affective Computing on Cyber Language

Shuang Huang,[1] Xuan Zhou,[2] Ke Xue,[3] Xiqiong Wan,[4] Zhenyi Yang,[5] Duo Xu,[6] Mirjana Ivanović,[7] and Xueer Yu[8,9]

[1] *Overseas Training Center, Shanghai International Studies University, Shanghai 200083, China*
[2] *School of Humanities and Social Science, Sichuan Conservatory of Music, Chengdu 610021, China*
[3] *School of Media & Design, Shanghai Jiaotong University, Shanghai 200240, China*
[4] *School of Mathematical Sciences, Fudan University, Shanghai 200433, China*
[5] *School of Software, Fudan University, Shanghai 200433, China*
[6] *Department of Arts Management, Tianjin Conservatory of Music, Tianjin 300171, China*
[7] *Department of Mathematics and Informatics, Faculty of Sciences, University of Novi Sad, 21000 Novi Sad, Serbia*
[8] *College of Arts and Science, Washington University in St. Louis, St. Louis, MO 63130, USA*
[9] *Marketing Department, J.L. Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA*

Correspondence should be addressed to Ke Xue; kxue@sjtu.edu.cn and Xiqiong Wan; xqwan@fudan.edu.cn

Received 5 February 2015; Accepted 8 April 2015

Academic Editor: Abdul Hanan Abdullah

Characterized by its customary symbol system and simple and vivid expression patterns, cyber language acts as not only a tool for convenient communication but also a carrier of abundant emotions and causes high attention in public opinion analysis, internet marketing, service feedback monitoring, and social emergency management. Based on our multidisciplinary research, this paper presents a classification of the emotional symbols in cyber language, analyzes the cognitive characteristics of different symbols, and puts forward a mechanism model to show the dominant neural activities in that process. Through the comparative study of Chinese, English, and Spanish, which are used by the largest population in the world, this paper discusses the expressive patterns of emotions in international cyber languages and proposes an intelligent method for affective computing on cyber language in a unified PAD (Pleasure-Arousal-Dominance) emotional space.

1. Introduction

In today's society, cyber space has become an important place for people to share information, exchange opinions, and communicate emotions. Due to its virtuality, autonomy, openness, inclusiveness and the high expressiveness owing to various technologies of new media, the language creativity of people has been inspired to the extreme, giving rise to the booming of cyber languages [1].

Professor Yu at Communication University of China pointed out that cyber language is a "unique natural language" commonly used in cyber space [2]. According to Ferdinand de Saussure's semiotic theory, Chinese scholars have classified cyber languages into the two categories of readable symbols and nonreadable symbols and studied their symbol system, ideographic features, and the formation rules [1, 3, 4]. However, there has not been a consistent definition of cyber language (network language, Internet language, or web language) so far. With the rapid development of modern communication and new media technologies, the Internet, the Internet of things, and the wireless communication network have been integrated into the omnipresent "ubiquitous network" which features increasingly varied expression patterns of cyber language, including icon, audio, video, and text as well as shapes, colors, and brightness. Based on all the findings from previous researches, we define cyber language as "a symbol system that people have agreed on and widely used in communication under the ubiquitous environment."

Cyber languages are very rich in the expression of emotions by either the simple assembly of readable and nonreadable symbols or the complexes of texts, icons, audio or video signs, and their hybrids [5]. Any changes in the component,

shape, color, layout, or presentation sequence may deliver different emotional messages. Cyber language is developing fast in all the world major languages such as Chinese, English, Japanese, German, French, and Spanish [1–3]. In the context of globalization, cyber language has brought new vitality to international language communication and strong capacity for expressing emotions. Therefore, research about emotional characteristics in cyber languages has attracted wide attention from linguistic study to public opinion analysis, internet marketing, service feedback monitoring, and social emergency management.

Affective computing, originally presented by Picard in 1997, indicates that emotional information can be perceived, processed, and computed by machine [6], which has been applied to cyber space in online opinion analysis [7], smart service design [8], psychological monitoring in ubiquitous learning [9], and dynamic emotion computation on vocal social media [10]. In the past decades, although great progress has been made with affective computing on natural language, there are still a lot of difficulties in dealing with cyber language due to its complexity and variability. Also, the cognition of emotional symbols in cyber language is closely related to the neural activities of human beings and affected by such factors as nationality and cultural background, which requires further multidisciplinary research on this issue.

Firstly gave a classification of emotional symbols in cyber language according to the Discovery Learning Theory and then analyzed the cognitive characteristics of different symbols. Based on our previous research findings, a mechanism model to show the dominant neural activities in that process was put forward. In order to analyze the expressive patterns of emotions in international cyber languages, a comparative study of Chinese, English, and Spanish languages was conducted in this paper, and finally an intelligent method was proposed for affective computing on the readable texts and nonreadable symbols in a unified PAD emotional space.

2. Emotional Symbols in Cyber Language and Neural Cognition

2.1. Classification of Emotional Symbols. The symbols that can be used to express emotions in cyber languages are very abundant and continuously innovative and include the simple assembly of readable and nonreadable symbols or the complexes of texts, icons, audio, video signs, and their hybrids [5].

According to the Discovery Learning Theory, the learning process is realized by the learner's cognitive representation, which refers to the mental process of turning perception of external substances into internal mental facts. The manner of cognitive mental representation will experience three stages as people grow up: first enactive representation, second iconic representation, and, third, symbolic representation, which show the sequence in human's cognition of different information types, that is, enactive information comes first and is followed by image information and then text information.

In order to study the impact of different types of information on emotional cognition in cyber languages, we classify the emotional symbols into six categories (ECSAGT) [5, 7]: enactive symbols, color symbols, structural symbols, audio symbols, graphic symbols, and text symbols, each of which delivers emotional messages by following certain encoding and commonly accepted rules.

2.2. Cognitive Characteristics of Emotional Symbols. The analysis of emotional symbols in cyber languages is related to the intention and expression of the information sender as well as the perception and cognition of the information receivers. The emotions of the sender and possible receivers are different, so we should determine that our target is to identify the sender's emotions from his presented symbols or to judge the activated emotions of the receivers by the information of those symbols, which will be evaluated based on the statistical significance [10]. In affective computing, we usually consider the latter.

According to researches in cognitive neuroscience, human emotions arise from the external signals, transmitted through peripheral sensory organs and the internal sensory pathways to the brain's limbic system where the rapid primary emotion is produced, followed by a relatively slow secondary emotion formed in the interaction of the higher cognitive limbic system and the cerebral cortex [11, 12]. This process is controlled by the emotional circuits of the human brain and will give rise to activation responses in corresponding brain regions.

Recent researches into human emotions have been well supported by updated experimental technologies such as EEG (electroencephalograph), ERPs (event-related potentials), fMRI (functional magnetic resonance imaging), and DTI (diffusion tensor imaging). In particular, the blood oxygenation level dependent functional magnetic resonance imaging (Bold-fMRI), with such advantages as being noninvasive, nontraumatic, and capable of locating accurately the activated brain areas, has been applied to the studies of language and emotion's neural mechanism [13–15].

In our previous research organized by Professor Dai et al. at Fudan University, we found that the cognitive responses to different types of symbols varied from one to another through experimental observation by EEG and fMRI [7]. For example, enactive, structural, color, and graphic symbols usually take less time and can give rise to the rapid primary emotions, which we call the primary emotional information. The semantic text symbols will take up more time while perceived by the advanced cortex in the brain. They usually generate the slower secondary emotions and belong to the secondary emotional information. Audio symbols contain both representational information and semantic information and therefore bear the characteristics of the primary and the secondary emotional information, but take more time resources than the first type. The primary emotional information is cross-cultural and independent of languages to a large extent. Once the cognitive rules of emotional symbols come into being, the secondary emotional information plays a vital role in expressing more in-depth emotion. The characteristics discussed above should be considered in the affective computing on the messages composed of mixed types of emotional symbols so as to reflect the dynamic cognitive responses to those symbols.

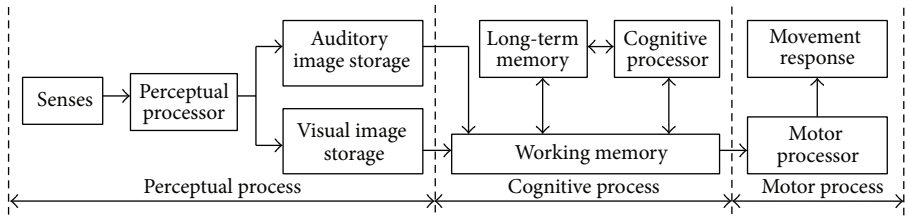


Figure 1: Information process in human processor model.

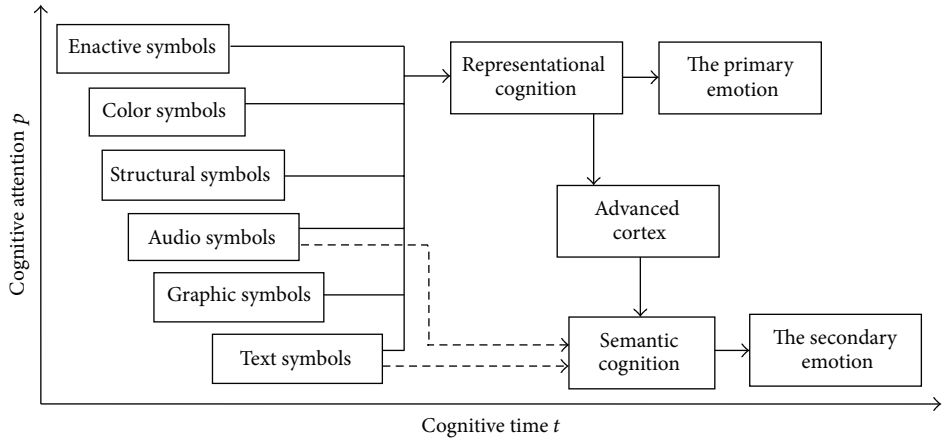


Figure 2: General cognitive characteristics of different types of emotional symbols in cyber language.

Table 1: Time parameters in the Perceptual Process and Cognitive Process.

Parameter	Mean	Range
Decay half-life of visual image storage	200 ms	90–1000 ms
Decay half-life of auditory storage	1500 ms	90–3500 ms
Perceptual processor cycle time	100 ms	50–200 ms
Decay half-life of working memory	7 sec	5–226 sec
Cognitive processor cycle time	70 ms	25–170 ms

As one of the essential issues in Principles of Visual Communication (PVC), the cognitive characteristic of visual constructs has been studied for many years [16]. Furthermore, researchers and engineers in the field of human computer interaction (HCI) have developed effective computational models to measure the time characteristics of different elements in that process [17, 18].

According to the human processor model (a.k.a. MHP) [19], the information process includes three subprocesses: Perceptual Process, Cognitive Process, and Motor Process, as shown in Figure 1. Time parameters in the Perceptual Process and Cognitive Process can be measured as in Table 1 [20].

From Table 1, we can find that the processing time of visual information is usually shorter than that of auditory information. However, in our study, the visual information involves enactive symbols, color symbols, structural symbols, graphic symbols, and symbols, which have different processing time. If only considering the simple action, color, and structure of those symbols in cyber language, the experiment showed that the faster orders of processing time are: enactive symbols, color symbols, structural symbols, audio symbols, graphic symbols, and text symbols in representational cognition [7]. Compared with the audio symbols, graphic symbols can be perceived faster but they need more processing time to be understood in the cognitive process. However, people can usually cognize the semantic meaning of audio symbols much faster than that of text symbols. Therefore, we offer a schematic diagram as Figure 2, which shows the general cognitive characteristics of different types of emotional symbols in cyber language.

2.3. Neural Mechanism Model. The brain mechanism of emotions has been systematically explored by scholars in the area of affective neuroscience [21–23]. In order to provide systemic guidance for analyzing the neural cognition of emotional symbols in cyber language, based on the summary analysis of existing theories and findings, we put forward a mechanism model to show the dominant neural activities in that process as in Figure 3 [5, 9].

The stimulus signals of emotional symbols will be delivered through the receiver's sensory pathways to his limbic system and produce the intuitive primary emotion and then generate the slower but more rational secondary emotion through the cognitive activities of the advanced cortex in the brain. Finally, the changes of emotion will lead to the physiological reactions which are perceived by the brain and form the specific emotional experience. Actually, the subjective assessment of an emotional symbol by the information receivers is the judgment of their emotional experiences.

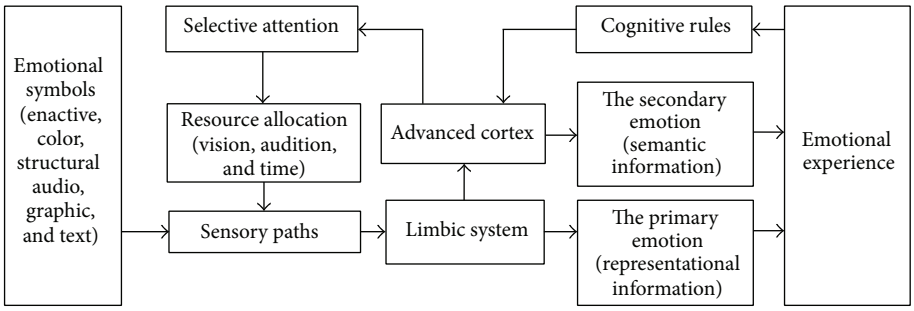


Figure 3: Neural cognition of emotional symbols in cyber language.

In that process, the advanced cortex of the brain regulates its selective attention on the sensory pathways, that is, it distributes the visual, auditory, and time resources autonomously. The selective attention depends on the information receiver's motivation, knowledge, memory, cognition and such advanced psychological activities as decision-making. After experiencing the same symbol multiple times, a cognitive rule will be created in the memory to produce the conventional responses to familiar symbols.

3. Expressive Pattern of Emotions in Cyber Language

3.1. Corpus and Affective Vocabulary. When we look at cyber language closely, we will find that it expresses emotions with readable words or nonreadable symbols that have specific sentiment orientation. They are in general defined as emotional words, which can be found in a large quantity in world's most used languages such as Chinese, English, and Spanish. For example [24], "Bro" is a short for brother. It sounds pretty warm while "lol" expresses joy and refers to "laugh out loud" or "crack up." In addition, there are also many special terms such as "pfffffff," a proud word which means whatever or as you like. "Tmr" is a short for "tomorrow" and refers to something that will be done tomorrow. "N00b" is a newbie or a green hand, usually used to furiously describe those who are clumsy at certain things. "W00t" stands for who or what and is used to describe something or somebody that is exciting or surprising. In some countries, people may use figures or symbols to express their feelings. For example, "1337" represents "elite," a person who is very competent. This number conveys feelings of surprise and joy. The number "56" means boring in English, aburrido in Spanish, and "无聊" in Chinese. The letter "D" is a big laugh. The symbol ":)" is a smile while ":(" means sadness.

A project carried out by the researchers of cyber language worldwide aims to build the corpus by collecting and sorting out frequently used cyber vocabulary and symbols such as General Inquirer, WordNet, and SentiwordNet. The corpus will include a large quantity of affective network terms, which are vital to the analysis of sentiment orientation of cyber language. A case in point is China's Hownet, which has collected 52,000 Chinese terms and 57,000 English words [24]. Among all the published ones, there are 219 words describing the intensity of emotions, 3,116 negative, 1,254 derogatory, 3,730 positive, 836 approbatory, and 38 propositional making a proposition. HowNet's Semantics Dictionary has also included a large collection of lexical semantic entries, each of which is composed of semantics and its description of a term. It offers guidance on how to analyze the above-mentioned affective expressions in a specific context.

The sentiment orientation affective cyber language presents is defined as sentiment polarity, which can be divided into three categories: positive, negative, and neutral. Every affective word's polarity and intensity correlates with emotional expression and cognition standards of cyber language in the specific context. The polarity and intensity of benchmark emotional words in the typical context can be identified through statistical studies of cyber language and has been covered by many corpuses, such as that of Chinese affective lexicon ontology collated and annotated by the Information Retrieval Laboratory of Dalian University of Technology. It provides thorough accounts from different perspectives of a word's part of speech, emotion category, intensity, and polarity and therefore offers important standards for the calculating of parameters related to affective cyber vocabulary.

Cyber languages are characterized by wide-ranging vocabulary that is profuse in sentiments and updated rapidly. The utilization of affective vocabulary is the basic method of the expression of emotions in cyber languages. Researchers worldwide have built the corpuses of cyber languages by collecting and sorting out frequently used vocabulary and symbols.

3.2. Expressive Pattern of Emotions. The emotional message of cyber language is decided not only by the affective vocabulary used in the sentence, but also by the expressive pattern in the whole sentence. Therefore, the same affective vocabulary can be completely opposite in meaning when expressed in a different pattern. For example, "When I just heard the news, I was quite upset. But after having lurked for a while, I found hikers were right about that, so just want to show up today to share my joy." There are two affective words in this sentence, the negative "upset" and the positive "joy" together with a concession word "but." Supposing P stands for the positive emotion word, N for the negative one and

Table 2: Commonly used connectors in Chinese, English, and Spanish.

Functions of connectors	Chinese	English	Spanish
Alternatives	① 不是…就是…	① Either…or…	① ni…ni…
	② 即不是…也不是…	② Neither…nor…	② no…tampoco…
	③ 或者	③ Or	③ y
	④ 以及	④ As well as	④ tambien
	⑤ 和, 与, 并且	⑤ And	⑤ y
	⑥ 和…都	⑥ Both…and…	⑥ ambos…y…
Cause and effect	① 因此	① Therefore	① por lo tanto
	② 所以	② So	② así que
	③ 总之, 因此	③ As a result	③ por consiguiente
	④ 由于	④ Because of	④ por
	⑤ 基于, 由于	⑤ Due to	⑤ gracias a
	⑥ 因为	⑥ Because	⑥ es que
Concession	① 还未, 仍然没有	① Yet	① aún
	② 但是, 但	② But	② pero
	③ 而, 正当	③ While	③ sino
	④ 相反地说, 而是	④ On the contrary	④ por el contrario
	⑤ 可是, 不过, 然而	⑤ However	⑤ sin embargo
	⑥ 然而	⑥ At the same time	⑥ mientras
Conclusion/summary	① 总之	① In a word	① en fin
	② 整体上, 大体	② On the whole	② en general
	③ 简单地说, 简言之	③ In brief	③ cortar el rollo
	④ 总结下就是说	④ To conclude	④ para concluir
	⑤ 总共, 总计	⑤ In all	⑤ en total
	⑥ 概括下说	⑥ To sum up	⑥ resumir
Examples	① 举例来说	① For example	① por ejemplo
	② 在那个案例上	② In that case	② en eso caso
	③ 解释下, 说明下	③ To illustrate	③ por ilustrar
	④ 一方面地讲	④ For one thing	④ por una parte
	⑤ 打比方说, 比如说	⑤ Such as	⑤ tal como
	⑥ 比如, 譬如	⑥ For instance	⑥ entre ellos

T for the concession, the emotion expression pattern of the above sentence can be generalized as the follows:

$$N + T + P \Longrightarrow P. \tag{1}$$

In cyber language, conjunctions are critical to the understanding of the emotional messages delivered in sentences and thereafter are important objects to be considered in the analysis of emotional structure and expression patters. Table 2 has included some commonly used connectors in Chinese, English, and Spanish [24].

Of course, a complete emotional expression pattern also involves degree, negative words, and punctuation marks. For example, the Chinese words "不很高兴," "不高兴," and "很不高兴" express unhappiness of very different degrees. Punctuation marks such as "!", "?", "…" and emoticons, in particular, demonstrate very distinct sentiment orientations. In addition, the sequence of affective words in a sentence will make a difference. For example, the English sentence "We are exhausted now; above all we are so happy for our success." adopts the pattern of "N + above all + $P \Rightarrow P$". The emotion of the whole sentence is determined by the phrase following "above all," which is "so happy." Without strict rules regulating cyber language expression, which is ever changing, we will have to use computers to automatically capture new entries and modify them with subjective cognition in order to build open corpuses of expressive patterns of emotions in sentences and analyze the messages delivered by them.

What remains a research issue is the emotional expression through some nonreadable symbols such as enactive symbols, audio symbols, structural symbols, color symbols, and graphic symbols. As with in text symbols, the most effective approach to affective computing on those emotional symbols is to establish an open and frequently updated knowledge library based on the ontology of nonreadable symbols, which is correlated with the cultural background, language context, and social environment and should be processed with emotional notations considering the language application environment [5, 9]. Among those symbols, the emotions in audio symbols are mainly reflected in the speed, intensity, pitch frequency, and spectral parameters of the audio signals and can be highly cross-cultural and cross

Table 3: Chinese version of the questionnaire.

Question	Emotion	−4	−3	−2	−1	0	1	2	3	4	Emotion
Q1	Angry										Activated
Q2	Wide-awake										Sleepy
⋮	⋮										⋮
Q12	Influential										Influenced

Q1: angry-activated; Q2: wide-awake-sleepy; Q3: controlled-controlling; Q4: friendly-scornful; Q5: calm-excited; Q6: dominant-submissive; Q7: cruel-joyful; Q8: interested-relaxed; Q9: guided-autonomous; Q10: excited-enraged; Q11: relaxed-hopeful; Q12: influential-influenced.

languages. In an online conversation, an high accuracy rate of emotion recognition can be achieved in the way of pattern recognition without even semantic analysis [10, 25–27].

4. Affective Computing on Cyber Language

4.1. Unified PAD Emotional Space. In order to be processed by the machine, the affective characteristics of an emotional symbol in cyber language should be described quantitatively. The rudimentary description is its positive or negative polarity with quantitative intensity. In most cases, researchers propose the "six big" types of emotions: anger, disgust, fear, joy, sadness, and surprise [28], which have been widely applied to the analysis of graphic, audio, and video signals. However, the emotional symbols in cyber language usually carry mixed affective characteristics and reflect dynamic changes in audio and video signals.

Mehrabian presented a 3D model which can describe any kinds of complicated emotions in a PAD emotional space [29, 30]. It includes three nearly independent continuous dimensions: Pleasure-Displeasure (P), Arousal-Nonarousal (A), and Dominance-Submissiveness (D). Experiment shows that all the known emotion states can be almost described in this space very well [10, 30]. So far, the PAD model has been successfully applied in a variety of areas, such as audio-visual speech synthesis [31], micro-blog sentiment analysis [32], and music emotion comparison [33]. The 3D dimensions of PAD provide a unified space for describing the mixed affective characteristics of all types of emotional symbols in cyber language as well as their dynamic changes. As well, any emotional state in the PAD space can also be described as the percentage rates of typical emotions based on a conversion metric function [10, 34].

Usually, the PAD values of commonly used emotional symbols should be firstly evaluated by subjective assessment according to the perception and cognition of the typical information receivers in the process of emotional notation and stored in the knowledge library based on the ontology, therefore providing the references for comprehensive computation by machine. In order to achieve precise and consistent results in subjective evaluation, Mehrabian designed an initial 34-item test questionnaire [29]. However, the application in practice indicates that the questionnaire should be designed according to the specific language due to the differences in language understanding and cultural backgrounds. Table 3 shows the Chinese version of the simplified 12-item questionnaire which was presented by the scholars from the Psychological Institute, Chinese Academy of Sciences [35].

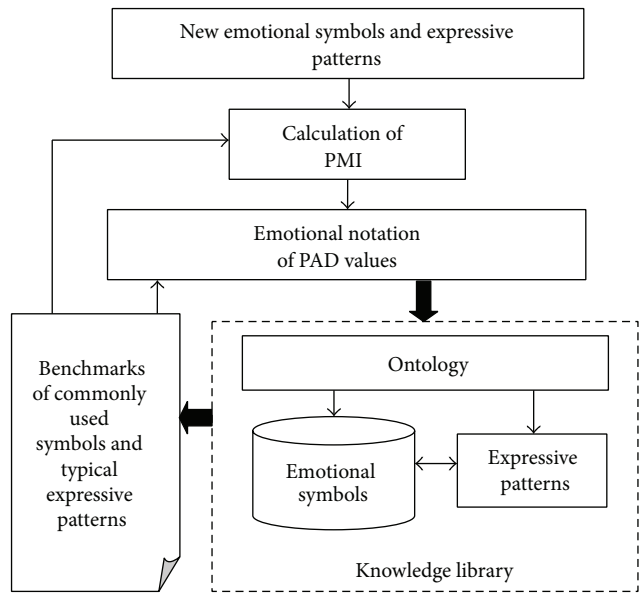


Figure 4: Knowledge library and emotional notation.

The assessment is based on what kind of feeling is more intense in each item. From the left to the right, it is scaled within the range from "−4" to "4." The scores in each item may be calculated and converted into the normalized values of P, A, and D [35].

4.2. Knowledge Library and Emotional Notation. The perception and cognition of emotional symbols in cyber languages are not only related to neural cognition but also affected by the information receivers' background and features. Therefore, we establish an open and frequently updated knowledge library as shown in Figure 4.

The emotional symbols as well as their expressive patterns are stored in this library based on the ontology. The commonly used symbols in the library should be firstly assigned with the emotional notation of PAD values as the benchmarks through the subjective assessment which has been discussed before in this paper.

In order to evaluate the values of the rest or a new emotional symbol, we adopted the PMI (Pointwise Mutual Information) method [36], which is based on the probabilities of the new symbol and its benchmarks in the knowledge library. For example, the improved HowNet-PMI algorithm

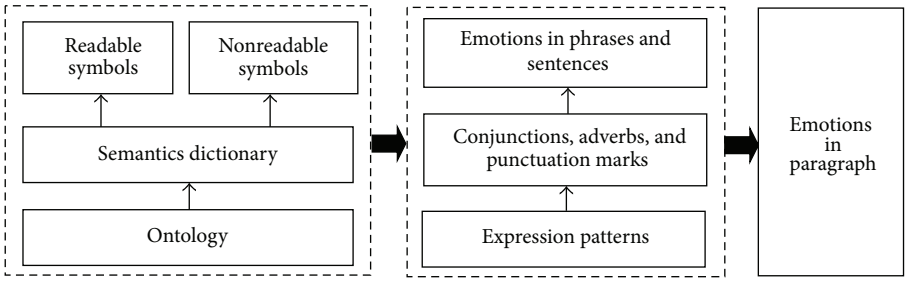


Figure 5: Semantic dictionary and expressive patterns in knowledge library.

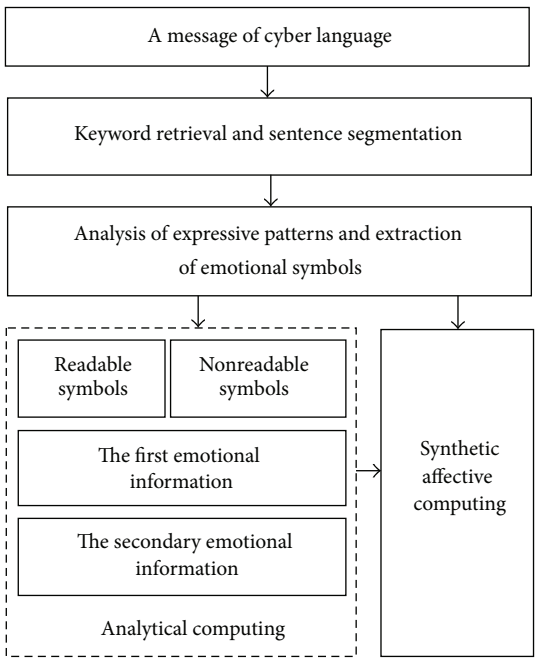


Figure 6: Basic process of affective computing on cyber language.

has been successfully applied to affective computing on Chinese cyber language [37].

As shown in Figure 5, in the knowledge library, we build up a semantic dictionary based on the ontology of international cyber languages, which includes both the readable and nonreadable symbols. The expressive patterns of emotions in sentences are represented by the knowledge such as the templates and rules which describe the commonly used structures along with the conjunctions, adverbs, and punctuation marks.

Based on the knowledge, affective computing can be carried out on a whole message which may contain one or more sentences with the additional nonreadable symbols.

4.3. Intelligent Computing Method. The basic process of affective computing on cyber language is shown as in Figure 6. It includes the following four steps [5]:

(1) The first step is Keyword Retrieval and Sentence Segmentation. A message of cyber language will be segmented into one or more sentences with possible additional nonreadable symbols for further processing by the structure analysis. Chinese, English, and Spanish have their different patterns in sentences, which can be mostly structured by the retrieved keywords such as conjunctions, adverbs, and punctuation marks.

(2) The second step is Analysis of Expressive Patterns and Extraction of Emotional Symbols. In this step, each segmented sentence needs to be broken up into a series of separate words by dedicated tools such as the Chinese word segmentation system NLPIR [38]. Hereafter, the expressive patterns of emotions in sentences with additional nonreadable symbols are analyzed, and all emotional symbols in the message are extracted based on the semantics dictionary, structural templates, and rules which are stored in the knowledge library.

(3) The third step is Analytical Computing. In this step, the readable and nonreadable emotional symbols are computed separately. In order to reflect the dynamic affective features in human cognitive processes, the enactive symbols such as flashing signs and video signs, as well as structural, color, and graphic symbols, are computed and ranged into the primary emotional information. The semantic text symbols are computed into the secondary emotional information. Audio symbols contained both the, especially, as discussed before in this paper, it contained the representational information and the semantic information. The former is related to vocal emotion only and can be computed by the LS-SRV estimator, which has been successfully applied to Wechat and QQ [10]. The later should be firstly converted into text sentences by a speech recognition tool and is afterwards computed similarly to the text symbols.

(4) The final step is Synthetic Affective Computing. The results by the step of Analytical Computing will hereafter be synthesized and adjusted based on the analysis of conventional expressive patterns, in order to reach a more accurate and comprehensive result. Dynamic affective features in the message of cyber language may be represented by the primary and the secondary emotions as well as the changing positions in a text sentence.

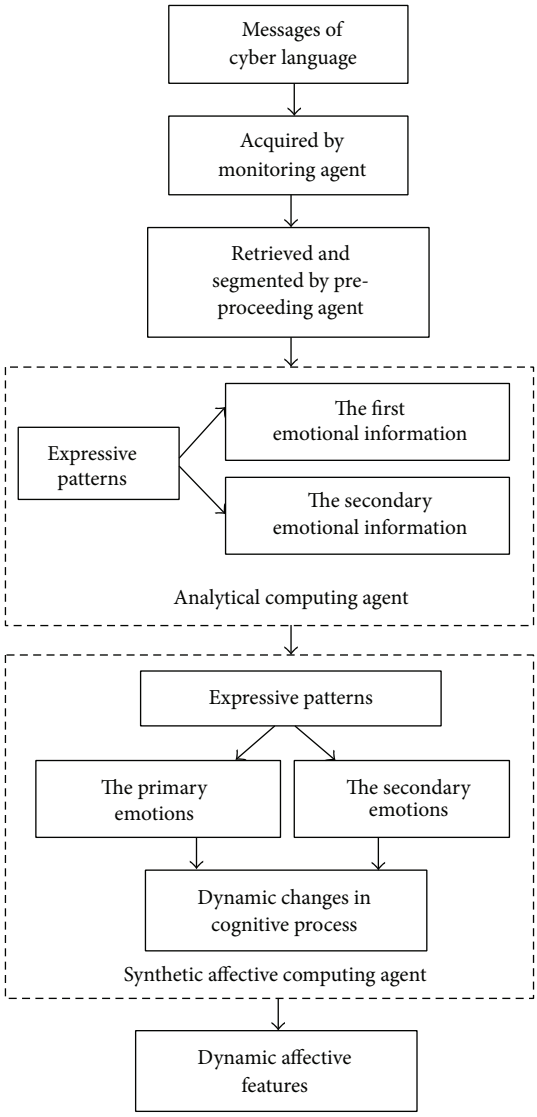


Figure 7: Intelligent computing method.

This process in Figure 6 can be implemented by the intelligent computing method as shown in Figure 7.

In the proposed method, the processing and computing tasks are accomplished by a multiagent system intelligently. It includes a Monitoring Agent, a Preprocessing Agent, an Analytical Computing Agent, and a Synthetic Affective Computing Agent. Different from the traditional computing on emotions, considering the characteristics of neural cognition, our method can give the results of the primary and the secondary emotions, respectively, and shows their dynamic changes. This is significant for the further analysis of emotion propagation through social media in cyber space [39–41].

5. Experiment and Result

Figure 8 is a piece of sports news reported by the famous Spanish newspaper *Mundo Deportivo* and the readers' online comments. There are total 72 comments on this news [24].

Figure 8: Sports news on *Mundo Deportivo* and the readers' online comments.

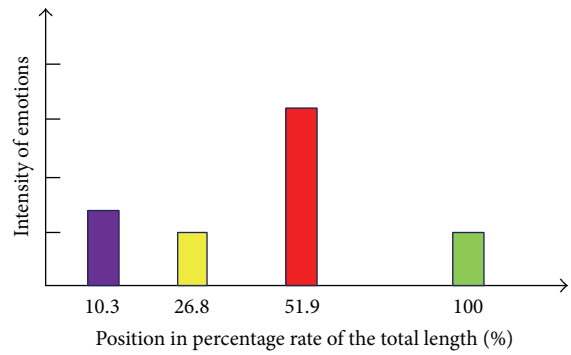


Figure 9: Dynamic changes of affective features in the online comments.

Once seeing this footage, the audiences are firstly attracted by the smiling pictures and videos of the news reporter and produce the primary joy emotion immediately the primary emotion. After computing on the online comments, we get the main types of emotions in different positions of the total comment text as shown in Figure 9, which reflect the dynamic changes of the secondary emotion in the online comments.

Figure 10 is a feeling description posted by a Chinese survivor on January 1, 2015, at 3:53, who escaped from the miserable Shanghai bund stampede which took place at 23:35 on December 31, 2014, and resulted in 36 deaths and 49 injuries. The survivor said: "Tonight's Bund was nothing I could've imagined because of the crowding and trampling accident. I was fortunate to have survived. I saw young lives perished in front of me, but I couldn't save them. They were put on stretchers and sent down to us one after another. We tried CPR for all, one, two, three... until we were all exhausted. Poor people, hopefully 120 and the hospital can do a better job. Have to thank the medical workers, foreigners as well as all the others that participated in the rescue efforts. We have tried our best ... (crying out loud)."

Figure 11 shows the PAD values of affective computing result on that description. It represents the strong intensity of the mixed emotions as well as their changes which include grief, despair, helplessness, and the gratefulness to the people who were involved in this rescue.

Figure 10: Feeling description by a survivor in Shanghai bund stampede.

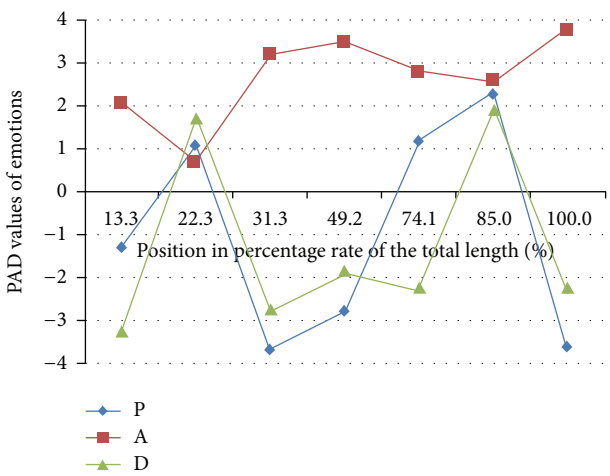


Figure 11: PAD values of affective computing result.

6. Conclusion and Discussion

With the rapid development and wide application of the Internet and ubiquitous networks, cyber space has provided people with a new virtual society and convenient living and working platform. Characterized by its customary symbol system and vivid expression patterns, cyber language not only acts as the tool for people to communicate in the cyber space, but also plays a vital role in affective exchange and emotional propagation as well as social psychology and behaviors and has caused high attention in many areas in recent years.

Due to the open, virtual, and dynamic language environment, affective computing on cyber language requires a systemic and interdisciplinary research. This paper presented a classification of the emotional symbols in cyber language and put forward a mechanism model to show the dominant neural activities in the cognitive process. Furthermore, after analyzing the expressive patterns of emotions in the languages of Chinese, English, and Spanish, this paper proposed an intelligent method for affective computing on cyber language in a unified PAD emotional space, which can deal with the multi symbol information and mixed emotions in a cyber message and show their dynamic changes according to the characteristics of the neural cognition process. Experimental results indicate that this method can reach an accuracy of more than 70% for the computing on text symbols and audio symbols and provide an effective approach to the application of a lot of areas such as public opinion analysis, internet marketing, service feedback monitoring, and social emergency management. However, the processing of the remaining nonreadable symbols had to be made by subjective evaluation in most cases.

In the future, the language ecosystem of cyber language and new media technologies will be ever changing and continuously updated. We suggest that future studies can be conducted in the following areas: (1) How to build an open and dynamic updated knowledge library to assist the affective computing by applying intelligent monitoring and big data mining techniques; (2) how to establish more thoroughly expressive emotional patterns and provide statistical fundamental parameters for the elaborate description of neural cognition by using advanced experimental observation techniques; and (3) how to explore a more effective method for computing on nonreadable symbols such as enactive and structural symbols.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported comprehensively by the National Natural Science Foundation of China (no. 91324010), the National Social Science Fund of China (14BXW046), the 2nd Regular Meeting Project of Science and Technology Cooperation between China and Serbia (no. 2-9/2013), Shanghai Decision Making and Advising Project (2014-GR-55), Shanghai Philosophy and Social Sciences Plan, China (no. 2014BGL022), Project of Scientific Research Fund of Sichuan Provincial Education Department (no. 15SB0219), and Project of Sichuan Conservatory of Music (CY2014173). Research work of this paper was accomplished by our interdisciplinary team based on the previous work [5, 7, 9]. Shuang Huang and Xuan Zhou are the joint first authors of this paper. Ke Xue and Xiqiong Wan are the joint corresponding authors.

References

[1] W. T. Chen, *The Study on Symbols Formation and Development of Cyber Language*, Northweast Normal University, Lanzhou, China, 2010.

[2] Y. G. Yu, *The Generality of Network Language*, China Economic Publishing House, Beijing, China, 2001.

[3] H. F. He, "Web language as viewed from perspective of symbolism," *Journal of Jianghan University (Humanities Sciences)*, vol. 22, no. 2, pp. 74–78, 2003.

[4] Y. F. Lu, *A Study on the Representing Principles of Netspeak Symbols*, Harbin Normal University, Harbin, China, 2013.

[5] S. Huang, X. Zhou, W. H. Dai, and M. Ivanović, "Neural cognition and intelligent computing on the emotional symbols of cyber language," in *Proceedings of the International Conference on Intelligent Computing and Internet of Things (ICIT '15)*, pp. 164–167, Harbin, China, January 2015.

[6] R. W. Picard, *Affective Computing*, MIT Press, Cambridge, Mass, USA, 1997.

[7] W. H. Dai, L. H. Huang, D. R. Xu et al., *Information Classification, Dissemination and Cognitive Model in Unconventional Emergency*, National Natural Science Foundation of China, Beijing, China, 2015.

[8] Y. G. Zhou, *On-Demand Smart Services Based on Context Awareness*, Fudan University, Shanghai, China, 2012.

[9] W. H. Dai, S. Huang, X. Zhou, X. E. Yu, M. Ivanović, and D. R. Xu, "Emotional intelligence system for ubiquitous smart foreign language education based on neural mechanism," *Journal of Information Technology Applications and Management*, vol. 21, no. 3, pp. 65–77, 2015.

[10] W. H. Dai, D. M. Han, Y. H. Dai, and D. R. Xu, "Emotion recognition and affective computing on vocal social media," *Information & Management*, 2015.

[11] A. R. Damasio, *Descartes Error: Emotion, Reason and the Human Brain*, Gosset/Putnam Press, New York, NY, USA, 1994.

[12] C. E. Izard, *The Psychology of Emotions*, Plenum Press, New York, NY, USA, 1991.

[13] Y. H. Wang, *An fMRI study of how a supra-modal decoding network exams interpretation [Ph.D. thesis]*, Tamkang University, Taipei, Taiwan, 2014.

[14] R. Horlings, *Emotion Recognition Using Brain Activity*, Delft University of Technology, Delft, The Netherlands, 2008.

[15] Y. H. Wang, X. H. Hu, W. H. Dai, J. Zhou, and T. Z. Kuo, "Vocal emotion of humanoid robots: a study from brain mechanism," *The Scientific World Journal*, vol. 2014, Article ID 216341, 7 pages, 2014.

[16] K. G. Barnhurst, M. Vari, and Í. Rodríguez, "Mapping visual studies in communication," *Journal of Communication*, vol. 54, no. 4, pp. 616–644, 2004.

[17] M. Bauerly and Y. L. Liu, "Computational modeling and experimental investigation of effects of compositional elements on interface and design aesthetics," *International Journal of Human-Computer Studies*, vol. 64, no. 8, pp. 670–682, 2006.

[18] S.-W. Hsiao and J.-R. Chou, "A Gestalt-like perceptual measure for home page design using a fuzzy entropy approach," *International Journal of Human Computer Studies*, vol. 64, no. 2, pp. 137–156, 2006.

[19] S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1986.

[20] 2015, http://en.wikipedia.org/wiki/Human_processor_model.

[21] A. R. Damasio, T. J. Grabowski, A. Bechara et al., "Subcortical and cortical brain activity during the feeling of self-generated emotions," *Nature Neuroscience*, vol. 3, no. 10, pp. 1049–1056, 2000.

[22] R. J. Davidson, D. Pizzagalli, J. B. Nitschke, and K. Putnam, "Depression: perspectives from affective neuroscience," *Annual Review of Psychology*, vol. 53, pp. 545–574, 2002.

[23] Y. J. Luo, T. T. Wu, and R. L. Gu, "Studies on neural correlates of emotion and cognition," *Bulletin of Chinese Academy of Sciences*, supplement 1, pp. 31–41, 2012.

[24] Z. Y. Yang, *The Method of Emotional Calculation and Research of Realizing Technology of International Network Language*, Fudan University, Shanghai, China, 2014.

[25] X. C. Jin, *Emotion Recognition Research Based on Speech Signals*, University of Science and Technology of China, Hefei, China, 2007.

[26] D. Ververidis and C. Kotropoulos, "Emotional speech recognition: resources, features, and methods," *Speech Communication*, vol. 48, no. 9, pp. 1162–1181, 2006.

[27] B. Li, *Research on Emotional Recognition in Multilingual Speech Signal*, University of Jinan, Jinan, China, 2010.

[28] P. Ekman and M. J. Power, *Handbook of Cognition and Emotion*, John Wiley & Sons, Sussex, NJ, USA, 1999.

[29] A. Mehrabian, "Framework for a comprehensive description and measurement of emotional states," *Genetic, Social, and General Psychology Monographs*, vol. 121, no. 3, pp. 339–361, 1995.

[30] A. Mehrabian, "Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.

[31] J. Jia, S. Zhang, F. B. Meng, Y. X. Wang, and L. H. Cai, "Emotional audio-visual speech synthesis based on PAD," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 3, pp. 570–582, 2011.

[32] H. T. Cao, *Chinese Microblog Sentiment Analysis Based on the PAD Model*, Dalian University of Technology, Dalian, China, 2013.

[33] K. F. MacDorman, S. Ough, and C.-C. Ho, "Automatic emotion prediction of song excerpts: index construction, algorithm design, and empirical comparison," *Journal of New Music Research*, vol. 36, no. 4, pp. 281–299, 2007.

[34] P. H. Sun and L. M. Tao, "Emotion measuring method in PAD emotional space," in *Proceedings of the 4th Harmonious Man-Machine Environment Joint Academic Conference*, pp. 638–645, 2008.

[35] Y. Lu, X. L. Fu, and L. M. Tao, "Emotion measurement based on PAD 3-D space," *Communications of the CCF*, vol. 6, no. 5, pp. 9–13, 2010.

[36] P. D. Turney and M. L. Littman, "Unsupervised learning of semantic orientation from a hundred-billion-word corpus," Tech. Rep. EGB-1094, National Research Council of Canada, 2002.

[37] X. D. Zhao, *The Simulation Research on Information Emotional Cognition and Dissemination in Public Safety Incidents*, Fudan University, Shanghai, China, 2014.

[38] NLPIR, 2014, http://ictclas.nlpir.org/.

[39] J. Bollen, H. Mao, and X. J. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[40] W. Dai, H. Hu, T. Wu, and Y. Dai, "Information spread of emergency events: path searching on social networks," *The Scientific World Journal*, vol. 2014, Article ID 179620, 7 pages, 2014.

[41] I. Kompatsiaris, D. Gatica-Perez, X. Xie, and J. Luo, "Special section on social media as sensors," *IEEE Transactions on Multimedia*, vol. 15, no. 6, pp. 1229–1230, 2013.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 698725, 6 pages
http://dx.doi.org/10.1155/2015/698725

Research Article

Neural Basis of Intrinsic Motivation: Evidence from Event-Related Potentials

Jia Jin,[1] Liping Yu,[1] and Qingguo Ma[2,3]

[1]*Business School, Ningbo University, Ningbo 315211, China*
[2]*School of Management, Zhejiang University, Hangzhou 310027, China*
[3]*Neuromanagement Lab, Zhejiang University, Hangzhou 310027, China*

Correspondence should be addressed to Qingguo Ma; maqingguo3669@zju.edu.cn

Received 14 January 2015; Accepted 5 April 2015

Academic Editor: Ye-Sho Chen

Human intrinsic motivation is of great importance in human behavior. However, although researchers have focused on this topic for decades, its neural basis was still unclear. The current study employed event-related potentials to investigate the neural disparity between an interesting stop-watch (SW) task and a boring watch-stop task (WS) to understand the neural mechanisms of intrinsic motivation. Our data showed that, in the cue priming stage, the cue of the SW task elicited smaller N2 amplitude than that of the WS task. Furthermore, in the outcome feedback stage, the outcome of the SW task induced smaller FRN amplitude and larger P300 amplitude than that of the WS task. These results suggested that human intrinsic motivation did exist and that it can be detected at the neural level. Furthermore, intrinsic motivation could be quantitatively indexed by the amplitude of ERP components, such as N2, FRN, and P300, in the cue priming stage or feedback stage. Quantitative measurements would also be convenient for intrinsic motivation to be added as a candidate social factor in the construction of a machine learning model.

1. Introduction

Human intrinsic motivation is concerned by both academic research and practical application for its great significance for human behavior. Therefore, it has gained considerable attention from scientists and educationalist for decades. However, for the intrinsic motivation it was always difficult to be directly measured and observed, and its explicit impact on human behavior was unclear which went against learning human's behavior.

In recent years, the rapid development of the neuroscience techniques made it possible for us to open the "black box" of our brain and observe people's neural responses directly. For the study of intrinsic motivation, researchers also considered if it was possible to probe it at brain level. Quirin et al. [1] employed functional magnetic resonance imaging (fMRI) to investigate the different motivation of power and affiliation, in which they found power-related versus affiliation-related social motivations had differential brain networks. In another study, Murayama et al. (2010) [2] also employed fMRI to investigate the neural evidence of interaction between intrinsic and extrinsic motivation at the spatial level. At the feedback stage, they found that BOLD signal in ventral striatum was prominently decreased when the extra reward for performance was removed at a later session of the task, while such a phenomenon was not observed in the control group where no performance-based monetary reward was provided for both sessions. These studies suggested that the variation of human intrinsic motivation could be reflected in brain activity which could not be early measured at behavior level.

In addition to fMRI, event-related potential (ERP) was another widely used neuroscience tool which can make up the temporal dynamic accuracy of fMRI. Therefore, in the current study we employed ERPs to compare the neural discrepancy of the two tasks with different levels of intrinsic fun throughout the whole task process. The purpose of our study was to explore the electrophysiological dynamics of the human intrinsic motivation through EEG recordings. According to previous study about ERPs and motivation, we supposed that three ERP components would appear in

the current experiment, N2 in the cue priming stage and FRN and P300 in the feedback stage.

The negative component N2 was reported peaking around 200 to 300 ms after the onset of a stimulus [3, 4]. Previous studies found that the visual N2 component was related to the deviation from the perception of the target cognitive control and action inhibition processes [4]. For example, Eimer (1993) [5] employed the go/no-go paradigm and conducted two experiments, in which the participants were asked to respond to a letter (go stimulus) but not to another (no-go stimulus). Results showed that the N2 enhancement elicited by the no-go stimuli was larger than that induced by the go stimuli, which reached its maximum in frontal areas. They suggested that this was because of the response mismatch and action inhibition. Subsequent studies also explained the larger nontarget N2 amplitude as subjects' inhibition of an anticipated response to the target [6].

FRN was another candidate component which would be found in the current study. It was reported in various tasks that FRN was related to the affective/motivation at the feedback stage [7, 8]. In an early study conducted by Gehring and Willoughby (2002) [8], they found a prominent differentiated FRN (d-FRN) toward the divergence of the loss gain feedback, which was suggested to reflect the subjective motivational and affective evaluation of the revealed outcome. Additionally, further studies also confirmed that the evaluative process indexed by FRN is sensitive to the motivational significance of ongoing event [7, 9–11]. For instance, in Ma et al.'s [10] 2014 work, they found that high effort could induce larger differentiated FRN responses to the reward and nonreward discrepancy across two experimental conditions. They suggested that this was because effort might increase subjective evaluation toward subsequent reward which was reflected in the FRN amplitude deflection.

In the outcome feedback stage, there always appeared another important ERP component, P300, which was always examined accompanying FRN. P300 was a positive ERP component peaking around 200–600 ms after the onset of feedback [12]. It was reported that P300 was sensitive to the magnitude [13, 14] and the valence of reward [11, 12, 15]. Furthermore, previous studies also agreed that P300 could also represent the attentional allocation and motivational/affective significance [10, 16, 17]. For instance, one of our recent studies [17] adopted a gambling task in the social context and was found independent of FRN; there was a general P300 divergence across agents of different degrees of closeness to the subjects which suggested that the valence effect of P300 could reflect motivational/affective implication of the outcome.

In the current study, in order to investigate the internal motivation of two tasks, we intended to compare electrophysiological response at the cue priming stage and feedback stage. According to the literature mentioned above, we expected that there would be a N2 component discrepancy in priming stage, in which the boring WS task would elicit a N2 enhancement compared with the interesting SW task, which suggested subjects' expectation to the interesting task. When it came to the feedback stage, we supposed that, compared with WS task, smaller FRN amplitude accompanied by larger P300 amplitude would appear in SW task, reflecting the higher subjective motivation to the interesting task's outcome.

2. Materials and Methods

2.1. Subjects. Sixteen healthy native Chinese graduate and undergraduate students (8 males), aged from 18 to 25 (mean age = 23.23; SD = 1.78) were enrolled. All of them with self-reported right-handedness and had normal or corrected-to-normal vision and did not have any history of neurological disorder or mental disease. Prior to the commencement of the experiment, informed consent was obtained from all participants. The study was also approved by the Internal Review Board of Zhejiang University Neuromanagement Lab.

2.2. Stimuli. The experiment included two blocks, and there were 45 trials in each block. Two tasks with different intrinsic motivation were adapted from Murayama et al.'s 2010 work [2]. In stop-watch (SW) task, the subjects were asked to stop an automatically started watch by pressing a button. They won the current trial only if the time of the watch finally fell within a specific deviation from 5 s time point. In order to make sure that the participants can succeed approximately 50% trials on average, a pilot study of thirty students was conducted before the formal experiment to confirm the time deviation. According to the result of the pilot study, the time duration of winning was determined as 70 ms deviation from 5 s time point. When it comes to the watch-stop (WS) task, the watch stopped automatically and the participants were only asked to simply press the button when it stopped. The stop timing for WS trials was varied between 4.2 and 5.8 seconds randomly, in purpose of matching the time duration of SW trials generally. There existed a 600–1000 ms randomized blank interval between trials. In each trial, a task cue was first presented for 2000 ms, indicating which task would be performed. After 600–1000 ms interval of cue onset, the task started and outcome of the performance was revealed for 2 s and interval across tasks was varied between 800 and 1200 ms. Stimuli were presented sequentially in the center of the CRT computer screen ($6.2^\circ \times 6.2^\circ$).

2.3. Procedure. In a shield room participants were comfortably seated 1 m away from a computer-controlled CRT monitor. Subjects were provided with a keypad to make their responses. They were instructed to complete one of the two tasks in each trial according to the cue instruction. The formal experiment started after a pilot practice. Participants were also asked to minimize body and muscle movements during the experiment. Stimuli, recording triggers, and responses were presented and recorded using E-Prime 2.0 software package (Psychology Software Tools, Pittsburgh, PA, USA).

2.4. EEG Recordings and Analyses. For the data recording, EEG was recorded with an electrode elastic cap with 64 Ag/AgCl electrodes according to the standard international 10–20 system and Neuroscan Synamp2 Amplifier (Scan 4.3.1, Neurosoft Labs, Inc., Virginia, USA). The sampling rate was 500 Hz and with band-pass 0.05–70 Hz. A frontal electrode

site between FPz and Fz was used for ground and left mastoid was chosen for reference. Electrooculogram (EOG) was also recorded from electrodes placed at 10 mm from the lateral canthi of both eyes (horizontal EOG) as well as above and below the left eye (vertical EOG). The experiment started when the electrode impedances were maintained below 5 kΩ.

For the data analysis, Neuroscan 4.5 software was used. The EOG artifacts were corrected offline for all subjects during preprocessing, which were corrected using the method initially proposed by Semlitsch et al. (1986) [18]. Trials containing amplifier clipping, bursts of electromyography activity, or peak-to-peak deflection exceeding ±100 μV were excluded from final analysis. Data was then transferred to the average of the left and right mastoids reference offline. ERPs were digitally filtered with a low pass filter at 30 Hz (24 dB/octave).

The EEG recordings were segmented for the epoch from 200 ms before the onset of target to 800 ms after the onset. The first pretarget of 200 ms was regarded as the baseline. In cue stage analysis, data was collapsed based on the two kinds of task cues. Based on visual observation of grand-average waveforms and previous ERP guidelines of Picton et al. (2000) [19], N2 component was analyzed. According to the scalp distribution of N2 and the previous studies [20, 21], we chose time range of 270–350 ms and selected nine electrode sites, namely, F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2, in frontal and central areas for statistical analysis. Repeated measure ANOVAs were conducted to examine the effect of N2 difference of the two task cues.

For the analysis of outcome feedback, there were three conditions, WS feedback and winning and failing results in SW task. Based on visual observation of grand-average waveforms and previous ERP reports on outcome feedback [7, 9], two ERP components, FRN and P300, were analyzed. According to the scalp distribution of FRN and the previous studies [7, 8], we chose time range of 160–200 ms and selected nine electrode sites, namely, F1,Fz, F2, FC1, FCz, FC2, C1, Cz, and C2, in frontal and central areas where it elicited the largest FRN amplitude, for statistical analysis. Similarly, we chose time window of 250–350 and nine electrode sites C1, Cz, C2, CP1, CPz, CP2, P1, Pz, and P2 for the analysis of P300. Similar repeated measure ANOVAs were also conducted for FRN and P300. The Greenhouse-Geisser [22] correction was applied in all statistical analyses when necessary (uncorrected *df* are reported with the ε and corrected P values), and the Bonferroni correction was used for multiple paired comparisons.

3. Results

As shown in Figure 1, repeated measure ANOVA results of N2 revealed significant main effect of cue category ($F(1, 15) = 6.252$, $P = 0.024$, $\eta^2 = 0.294$) while the main effect of electrodes ($F(8, 120) = 2.200$, $P = 0.093$, $\varepsilon = 0.419$) and interaction effect of cue and electrodes were not observed ($F(8, 120) < 1$). The mean amplitude of N2 showed cue of WS task (mean = 1.374 μV, SD = 1.256) elicited a larger N2 (negative polarity, smaller voltage means larger amplitude) amplitude than that of SW task (mean = 3.212 μV, SD = 1.154).

The general waveform of outcome feedback was shown in Figure 2. Repeated measure ANOVA results of FRN showed significant main effect of outcome valence ($F(2, 30) = 38.938$, $P = 0.000$, $\eta^2 = 0.722$). Pairwise t-test showed that the winning trials (mean = 7.916, SD = 1.004) elicited smaller FRN (negative polarity, smaller voltage means larger amplitude) amplitude than that of failing trials ($P < 0.001$, mean = 7.916, and SD = 1.004) and WS trials ($P < 0.001$, mean = 7.916, and SD = 1.004) while failing trials also showed a smaller FRN amplitude than WS trials ($P = 0.017$). On the other hand, the results of P300 also showed a similar effect. The main effect of P300 was observed ($F(2, 30) = 36.061$, $P = 0.000$, $\eta^2 = 0.706$) and pairwise t-test also showed the winning trials (mean = 14.575, SD = 1.095) elicited larger P300 (positive polarity, larger voltage means larger amplitude) amplitude than that of failing trials ($P < 0.001$, mean = 11.216, and SD = 1.468) and WS trials ($P < 0.001$, mean = 72.896, and SD = 0.705) while loss trials also showed a larger P300 amplitude than WS trials ($P = 0.049$).

4. Discussion

This study was carried out to explore the temporal dynamics of human intrinsic motivation, which is an important facet of human behavior. We investigated how a particular task affects the subjects' intrinsic motivation by giving an interesting stop-watch (SW) task with intrinsic fun and a boring watch-stop (WS) task.

Our data showed a prominent N2 discrepancy between two task cues, suggesting the expectation of participants in performing the interesting SW task. According to the N2 literature mentioned, N2 amplitude represented mismatch and action inhibition. In the current study, the two tasks that the participants faced were of different intrinsic fun. Our results showed that N2 amplitude was enhanced when the upcoming task was the boring one, suggesting a mismatch between the expectation of the task and the actual presented task. Therefore, N2 may be a candidate index of intrinsic motivation. In addition, the current study also revealed that N2 can reflect not only the mismatch between target and nontarget as suggested by previous studies [4, 23, 24] but also the mismatch between the actual presented stimuli and their expected stimuli.

In the following feedback stage, we found that the outcome of the WS task induced larger FRN amplitude and smaller P300 amplitude than those of the SW task. Furthermore, failing trials in the SW task elicited larger FRN and smaller P300 amplitude than winning trials. These results indicated that subjective valuation of outcome was decreased in the WS task and was even lower than the failing feedback of the SW task, which is in accordance to previous findings that FRN and P300 could reflect subjects' affective/motivational evaluation of outcome. As no extrinsic incentives were given, the only source of human motivation came from the task itself, and people always showed higher intrinsic motivation to the interesting task. Therefore, a potential mechanism was that higher motivation led to higher affective evaluation toward outcome information. For the interesting SW task with higher intrinsic motivation,

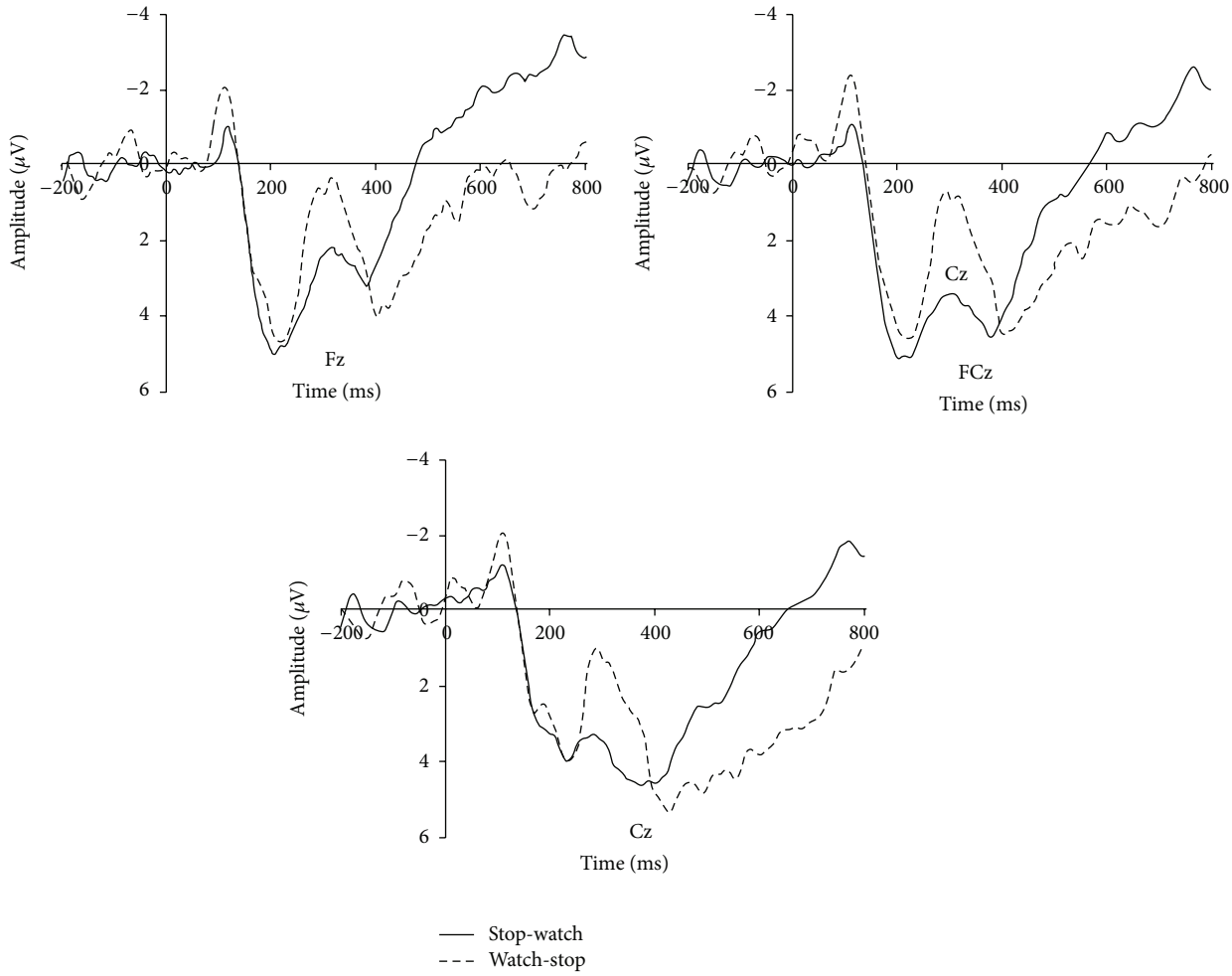


Figure 1: N2 results. For illustrative purpose, grand-average ERP waveforms of N2 from three frontal midline electrodes (Fz, FCz, and Cz) were plotted as a function of conditions.

the outcome of the task was of higher affective significance, and the FRN effect decreased, accompanied by an increased P300 effect. Moreover, recent studies indicated that the FRN amplitude was positively correlated with the activation of reward-related regions in the brain, including the ventral striatum [25, 26]. Therefore, carrying out an interesting task was a reward in itself, even though no extrinsic reward was given. The participants considered the interesting task as fun, whereas they considered the boring task as a request to complete the task.

Meanwhile, a prominent effect for gain-loss discrepancy in the SW task was present. FRN deflection loomed smaller in the winning condition than in the failing condition, suggesting that winning feedback was of higher evaluation than that of the failing one. Furthermore, P300 reflected the valence effect of the stimuli. These results were in accordance with previous findings [12], which can also be explained by the subjects' higher evaluation of winning outcomes than that of failing outcomes. Previous studies always measured the intrinsic motivation at free-choice stage on behavioral level after participants finished the given task [27, 28] while the current study measured intrinsic motivation on brain level during the processing of tasks. Compared with the previous way, the current experiment considered amplitude of endogenous ERP component as an index of intrinsic motivation which was more objective and accurate.

The social attribute of humans was always less engrossed in studies of machine learning. As humans, we would be tired, interested, or not interested. These social factors can largely influence our behavior. Therefore, human motivation should be factored in when imitating human behavior. The current results revealed that the dynamic shifting of human intrinsic motivation from a task can reflect in the deflection of specific ERP components, such as N2, FRN, or P300. Therefore, in a machine learning model, components related to motivation may be a candidate factor of sociality.

To sum up, this study investigated the neural mechanism of intrinsic human motivation by comparing an interesting SW task and a boring WS task. The participants showed reduced N2 amplitude in the cue priming stage when the SW cue appeared, whereas, in the feedback stage, the feedback of SW task elicited reduced FRN amplitude and enhanced P300 amplitude. These results provide evidence for the existence of intrinsic motivation through electrophysiological activity on

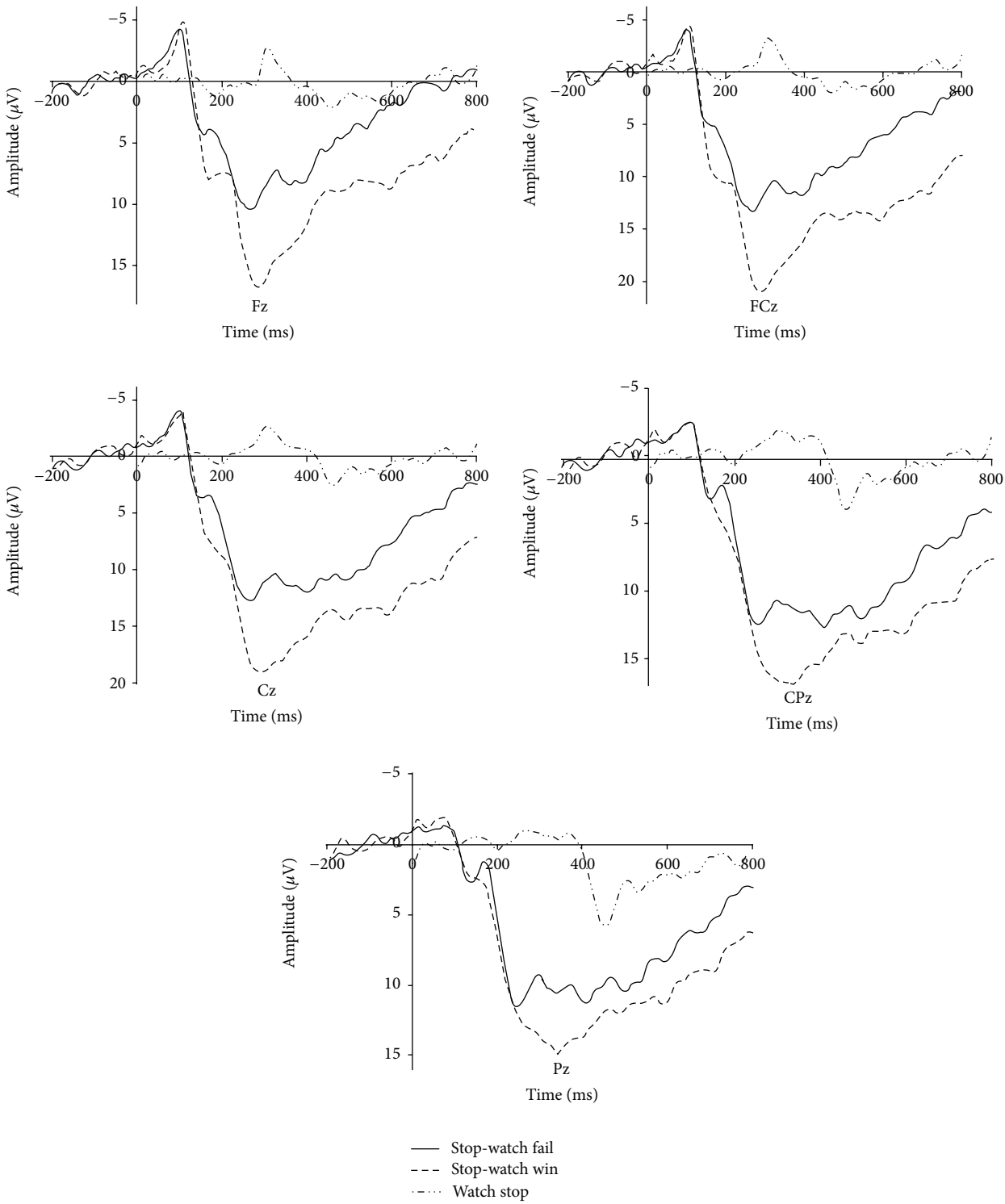


Figure 2: FRN and P300 results. For illustrative purpose, grand-average ERP waveforms of FRN from three frontal midline electrodes (Fz, FCz, and Cz) and P300 from two parietal electrodes (Cz, CPz, and Pz) were plotted as a function of conditions.

brain level and compared the degree of intrinsic motivation quantitatively. Quantitatively measured intrinsic motivation may also be a candidate social factor in a machine learning model.

Conflict of Interests

There is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by Grant 71371167 from the National Natural Science Foundation, Grant 2013GXS2D025 (Marine economy development strategy of Zhejiang providence) from a key project of National Soft Science, Grants XKW14D2006 and XYW15001 from Ningbo University and sponsored by K. C. Wong Magna Fund in Ningbo University. The authors thank Wenwei Qiu for stimuli programming.

References

[1] M. Quirin, F. Meyer, N. Heise et al., "Neural correlates of social motivation: an fMRI study on power versus affiliation," *International Journal of Psychophysiology*, vol. 88, no. 3, pp. 289–295, 2013.

[2] K. Murayama, M. Matsumoto, K. Izuma, and K. Matsumoto, "Neural basis of the undermining effect of monetary reward on intrinsic motivation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 49, pp. 20911–20916, 2010.

[3] E. Jodo and Y. Kayama, "Relation of a negative ERP component to response inhibition in a Go/No-go task," *Electroencephalography and Clinical Neurophysiology*, vol. 82, no. 6, pp. 477–482, 1992.

[4] J. R. Folstein and C. van Petten, "Influence of cognitive control and mismatch on the N2 component of the ERP: a review," *Psychophysiology*, vol. 45, no. 1, pp. 152–170, 2008.

[5] M. Eimer, "Effects of attention and stimulus probability on ERPs in a Go/Nogo task," *Biological Psychology*, vol. 35, no. 2, pp. 123–138, 1993.

[6] K. J. Bruin and A. A. Wijers, "Inhibition, response mode, and stimulus probability: a comparative event-related potential study," *Clinical Neurophysiology*, vol. 113, no. 7, pp. 1172–1182, 2002.

[7] Y. Leng and X. L. Zhou, "Modulation of the brain activity in outcome evaluation by interpersonal relationship: an ERP study," *Neuropsychologia*, vol. 48, no. 2, pp. 448–455, 2010.

[8] W. J. Gehring and A. R. Willoughby, "The medial frontal cortex and the rapid processing of monetary gains and losses," *Science*, vol. 295, no. 5563, pp. 2279–2282, 2002.

[9] Q. Ma, J. Jin, L. Meng, and Q. Shen, "The dark side of monetary incentive: how does extrinsic reward crowd out intrinsic motivation," *Neuroreport*, vol. 25, no. 3, pp. 194–198, 2014.

[10] Q. Ma, L. Meng, L. Wang, and Q. Shen, "I endeavor to make it: effort increases valuation of subsequent monetary reward," *Behavioural Brain Research*, vol. 261, pp. 1–7, 2014.

[11] N. Yeung, C. B. Holroyd, and J. D. Cohen, "ERP correlates of feedback and reward processing in the presence and absence of response choice," *Cerebral Cortex*, vol. 15, no. 5, pp. 535–544, 2005.

[12] Y. Wu and X. Zhou, "The P300 and reward valence, magnitude, and expectancy in outcome evaluation," *Brain Research*, vol. 1286, pp. 114–122, 2009.

[13] A. Sato, A. Yasuda, H. Ohira et al., "Effects of value and reward magnitude on feedback negativity and P300," *NeuroReport*, vol. 16, no. 4, pp. 407–411, 2005.

[14] N. Yeung and A. G. Sanfey, "Independent coding of reward magnitude and valence in the human brain," *Journal of Neuroscience*, vol. 24, no. 28, pp. 6258–6264, 2004.

[15] G. Hajcak, J. S. Moser, N. Yeung, and R. F. Simons, "On the ERN and the significance of errors," *Psychophysiology*, vol. 42, no. 2, pp. 151–160, 2005.

[16] S. Nieuwenhuis, G. Aston-Jones, and J. D. Cohen, "Decision making, the P3, and the locus coeruleus-norepinephrine system," *Psychological Bulletin*, vol. 131, no. 4, pp. 510–532, 2005.

[17] Q. G. Ma, Q. Shen, Q. Xu, D. D. Li, L. C. Shu, and B. Weber, "Empathic responses to others' gains and losses: an electrophysiological investigation," *NeuroImage*, vol. 54, no. 3, pp. 2472–2480, 2011.

[18] H. V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, "A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP," *Psychophysiology*, vol. 23, no. 6, pp. 695–703, 1986.

[19] T. W. Picton, S. Bentin, P. Berg et al., "Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria," *Psychophysiology*, vol. 37, no. 2, pp. 127–152, 2000.

[20] S. Righi, L. Mecacci, and M. P. Viggiano, "Anxiety, cognitive self-evaluation and performance: ERP correlates," *Journal of Anxiety Disorders*, vol. 23, no. 8, pp. 1132–1138, 2009.

[21] A. Lavric, D. A. Pizzagalli, and S. Forstmeier, "When 'go' and 'nogo' are equally frequent: ERP components and cortical tomography," *European Journal of Neuroscience*, vol. 20, no. 9, pp. 2483–2488, 2004.

[22] S. W. Greenhouse and S. Geisser, "On methods in the analysis of profile data," *Psychometrika*, vol. 24, pp. 95–112, 1959.

[23] F. Klostermann, M. Wahl, F. Marzinzik, G.-H. Schneider, A. Kupsch, and G. Curio, "Mental chronometry of target detection: human thalamus leads cortex," *Brain*, vol. 129, no. 4, pp. 923–931, 2006.

[24] C. L. Dickter and B. D. Bartholow, "Ingroup categorization and response conflict: interactive effects of target race, flanker compatibility, and infrequency on N2 amplitude," *Psychophysiology*, vol. 47, no. 3, pp. 596–601, 2010.

[25] J. M. Carlson, D. Foti, L. R. Mujica-Parodi, E. Harmon-Jones, and G. Hajcak, "Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study," *NeuroImage*, vol. 57, no. 4, pp. 1608–1616, 2011.

[26] T. F. Münte, M. Heldmann, H. Hinrichs et al., "Nucleus accumbens is involved in human action monitoring: evidence from invasive electrophysiological recordings," *Frontiers in Human Neuroscience*, vol. 1, article 11, 2008.

[27] E. L. Deci, "Effects of externally mediated rewards on intrinsic motivation," *Journal of Personality and Social Psychology*, vol. 18, no. 1, pp. 105–115, 1971.

[28] E. L. Deci, R. Koestner, and R. M. Ryan, "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation," *Psychological Bulletin*, vol. 125, no. 6, pp. 627–668, 1999.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 489793, 10 pages
http://dx.doi.org/10.1155/2015/489793

Research Article

Intelligent Context-Aware and Adaptive Interface for Mobile LBS

Jiangfan Feng[1,2] and Yanhong Liu[1]

[1] *College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[2] *Key Laboratory of Instrument Science and Dynamic Test, North University of China, Taiyuan 030051, China*

Correspondence should be addressed to Jiangfan Feng; fengjf@cqupt.edu.cn

Received 5 November 2014; Revised 27 January 2015; Accepted 4 February 2015

Academic Editor: Dongrong Xu

Context-aware user interface plays an important role in many human-computer Interaction tasks of location based services. Although spatial models for context-aware systems have been studied extensively, how to locate specific spatial information for users is still not well resolved, which is important in the mobile environment where location based services users are impeded by device limitations. Better context-aware human-computer interaction models of mobile location based services are needed not just to predict performance outcomes, such as whether people will be able to find the information needed to complete a human-computer interaction task, but to understand human processes that interact in spatial query, which will in turn inform the detailed design of better user interfaces in mobile location based services. In this study, a context-aware adaptive model for mobile location based services interface is proposed, which contains three major sections: purpose, adjustment, and adaptation. Based on this model we try to describe the process of user operation and interface adaptation clearly through the dynamic interaction between users and the interface. Then we show how the model applies users' demands in a complicated environment and suggested the feasibility by the experimental results.

1. Introduction

USA Scholar Schlit first proposed three target location based services (LBS) in 1994: spatial information, social information, and resources nearby. Nowadays, mobile LBS interface has significant influence, while a mass of users carry their mobile terminals to require location services such as the wish to find a better route all over. Under this circumstance, problems of traditional mobile LBS interface are revealed which lack of ability to adapt users' demands initiative. Context-aware human-computer interface makes mobile LBS interaction more natural and efficient which is able to adapt to different users' characters and requirements by taking advantage of usable information about users' tasks (e.g., locations and preferences of users, experiences, and surrounding environments). Many of these adaptive interfaces serve specific users, such as the interface which Cao designed for children that apply cartoon-icon in 2007 [1] and the user interface adaptation proposed by Zouhaier et al. which is based on context awareness for disabled people in 2013 [2], for the characters of specific users are obvious. And these researches are based on the human-computer interaction model which describes the characteristics of the interaction process between human and machine. Some early researches of adaptive interface model are based on user features, like Rich proposed a modeling method which classifies the users based on their background and then provide different services [3]. User modeling concentrated not only on users' cognitive or reason, such as knowledge, goals, and planning [4], but also on emotional and personality [5, 6]. In the field of mobile LBS, Shi and Bian developed an adaptive expression of spatial information and the adaptation policy of the interaction elements on LBS interface in 2007 [7] and Lathia et al. proposed state-of-the-art advanced traveler information system (ATIS) which can adapt to users' environment and activities in 2012 [8]. However, there are certain problems and shortcomings in the current study.

On the one hand, most of context-aware adaptive interfaces are designed for specific tasks or applications, and many researchers construct different models to satisfy users' various

requirements that are ignoring the diversity of one person. Classification for users in one respect cannot represent the features on their other aspects. On the other hand, context-aware information is always used to predict the usable interface but ignores its effect in the dynamic interaction process. We maybe consider other aspects such as preference when we recommend a suitable interface to a user built on his/her cognitive ability. The challenge of adaptive interface in mobile LBS is not simply to provide users information whenever and wherever but also to provide appropriate information for users when they need. The current research on the interface adaptation lacks the exploration of user dynamic interactive behavior. When the passage of time, task, and context information change, then the content of adaptation changes. According to these problems, this paper proposed dynamic adaptive model and presents a corresponding method.

In view of the above problems, this paper presents a context-aware adaptive interface for mobile LBS. At first, we establish a user model which has better generalization and differentiation degree based on users' basic characteristic and the behavior characteristic, and then we match the user model with the refine interface element modules; proposed adaptive interface modeling method and system structure combine with dynamic interaction behavior. At last, we explain the adaptive process through a scenario. In the following Related Works section the feasibility and applicability of context-aware interface to be adapted to users in some solutions are discussed.

2. Related Works

This section shows the focus that we described below within context-aware adaptive interface for mobile LBS existing literature. There are three parts that we discussed: context-aware technology, adaptive user interface, and adaptive spatial information.

2.1. Context-Aware Information. People often naturally used implicit information to make the content rich when there is a process of human to human interaction for they understand the situation of each other while for computers it seems difficult to master this skill in comparison. Therefore, context-aware technology was used widely in order to attain the purpose of natural interaction. Context is determined by Merriam-Webster's collegiate dictionary as "the interrelated conditions in which something exists or occurs." To put it more specifically, Schilit and Theimer [9] proposed that context contains location and identities of nearby people and objects in 1994, and in 1997 Brown et al. [10] added time of the day, season of the year, and temperature to the original definition. Up to now, context is broadening to a comprehensive concept including task context, user context, and circumstance context. Generally speaking, context based on mobile phones can be divided into three parts as follows [11]:

(1) user environment: location, preference, experiment, social relations, and so forth;

(2) mobile environment: device suitable for users to input or display, network, Bluetooth, and so forth;

(3) physical environment: weather, date, noisy, and so forth.

Moreover, context-aware technology has the capability to sense, detect, and grab the environment around users and get the dynamic changes to speculate their behavior [12].

Context-aware technology plays an important role in mobile terminals which equip a rich set of sensors (e.g., camera, accelerometers, GPS, digital compass, gyroscope, ambient light sensors, proximity sensors, multitouch panels, and microphone) [13]; it also enriches the function of GIS to provide users a variety of services. Tomitsch et al. discussed the context of human actions in public space and how they fed back [14]. Lathia et al. [8] proposed mobile traveler information system which can become personalized services based on explicit preferences. J. Karat and C.-M. Karat [15] proposed context-aware route recognition approach to improve the accuracy of routing recognition. Abowd et al. proposed a mobile context-aware tour guide in 1997 [16]; Cai specifies a semantic model which combined with context and demonstrates how this model supports contextualized interpretation of vague spatial concepts during human-GIS interactions in 2007 [17]. Chung and Schmandt proposed a mobile user-aware route planner which can learn a user's everyday routes and provides directions from locations along those routes in 2008 [18].

2.2. Mobile Adaptive Interface. Human-computer interface (HCI), which is also known as the user interface, is media for the exchange of information between user and computer. The traditional design methods consider the efficiency problem of using rarely, and the traditional interface can only adapt to a few people, but also cannot meet the requirements for one person in different periods with the fixed user interface designed according to users' average level while the computer used popular and user group became more and more widely used. Adaptive user interface (AUI) which can adjust itself to fit a user or a task [19] emerged and developed fast while the requirement of omnipresent computing challenge traditional interface emerged and increased. Earlier in the research of an adaptive user interface, it requires three models: system model, user model, and the interaction model [20]. The system model describes the characteristics of the system that can be changed, such as the system to be able to adaptive. Acquisition and application of the user model are the foundation of an adaptive user interface to make the system adapt to the individual user behavior. Interaction model defines how the system is modified, and what it can adapt to. Above all, the degree of adaptability in the adaptive process depends on the user model which describes users' knowledge that can be utilized to facilitate human-computer interaction.

Many researches take advantage of adaptation to provide personalized services for users such as helping users to obtain information, giving users a recommendation, tailoring information for users, or providing help. Yoon et al. proposed an adaptive mixture-of-experts model to solve the complexity

and personality problem of multiuser interface in 2012 [21]. And Cheng and Liu developed an adaptive recommendation system that inferred users' preferences and adjusted the user interfaces [22]. Wang et al. presented an automatic approach which helps users who suffer from visual impairment to make use of online map with independent access to geographic direction [23]. Sulaiman and Sohaimi [24] discuss a possible interface which is simple enough for older users through analyzing the situation of using a mobile phone.

2.3. Adaptive Spatial Information. Nowadays, people can carry mobile devices everywhere and every time with the development and extensive application of mobile communication and internet technology. In this case, a large amount of requirements concentrated on the interests of users themselves: environment information such as recommending interest point to users by acquiring the users' location. We often need to consider the personal issues for spatial information used by more and more users. In particular, the users which have mobile phones with different running speed needed different degrees of information presentation, and the users with unique moving speed needed different scale. In addition, different users request different aspects of spatial information; for example, tourists pay attention to scenic spots while drivers follow with road conditions. Another adaptive problem is how to determine quantity of information displayed on the screen. Many navigation charts are based on accessibility to display all the details when facing the problem. In fact, showing the map too detailed not only is difficult to understand and display but also makes the user focus on useless information which limits the effectiveness. Therefore, mobile adaptive visualization of spatial information must be based on users' needs to provide details step by step [25] such as providing user detailed information when he/she amplifies the map gradually.

In summary, the key elements of adaptive spatial information are related with the user, the mobile terminal, and the environment. User aspects include user background (such as physiological differences, preference differences, and cognitive differences), user location (position, speed, and direction), and user requirements. Environment aspects include basic information (such as temperature, weather, date, and time), and related users' information means the information of the other users which related to the users' tasks. And the effects of system contain transmission speed, information receiving rate, and so on.

Adaptive interface has been extensively studied, but the self-learning user interaction lacks. Here we adopt the use of adaptive user interface model to predict user's intent effectively with their spatial experience. Moreover, the proposed use of experience awareness assists in the prediction of user's intent while satisfying the early fire detection requirement.

3. Adaptive User Interface Model

Cognitive psychology regards people as an information processing system, and people often have different action according to the environment, cognition, and personality tendency differences to reach the established goals [26]. There are interaction spaces between human, computer, and environment. If the information presented on the interface can adapt to the user's cognitive psychology and personality traits, users will complete the task quickly for reducing the user operations.

Adaptive user interface model consists of the following components: user model (UM), task model (TM), interaction model (IM), domain model (DM), environment model (EM), and presentation model (PM).

3.1. User Model. The user model needs to abstract the individual differences of users which may relate with personalized service for no two users are identical. The interface style may be affected by personality or preference differences while the expression mode of interface may be affected by cognitive or physiological differences. Here we define the user model (UM) as a collection: UM = {*User ID, Knowledge, Physiology, Inclination*}. *User ID* represents a unique identifier of user. And we divide users' background into three parts: *Knowledge*, *Physiology*, and *Inclination*. *Knowledge* summarizes the knowledge level of the user, *Physiology* is on behalf of users' physiological characteristics, and *Inclination* represents the subjective desire in every aspects.

At first, people always associate academic quantification when it comes to knowledge, but we use the concept to represent users' skill level including proficiency in interface using and cognitive ability on the map. Knowledge = {*Education, Professional level, Proficiency level*}. And we can see the differences existed between expert users in a field on software proficiency through the description. Second, factors of physical abilities cannot be ignored in the interaction. Physiology = {*Age, Sex, Health*}. And last, Inclination = {*Occupation, Personality, Habit, Preference*}; these factors can influence the choice tendency of users.

3.2. Interface Static Elements. The task model was divided into abstract task model and specific task model. Catch and abstract the users' needs and described the needs as abstract tasks. Describe the interactive behavior in the system and the dynamic behavior in the process of the interaction. The task is an activity which is used in order to satisfy the user's goals. We can abstract task as TM = {*Operation, Object, TC*}. Operation means a task operation which act on an object. The object means an object which needs to operate. TC represents the task context. The task model can be embodied into STM = {*tID, operationType, dataItem, dataType, C*}. *tID* means the task ID which needs to complete. operationType is the type of interaction operation such as read, write, or command. dataItem is a data item which consists of data ID, data attributes, and data value. dataType means the data type of the operate data. *C* is a set of constraints to the data item. Specify the task context to the constraint of the data and data manipulation.

The domain model (DM) is defined as DM = {*Object, Attribute, Contact, Time*}; Object is a set of interface object at a specific time. The interface object is changing when the environment changes or user task changes; therefore we use

Time which is corresponding with the Time of TM and EC to distinguish.

The interface model (IM) describes the interface, and express the various controls manipulation in the dynamic interactive process. Adjust the user interface components and structure according to the specific task which is analyzed through the task model. The interface model can be abstracted as IM = {*Control, controlConstraint, controlRelation*}. Control contains control ID and control attribute. *controlConstraint* means the constraints of the controls on the interface. *controlRelation* means the relationship of the controls on the interface.

The presentation model (PM) is defined as PM = {*Module, MEC, MUC, MTC*}; Module is a collection of interface components. MEC, MUC, and MTC are component properties under the influence of the environmental context, user context, and task context.

3.3. Interactive Context. We can divide interactive context (IC) into three parts: environment context (EC), user context (UC), and task context (TC). Interactive environment context can be divided into user environment and equipment environment, where the user environment includes time, place, and weather and equipment environment includes transmission rate, and resolution. EC includes the user environment and the device environment. The user environment includes perimeter environment which may affect operations of the user and user's own context environment. Device environment includes transmission rate and resolution. EC = {*DE, UE*}; DE means the device environment and UE means the user environment. DE = {*MS, SO, TR*}. MS means the movement speed of the device; SO is the screen orientation, and TR is transmission rate of the device. UE = {*Loc, Sur, Soc, Act*}. Loc means the location information of the user. Sur represents the surrounding users of the user. Soc means the social information of the user which can be obtained from the social software open interface. Act is the information which is provided by the previous operation.

3.4. Interaction Methods. A large amount of contextual information and the changeable situation requires a decision mechanism to determine what kind of context information is needed and when the information can be used. The decision mechanism has three points which must be paid attention to: selecting the appropriate context, allocating context priority level reasonable, and having the dynamic adaptability (environment-stimulation, task-stimulation, etc.).

The most important in the interaction process is to understand the purpose, and then we must reduce the problems caused in the process of achieving this goal; the last is to let the user interface elements fit the user like Figure 2.

The interaction model adjusts interface mode constantly when some factors dynamic changes and stimulates.

4. Interaction Strategy

The adaptive user interface framework which based on the GIS interface model is presented in the Figure 3. The Figure 3 describes the process of context information collection and the eventual capture, and explains how to analysis and handle events through the certain reasoning mechanism. According to psychology related research achievements, people will divide continuous events into several activities in perceptual according to kinds of characteristics. Individual differences in cognitive flexibility may underlie a variety of different user behaviors [27]. And users in the same activities tend to repeat steps which can reduce the user experience.

Combine the knowledge base to identify the user's interaction patterns and predict the most likely interaction behavior candidates of the user; then the adaptive recommendation results appeared on the interface layer. The adaptive user interface framework interconnected between layer and layer; the model layer determines the need of context information collection, the adaptive layer is used to realize the user interface adaptive function and scheduling, and the interface layer is used to present the results of self-adaptation.

The adaptive decision-making mechanism is implemented by capturing the user interaction sequence. When the same user completes a task on the mobile GIS interface, the same action sequence is often repeated. So we can predict the future behavior of the user through the interaction sequence judgment, storage, and matching. The tasks which were completed by the GIS human-computer interface can be refined, and the user operation of every subtask will have certain regularity which also contains its unique personalized information. User actions can be refined into lower levels of atomic operations, such as a button click, an operation of input box, and the map zoom event. We can describe a user action as A = (action object, action type). Several user actions compose an action sequence, which can be collected and matched to predict the most likely next step of users. Then adjust the interface elements dynamically and achieve the goal of continuously reducing the user operation.

In the process of matching the sequence, we judge the next action according to the front action, but the longer the length of the sequence matching, often the better the results. Therefore, we should choose more suitable length sequence to match. Defining the average length probability of the match sequence is $L(a \mid s) = l_t(a, s)/\sum_i l_t(a_i, s)$. In addition, the happening of the action may also be related to the other action and not just related to the matching of model length, so we use frequency of action occurrence $P(a \mid s) = f(a, s)/\sum_i f(a_i, s)$ to describe the action occurrence probability. We use the action prediction $R_t(a, s) = L(a \mid s) + P(a \mid s)$ to determine the action occurring possibility.

The interactive action sequences appearing occasionally can be seen as preinteraction pattern, which occurred repeatedly will be put into the pattern library. The current interaction sequence is matching with the action sequences library, and the matching starts from the current action and increases in length gradually under the context environment. Obtain all the forecast candidate set; then choose the action which has highest action evaluation $R_{\max}$ as the prediction results like Figure 4. E means the new action set and C means the existing action set.

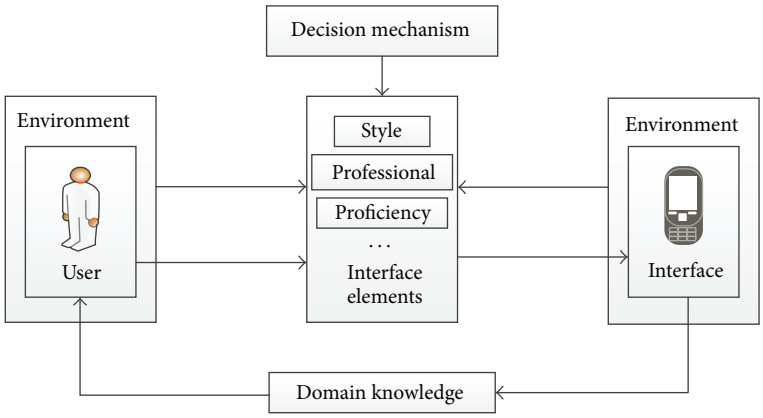


Figure 1: The environment of interaction.

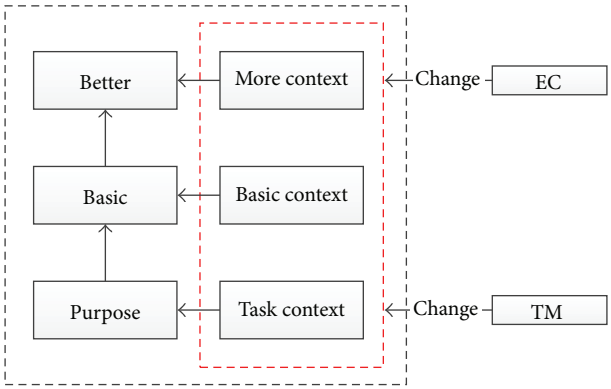


Figure 2: User interface adaptive elements.

5. The Modeling Method

In the third part of the paper the adaptive user interface model is put forward and this section will illustrate the construction process from the abstract model to the specific model. Create an adaptive interface model on the basis of the user model, extract the information from the user model to form the domain model, and then extract the management tasks in the field of domain model to form the task model.

The first step is to build a user model. The user ID here refers to the account of each user in the system which is used to record different user information. Knowledge here means the proficiency of the user used navigation software and professional level of the user (Figure 1). Different levels of users can lead to different operations. And the Physiology refers to the aspects of users' age and gender differences. We will comprehensive considering these aspects in the experimental personnel selection. Inclination information is recorded automatically in the system.

Domain model which is considered from the user model needs to list managerial entity objects and analyze the properties of these entity objects. For example, when a user finds the route, entities in the domain include buttons, input box, and map, and the interface elements of other services may be involved for different users. The relationship between the interface elements is the ordinal relation in the operating sequence which is described in the adaptive strategies.

Each interact action corresponds to a task in the task model, such as the switch interface, input box operation, and determining button click. For example, a user wanted to find the point of restaurant; the description of the scenario in the user model can be defined as US = ⟨(UserID1, Preference), (Device1, UserEn1)⟩, UserEn1 = ⟨Location, restaurants, Pre-operation⟩; the corresponding task model can be defined as Task = ⟨AT1, AT2, AT3⟩, AT1 = ⟨Click, ButtonStart, TC1⟩, AT2 = ⟨Click, ButtonSelect, TC2⟩, AT3 = ⟨Zoom, Map, TC3⟩, TC = ⟨Pre-operation, OPtime, Null⟩.

The building of interaction model according to each task of the set of events in the task model described the atomic operations in the user interface such as clicking on and long press and described the corresponding commands of interacting objects, such as a jump and zooming.

6. Experiment and Analysis

6.1. The Scene. In order to verify the result of the study, the following scenario is designed to verify that the adaptive method is easy to use. And then we evaluate some values which are measurable.

A Traditional Route and POI Searching Is as Follow

(1) User enters the application.
(2) Go to the weather page to check the weather.
(3) Go back to the main page.
(4) Click the button to enter the route searching page.
(5) Enter the locations to which the user wants to go.
(6) Click the button to choose transportation.
(7) Click the button to display the route.
(8) Complete the route lookup.
(9) Click the button to enter the POI selection page.
(10) Choose a specific point of interest.
(11) The points of interest appeared on the map.

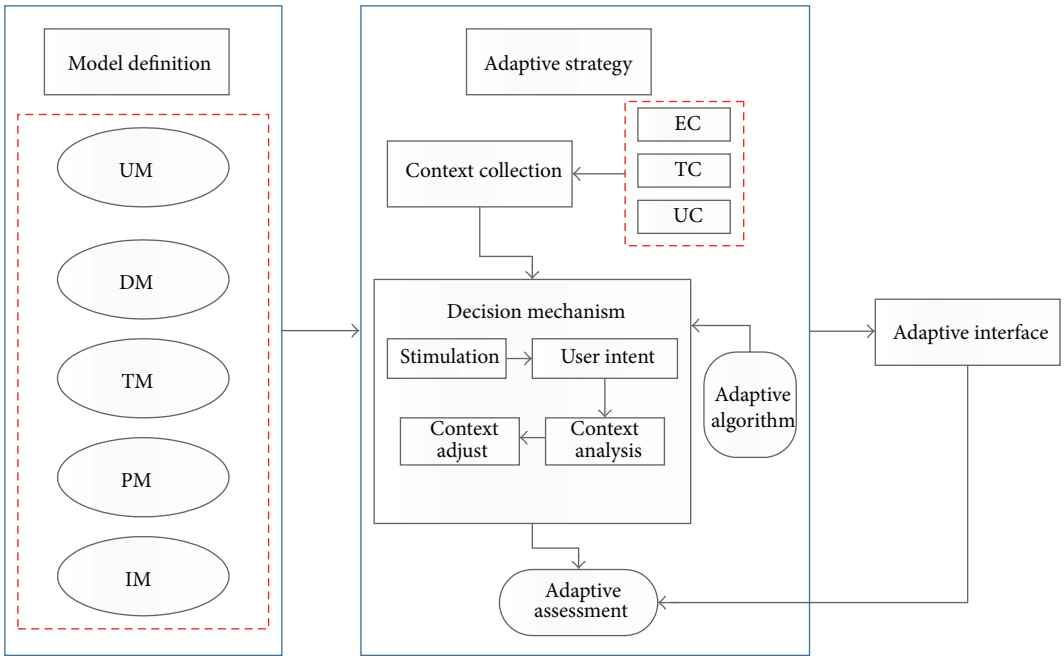


Figure 3: Mainly adaptive system.

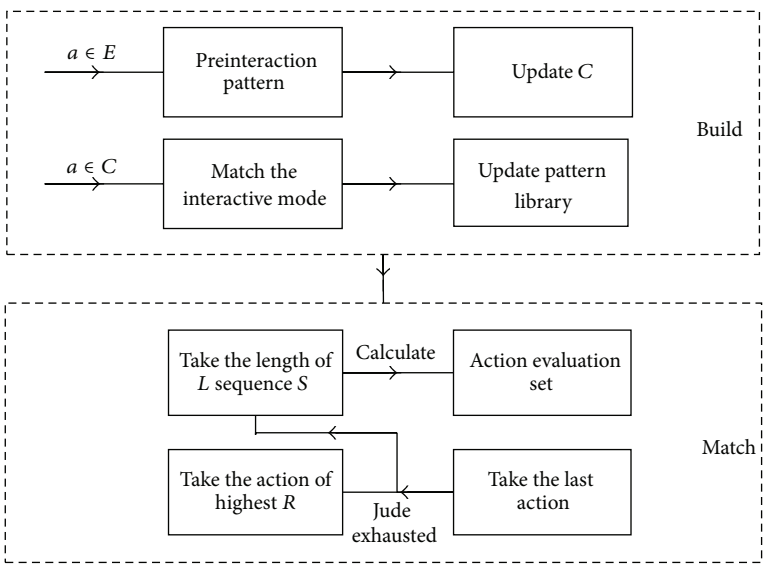


Figure 4: Sequence matching.

(12) Choose one of them to show the route.

(13) Complete the POI searching.

A Traditional Map Operation Is as Follows

(1) User logs into the application.

(2) User clicks the search button.

(3) Enter the search site in the input box.

(4) Click the confirm button.

(5) Zoom in to check the location

(6) Zoom in once again.

(7) Click on the map sign to check the specific location name.

Context-Aware Adaptive Interface for Route and POI Searching Is as Follows

(1) User clicks the traffic mode button (preferences record) to enter the application.

(2) Choose the traffic mode (advice according to the weather condition) and enter the application.

Table 1: Context-aware adaptive interface scene design.

User	System
(1) User clicks the traffic mode button.	(1) Display "bad weather"; recommend another traffic mode.
(2) User chooses traffic mode and clicks the button.	(2) Display the location input box.
(3) User inputs the location and clicks the OK button.	(3) Show the route and the interest point button.
(4) User clicks the interest point button.	(4) Display the points of interest.
(5) User clicks the interest point.	(5) Display the route.
(6) User logs into the application.	(6) Record the user's operation.
(7) User clicks the search button.	(7) Enter the search interface.
(8) Enter the search site in the input box.	(8) Operation sequence matching and display location markers directly.
(9) Zoom in to check the location.	(9) Operation sequence matching and zoom in again.
(10) Click the map sign.	(10) Display the specific location name.

(3) Enter the locations to which the user wants to go.

(4) Click the button to display the route.

(5) Complete the route lookup.

(6) Click the button (preferences record) to show the points of interest.

(7) Choose one of them to show the route.

(8) Complete the POI searching.

Context-Aware Adaptive Interface for Map Operation Is as Follows

(1) User logs into the application.

(2) User clicks the search button.

(3) Enter the search site in the input box.

(4) Zoom in to check the location

(5) Click on the map sign to check the specific location name.

Analysis of these scenarios can be found that, the traditional routes and POI searching need more options, for more user active choices, which will produce more returns and select operation. Here is the design of context-aware adaptive interface which is user centered (Table 1).

Figure 5 shows the route searching interface. The application icon is displayed as traffic mode according to user's preferences. It will prompt the weather condition when user clicks the button and give advice. Click on the icon to enter the location input interface. The context-aware adaption can reduce the traffic mode selection operation and also give advice according to the environment actively.

Figure 6 shows the adaption of interesting point searching. In the process of user walking, finding interesting point and giving corresponding button to users according to the preferences can meet users' requirements more easily. Click the button to access the route.

Figure 7 shows that the system recorded the user action sequence and forecasted the next steps. The operation sequence frequently appearing of the user in this scenario is $\langle 1, \text{liu}, \text{searchBtn}, \text{click}, \text{time1}\rangle$, $\langle 2, \text{liu}, \text{Inputbox}, \text{input}, \text{time2}\rangle$, $\langle 3, \text{liu}, \text{OkBtn}, \text{click}, \text{time3}\rangle$, $\langle 4, \text{liu}, \text{Map}, \text{zoom in}, \text{time4}\rangle$, $\langle 5, \text{liu}, \text{Map}, \text{zoom in}, \text{time5}\rangle$, $\langle 6, \text{liu}, \text{Map}, \text{click}, \text{time5}\rangle$. The system matched the first three operations and then predicted the next operation and adjusted the interface automatically. The first figure shows entering the search interface after clicking the search button, the second figure shows the amplifying map automatically after clicking the Ok button, and the third figure shows the location information after clicking the site.

We reflect the dynamic adaptive from two aspects mainly from this experiment: the choice of transportation mode and the user's interest concerns.

6.2. User Evaluation and Analysis. We choose some test users with certain discrimination and finish the appointed tasks. We choose 50 testers by taking the differences of users into consideration in the user model. We choose half of the testers' education level such that it is above the average and the other half is below the average. Including the testers, the proficiency can be divided into skilled, general, and strange. The sex ratio is 1.5 : 1 and age distribution from 20 to 60 years old, which were randomly selected.

The international organization for standardization (ISO) includes the usability evaluation factors of a product which fixed tasks in a specific environment which are effectiveness, interaction efficiency, and user satisfaction. The effectiveness is used to judge whether it can achieve certain functions and interface supports the corresponding function. There are two functions of the testing interface: providing the transportation recommended automatically and providing the interest recommendation when the user finds route. The user satisfaction is the subjective satisfaction of the user interface. The evaluations of these two aspects are assessed by the user survey. Interaction efficiency is decided by error rate, completion time, being easy to learn, and being easy to use. We can record the error time and completion time of two tasks and obtain testers' evaluations about being memorable, easy to learn, and easy to use and the efficiency. We let the testers to complete two contrast tasks under the same condition and get some pairs of observe values. Analyze these values to draw inferences. The difference result which is gotten from same testers in the same environment can be regarded as the differences made by different system. We

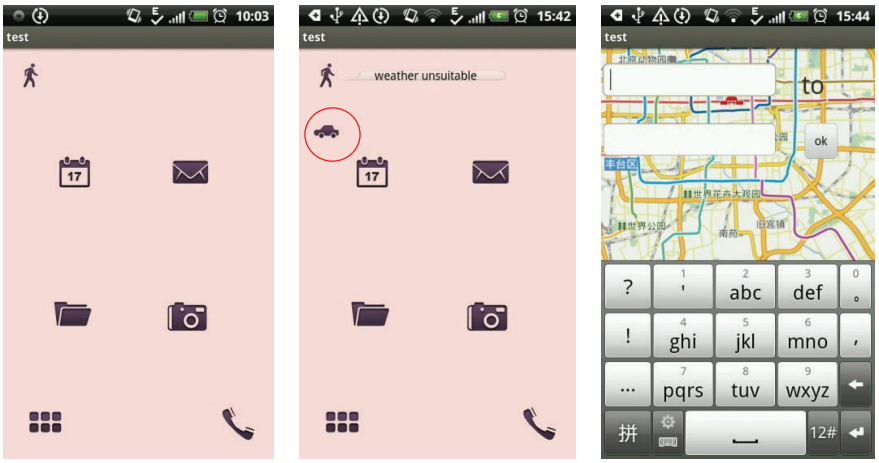


Figure 5: Route searching.

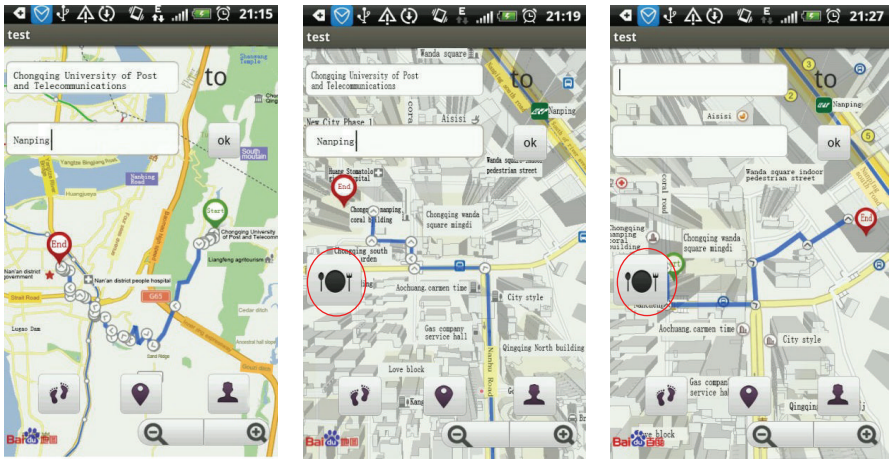


Figure 6: POI searching.

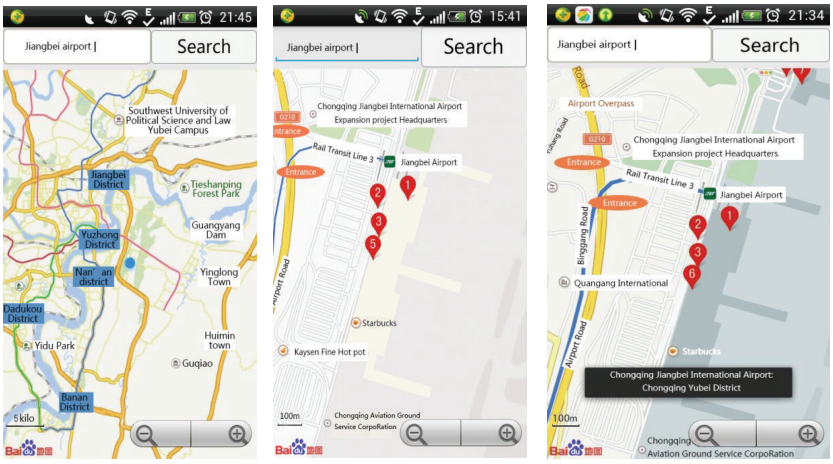


Figure 7: Map operation.

use general navigation system to do the comparison test and obtain independent observations in pairs.

X_i represents the time spent by comparison system to complete a task, and Y_i represents the time spent by test system to complete a task. Suppose there are n pairs of independent observations: (X_i, Y_i) $(1 \leq i \leq n)$. $D_i = X_i - Y_i$ $(1 \leq i \leq n)$ is the difference of X_i and Y_i; observations of these values' sample mean and sample variances are recorded as $\overline{d}$ and s_d^2. The rejection region is $|t| = |\overline{d}/(s_d/\sqrt{n})| \geq t_{a/2}(n-1)$.

The first task is to choose a vehicle and find the route to reach somewhere. The second task is finding interest point around. And the third task is the map operation. At the beginning of the experiment stage let the users be familiar with the system and task. Record the time of each task completed and the time of total task completed in the course of the experiment. Count the total number of errors which emerged in the processes of the tests using. And let each test personnel complete the questionnaire at the end of the experiment. The time of users, which are familiar with this kind software, to complete the first task is usually about 23 s, and the time of users to find the corresponding route by using the system which records the users' selection and offers suggestions is about 16 s. These times are average values of testers. For the first task $\overline{d} \approx 6$, $s_d \approx 40.714$, $t_{0.05}(49) = 1.6794$, $|t| \approx 1.032$ which are out of the rejection region. For the second task $\overline{d} \approx 7$, $s_d \approx 26.589$, $t_{0.05}(49) = 1.6794$, $|t| = 1.8429$ which fall into the rejection region and $\overline{d} \approx 5$, $s_d \approx 20.533$, $t_{0.05}(49) = 1.6794$, $|t| = 1.7046$ which fall into the rejection. The fundamental task needs less average time, but the advantage is not obvious through the calculation. In the second task and the third task, operating time reduces significantly through the calculation. The error rate has little difference between the contrast system and the test system. We can see that other factors are higher than contrast system except memorability. In addition, efficiency of the test system improves obviously.

Then we get the user's satisfaction degree of the interface on the aspects like being memorable, easy to learn, and easy to use and efficiency through the questionnaire survey. These 50 testers give the scores of two systems using experience and get the average of each index, respectively, like Figure 8. We can see that memorability of the traditional human-computer interface is higher than the adaptive interface for elements of adaptive human-computer interface are changeable. On the other aspect, the scores of text system are higher than the contrast system which efficiency is greatly improved. To illustrate the adaptation of the text system is greatly improved.

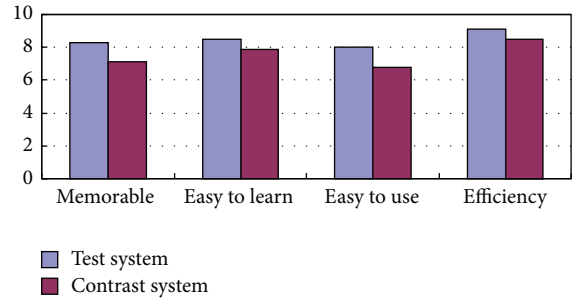


Figure 8: Questionnaire statistics.

7. Conclusion

This paper proposes a context-aware adaptive human-computer interface model for mobile LBS which is based on the user model and described in three aspects: static composed elements, dynamic interactive behavior, and adaptive strategy. The adaptive user interface proposed in this paper has advantages compared with traditional adaptive user interface as follows: (1) avoiding the limitation of the traditional adaptive user interface caused by user classification and achieving the adaptation according to the combination of each user's habits and external experiments; (2) paying more attention to the dynamic interaction process and adjusting the user interface in the interactive process more in line with the real-time interaction; (3) using the context information dynamically and then making the context information using more effective.

The adaptive system based on the model in this paper has some deficiency; for example, the range of adaption should be extended, and also there are limitations of the current research to gain more effective information on user knowledge, ability, and so on. In the future, we will further optimize the stimulus-judgment method, more effectively use context information, enhance the adaptive result, and improve the interface layout mechanism to reach the goal of smooth and natural interface adaptive.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The work is supported by the National Nature Science Foundation of China (41201378), the Science Foundation of China for Postdoctors (no. 2014M561212), the Natural Science Foundation Project of Chongqing (cstc2014jcyjA00027), and the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ120526).

References

[1] X. Cao, "Designing mobile phone interface with children," in *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI EA '07)*, pp. 2309–2314, San Jose, Calif, USA, April-May 2007.

[2] L. Zouhaier, Y. B. Hlaoui, and L. Jemni Ben Ayed, "Building adaptive accessible context-aware for user interface tailored to disable users," in *Proceedings of 37th IEEE Annual Computer Software and Applications Conference Workshops (COMPSACW '13)*, pp. 157–162, 2013.

[3] E. Rich, "User modeling via stereotypes," *Cognitive Science*, vol. 3, no. 4, pp. 329–354, 1979.

[4] A. Kobsa, "Generic user modeling systems," *User Modeling and User-Adapted Interaction*, vol. 11, no. 1-2, pp. 49–63, 2001.

[5] I. Poggi, C. Pelachaud, and B. de Carolis, "To display or not to display? Towards the architecture of a reflexive agent," in *Proceedings of the 2nd Workshop on Attitude, Personality and Emotions in User-Adapted Interaction (UM '01)*, 2001.

[6] Y.-T. Fan, D.-X. Teng, H.-Y. Yang, C.-X. Ma, G.-Z. Dai, and H.-A. Wang, "An adaptive user interface model based on experience awareness," *Jisuanji Xuebao/Chinese Journal of Computers*, vol. 34, no. 11, pp. 2211–2223, 2011.

[7] X. Shi and F. Bian, "Research on AHCI specific human-computer interface of GIS," *Geomatics and Information Science of Wuhan University*, vol. 32, no. 10, pp. 924–927, 2007.

[8] N. Lathia, L. Capra, D. Magliocchetti et al., "Personalizing mobile travel information services," *Procedia—Social and Behavioral Sciences*, vol. 48, pp. 1195–1204, 2012.

[9] B. N. Schilit and M. M. Theimer, "Disseminating active map information to mobile hosts," *IEEE Network*, vol. 8, no. 5, pp. 22–32, 1994.

[10] P. J. Brown, J. D. Bovey, and X. Chen, "Context-aware applications: from the laboratory to the marketplace," *IEEE Personal Communications*, vol. 4, no. 5, pp. 58–64, 1997.

[11] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Handheld and Ubiquitous Computing: First International Symposium, HUC '99 Karlsruhe, Germany, September 27–29, 1999 Proceedings*, vol. 1707 of *Lecture Notes in Computer Science*, pp. 304–307, Springer, Berlin, Germany, 1999.

[12] W. Liu, X. Li, and D. Huang, "A survey on context awareness," in *Proceedings of the International Conference on Computer Science and Service System (CSSS '11)*, pp. 144–147, June 2011.

[13] D. Chu, N. D. Lane, T. T.-T. Lai et al., "Balancing energy, latency and accuracy for mobile sensor data classification," in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems (SenSys '11)*, pp. 54–67, Seattle, Wash, USA, November 2011.

[14] M. Tomitsch, C. Ackad, O. Dawson, L. Hespanhol, and J. Kay, "Who cares about the Content? An analysis of playful behaviour at a public display," in *Proceedings of the International Symposium on Pervasive Displays (PerDis '14)*, p. 160, 2014.

[15] J. Karat and C.-M. Karat, "The evolution of user-centered focus in the human-computer interaction field," *IBM Systems Journal*, vol. 42, no. 4, pp. 532–541, 2003.

[16] G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper, and M. Pinkerton, "Cyberguide: a mobile context-aware tour guide," *Wireless Networks*, vol. 3, no. 5, pp. 421–433, 1997.

[17] G. Cai, "Contextualization of geospatial database semantics for human-GIS interaction," *GeoInformatica*, vol. 11, no. 2, pp. 217–237, 2007.

[18] J. Chung and C. Schmandt, "Going my way: a user-aware route planner," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp. 1899–1902, ACM, New York, NY, USA, April 2009.

[19] E. Ross, "Intelligent User Interfaces: Survey and Research Directions," 2000, http://www.cs.bris.ac.uk.

[20] D. Benyon, "Adaptive systems: a solution to usability problems," *User Modelling and User-Adapted Interaction*, vol. 3, no. 1, pp. 65–87, 1993.

[21] J.-W. Yoon, S.-I. Yang, and S.-B. Cho, "Adaptive mixture-of-experts models for data glove interface with multiple users," *Expert Systems with Applications*, vol. 39, no. 5, pp. 4898–4907, 2012.

[22] S. Cheng and Y. Liu, "Eye-tracking based adaptive user interface: implicit human-computer interaction for preference indication," *Journal on Multimodal User Interfaces*, vol. 5, no. 1-2, pp. 77–84, 2012.

[23] Z. Wang, N. Li, and B. Li, "Fast and independent access to map directions for people who are blind," *Interacting with Computers*, vol. 24, no. 2, pp. 91–106, 2012.

[24] S. Sulaiman and I. S. Sohaimi, "An investigation to obtain a simple mobile phone interface for older adults," in *Proceedings of the International Conference on Intelligent and Advanced Systems (ICIAS '10)*, pp. 1–4, June 2010.

[25] W. A. E. F. Abdelsalam, *Maintaining quality of service for adaptive mobile map clients [M.S. thesis]*, University of Waterloo, Ontario, Canada, 2001.

[26] A. Dillon and C. Watson, "User analysis in HCI - The historical lessons from individual differences research," *International Journal of Human Computer Studies*, vol. 45, no. 6, pp. 619–637, 1996.

[27] A. G. Christian and J. F. Ivonne, "A new behavioral measure of cognitive flexibility," in *Engineering Psychology and Cognitive Ergonomics. Understanding Human Cognition: 10th International Conference, EPCE 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21–26, 2013, Proceedings, Part I*, vol. 8019 of *Lecture Notes in Computer Science*, pp. 297–306, Springer, Berlin, Germany, 2013.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 531650, 12 pages
http://dx.doi.org/10.1155/2015/531650

Research Article

The Large Scale Machine Learning in an Artificial Society: Prediction of the Ebola Outbreak in Beijing

Peng Zhang, Bin Chen, Liang Ma, Zhen Li, Zhichao Song, Wei Duan, and Xiaogang Qiu

College of Information and Management, National University of Defense Technology, Changsha 410073, China

Correspondence should be addressed to Peng Zhang; zhangpeng_yes@163.com

Received 4 January 2015; Revised 22 March 2015; Accepted 31 March 2015

Academic Editor: Dongrong Xu

Ebola virus disease (EVD) distinguishes its feature as high infectivity and mortality. Thus, it is urgent for governments to draw up emergency plans against Ebola. However, it is hard to predict the possible epidemic situations in practice. Luckily, in recent years, computational experiments based on artificial society appeared, providing a new approach to study the propagation of EVD and analyze the corresponding interventions. Therefore, the rationality of artificial society is the key to the accuracy and reliability of experiment results. Individuals' behaviors along with travel mode directly affect the propagation among individuals. Firstly, artificial Beijing is reconstructed based on geodemographics and machine learning is involved to optimize individuals' behaviors. Meanwhile, Ebola course model and propagation model are built, according to the parameters in West Africa. Subsequently, propagation mechanism of EVD is analyzed, epidemic scenario is predicted, and corresponding interventions are presented. Finally, by simulating the emergency responses of Chinese government, the conclusion is finally drawn that Ebola is impossible to outbreak in large scale in the city of Beijing.

1. Introduction

Ebola epidemic in West Africa has aroused high concern and begun to spread to other regions recently. EVD spreads through body fluids, with high infectivity and mortality [1]. Up to November 26, 2014, about 15935 infections along with 5689 death cases have been reported. World Health Organization (WHO) declared that all countries would pay attention to Ebola emergency and provide necessary medical aids to these countries such as Guinea, Liberia, Nigeria, and Sierra Leone [2]. Up to now, the battle against EVD is ongoing and many governments have made emergency plans. Moreover, vaccines against EVD are under test, and it will come into use [3]. Recently, the spokesman of Health Ministry has declared that Ebola would not outbreak in China in large scale, though the imported risk of EVD exists.

In China, dating back to the outbreak of Wenchuan earthquake in 2008, Natural Science Foundation of China (NSFC) has already carried out a major research plan of unconventional emergency management. Supported by NSFC, National University of Defense Technology (NUDT) has established the platform of computational experiments [4]. Meanwhile, artificial Beijing is generated according to the geodemographics data [5]. Additionally, confirmatory experiments, such as the propagation of H1N1 pandemic influenza, have proved the validation of artificial Beijing [6]. However, it has also exposed some drawbacks. Fixed behavior mode especially restricts the heterogeneity and self-adaption of individuals. For example, contacts always occur among the minority and cannot simulate the actual sense of epidemic. Artificial Beijing is a dynamic system where individuals' behaviors are continually evolving. Therefore, multiagent learning is involved to optimize the behaviors and travelling mode. It endows individuals with the abilities to adapt to the virtual city according to previous knowledge and current situation.

Since the outbreak of Ebola in West Africa, many significant works have been done to explore the propagation mechanisms and corresponding interventions. Some foreign scholars have distilled the propagation parameters of EDV using the first-hand data by investigation [7]. Simultaneously, domestic scholars have also predicted the outbreak of Ebola

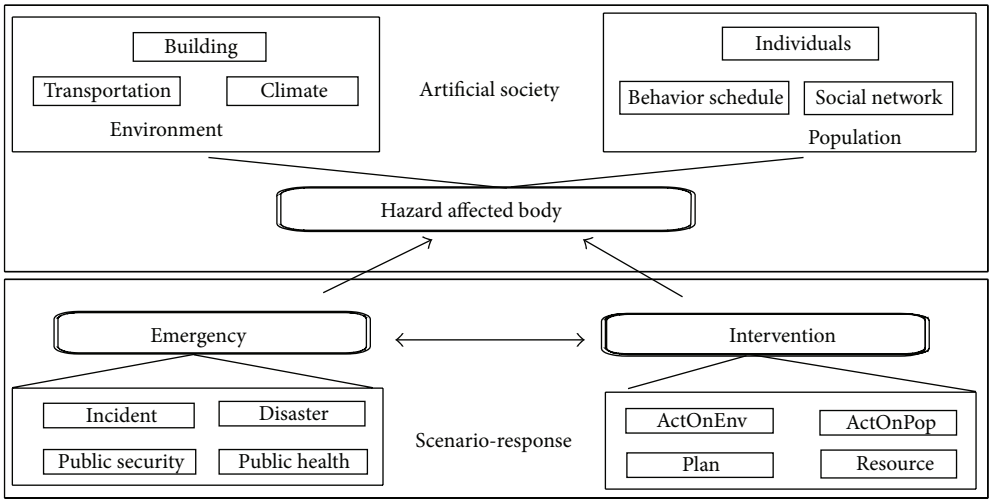


Figure 1: The relationship between PSTT theory and ACP approach.

in China by analytic method [8]. However, they always neglect the actual interaction conditions among individuals, which always led to the amazingly increased infections and death cases. In our study, the prediction of Ebola emergency is based on artificial Beijing, where social networks, individuals' behaviors, and environment factors are considered at the same time. Subsequently, the occupations of infection cases are summarized, the infection locations are analyzed, and interventions such as isolation and immunization are also discussed. Noncontact infections especially are analyzed in our design. In addition, the infections of medical workers and the corresponding interventions are also discussed. Finally, four levels of emergency responses are simulated, according to the current emergency plans of Chinese government.

2. Reconstruction and Optimization of Artificial Beijing

2.1. Reconstruction of Artificial Beijing. Emergencies, hazard affected bodies and interventions are viewed as the core parts of the public security triangular theory (PSTT) [9]. In addition, according to ACP approach, artificial society is the foundation of computational experiments [10]. As shown in Figure 1, hazard affected bodies are just the components of artificial society, while emergencies and interventions are the cores of scenario-response theory [11]. Therefore, population is the foundation of artificial society, and environment provides the places for activities. Generally, emergencies always break the internal balance of artificial society, and then interventions will revise the states to normal.

In artificial Beijing, about 19.6 millions of individuals and 8 millions of buildings are generated. Based on census data, households are generated and individuals are endowed with social roles such as infant, student, worker, elder, and the unemployed. What is more, multiple social networks are involved including family, classmate, neighborhood, and coworker. Simultaneously, types of environment are designed such as house buildings, workplaces, educational institutions, consumption locations, entertainment locations, and medical institutions. According to the social roles, behavior schedules of individuals are designed. As shown in Table 1, basic schedules of student specify the location, duration, and probability of each activity [12]. p_i in the table means the action probability in the relevant period. In addition, correlative locations are assigned for each individual. For instance, the correlative locations of a student include dormitory, classroom, library, playground, and restaurant. The detailed generation process of artificial Beijing has been discussed in literature [13], contributed by another member of our team.

Subsequently, domain-oriented computational experiments are supported by artificial Beijing, such as epidemic propagation [14], rumor spreading, and traffic evacuation. In the study of Ebola epidemic, it is necessary to expand the corresponding attributes of population and environment. Therefore, typical occupations are designed to simulate the main populations in the virtual city, including medical workers, students, workers, and retirees. Additionally, individual attributes such as age, occupation, and health state are involved. Simultaneously, typical environments such as residential buildings, hospitals, and restaurants are considered. Residential buildings are viewed as the main areas of Ebola propagation because of high contact frequency among families. Hospitals are the places for treatment and isolation, while restaurants provide the places to establish the temporary group with weak links [6], where EVD spreads by noncontact infections. Additionally, the capacity and contact frequency of environment are also considered.

2.2. Optimization on Individuals' Behaviors Based on Machine Learning. As previously discussed, basic schedules define the daily activities of individuals. However, some problems are exposed in artificial Beijing. On the one hand, large scale individual-based simulation brings amazing overhead in computing and communication. Moreover, interventions will also directly bring extra-cost by replanning the schedules of individuals dynamically. On the other hand, it is hard

Table 1: Basic schedules of student in a working day.

Duration (Δt)	Activity	Location	Probability
00:00–06:00	Sleep	Dormitory	$p0$ (1.00)
06:00–08:00	Breakfast	Restaurant	$p1$ (0.68)
	Sports-breakfast	Playground-restaurant	$1 - p1$ (0.32)
08:00–12:00	Class	Classroom	$p2$ (0.77)
	Study	Library	$1 - p2$ (0.23)
12:00–14:00	Lunch	Restaurant	$p3$ (0.90)
	Lunch-rest	Restaurant/dormitory	$1 - p3$ (0.10)
14:00–18:00	Class	Classroom	$p4$ (0.77)
	Study	Library	$1 - p4$ (0.23)
18:00–20:00	Dinner	Restaurant	$p5$ (0.63)
	Dinner-sports	Restaurant/playground	$1 - p5$ (0.37)
20:00–22:00	Rest	Home	$p6$ (0.63)
	Study	Classroom	$1 - p6$ (0.37)
22:00–24:00	Sleep	Home	$p7$ (1.00)

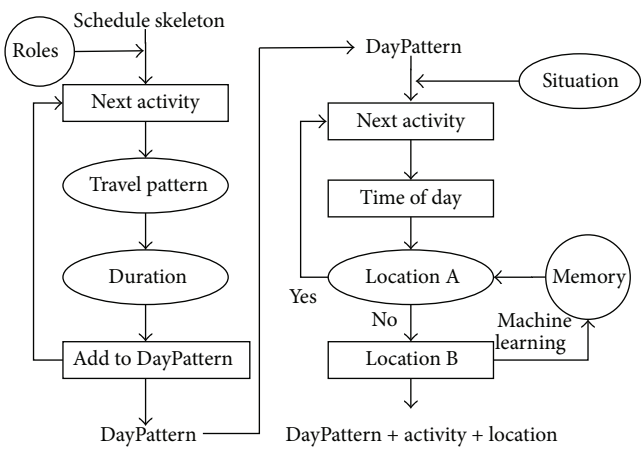


Figure 2: Schedule, DayPattern, and activity.

to depict the activities of individuals, which is associated with the reasonability of artificial Beijing. Furthermore, it will affect the interactions among individuals, directly related to the accuracy and reliability of the prediction of Ebola epidemic. Therefore, two ways are outlined to optimize the performance: (1) improving parallel engine by introducing new technologies and algorithms and (2) optimizing the behaviors of individuals and the structure of artificial Beijing. In this section, detailed optimization on individuals' behaviors will be discussed, which not only decreases the computing consume, but also improves the reliability of the virtual city.

In our study, behaviors schedules are replanned every day according to the history operation and current situation. As shown in Figure 2, *DayPattern* and *activities* compose the skeleton of behaviors schedules. *DayPattern* is designed according to the social roles of individuals, which contains activity items, durations, and the travel patterns between two *activities*. For instance, the *DayPattern* of a student may consist of breakfast, study, lunch, rest, sport, dinner, entertainment, and sleep. Subsequently, detailed activities are specified according to *Daypattern* and current situation. Usually, each *activity* corresponds to a location (l) and duration (Δt), stored in the memory (M) of computer. It is priority for an individual to take an activity in the memory and choose the predesigned location. However, an individual needs to replan its *activities* sometimes. Most of *activities* will especially be replanned under emergency situation. In addition, an individual just needs to choose another location for *activities* in case that the current location in the memory is unreachable.

It provides an individual with the ability to adapt to the environment and cooperate with each other. However, two kinds of difficulties are presented: the balance between exploring and choosing (an individual needs to explore information from not only environments but also other individuals) and the dimension of computing (it involves quantity discrete states and brings large consume in computing).

In our study, machine learning is introduced to solve the replanned problems, which will be discussed in detail. First of all, a location is considered feasible if the following condition is met:

$$\exists l \in G_l, \quad l \in G\{a(\tau)\}, \tag{1}$$

where τ is an index of activities in a given schedule S, G_l is the set of facility at location l, and $G\{a(\tau)\}$ is the set of facilities compatible with activities of type $a(\tau)$ [15]. In our study, l_N is the maximum volume of l, and l_n is the current volume. If an individual enters l where l_N equals l_n, it will search for new locations around him. In addition, Δt represents the duration time of an activity, which should be in the window of the opening time (l_T) at location l.

Let L be the choice set for the given activity defined by (1) and let t_k be the travel time to the location $l(k) \in R_L$; then $R_L \subseteq L$ is the subset of locations reachable. If $t_1 < t_2 < \cdots < t_m$ is the ranking of t_k, then the heuristics of machine learning are listed as follows:

(H1): if $l(k) \in R_L$, then rank $t_1 < t_2 < \cdots < t_m$;

(H2): if $l_N = l_n$, then remove $l(k)$ from R_L;

(H3): if $\Delta t \not\subseteq l_T$, then remove $l(k)$ from R_L;

(H4): if $(l(k) \in R_L) \bigwedge (l(k') \in R_L \mid t_k = \min t_{k'}) \bigwedge (\Delta t \subseteq l_T)$, then choose $l(k)$.

The algorithm starts from (H1). If $l'(k)$ is full or $l'(k)$ is closed, then $l'(k)$ will be removed from R_L and the algorithm returns to (H1). (H4) illustrates that individuals would choose the feasible location with the minimal travel time.

Let $\langle \tau, \Delta t, l \rangle$ be an activity of individual (j); then its behavior schedules (S_j) are formalized as

$$S_j = \sum_{i=1}^{m} \left\langle \tau_{ji}, \Delta t_{ji}, l_{ji} \right\rangle, \tag{2}$$

where m is the total of activities and $\sum_{i=1}^{m} \Delta t_{ji} = 24\,\text{h}$. Subsequently, if S represents the whole schedules of individuals, n is the total of individuals, and then it is formalized as follows:

$$S = \sum_{j}^{n} S_j = \sum_{j}^{n} \sum_{i=1}^{m} \left\langle \tau_{ji}, \Delta t_{ji}, l_{ji} \right\rangle. \tag{3}$$

Finally, S is stored in the memory of computer, and decreases the computing consume in operation. This algorithm is suitable for the large scale individual-based system. First, it just needs to update minority individuals in each step. Actually, only minority of activities is replanned under emergency or in case that locations are unreachable. Second, it is efficient to specify the daily activities of individuals, because most activities along with the locations are restored in the memory.

After continual optimization, the operation of artificial society and the behaviors of individuals would be reasonable. In our design, behavior schedules are designed based on history operation and current situation. Optimized schedules provide individuals with more freedom to adapt to environment, and it is especially important for them to react under emergency. Moreover, dynamic schedules are able to simulate their behaviors in emergency management. For example, an individual will keep far away from the epidemic areas if epidemic outbreaks, while he may be isolated at home under interventions.

3. Models

The main characteristics of EVD include short incubation, high mortality, and fluid shift. In this section, the course model along with propagation model will be established.

3.1. Ebola Course Model. According to the epidemiology classification, individuals are divided into 4 categories: the susceptible, the exposed, the infected, and the removed [16]. As for EVD, susceptible and exposed individuals have no infectivity, while infected individuals are able to infect others. Removed individuals include death ones and discharged ones. Assuming that the total number is N, four categories of individuals at time t is $S(t)$, $E(t)$, $I(t)$, and $R(t)$, and then $S(t) + E(t) + I(t) + R(t) = N$,

$$S \xrightarrow{\alpha SI} E \xrightarrow{\beta E} I \xrightarrow{\gamma I} R. \tag{4}$$

The dynamic process of disease course is shown in (4). αSI especially is the infection probability of SEIR model. Exposed individuals could be transferred to the infected at the rate of β in a unit time, while removed rate from the infected is γ. Subsequently, it is easy to get the conclusion that the incubation period is $1/\beta$, and the infectious period is $1/\gamma$. In addition, γ/β is viewed as the reproductive number (R_0) of an infectious disease. R_0 is always defined as the expected number of secondary infectious cases generated by an average infectious case in an entirely susceptible population [17]. R_0 could be expressed as $R_0 = kbD$. Where k is the contact times for each infected individual in unit time, b is the infection probability for per contact between infected and susceptible individuals, and D is the mean duration of infection. To control the epidemic situation, R must be maintained below 1 by interventions.

Ebola propagation parameters in West Africa have been revised by Chinese Academy of Sciences, as shown in Table 2. The average durations of exposed period and infected period are 10.2 and 5 separately; therefore, β and γ are set as 0.098 and 0.2. Then the average reproductive number R_0 is 2.041, which equals γ/β. Additionally, the value is coincidence to the statics data ($R_0 \in [1.4, 2.26]$) in West Africa.

In our design, the course model of Ebola is built based on these parameters above. The exposed period lasts 2 to 21 days with no or little infectivity, while the infected period lasts 1 to 10 days with intensity infectivity. In addition, the mortality is set as 70%, according to the WHO report [8]. In the last period, the recovering time conforms to $U(0, 10)$ days, while the death time conforms to $U(5, 17)$ days. The course model and the propagation parameters are shown in Figure 3.

Table 2: Ebola propagation parameters in West Africa.

Parameter	Mean	sd	Variance
Incubation period	10.20	6.00	36.00
Infectious period	5.00	4.70	22.09
Admission to death	4.20	6.40	40.96
Admission to discharge	11.80	6.10	37.21
Generation time	15.30	9.30	86.49

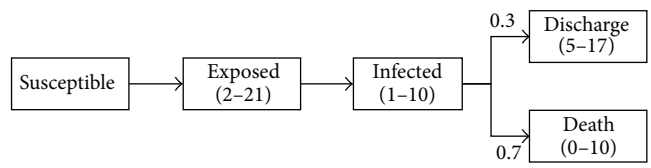


Figure 3: Ebola course model.

3.2. Propagation Model. As previously discussed, the main manner of propagating EVD is biologic shift. In our study, it is divided into contact propagation and noncontact propagation. Contact propagation spreads EVD by contacts between

Table 3: Detailed parameters of different buildings.

Building type	Capacity	Contact frequency	Survival time
Home	10	20	24 h
School	1000	30	12 h
Factory	1000	10	6 h
Restaurant	200	50	18 h
Hospital	1000	10	24 h

Table 4: The verification of infection probability.

IP	R_0	Dt	Comment
0.005	1.7141	40	Failed
0.008	1.9107	27	Failed
0.01	2.2108	21	Selected
0.02	3.1105	11	Discard
0.05	4.7008	5	Discard

the infected and susceptible ones. In the process, contact infection probability (CIP) and contact time (CT) are discussed. According to the literature [3], EVD may survive several hours outside the bodies, and one may suffer from the virus according to polluted materials used by patients. Noncontact propagation describes the infections through polluted materials or buildings. For example, an infected individual A has ever stayed in a building and B is probably infected once entering into this building. Additionally, the noncontact infection probability (NCIP) is relevant to CIP, and it decreases along with the time (t) elapsing. Once no patient enters into the building, the infectivity of environment will be weakened after the duration time (T). In our study, exponential distribution with the weakening parameter α is introduced to depict the process, and α is predefined as 1. Assuming that CIP is x_1, NCIP is x_2, and their relationship is formalized as

$$x_2 = x_1 * e^{-\alpha(T-t)/T}, \quad t \in [0, T]. \tag{5}$$

In our design, contact infections mainly consider the contact frequency of individuals in different types of environments, while noncontact infections mainly consider the survival time of Ebola virus in special environment. As shown in Table 3, the capacity, contact frequency, and survival time of Ebola in different environments are presented.

In addition, both CIP and NCIP are related to self-protection (SP) levels. SP describes the immunity levels of individuals, associated with the usage of antibiotics, prevention broadcast and physical condition. In a sense, these factors are also positive correlation to SP. If SP is set as y, the actual infection probability (AIP) will be shown as follows:

$$\begin{aligned} x_{\mathrm{AIP1}} &= x_1 * (1 - y), \quad y \in [0, 1], \\ x_{\mathrm{AIP2}} &= x_2 * (1 - y), \quad y \in [0, 1], \end{aligned} \tag{6}$$

where x_{AIP1} is the actual infection probability of contact infection, while x_{AIP2} is the actual infection probability of noncontact infection. Of course, the values of SP are different among individuals. Especially, medical workers are designed with high SP.

4. Prediction and Analysis of EVD Epidemic

The scenario of Ebola propagation is set as follows. An Ebola carrier entered into the city of Beijing, and the patient was not isolated immediately as the unobvious symptoms in incubation period. Once symptoms have been exposed for a few days, the cross infections would cause the outbreak of Ebola epidemic. The main tasks of this experiment are listed as follows: (1) the propagation among typical occupations is analyzed, and corresponding measures are also discussed; (2) the infections in typical environments along with corresponding interventions are analyzed; (3) the roles of vaccines, antibiotics, and treatments are discussed; (4) the impacts on population and environment are analyzed, and the roles of government are also discussed.

4.1. Prediction of EVD Epidemic. As previously discussed, the infectivity is determined by infection probability and contact frequency. Contact frequency (CF) is predefined in artificial Beijing, while the infection probability (IP) will be confirmed by the propagation parameters in West Africa. Actually, the infection probability is just the internal parameter for prediction, which is different from that in medical science. The estimated reproduction numbers ($R_0 \in [1.4, 2.26]$) and doubling times (Dt $\in [15.7, 30.2]$) are gained, according to the literature [8]. In artificial Beijing, the infection probability will be gained in the preparation experiment according to R_0 and Dt.

In the preparation experiment, the basic infection probability is verified by comparing simulation result with theoretical value. Considering the differences among occupations, four root infectious cases are supposed, including a worker, a doctor, a student, and a retiree. Subsequently, 100 days of propagation is simulated at different infection probabilities. The interaction experiment refers to millions of individuals and leads to huge consumption in communication and computation. The operation environment consists of 24 cores and 128 GB memories. It will take 1.5 hours to simulate the whole process, if simulation step is set as 10 minutes. Each sample will run 10 times and calculate the average R_0. The detailed parameters are listed in Table 4.

Actually, low infection probability always results in failure of propagation, and thus it is hard to simulate the outbreak of epidemic. In contrast, high infection probability always results in high infection velocity and amazingly increased infection cases. By comparison and analysis, the infection probability is finally set as 0.01, where R_0 is 2.2108 and Dt is 21. Although R_0 is a little higher than the average value in West Africa, it is still reasonable to predict the epidemic spreading in Beijing. On the one hand, effective contacts are always restricted by many factors, which would lead to low infections; on the other hand, the population density of Beijing is higher than that of Africa.

Based on current infection probability, the infection scenario in 100 days is simulated. The total infections will reach

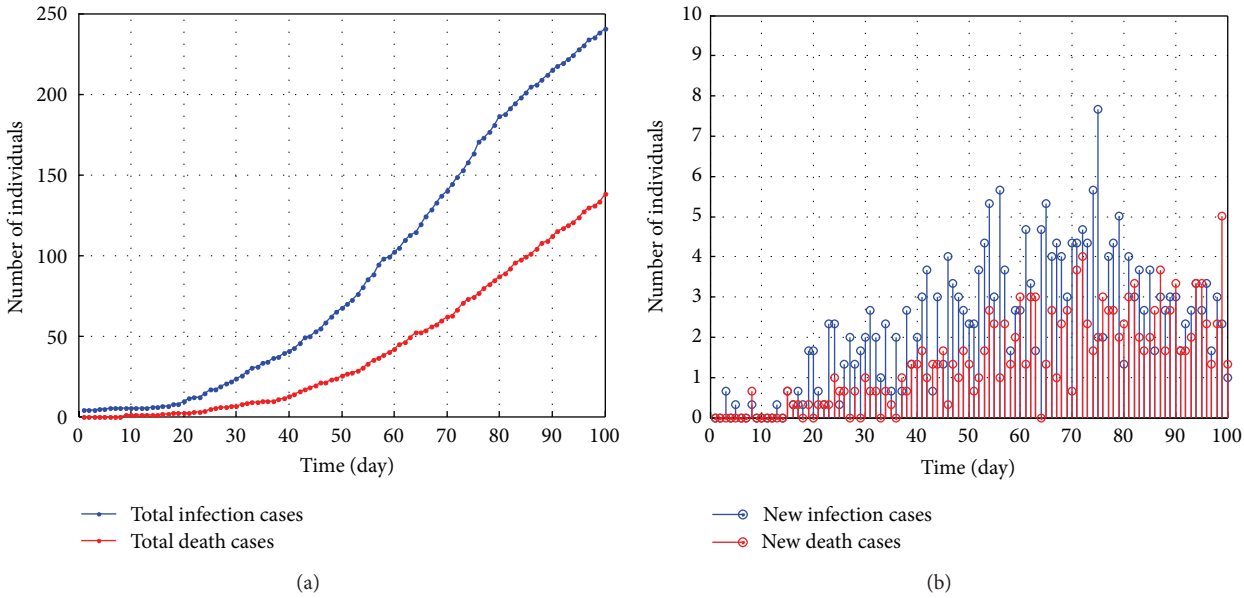


Figure 4: Total infection cases and new cases in 100 days under no interventions: (a) total infection cases and death cases; (b) new infection cases and death cases.

240 in the 100th day along with 140 death cases. Meanwhile, both new infections and death cases are at the rate in single digits every day. As shown in Figure 4, the growth slows down in the 75th day. Actually, infection cases have reached saturation in the initial social networks. For instance, an infectious student has infected his families, friends, and classmates as much as possible during this period. Meanwhile, these individuals just compose a relative isolated social network. Once the infection transfers to another group, new growth will emerge.

Moreover, the situations in 180 days and 240 days are also predicted. As shown in Figure 5, the infections emerge exponentially in the 150th day. The infection cases will reach 10 thousands in the 180th day, and new cases are about one hundred per day. Subsequently, the infection cases will reach 68391 in the 240th day, and new cases are about 1 thousand per day. Of course, the result is gained in the absence of interventions and it is impossible in practice.

4.2. Experiment Analysis. As previously discussed, the incubation and infectious periods are 10.2 and 5 separately. Therefore, the average period of propagation is 15.2, which equals the sum of 10.2 and 5. By computing, the average generations (AG) in 100 days, 180 days, and 240 days are 6.58, 11.8, and 15.79, respectively. Additionally, the average reproductive number (R_0) is 2.041. Subsequently, the infected cases at different times can be calculated. As shown in Table 5, the simulation results (SR) and the theoretical infections (TI) are basically in the same magnitude. Of course, the simulation infections are smaller than theoretical values. However, it is rational because of the rigorous interaction conditions in experiment. In a word, it predicts the possible epidemic situations in a long time, and the simulation result is reliable in a sense.

Table 5: Comparing with the theoretical and simulation infections.

Time (day)	AG	TI	SR
100	6.58	437	240
180	11.84	18647	10222
240	15.79	312330	68391

In the prediction, detailed infections of occupations, environments, and infection manners are also analyzed. As shown in Figure 6(a), residential buildings are the main places for propagation, which take the proportion of 51%. It shows that families are the most possible infections if no interventions are taken. As shown in Figure 6(b), medical workers are always at high-risk environment, and infected medical workers take the proportion of 15% although they only take a small part in the whole population. As shown in Figure 6(c), noncontact infection cases take the proportion of 4%, which mainly takes place in restaurants or hospitals. In addition, the proportions of different propagation generations are also shown in Figure 6(d).

5. Sensitivity Experiments

In the section, a series of sensitivity experiments are presented to analyze the validity of interventions against Ebola. Traditional interventions mainly include isolation of symptomatic cases, observation of asymptomatic individuals, and inoculation on focus groups. In our design, the ratio of seeing a doctor (RSD), the time of seeing a doctor (TSD), isolation ratio (ISR), immunization ratio (IMR), disinfection degree (DD), and self-protection (SP) are analyzed and discussed in detail. According to the experiences with H1N1 and SARS, the infection cases are at the level of hundreds. Therefore,

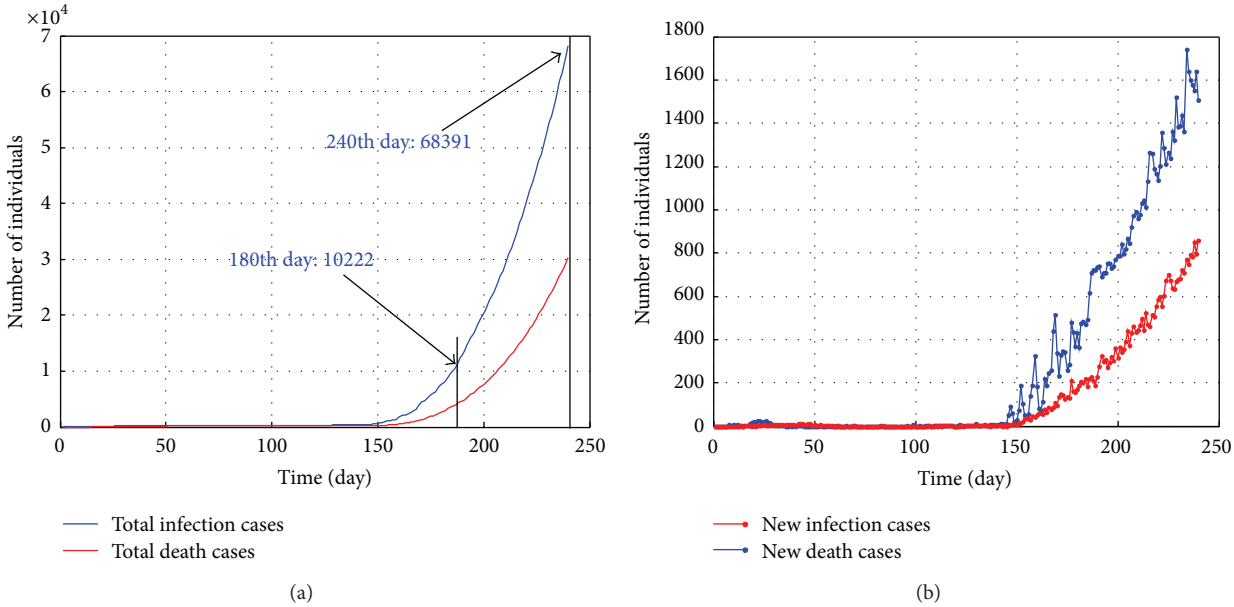


Figure 5: Total infection cases and new cases in 180 days and 240 days under no interventions: (a) total infection cases and death cases; (b) new infection cases and death cases.

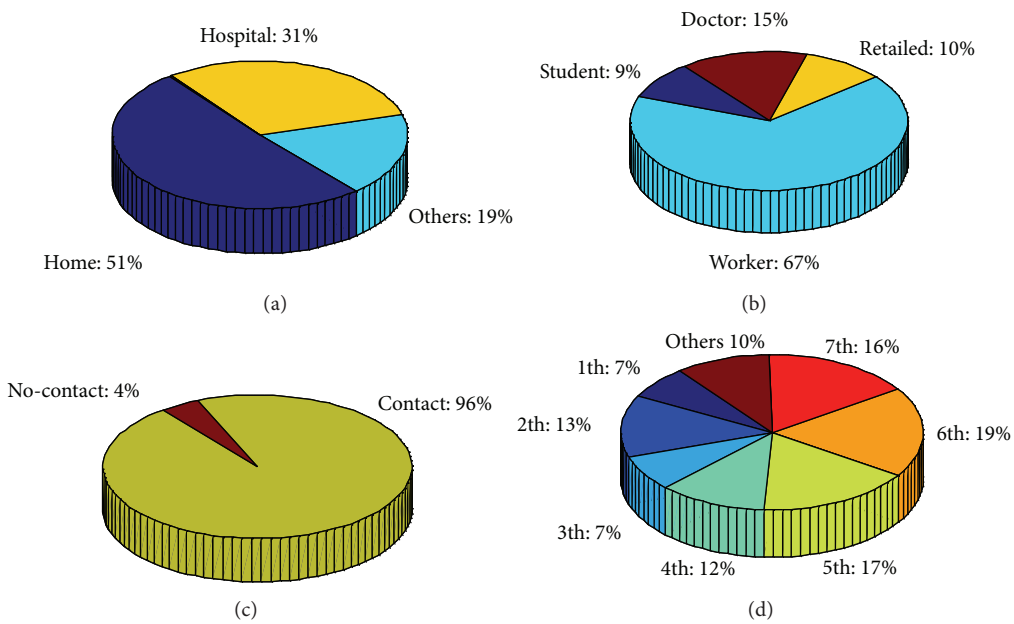


Figure 6: Detailed infections in 100 days under no interventions: (a) infection locations at home, hospital, and other places; (b) the occupation type of infections in worker, student, doctor, and retailed individuals; (c) the proportion between noncontact infections and contact infections; (d) the proportion of different propagation generations.

sensitivity experiments just simulate the epidemic spreading in 100 days and then analyze the validity of interventions.

5.1. The Ratio of Seeing a Doctor. The ratio of seeing a doctor is able to affect the whole infections directly. Once an exposed patient contacts with other individuals, it is possible to spread EVD at the same time. In addition, the polluted materials also have infectivity since EDV is able to survive for several hours outside the host. Assuming that TSD is the 1st day, RSD are set as 0.5, 0.7, and 0.9, respectively, and the corresponding infections are analyzed.

As shown in Figure 7, high ratio of seeing a doctor always leads to low infection cases. If RSD reaches 0.9, the infection cases are less than 20 during 100 days, while the number will increase to 70 if the ratio is 0.5. Therefore, it is necessary to increase the RSD as much as possible.

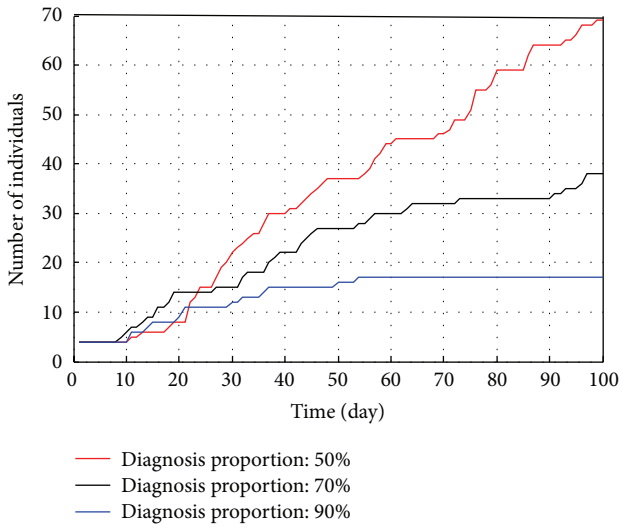


Figure 7: Total infections under different RSD.

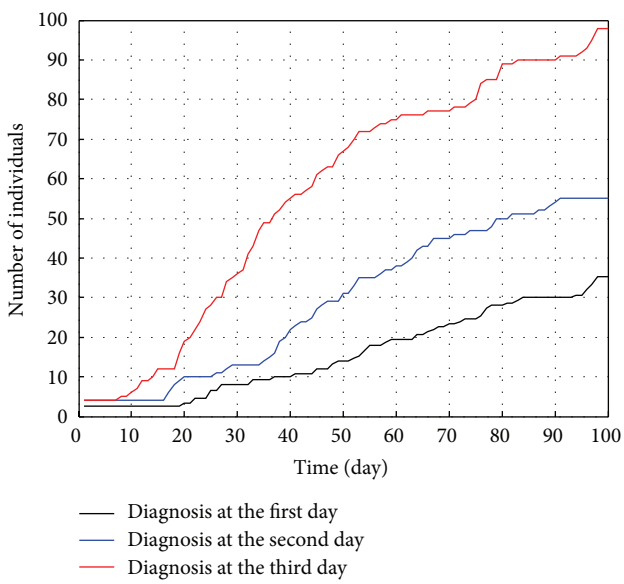


Figure 8: Total infection cases under different TSD.

5.2. The Time of Seeing a Doctor. Theoretically, early isolation and treatment will obtain better effects. In practice, EVD carriers are always diagnosed after the symptoms emerge. However, it is the exposed period that leads to the propagation between the infected and susceptible ones. Therefore, it is key to diagnosing Ebola carriers as early as possible. Experimental hypothesis is that there is only one initial carrier, and all subsequent infections will be sent to hospital. In our design, TSD is in the first day, second day and third day after the symptom exposes while RSD is set as 70%.

As shown in Figure 8, the sensitivity of TSD is demonstrated. If all patients are sent to hospital in the 1st day, the number of total infection cases is about 35, the 2nd day is 55, and the 3rd day soars to 98. Then it is easy to draw the conclusion that immediate treatments are necessary under the emergency of Ebola. In practice, patients are always sent to hospital in the 2nd day because they do not pay enough attention to the symptoms originally.

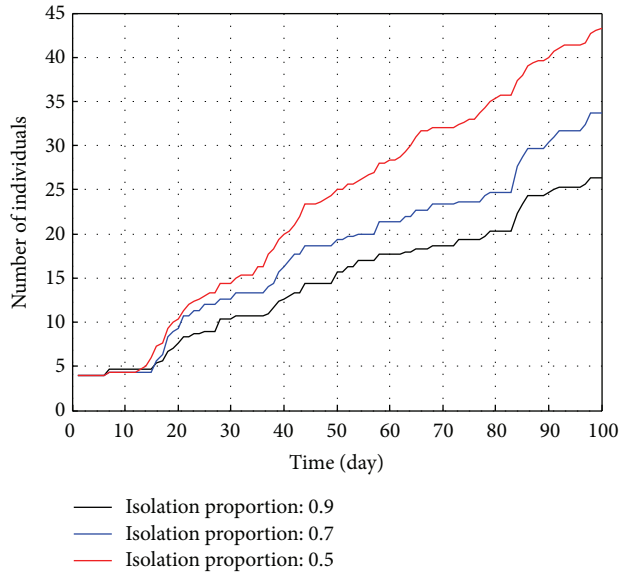


Figure 9: Total infections under different ISR.

5.3. Isolation of Potential Infected Individuals. In our study, two ways of infecting Ebola are outlined: indirect contact and direct contact. Although there are no obvious symptoms and infectivity in incubation, it is also necessary to isolate the potential infected individuals, as Ebola carriers are hard to distinguish. Once a large number of Ebola carriers transfer to the exposed, epidemic will outbreak in large scale. Therefore, it is necessary to isolate them who have contacted with diagnose patients. Simultaneously, potential infections by noncontact style would also be isolated. For instance, individuals entering the infectious environment also need to be isolated.

Assume that TSD is the 2nd day, RSD is 0.7, and the differences are the ISR of potential infected individuals. As shown in Figure 9, ISR are set as 0.5, 0.7, and 0.9, and the epidemic situations are also analyzed separately. The result illustrates that isolating focus individuals will reduce the infection cases. The infection cases are 43 when ISR is 0.5, while the total infections have decreased to 26 if ISR is 0.9.

5.4. Immunization on Potential Infections. Although effective vaccines are still under test, there is no doubt that they would play an important role in the subsequent fighting against EVD. The experimental hypothesis is that effective vaccines are under manufacture and the amount is adequate. Assume that TSD is the 1st day, RSD is 0.9, ISR is 0.9, and then IMR on potential infections are discussed. As a rule of thumb, only minority of individuals may receive vaccines. Once immunization is taken, individuals will never be infected by others.

In this work, IMR are set as 10%, 20%, and 30%, and the corresponding infections are analyzed separately. As shown in Table 6, the infection cases have almost reduced half if IMR

Table 6: Infection and death cases under different IMR.

Inoculation ratio	Infection case	Death case
10%	210	143
20%	168	98
30%	116	72

Table 7: Average TI and NCI under different DD.

DD	TI	NCI
0.1	230	9.8
0.3	215	8.1
0.5	197	6.9
0.7	181	5.1
0.9	168	3.2

Table 8: Average TI and infected MW under different SP.

SP	TI	Infected MW
0.2	216	31.05
0.4	203	24.78
0.6	184	15.22
0.8	157	10.36

is 30%. Vaccines slow down the spreading trend obviously; however, the cost of vaccines is always expensive and the amount is always limited. Therefore, it needs to inoculate the focus groups accurately such as medical workers and potential infections.

5.5. Interventions on Noncontact Infections. In this part, noncontact propagation in public places is discussed. For instance, in restaurant, noncontact infections are viewed as the main manner because cross infection through dinnerware is serious. The infectivity of environment is determined by virus dose and vitality, which may be weakened by disinfection. Actually, disinfection can decrease the virus dose and reduce the virus vitality at the same time.

As shown in Table 7, noncontact infections (NCI) under different DD are analyzed. DD are set as 0.1, 0.3, 0.5, 0.7, and 0.9 separately, and the result shows that NCI decreases along with the increasing of DD. Moreover, it also reduces the total infections (TI) indirectly. Therefore, it is effective to improve DD in epidemic spreading.

5.6. Interventions on Medical Workers. According to the WHO report [8], medical workers (MW) account for a high proportion in the whole cases for the frequent contacts with Ebola patients. Additionally, EDV is able to survive for a period in environment due to halfway disinfection and lead to new infections in the manner of noncontact. To study the infection of MW, two propagation manners are analyzed at the same time. In the absence of interventions, SP is set as 0.2, 0.4, 0.6, and 0.8 separately, and the validity of interventions is discussed.

As shown in Table 8, the infections are associated with SP. The infected cases of MW are about 10 if SP is 0.8 while the number has increased to 30 if SP is 0.2. Moreover, these infections would infect others with different occupations and lead to the sharp increase of Ebola cases.

6. Four Levels of Response Strategies on Ebola Propagation

6.1. Four Levels of Emergency Response Strategies. According to the emergency responses of Chinese government, there are 4 levels of response strategies against epidemic spreading. Level 4 is the weakest, including disinfection, hospital watch, and treatment; level 3 adds the isolation of familiar-contact persons and preparation of vaccines; level 2 includes trace isolation, suspending classes or works, and inoculation in small scale; level 1 is the most rigorous, in which inoculation is taken, classes are suspended, and factories are shut in large scale. In our design, different levels of response strategies are simulated, and combined interventions are estimated. As shown in Table 9, isolation, immunization, disinfection, self-protection, and diagnose factors are analyzed simultaneously. Where, LT is the time of loading interventions, TSD is the time between symptoms emerging and disease confirmed, RSD is the ratio of visiting a doctor, ISR is the isolation ratio, IMR is the immunization rate, SP is the self-protection, and DD is the disinfection degree.

As shown in Figure 10, epidemic situation will be under control by loading response strategies at level 2. Once taking the strongest interventions at level 1, new infection cases will decrease in a short time. Generally, rigorous interventions against Ebola will achieve better results at the cost of social orders and public resources. Luckily, since the outbreak of SARS in 2003, Chinese government has established the emergency management system, which has passed through the trail of H1N1 influenza. Although it is difficult to match all experiment parameters to the actual situation consistently, in a sense, four levels of response strategies are able to delineate the epidemic spreading under different interventions.

6.2. Result and Analysis. As previously discussed, response strategy at level 2 is the priority selection if a single Ebola case emerges. Under serious situation, response strategy at level 1 will be taken subsequently. The result shows that the epidemic will be under control if a single Ebola infection case emerges in the city of Beijing.

In our study, the results are credible to some degree, and the detailed analysis is listed as follows. Firstly, the construction of artificial Beijing is directly supported by NSFC. With the assistance of related institutions, the generation of individuals and buildings and their distributions meet the statistical features. Moreover, the modeling work of individuals' behaviors and social networks integrates some positive results of other research teams such as Fudan University. In addition, the reasonability of artificial Beijing has been verified in the case of H1N1 influenza, which is affirmed by NSFC. Secondly, disease model is built according to the propagation parameters in West Africa. Although these parameters may be different from that in Beijing, it has been yet revised by Chinese Academy of Sciences. Moreover, interventions are

Table 9: Detailed design of four levels of response strategies.

Response level	LT (day)	TSD (day)	RSD (%)	ISR (%)	IMR (%)	SP (%)	DD (%)
Level 4	15	4	0.5	0	0	0	0
Level 3	10	3	0.5	0.5	0	0.2	0
Level 2	7	2	0.7	0.7	0.1	0.5	0.6
Level 1	3	1	0.9	0.9	0.3	0.8	0.8

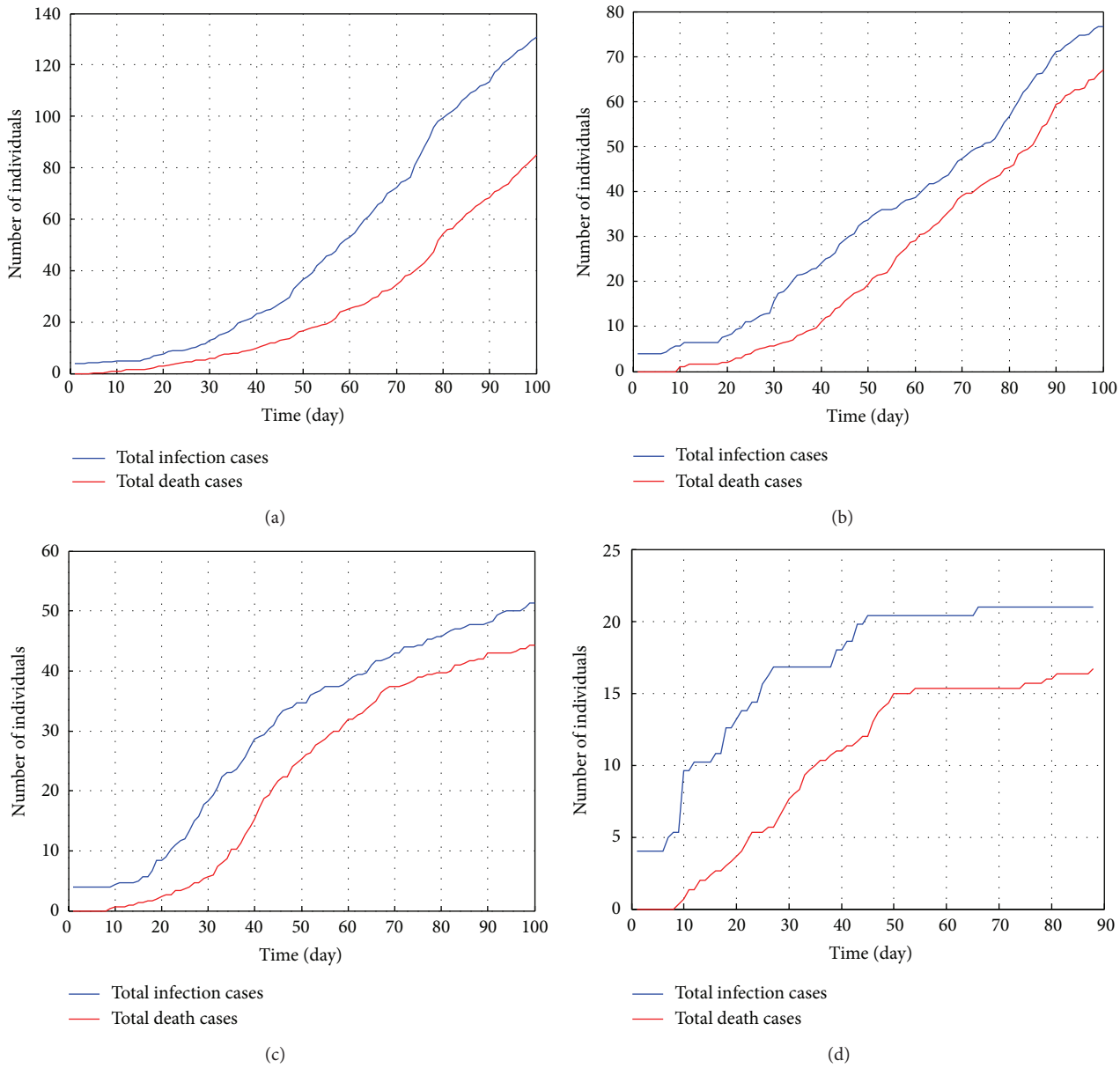


Figure 10: Total infection case and death case under different levels of response strategies: (a) the 4th level; (b) the 3rd level; (c) the 2nd level; (d) the 1st level.

designed according to the emergency response plans, and the simulation result is coincident with the conclusions of corresponding researches. Thirdly, quantitative analysis also testifies the reasonability of simulation result. As previously discussed, R_0 is the key parameter in disease propagation. Once R_0 is bigger than 1, the infection disease will spread; once R_0 is less than 1, the epidemic situation will be under control. In a word, the goal of interventions is to reduce R_0 from the current value to below 1. In our experiments, the current R_0 of 4 levels of response strategies are 1.82, 1.43, 0.93,

and 0.48 separately at the 100th day. Therefore, it is reasonable to control the epidemic situation if rigorous interventions are taken immediately.

7. Conclusion and Discussion

7.1. Conclusion. The main work of this paper is summarized as follows. Firstly, artificial Beijing is reconstructed to meet the demand of epidemic spreading. In addition, individuals' behaviors are optimized by the technology of machine learning. Secondly, Ebola course model and the propagation model are also built according to the propagation parameters in West Africa. Thirdly, the epidemic situations of Ebola influenza in 100 days, 180 days, and 240 days are predicted, and the infection cases among different occupations and environments are analyzed separately.

In this paper, the propagation process of EVD and its corresponding interventions are simulated based on artificial Beijing. Moreover, diagnosis, isolation, immunization, self-protection, and disinfection are discussed. In terms of the features of fluid shift, two manners of infection (contact and noncontact) are also analyzed. Finally, it is rational to get the conclusion that it is impossible to bring the outbreak in large scale, though Ebola imported risk shall exist.

7.2. Discussion. This paper proposed a new method to study the disease propagation based on virtual city. The contributions of our work are outlined as follows. Firstly, artificial Beijing is generated based on demographics. Generally, some agent-based systems are always built based on simple rules, which cannot simulate the real activities of individuals. In our study, the distributions of population, environment, and social networks comply with the statistical features. Secondly, individuals' behaviors are optimized by machine learning. In our design, individuals are feasible to adjust their behaviors under special situation by replanning their schedules. Simultaneously, the optimization improves the computing performance, which involves large scale entities. Thirdly, Ebola's propagation process is simulated through the contacts among individuals. In our study, the course period and propagation characteristics of disease are two main factors in epidemic spreading. For any diseases, if we build their disease course model and propagation model, it is easy to simulate and analyze the process of epidemic spreading.

It is a scientific method with universality. With the increasing complexity of social systems, studying these problems in a traditional way becomes impossible. Although survey and qualitative analysis are able to interpret some phenomena, we can hardly explore the immanent reasons within phenomena. In our study, artificial society is a dynamic and involving virtual system, which would provide basic environment for kinds of complex experiments in social, economic, and military fields. On the one hand, it allows users to expand the attributes of existing entities to satisfy the requirement of experiments. On the other hand, we are able to cultivate the virtual society toward the desired direction and support the corresponding researches. For instance, it is necessary to cultivate the traveling mode of individuals to study personal evacuation under emergency situation. Naturally, the reliability of result mainly depends on the rationality of virtual society and the accuracy of domain models. Although solving these problems perfectly is difficult, it is still significant to build valid models and optimize the artificial society continually. In a word, this approach is universal in a sense, and it will play an especially important role in emergency management.

However, several aspects in our work still need to be improved or optimized. (1) Artificial Beijing needs to be validated constantly. Although the generation process is based on geodemographics, it is still hard to depict the actual interactions among individuals and environment. Following aspects of the virtual city also need to be improved and optimized, such as social networks, the mapping between environments and individuals, and the individuals' behaviors. Although some work has been made, it is limited to depict the behaviors perfectly. (2) Ebola parameters need to be revised constantly. Although disease models are built according to the parameters in West Africa, the differences between them cannot be ignored. Therefore, the error inevitably exists in predicting the epidemic spreading in Beijing. For instance, many cases in Africa are infected by contacting the corpse of Ebola patients, while this funeral custom is nonexistent in China. (3) Four levels of response strategies cannot match well actual situations. The motivation is just to simulate and valuate the response interventions at different degrees, and it is hard to contain all possible factors. In addition, social costs of interventions are not considered in our design. (4) Some experimental parameters are not supported by data. For instance, contact frequency and infection probability are set by the rule of thumb. Although it is significant for predicting the epidemic situation, there is really no means in medical field. In addition, it is hard to carry out sensitivity experiments in practice. Once epidemic outbreaks, interventions tend to be taken in group, but not singly.

Although some aspects need to be consummated and strengthened in our work, the predicting experiment is still significant in practice. Firstly, infections cases are gained by individuals' contacts, and it is valid to forecast and analyze the epidemic situation. Secondly, sensitivity experiments are taken to analyze the roles of key factors in interventions. Thirdly, different levels of response plans are designed, which are significant for decision makers to estimate epidemic situations and take proper actions.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by National Nature Science Foundation of China under Grant nos. 91024030, 71303252, 61403402, 41201544, and 71373282.

References

[1] T. E. Fletcher, R. A. Fowler, and N. J. Beeching, "Understanding organ dysfunction in Ebola virus disease," *Intensive Care Medicine*, vol. 40, no. 12, pp. 1936–1939, 2014.

[2] N. Hawkes, "Ebola outbreak is a public health emergency of international concern, WHO warns," *The British Medical Journal*, vol. 349, Article ID g5089, 2014.

[3] J. J. Farrar and P. Piot, "The Ebola emergency—immediate action, ongoing strategy," *The New England Journal of Medicine*, vol. 371, no. 16, pp. 1545–1546, 2014.

[4] Z. Peng, F. Z. Chen, C. Bin, M. R. Qing, and Q. X. Gang, "A framework of computational experiment within special artificial society scene," *International Journal of Modeling and Optimization*, vol. 4, no. 2, pp. 104–110, 2014.

[5] Y. Ge, L. Liu, X. Qiu, H. Song, Y. Wang, and K. Huang, "A framework of multilayer social networks for communication behavior with agent-based modeling," *Simulation*, vol. 89, no. 7, pp. 810–828, 2013.

[6] Z. Cheng, X. Qiu, P. Zhang, and R. Meng, "An agent-based artificial transportation system framework for H1N1 transmission simulation," in *System Simulation and Scientific Computing*, Communications in Computer and Information Science, pp. 313–321, Springer, Berlin, Germany, 2012.

[7] WHO Ebola Response Team, "Ebola virus disease in west Africa—the first 9 months of the epidemic and forward projections," *The New England Journal of Medicine*, vol. 371, pp. 1481–1495, 2014.

[8] H. Jiang, "Simulating the Ebola virus disease transmission and outbreak in china using computational epidemiologic model," *Chinese Journal of Emergency Medicine*, vol. 23, no. 9, pp. 974–978, 2014.

[9] W. Fan, Y. Liu, and W. Wang, "Triangular framework and '4+1' methodology for public security science and technology," *Science & Technology Review*, vol. 27, no. 6, pp. 1–3, 2009.

[10] F. Wang, "Artificial societies, computational experiments, and parallel systems: a discussion on computational theory of complex social economic systems," *Complex Systems and Complexity Science*, vol. 1, no. 4, pp. 25–35, 2004.

[11] M. Tang, X. Mao, H. Zhou, and X. Tan, "An approach to modelling city-scale artificial society based-on organization metaphor," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 953–958, October 2012.

[12] B. Chen, L. Zhang, G. Guo, and X. Qiu, "KD-ACP: a software framework for social computing in emergency management," *Mathematical Problems in Engineering*. In press.

[13] Y. Ge, *An Agent-Based Framework and Generation of an Artificial Society*, National University of Defense Technology, Changsha, China, 2014.

[14] B. Chen, Y. Ge, L. Zhang, Y. Zhang, Z. Zhong, and X. Liu, "A modeling and experiment framework for the emergency management in AHC transmission," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 897532, 18 pages, 2014.

[15] T. A. Arentze and H. J. P. Timmermans, "A learning-based transportation oriented simulation system," *Transportation Research Part B: Methodological*, vol. 38, no. 7, pp. 613–633, 2004.

[16] X. Wang, Z. Cao, P. Y. Peng, and D. Zheng, "Prediction of the epidemic trends of pandemic H1N1 2009 using SEIR model," *International Journal of Virology*, vol. 18, no. 6, pp. 161–166, 2011.

[17] M. Lipsitch, T. Cohen, B. Cooper et al., "Transmission dynamics and control of severe acute respiratory syndrome," *Science*, vol. 300, no. 5627, pp. 1966–1970, 2003.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 401024, 11 pages
http://dx.doi.org/10.1155/2015/401024

Research Article

Exploiting Language Models to Classify Events from Twitter

Duc-Thuan Vo,[1] Vo Thuan Hai,[2] and Cheol-Young Ock[1]

[1] *School of Electrical Engineering, University of Ulsan, 93 Daehak-ro, Nam-gu, Ulsan 680-749, Republic of Korea*
[2] *Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 156-743, Republic of Korea*

Correspondence should be addressed to Cheol-Young Ock; okcy@ulsan.ac.kr

Received 23 November 2014; Accepted 2 January 2015

Academic Editor: Weihui Dai

Classifying events is challenging in Twitter because tweets texts have a large amount of temporal data with a lot of noise and various kinds of topics. In this paper, we propose a method to classify events from Twitter. We firstly find the distinguishing terms between tweets in events and measure their similarities with learning language models such as ConceptNet and a latent Dirichlet allocation method for selectional preferences (LDA-SP), which have been widely studied based on large text corpora within computational linguistic relations. The relationship of term words in tweets will be discovered by checking them under each model. We then proposed a method to compute the similarity between tweets based on tweets' features including common term words and relationships among their distinguishing term words. It will be explicit and convenient for applying to k-nearest neighbor techniques for classification. We carefully applied experiments on the Edinburgh Twitter Corpus to show that our method achieves competitive results for classifying events.

1. Introduction

Twitter (https://twitter.com/) is a social networking application that allows people to microblog about a broad range of topics. Users of Twitter post short text, called "tweets" (about 140 characters), on a variety of topics as news events and pop culture, to mundane daily events and spam. Recently, Twitter has grown over 200 million active users producing over 200 million tweets per day. Twitter is a popular microblogging and social networking service that presents many opportunities for researches in natural language processing (NLP) and machine learning [1–6]. Locke and Martin [5] and Liu et al. [4] train a classifier to recognized entities based on annotated Twitter data for Named Entity Recognition (NER). Some research has explored Part of Speech (PoS) tagging [3], geographical variation in language found on Twitter [2], modeling informal conversations [1], and also applying NLP techniques to help crisis workers with the flood of information following natural disasters [6]. Benson et al. [7] applied distant supervision to train a relation extractor to recognize artists and venues mentioned within tweets of users who list their location.

Classifying events in Twitter is a difficult task that focuses on the automatic identification and classification of various types of events in tweet texts. In Twitter, events are topics that often draw public attention, for example, football matches or natural disasters. Several approaches have been proposed to classify events for detection such as wave analysis [8, 9], topic model approach based on latent Dirichlet allocation [10], hierarchical Dirichlet processes [11], and text classification and clustering [12]. Kireyev et al. [8] explored the use of topics models for analysis of disaster-related Twitter data. Sakaki et al. [12] investigated the real-time interaction of events such as earthquakes in Twitter and proposed an algorithm to monitor tweets and to detect target events. However, existing approaches encounter failures from in either latent topics detection or analyzing terms relationships. Because topic model techniques [13–15] have only focused on how to list set of relevant words into a group (called topic) it is missed on analyzing relations between topics. Considering tweets have been discussed in two events shown in Table 1, we are easy to recognize that T_1 and T_2 are discussed in event 1 and T_4 and T_5 are discussed in event 2. However, if using topic models the system will group T_1, T_2, and T_3 in the same event category

Table 1: Some samples of discussed tweets in two events.

Category	Tweets	Relatedness with event
Event 1	T_1: Amy Winehouse has passed away aged 27.	Yes
	T_2: Amy Winehouse found dead at her home in North London.	Yes
	T_3: Nelson Mandela, who led the peaceful transition from white-only rule, has died aged 95.	No
Event 2	T_4: plane crash kills majority of KHL team Lokomotiv.	Yes
	T_5: plane crash in Russia kills 36 or 37 assumed to be hockey player.	Yes
	T_6: plane crash, helicopter, was in Moscow with 2 dead.	No

even T_3 does not belong to the event because set of relation words as <"*passed away*," "*dead*," "*died*"> in these tweets is in the same topic model. Likewise, T_6 will be grouped into event 2 with T_4 and T_5 together even if T_6 does not belong to this event because sets of relation words as <"*plane*," "*crash*," "*helicopter*">, <"*Russia*," "*KHL team*," "*Lokomotiv*," "*hockey*">, and <"*kills*", "*dead*"> in these tweets are within the same topic models, respectively. Due to limitations in using topic models, we therefore propose the method to exploit language models having relations reference to not only analyze topics but also analyze relatedness of event in tweets to overcome these problems.

In this paper, we investigate the use of generative and discriminate models for identifying the relationship of objects in tweets that describe one or more instance of a specified event type. We adapt language modeling approaches that capture how descriptions of event instances in text are likely to be generated. Our method will find the distinguishing term words between tweets and examining them with a series of relationships, extracted by language models such as ConceptNet [16] and LDA-SP [17]. These language models have been widely studied based on large text corpora within computational linguistic relations. Hence the relationship among distinguishing terms and common terms between tweets becomes clear to measure their similarity by examining them under each model. Measuring similarity between tweets is explicit and convenient to apply it in the classifier algorithms, such as SVM and k-nearest neighbor (kNN), to classify events in Twitter.

The rest of this paper is structured as follows. Section 2 presents related work that refers to research on event detection. In Section 3, we discuss exploiting language models. In addition, we present a method to calculate the similarity between tweets for event classification. In the next following section, experiments that are applied to the Edinburgh Twitter Corpus for event classification are presented and discussed. Section 5 ends with conclusions and future work.

2. Related Work

Several applications have detected events in Web to apply to weblogs [18–20], news stories [21, 22], or scientific journal collections [23]. Glance et al. [19] presented the application of data mining, information extraction, and NLP algorithms for event detection across a subset of approximately 100,000 weblogs. They implemented a trend searching system that provides a way to estimate the relative buzz of word of mouth for given topics over time. Nallapati et al. [22] attempted to capture the rich structure of events and their dependencies on a news topic through event models by recognizing events and their dependencies on event threading. Besides the standard word for based features, their approaches took into account novel features such as the temporal locality of stories for event recognition. Besides that, some researches [24–27] have analyzed social network to search or detect emergency events on the internet. Dai et al. [25] presented a cycle model to describe the internet spreading process of emergency events which applied the Tobit model by analyzing social psychological impacts. Hu et al. [27] analyzed historical attributes then combined with HowNet polarity and sentiment words on microblog which has network information transmission of social emergency events. And, they then provided the important guidance in the analysis of microblog information dissemination that has relatedness with social emergency events on internet. Meanwhile, Dai et al. [24] proposed a method to search the shortest paths of emergency events through IBF algorithm by analyzing social network.

Some research has focused on summarizing Twitter posts for detecting events [28–31]. Harabagiu and Hickl [28] focused on the summarization of microblog posts relating to complex world events. To summarize, they captured event structure information from tweets and user behavior information relevant to a topic. Takamura et al. [31] summarized Japanese Twitter posts on soccer games during the time when people provide comments and expressed opinions on the timeline of a game's progress. They represented user actions in terms of retweets, responses, and quoted tweets. In particular, Sharifi et al. [30] detected events in Twitter by summarizing trending topics using a collection of a large number of posts on a topic. They created summaries in various ways and evaluate those using metrics for automatic summary evaluation.

Recently, several approaches have been proposed to detect events from tweets using topic model approach [8, 10, 12]. Kireyev et al. [8] explored the use of topic models for the analysis of disaster-related Twitter data. Becker et al. [32] and Popescu et al. [33] investigate discovering clusters of related words or tweets which correspond to events in progress. Sakaki et al. [12] investigated the real-time interaction of events in Twitter such as earthquakes and propose an algorithm to monitor tweets and to detect a target event. Diao et al. [10] attempted to find topics with bursty patterns on microblogs; they proposed a topic model that simultaneously captures two observations such as posts published around the same time and posts published by the same user. However, existing approaches have still met with failure in either

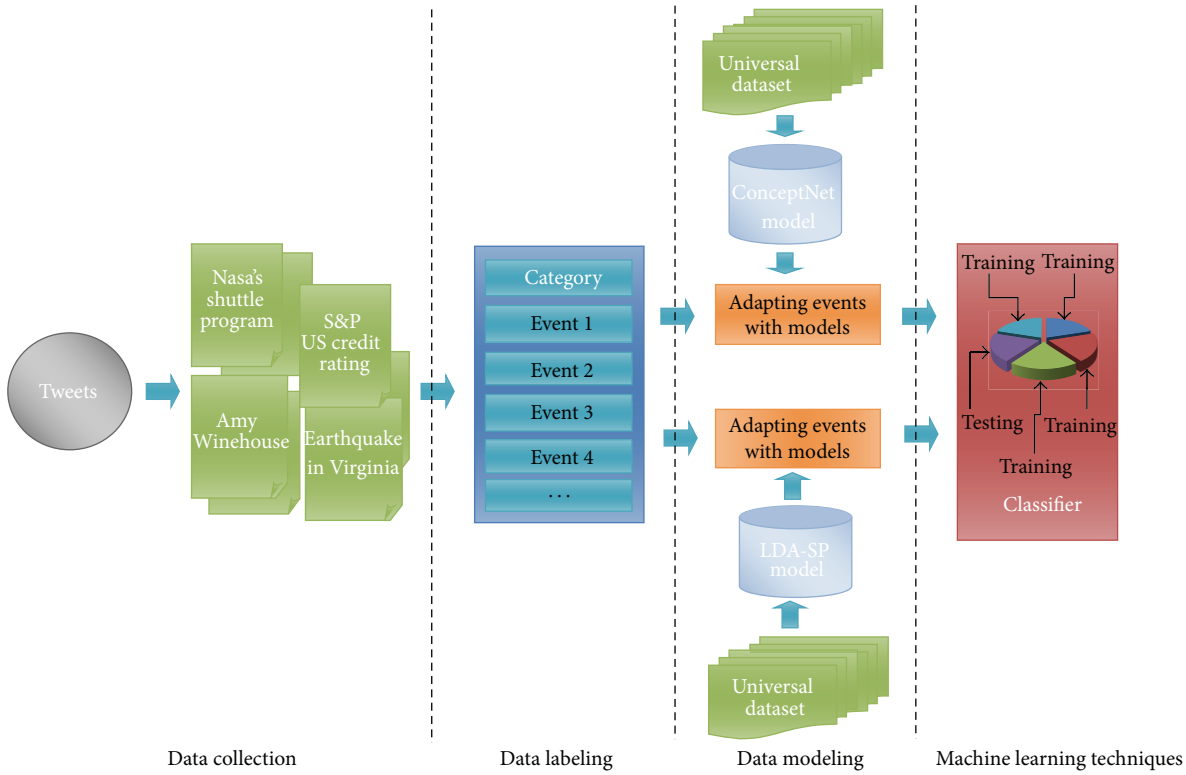


Figure 1: Proposed method.

latent topic detection or analyzing relationship terms, because tweets messages usually contain very limited common words in topics. Therefore, in this paper we propose a method to discover the relationship of objects in tweets by exploiting language models used to compare each of the snippets indirectly for classifying events in Twitter.

3. Exploiting Language Models to Classify Events

In this paper, we investigate the use of generative and discriminate models for identifying the relationship among objects in tweets that describe one or more instances of a specified event type. We adapt language modeling approaches that capture how descriptions of event instances in text are likely to be generated. We use language models to select plausible relationships between term words in tweets such as the relationship of "Object-Object" or "Object-relation-Object," which aim to detect the relatedness of an event in tweets. We assume that the data collection of language models contains suitable knowledge on the relationships among term words to discover the elemental relationship among tweets with a statistical analysis to classify events. We explore two types of language models that have obtained high correlation with human judgment such as ConceptNet and LDA-SP. These models are used for calculating the similarity of a pairwise of tweets for detecting events. The relationship between the discriminate term words of the tweets will be discovered by checking their relatedness under pairs of relations. In addition, the similarity between tweets is computed based on their common term words and the relationship between their discriminate term words. It is intuitive and convenient to apply it in classifier algorithms to classify events in Twitter. The general proposed method consists of four stages as (1) data collection, (2) labeling stage, (3) data modeling, and (4) machine learning shown in Figure 1. Stages 1 and 2 will be discussed in Section 4.1; stage 3 will discussed in Section 3; and state 4 will be discussed in Sections 3.3 and 4.2.

3.1. ConceptNet Model. To model the "Object-Object" relationships in tweets, we consider the ConceptNet [16] model. It is a large semantic graph containing concepts and the relations between them. It includes everyday basic, cultural, and scientific knowledge, which has been automatically, extracted from the internet using predefined rules. In this work, we use the most current version ConceptNet 5. As it is mined from free text using rules, the database has uncontrolled vocabulary and contains many false/nonsense statements. ConceptNet contains 24 relations with over 11 million pairs of relation. For example, "Nasa is located in United States" is presented as AtLocation ("Nasa", "United States") in ConceptNet model. Table 2(a) shows list of 24 relations, and Table 2(b) shows samples of four relations as MadedOf, AtLocation, MotivedbyGoad, and RecievesAction. Speer and Havasi [16] provide more details of the model in their paper. We first examine all relations in the ConceptNet 5 database (http://conceptnet5.media.mit.edu/) and define which are relevant

Table 2: ConceptNet model. (a) List of relations. (b) Samples of extracted relations.

(a)

MotivatedByGoal; CausesDesire; WordNet/ParticipleOf; MemberOf; HasA; NotDesires; UsedFor; AtLocation; Entails; DefinedAs; InstanceOf; HasPainIntensity; ReceivesAction; SimilarTo; RelatedTo; NotHasProperty; PartOf; HasLastSubevent; TranslationOf; HasProperty; NotHasA; CapableOf; WordNet/adverbPertainsTo; NotCapableOf; LocationOfAction; SimilarSize; HasPainCharater; HasContext; NotMadeOf; HasFirstSubevent; SymbolOf; LocatedNear; NotUsedFor; ObstructedBy; Desires; DerivedFrom; HasSubevent; MadeOf; Antonym; CreatedBy; Attribute; DesireOf; IsA; Causes

(b)

MadeOf		AtLocation		MotivatedByGoal		ReceivesAction	
Atomic bomb	Uranium	Nasa	United states	Fight war	Freedom	Bacteria	Kill
Computer	Silicon	Alcoa	Pittsburgh	Get drunk	Forget life	Army tank	Warfare
Gas	Oil	Tv channel	Russia	Pen	Write letter	Bread	Cook
Song	Music	Aozora bank	Japan	Join army	Defend country	Candle	Burn for light
Person	Live cell	Apartheid	Mall	Kill	Hate someone	Tomato	Squash
Light	Energy	Golden gate	Bridge	Live life	Pleasure	Tobacco	Chew
Carton	Wax paper	Art	Gallery	Sing	Performance	Supply	Store
Chocolate	Cocoa bean	Audience	Theatre	Socialize	Be popular	Ruby	Polish
Telephone	Electronics	Crab	Coastal area	Study	Concentrate	Money	Loan
Window	Glass	Handgun	Army	Visit museum	See history	Life	Save
...	...	...	...	...	...	...	...

to relations in target events by keywords matching (in experiments) to extract relations.

3.2. LDA-SP Model. To model the "Object-relation-Object" relationships in tweets, we adapt the LDA-SP model [17], which has been used for the selectional preference task in order to obtain the conditional probabilities of two objects in a relation. In particular, the LDA-SP, using LinkLDA [34], is an extension of latent Dirichlet allocation (LDA) [13] which simultaneously models two sets of distributions for each topic. The generative graphical model of LDA versus LDA-SP is depicted in Figure 2. In LDA-SP, they presented a series of topic models, at which objects belonged to them, for the task of computing selectional preferences. These models vary in terms of independence between $Topic_i$ and $Topic_j$ that is assumed. These two sets represent the two arguments for the relation $R\ (Topic_i, Topic_j)$. Each topic contains a list of relation words. Each relation, R, is generated by picking up over the same distribution, which keeps two different topics, $Topic_i$ and $Topic_j$, sharing the same relation (Figure 2(b)). The LDA-SP is able to capture information about the pairs of topics that commonly cooccur. To model the relations with LDA-SP, we also follow the data preparation in [21], which was automatically extracted by TextRunner [35] from 500 million Web pages. This resulted in a vocabulary of about 32,000 noun phrases, a set of about 2.4 million tuples with 601 topics in our generalization corpus. Some samples of topics extracted through LDA-SP are illustrated in Table 3.

3.3. Similarity Measures in Tweets. Classifying events in tweets from Twitter is a very challenging task because a very few words cooccur in tweets. Intuitively, the problem can be solved by exploring the relationships between tweets well; the intrinsic relationship among words may be discovered with a thesaurus. Hence, we present a method to discover the intrinsic relationships between objects based on statistical analysis of language models and then gain the similarity between tweets accordingly. We consider two types of relationships in tweets such as "Object-Object" and "Object-relation-Object."

"*Object-Object*". The event "*Death of Amy Winehouse*" is posted in tweets T_1, T_2, and T_3 shown in Figure 3. Traditional methods can only find one cooccurring term, "*Amy Winehouse*," in the tweets after removing stop words. However, if we analyze and compare the relatedness between the pairs <"*Singer*"-"*Amy Winehouse*">, <"*Amy Winehouse*"-"*passed away*"> and <"*Amy Winehouse*"-"*dead*">, and <"*Amy Winehouse*"-"*R.I.P.*">, closer relationships will be exposed: "*Object-Object*" as "$Topic_1$-$Topic_2$" where a set of terms {"*Singer*"; "*Amy Winehouse*"} is in $Topic_1$ and a set of terms {"*death*", "*passed away*", "*R.I.P.*"} is in $Topic_2$.

"*Object-Relation-Object*." The event "plane carrying Russian hockey team Lokomotiv crashes" is posted in T_4, T_5, and T_6 shown in Figure 4. We can discover the relationship between "Object-relation-Object" such as <"*Plane*"-"*crash*"-"*KHL team Lokomotiv*">, <"*Plane*"-"*crash*"-"*Russia*">, and <"*Plan*"-"*crash*"-"*KHL team*">. This also exhibits the closer relationships "*Object-relation-Object*" as "$Topic_3$-*crash*-$Topic_4$" where the term {"*plane*"} belongs to $Topic_3$ and a set of terms {"*russia*", "*khl team lokomotiv*", "*hockey*", "*khl team*"} belongs to $Topic_4$.

Our method extracts relation tuples from language models such as ConceptNet and LDA-SP. We treat all tweets from Twitter that are contained in the collection equally and then perform to match models of tuples generated from ConceptNet and LDA-SP with them. Hence, if we can discover relation tuples as "third-party" for both tweets and calculate the similarity between the two tweets by comparing the

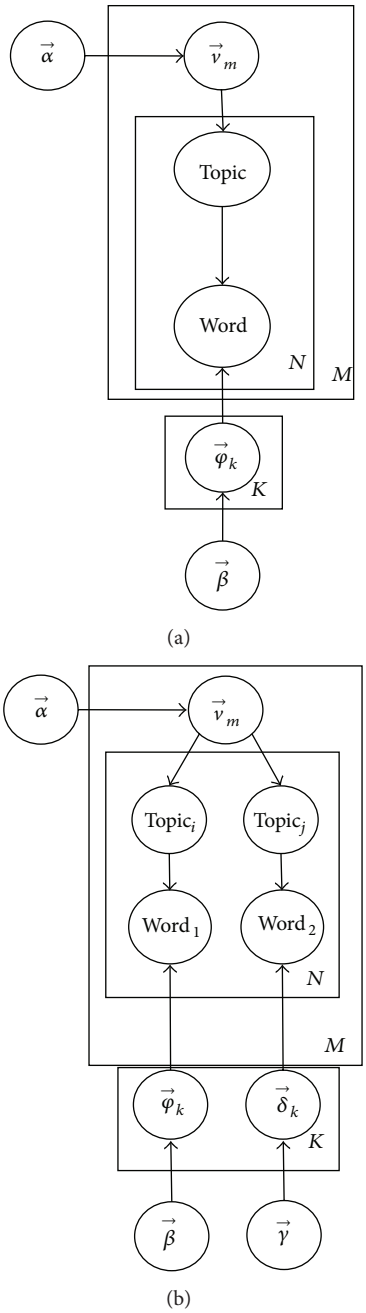


Figure 2: Graphical model of LDA model (a) versus LDA-SP (b).

distinguishing term words with these tuples, we may find the real relationship underlying the two tweets. We assume that the data collection language models contain sufficient knowledge about the relationships among term words, from which we can find the elemental relationship among tweets.

For computing similarity between tweets, we derive a set of relations, $R_i = (object_m, object_n)$ matched from language models and tweets combining with Bag-of-Words. Considering two original tweets, d_1 and d_2, in data collection D, we check with $object_m$, $object_n$ existing in each tweet which match with relation tuples $R_i = (object_m, object_n)$ extracted from ConceptNet model. In using LDA-SP, we exam not only relations but also $object_m$, $object_n$ existing in each tweet and then match them with relation tuples $R_i = (object_m, object_n)$ generated from LDA-SP. We then replace matched objects in tweets by relation tuples from language models. Thus, the relationship between the distinguishing terms of the tweets can be discovered by examining their relatedness under pairs of relations by "third-party." We consider calculating the similarity between two tweets based on their common terms and the relationship between their distinguishing terms. To calculate the similarity between two tweets in an event category, we represent them as vectors:

$$\begin{aligned} d_1 &= (w_1, w_2, \ldots, w_n) \\ d_2 &= (w_1, w_2, \ldots, w_n), \end{aligned} \tag{1}$$

where w_i is the weight of the ith feature in the vector of d_j and is defined by the tf-$i$df measure as follows:

$$w_i = \text{tf}_{ij} \times \log_2 \left(\frac{M}{\text{df}_j} \right), \tag{2}$$

where M is the total number of documents in the collection, df_j is the document frequency, that is, the number of documents in which term w_i occurs, tf_{ij} is the term frequency of term w_i in document d_j, and tf_{ij} is simply the number of occurrences of term w_i in document d_j.

With the relationship between the two distinguishing term words on a diversity of assigned model tuples, we can calculate the similarity of vectors d_1 and d_2 with the cosine method shown in

$$\text{sim}(d_1, d_2) = \cos(d_1, d_2) = \frac{d_1 \cdot d_2}{|d_1|\,|d_2|} = \frac{\sum_{i=n} w_{i1} w_{i2}}{\sqrt{\sum_{i=n} w_{i1}^2}\sqrt{\sum_{i=n} w_{i2}^2}}. \tag{3}$$

For classifying events from tweets, many classifiers first need to calculate the similarity between tweets. kNN is one of the best methods of similarity calculation and selection of a proper number of neighbors. Therefore, it is intuitive and convenient to apply similarity calculation between tweets to kNN for classifying events. If our proposed method can calculate the similarity among tweets more accuracy, the kNN will select more appropriate neighbors for a test case and the classification performance of kNN will be higher than original tf-$i$df, since the performance of kNN based on the similarity measuring method outperforms other methods with tf-$i$df measure. We conclude that the proposed method is more effective on calculating tweets similarity to classify events. The result will be discussed in more detail in experimentation section.

4. Experimentation

4.1. Experimental Datasets and Evaluation Measures. We have conducted experiments on the Edinburgh Twitter Corpus [36], a collection of events in Twitter, for event classification. The corpus contains 3034 tweet IDs spread into 27 event

Table 3: LDA-SP model. (a) Samples of list topics; (b) sample of list relations.

(a)

Topic 2: driver; civilians; soldiers; motorcyclist; teenager; thomas; policeman; us soldier; smith; george; motoris; father; ...
Topic 19: china; business line; japan; government; israel; judge; india; iran; russia; democrats; court; lawmakers; ...
Topic 106: myspacetv videos; the trail; leaves; moses; the curtain; stones; santa; flowers; victim; posters; stars; flames; ...
Topic 114: britney spears; paris hilton; angelina jolie; tom cruise; the actress; lindsay lohan; amy winehouse; singer; ...
Topic 116: fire; violence; the war; the storm; fighting; katrina; explosion; tornado; earthquake; civil war; dead; heaven; ...
Topic 171: john; david; mark; mike; steve; bill; michael; peter; scott; smith; johnson; brown; executive; robert; jeff; brian; ...
Topic 251: police; group; team; company; day; year; case; report; miller; officials; king; wilson; story; news; friday; ...
Topic 286: article; report; author; court; bible; story; letter; paul; reuters; researchers; statement; respondents; ...
Topic 390: car; train; vehicle; bus; fingers; truck; boat; plane; river; route; traffic; driver; aircraft; train; track; bike; ...
Topic 428: airplane; aircraft; pilot; sparks; birds; crew; terrorists; nasa; people; passengers; the captain; bullets; the jet; ...
Topic 433: family; couple; mary; sarah; thomas; elizabeth; margaret; jesus; jane; matt; martin; daniel; frank; anna; nancy; ...
Topic 454: game; operation; experiment; treatment; procedure; scenario; victim; exercise: measurement; error; idea; ...
Topic 525: bush; the president; president bush; jesus; paul; the minister; clinton; smith; george w. bush; obama; ...
Topic 561: the world; christians; mulims; no matter; americans; jews; catholics; normoms; the chinese; hindus; ...
Topic 570: the sun; light; the moon; the beam; earth; mars; venus; laser; darkness; stars; jupiter; a hush; radiation; ...

(b)

Relations	Relationship of topics (Topic_i-Topic_j)
Be cite	525–561; 251–286; 286–251; 251-251; 542–251; 371–286; 542–371; 542–286; 251–162; 134–286; 162–286; 371–251; 286–162; 286–171; 542–454; 286–538; 454–286; 286–10; 134–24; 538–286; 285-286; 575–454; 572–286; 328–286; 19–454; ...
Blame on	116–428; 329–531; 116–531; 329-329; 329–116; 116–584; 329–584; 584–531; 314–531; 116–329; 480–531; 171–116; 116–160; 239–584; 458–531; 404–531; 584–116; 196–116; 531–458; 584-584; 531–116; 196–531; 176–531; 545–147; 171–2; ...
Crash into	428–287; 428–571 390–106; 428–139; 428–390; 428-428; 390–139; 390-390; 390–287; 390–428; 428–570; 390–570; 139–106; 139–428; 139-139; 428–328; 287–106; 139–390; 390–328; 139–287; 428–374; 390–374; 287–139; 570–287; 106–428; ...
Spot in	114–433; 433-433; 116–525; 114–287; 287–433; 114–570; 405–433; 433–405; 251–433; 114-114; 223–433; 570–433; 433–570; 114–132; 287–405; 114–251; 543–433; 230–433; 223–570; 114–424; 433–287; 433–114; 570-570; 433–132; 223–279; ...

categories. Currently, some tweets in the dataset are deleted or lost from Twitter. We developed a tool using Twitter API (http://twitter4j.org) to collected documents including tweets, retweets, responses, and quoted tweets; we then filtered documents to guarantee that each event category contains at least 70 tweets. After the removal of noise and stop words, each word is stemmed into its root form. Table 4 shows the rest of nine significant event categories with checked mark for experiments as event 1, event 6, event 7, event 9, event 13, event 14, event 15, event 16, and event 21.

In this study, experiments are evaluated based on the precision, recall, and F-measure with our proposed method. The precision, recall and F-Measure are the evaluation metrics often used to rate the information retrieval system's performance. Precision is the number of correct results divided by the total number of returned responses; recall is the number of correct results divided by the number of results that should have been returned and F-measure is used to balance between the recall and precision as follows:

$$\text{Precision} = \frac{\text{number\_of\_correct\_responses}}{\text{number\_of\_responses}},$$

$$\text{Recall} = \frac{\text{number\_of\_correct\_responses}}{\text{number\_of\_corrects}},$$

$$F\text{-measure} = \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}}. \tag{4}$$

4.2. Experiments and Comparison. Checking similarity between tweets before experiments, we select some samples of tweets from experimental datasets as shown in Table 1. We used the tf-*i*df combined with the similarity functions to compare performance before and after using language models. Note that T_1 and T_2 were discussed in the same event; T_4 and T_5 were also discussed in the same event. And two pairs of tweets are, respectively, to calculate similarity with stop words removal. The result depicted in Table 5 shows that the tweets using ConceptNet and LDA-SP increase the similarity of questions from the same category. Moreover, if the tweets did not belong to target event like T_3 and T_6, the method will reduce the similarity measure that helps system performance of classifying efficiently.

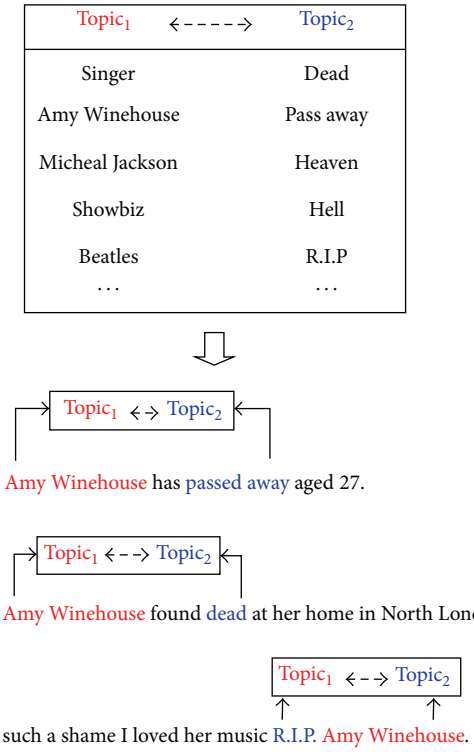


Figure 3: Relationship "*Topic*$_1$-*Topic*$_2$" in tweets of event "Death of Amy Winehouse."

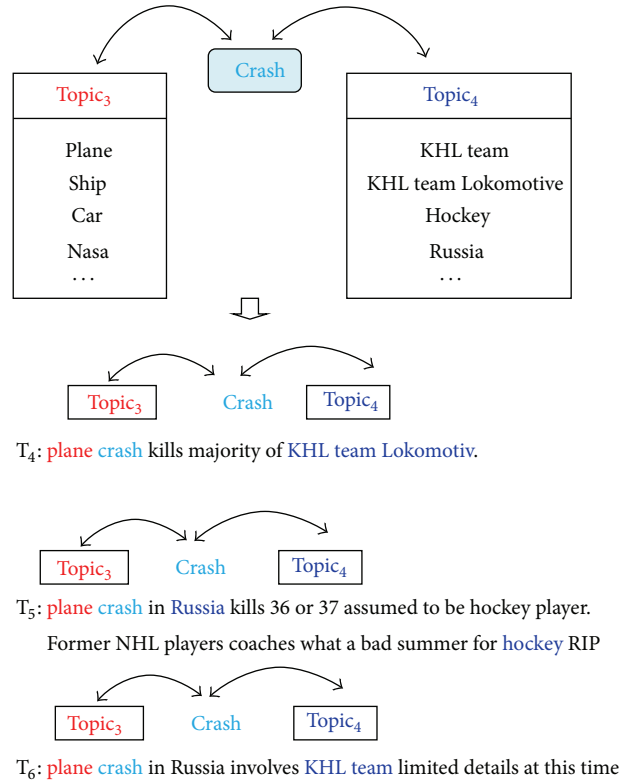


Figure 4: Relationship "*Topic*$_3$-*relation*-*Topic*$_4$" in tweets of event "plane carrying Russian hockey team Lokomotiv crashes."

To classify events, 70% of the tweets for each category are randomly selected for training, and the rest is for testing. In our experiments, we compare the performance of four classifiers implemented as follows: (1) baseline *k*NN (without language model); (2) baseline SVM; and the *k*NN method combining our proposed methods (3) *k*NN-M1 (*k*NN with language model ConceptNet) and (4) *k*NN-M2 (*k*NN with language model LDA-SP). The SVM is also constructed using the tf-*i*df method to weight each vector component of the tweet and is used as second baseline for comparison with our proposed methods. We chose SVM because of a powerful and robust method for text classification [37–39]. The evaluation follows 5-fold cross validation schema. Table 6 shows the performance results applied to 7 categories of events from Twitter. The bold numbers show the best F-measure of each event in four methods. For instance, the system obtained the highest F-measure of 85.3% in event 1 with method *k*NN-M2. Method *k*NN-M1 yielded better F-measure results in most of the event categories: event 6, event 7, event 9, event 14, event 15, and event 16. And, method *k*NN-M2 achieved better F-measure result in three categories: event 1, event 13, and event 21.

The overall performance comparison is presented in Figure 5. We can see that the performance of *k*NN-M1 outperforms *k*NN-M2, SVM, and *k*NN. Both of our proposed methods are also higher than the baselines, *k*NN and SVM, in most of the performance metrics. In the overall results, *k*NN-M1, *k*NN-M2, SVM, and *k*NN obtained an F-measure of 85%, 84.7%, 78.4%, and 76.8%, respectively.

4.3. Discussions. We believe that effective performance of our proposed methods is result of the following reasons.

First, noise and exclamative and repeated texts usually occur in the tweets of each event. The following are examples of such tweets. T$_1$: "*Sad day Sky sources now confirming Amy Winehouse is dead A musical legend who died way too young in my opinion*," T$_2$: "*Amy Winehouse found dead in her London flat according to sky news*," and T$_3$: "*Hmm...omg...gruuu Amy Winehouse is dead not totally surprised though ohhh*." We can observe that {"*Amy Winehouse*"; "*dead*"} is repeated text, {"*gruuu*"; "*ohhh*"} is noise text, and {"*Hmm*"; "*omg*"} is exclamative text. The repeated text will result in a positive value in the similarity measure; however, noise and exclamative texts will result in a negative value in the similarity measure. For preprocessing, stop words had been removed by a defined list of stop words automatically. However, we had checked and revised noise texts manually if they do not belong to list of stop words. For example, a lot of words "deadddddd" will be revised into "dead," or {"RIP," "R I P"} will be revised into "R.I.P."

The second reason we believe our method had effective performance is that quality universal datasets are used to build language models. In this study, more than five billion relation records extracted from Concept are used to build the models. In addition, models from LDA-SP are built by extracting 2.4 million tuples of relations and 601 topics. Furthermore, ConceptNet is a graphical relationship model which uses predefined rules. However, LDA-SP still has some

Table 4: Experimental datasets.

Category	Description	Number of tweets	Checked
Event 1	Death of Amy Winehouse	774	√
Event 2	Space shuttle Atlantis lands safely, ending NASA's space shuttle program	45	
Event 3	Betty Ford dies	8	
Event 4	Richard Bowes, victim of London riots, dies in hospital	27	
Event 5	Flight Noar Linhas Aereas 4896 crashes, all 16 passengers dead	9	
Event 6	S&P downgrades US credit rating	275	√
Event 7	US increases debt ceiling	73	√
Event 8	Terrorist attack in Delhi	40	
Event 9	Earthquake in Virginia	271	√
Event 10	Trevor Ellis (first victim of London riots) dies	63	
Event 11	Goran Hadzic, Yugoslavian war criminal, arrested	2	
Event 12	India and Bangladesh sign a peace pact	3	
Event 13	Plane carrying Russian hockey team Lokomotiv crashes, 44 dead	225	√
Event 14	Explosion in French nuclear power plant Marcoule	137	√
Event 15	NASA announces discovery of water on Mars	110	√
Event 16	Google announces plans to buy Motorola Mobility	130	√
Event 17	Car bomb explodes in Oslo, Norway	21	
Event 18	Gunman opens fire in children's camp on Utoya island, Norway	28	
Event 19	First artificial organ transplant	16	
Event 20	Petrol pipeline explosion in Kenya	27	
Event 21	Famine declared in Somalia	71	√
Event 22	South Sudan declares independence	26	
Event 23	South Sudan becomes a UN member state	7	
Event 24	Three men die in riots in Birmingham	12	
Event 25	Riots break out in Tottenham	19	
Event 26	Rebels capture Tripoli international airport, Libya	4	
Event 27	Ferry sinks in Zanzibar, around 200 dead	21	

Table 5: Sample of similarities calculated by the proposed methods and the tf-*i*df method.

Tweets	tf-*i*df	tf-*i*df + ConceptNet	tf-*i*df + LDA-SP
T_1: Amy Winehouse has passed away aged 27. T_2: Amy Winehouse found death at her home in North London.	0.16	0.365	0.4
T_1: Amy Winehouse has passed away aged 27. T_3: Nelson Mandela, who led the peaceful transition from white-only rule, has died aged 95.	0.123	0.078	0.084
T_2: Amy Winehouse found death at her home in North London. T_4: plane crash kills majority of KHL team Lokomotiv.	0	0	0
T_4: plane crash kills majority of KHL team Lokomotiv. T_5: plane crash in Russia kills 36 or 37 assumed to be hockey player.	0.433	0.452	0.468
T_5: plane crash in Russia kills 36 or 37 assumed to be hockey player. T_6: plane crash, helicopter, was in Moscow with 2 dead.	0.272	0.146	0.104

errors [17] in computing word statistics. In the experiment results, performance of ConceptNet is better than LDA-SP.

The third reason believed to be behind our method's effective performance is that the models extracted from LDA-SP are intensely analyzed compared to ConceptNet for relationship. However ConceptNet obtained better performance results. Texts from tweets are incomplete sentences that result in failures in grammar parsing for analyzing relation. We did not include grammar parsing for analyzing tweets based on LDA-SP model. Therefore, ConceptNet exhibits a better performance for classifying events from Twitter than LDA-SP.

Table 6: Experimental results.

Category	*k*NN			SVM			*k*NN-M1 (ours)			*k*NN-M2 (ours)		
	P	*R*	*F*	*P*	*R*	*F*	*P*	*R*	*F*	*P*	*R*	*F*
Event 1	76.3	71.6	73.8	75.2	75.5	75.3	86.1	77.5	81.6	88.2	82.6	**85.3**
Event 6	84.6	85.4	84.9	86.9	87.2	87.1	91.1	89.4	**90.2**	89.1	86.4	87.7
Event 7	78.9	72.3	75.5	80.4	76.2	78.2	87.5	82.3	**84.8**	82.4	78.9	80.6
Event 9	83.9	78.8	81.3	85.5	80.2	82.3	93.8	92.9	**93.4**	87.2	83.3	85.2
Event 13	83.6	72.4	77.5	82.8	75.6	79.1	86.2	80.5	83.3	87.3	82.6	**84.9**
Event 14	70.1	67.8	68.9	71.6	70.0	70.8	85.2	78.7	**81.8**	83.8	74.3	78.8
Event 15	79.3	71.5	75.2	81.0	70.8	75.6	90.1	87.9	**88.9**	88.8	85.8	87.3
Event 16	80.5	72.4	76.2	82.5	73.1	77.5	85.7	80.0	**82.8**	85.5	79.6	82.5
Event 21	81.6	74.1	77.7	82.4	76.8	79.5	83.9	77.8	80.7	85.4	77.1	**81.0**
Overall	79.5	74.4	76.8	79.9	77.0	78.4	87.9	82.4	**85.0**	87.4	82.3	84.7

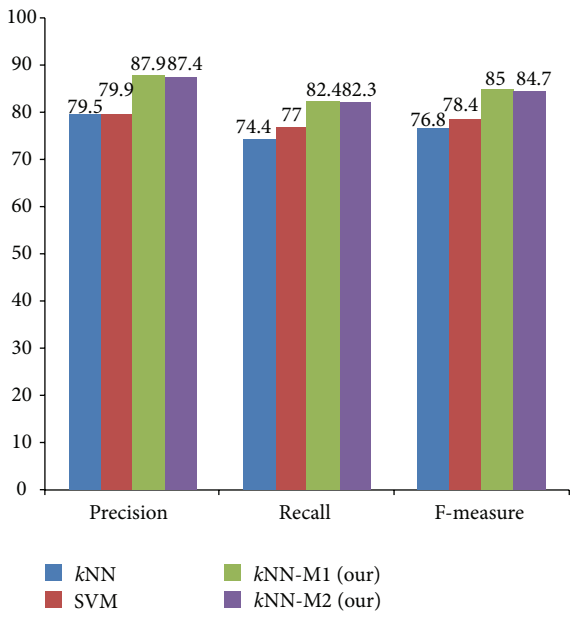


Figure 5: Overall performance comparisons.

5. Conclusion and Future Work

We have presented methods to classify events from Twitter. We first find the distinguishing terms between tweets in events and calculate their similarity with learning language models: LDA-SP and ConceptNet. Next, we discover the relationship between the distinguishing terms of the tweets by examining them under each model. Then, we calculate the similarity between two tweets based on their common terms and the relationship between their distinguishing terms. The outcomes make it convenient to apply *k*NN techniques to classify events in Twitter. As a result, our approach obtained better performance results with both ConceptNet and LDA-SP than other methods.

Regarding future work, the research has been suggested with attractive aspects to improve as follows. First, this approach can be considered for future work, including it with a larger corpus and experimenting with other event types. Second, we will continue to investigate how to apply grammar parsing in tweets so that we can analyze deeply relationships to serve for classifying events. Finally, the research can be applied unsupervised learning with semantic similarity models as pointwise mutual information (PMI) [40, 41] and latent semantic analysis (LSA) [42, 43].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A2006906). The authors would like to thank the anonymous reviewers for their constructive comments and suggestions on the paper.

References

[1] N. Dalvi, R. Kumar, and B. Pang, "Object matching in tweets with spatial models," in *Proceedings of the 5th ACM International Conference on Web Search and Data Mining (WSDM '12)*, pp. 43–52, February 2012.

[2] J. Eisenstein, N. A. Smith, and E. P. Xing, "Discovering sociolinguistic associations with structured sparsity," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT '11)*, 2011.

[3] K. Gimpel, N. Schneider, B. O'Connor et al., "Part-of-speech tagging for twitter: annotation, features, and experiments," in *Proceedings of 51st Annual Meeting of the Association for Computational Linguistics (ACL '11)*, pp. 42–47, June 2011.

[4] X. Liu, S. Zhang, F. Wei, and M. Zhou, "Recognizing named entities in tweets," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT '11)*, pp. 359–367, June 2011.

[5] B. Locke and J. Martin, *Named entity recognition: adapting to microbliogging [Senior thesis]*, University of Colorado, 2009.

[6] G. Neubig, Y. Matsubayashi, M. Hagiwara, and K. Murakami, "Safety information mining—what can NLP do in a disaster," in

Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP '11), pp. 965–973, 2011.

[7] E. Benson, A. Haghighi, and R. Barzilay, "Event discovery in social media feeds," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT '11)*, pp. 389–398, Portland, Ore, USA, June 2011.

[8] K. Kireyev, L. Palen, and K. Anderson, "Applications of topics models to analysis of disaster-related twitter data," in *Proceedings of the NIPS Workshop on Applications for Topic Models: Text and Beyond*, December 2009.

[9] J. Weng and B. S. Lee, "Event detection in twitter," in *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM '11)*, Barcelona, Spain, July 2011.

[10] Q. Diao, J. Jiang, F. Zhu, and P.-E. Lim, "Finding bursty topics from microblogs," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers (ACL'12)*, pp. 536–544, Jeju, South Korea, 2012.

[11] Z. J. Gao, Y. Song, S. Liu et al., "Tracking and connecting topics via incremental hierarchical Dirichlet processes," in *Proceedings of the 11th IEEE International Conference on Data Mining (ICDM '11)*, pp. 1056–1061, Vancouver, Canada, December 2011.

[12] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors," in *Proceedings of the 19th International World Wide Web Conference (WWW '10)*, pp. 851–860, New York, NY, USA, April 2010.

[13] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. 4-5, pp. 993–1022, 2003.

[14] X.-H. Phan, C.-T. Nguyen, D.-T. Le, L.-M. Nguyen, S. Horiguchi, and Q.-T. Ha, "A hidden topic-based framework toward building applications with short web documents," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 7, pp. 961–976, 2011.

[15] D. T. Vo and C. Y. Ock, "Extraction of semantic relation based on feature vector from Wikipedia," in *PRICAI 2012: Trends in Artificial Intelligence: Proceedings of the 12th Pacific Rim International Conference on Artificial Intelligence, Kuching, Malaysia, September 3–7, 2012*, vol. 7458 of *Lecture Notes in Computer Science*, pp. 814–819, Springer, Berlin, Germany, 2012.

[16] R. Speer and C. Havasi, "Conceptnet 5: a large semantic network for relational knowledge," in *The People's Web Meets NLP*, Theory and Applications of Natural Language Processing, pp. 161–176, Springer, Berlin, Germany, 2013.

[17] A. Ritter, Mausam, and O. Etzioni, "A latent dirichlet allocation method for selectional preferences," in *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL '10)*, pp. 424–434, July 2010.

[18] M. Cheong and V. Lee, "Integrating web-based intelligence retrieval and decision-making from the twitter trends knowledge base," in *Proceedings of the 2nd ACM Workshop on Social Web Search and Mining (SWSM '09), Co-located with the 18th ACM International Conference on Information and Knowledge Management (CIKM '09)*, pp. 1–8, November 2009.

[19] N. Glance, M. Hurst, and T. Tomokiyo, "Blogpulse: automated trend discovery for weblogs," in *Proceedings of the WWW Workshop on the Weblogging Ecosystem: Aggregation, Analysis, and Dynamics*, 2004.

[20] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, "Information diffusion through blogspace," in *Proceedings of the 13th International Conference on the World Wide Web (WWW '04)*, pp. 491–501, 2004.

[21] J. Allan, R. Papka, and V. Lavrenko, "On-line new event detection and tracking," in *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '98)*, pp. 37–45, 1998.

[22] R. Nallapati, A. Feng, F. Peng, and J. Allan, "Event threading within news topics," in *Proceedings of the 13th ACM Conference on Information and Knowledge Management (CIKM '04)*, pp. 446–453, November 2004.

[23] B. Shaparenko, R. Caruana, J. Gehrke, and T. Joachims, "Identifying temporal patterns and key players in document collections," in *Proceedings of the IEEE ICDM Workshop on Temporal Data Mining: Algorithms, Theory, and Applications (TDM '05)*, pp. 165–174, 2005.

[24] W. H. Dai, H. Z. Hu, T. Wu, and Y. H. Dai, "Information spread of emergency events: path searching on social networks," *The Scientific World Journal*, vol. 2014, Article ID 179620, 7 pages, 2014.

[25] W. H. Dai, X. Q. Wan, and X. Y. Liu, "Emergency event: internet spread, psychological impacts and emergency management," *Journal of Computers*, vol. 6, no. 8, pp. 1748–1755, 2011.

[26] H. Z. Hu, D. Wang, W. H. Dai, and L. H. Huang, "Psychology and behavior mechanism of micro-blog information spreading," *African Journal of Business Management*, vol. 6, no. 35, pp. 9797–9807, 2012.

[27] X. H. Hu, T. Mu, W. H. Dai, H. Z. Hu, and G. H. Dai, "Analysis of browsing behaviors with ant colony clustering algorithm," *Journal of Computers*, vol. 7, no. 12, pp. 3096–3102, 2012.

[28] S. Harabagiu and A. Hickl, "Relevance modeling for microblog summarization," in *Proceedings of the 5th International Conference Webblogs and Social Media (ICWSM '11)*, pp. 514–517, Barcelona, Spain, July 2011.

[29] D. Inouye and J. K. Kalita, "Comparing twitter summarization algorithms for multiple post summaries," in *Proceedings of the IEEE 3rd International Conference on Privacy, Security, Risk and Trust (PASSAT) and IEEE 3rd Inernational Conference on Social Computing (SocialCom '11)*, pp. 298–306, Boston, Mass, USA, October 2011.

[30] B. Sharifi, M. Hutton, and J. Kalita, "Automatic summarization of twitter topics," in *Proceedings of the National Workshop Design and Analysis of Algorithm*, pp. 121–128, Assam, India, 2010.

[31] H. Takamura, H. Yokono, and M. Okumura, "Summarizing a document stream," in *Advances in Information Retrieval*, vol. 6611 of *Lecture Notes in Computer Science*, pp. 177–188, Springer, New York, NY, USA, 2011.

[32] H. Becker, M. Naaman, and L. Gravano, "Beyond trending topics: real-world event identification on twitter," in *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM '11)*, 2011.

[33] A. M. Popescu, M. Pennacchiotti, and D. A. Paranjpe, "Extracting events and event descriptions from Twitter," in *Proceedings of the 20th International Conference Companion on World Wide Web (WWW '11)*, pp. 105–106, April 2011.

[34] E. Erosheva, S. Fienberg, and J. Lafferty, "Mixed-membership models of scientific publications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 1, pp. 5220–5227, 2004.

[35] M. Banko and O. Etzioni, "The tradeoffs between open and traditional relation extraction," in *Proceedings of the ACL '08: HLT*, 2008.

[36] S. Petrović, M. Osborne, and V. Lavrenko, "The Edinburgh twitter corpus," in *Proceedings of the NAACL HLT 2010 Workshop on*

Computational Linguistics in a World of Social Media (WSA '10), June 2010.

[37] T. Joachims, *Learning to classify text using support vector machines [Ph.D. dissertation]*, Kluwer Academic Publishers, 2002.

[38] D. T. Vo and C. Y. Ock, "Learning to classify short text from scientific documents using topic models with various types of knowledge," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1684–1698, 2015.

[39] W. Zhang, T. Yoshida, and X. Tang, "Text classification based on multi-word with support vector machine," *Knowledge-Based Systems*, vol. 21, no. 8, pp. 879–886, 2008.

[40] K. W. Church and P. Hanks, "Word association norms, mutual information, and lexicography," *Computational linguistics*, vol. 16, no. 1, pp. 22–29, 1990.

[41] B. Gerloff, "Normalized (pointwise) mutual information in collocation extraction," in *Proceedings of the Biennial GSCL Conference*, 2009.

[42] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, vol. 41, no. 6, pp. 391–407, 1990.

[43] T. K. Landauer, P. W. Foltz, and D. Laham, "An introduction to latent semantic analysis," *Discourse Processes*, vol. 25, no. 2-3, pp. 259–284, 1998.

Hindawi Publishing Corporation
Computational Intelligence and Neuroscience
Volume 2015, Article ID 123028, 10 pages
http://dx.doi.org/10.1155/2015/123028

Research Article

Design of Automatic Extraction Algorithm of Knowledge Points for MOOCs

Haijian Chen,[1,2] Dongmei Han,[1,3] Yonghui Dai,[1] and Lina Zhao[1,4]

[1] *School of Information Management and Engineering, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, China*
[2] *School of Open Education, Shanghai Open University, 288 GuoShun Road, Shanghai 200433, China*
[3] *Shanghai Financial Information Technology Key Research Laboratory, 777 Guoding Road, Shanghai 200433, China*
[4] *School of Information Management, Shanghai Finance University, 995 Shangchuan Road, Shanghai 200433, China*

Correspondence should be addressed to Yonghui Dai; dyh822@163.com

Received 7 August 2014; Accepted 27 September 2014

Academic Editor: Weihui Dai

In recent years, Massive Open Online Courses (MOOCs) are very popular among college students and have a powerful impact on academic institutions. In the MOOCs environment, knowledge discovery and knowledge sharing are very important, which currently are often achieved by ontology techniques. In building ontology, automatic extraction technology is crucial. Because the general methods of text mining algorithm do not have obvious effect on online course, we designed automatic extracting course knowledge points (AECKP) algorithm for online course. It includes document classification, Chinese word segmentation, and POS tagging for each document. Vector Space Model (VSM) is used to calculate similarity and design the weight to optimize the TF-IDF algorithm output values, and the higher scores will be selected as knowledge points. Course documents of "C programming language" are selected for the experiment in this study. The results show that the proposed approach can achieve satisfactory accuracy rate and recall rate.

1. Introduction

Massive Open Online Courses (MOOCs) have played a great role in the process of construction of learning society [1]. With a rapid development of more than ten years of online learning, online learning resources have been seriously overloaded, and it is difficult for a learner to find suitable learning resources for his own learning resources [2]. Therefore, how to realize the knowledge sharing and knowledge discovery in MOOCs era has attracted the attention of experts in the field of education. The ontology technology is one of the effective ways to solve the knowledge sharing and knowledge discovery, more and more scholars apply it to MOOCs in recent years, and ontology construction has become a hot spot research. At present, most of the construction of domain ontology has to be done manually, which is using a plain document editor or ontology editing tools (such as protégé, Swoop, Ontolingua, and OntoEdit) to add one by one manually. Protégé is a very popular and useful tool [3, 4]. Obviously, this method is not only time-consuming, error prone and difficult to update, but also it needs the participation of experts in the field. The most important aspect is that the manual construction of ontology is inefficient, and it is hard to be popularized. Ontology learning usually use ontology engineering, machine learning technology, statistics and principles of many other subjects to realize the automatically or semiautomatically construction of ontology [5]. By ontology learning, concepts and classifications can be extracted from a variety of nonstructured document [6]. Automatic construction of ontology will greatly improve the development process of semantic ontology and easy to achieve knowledge discovery and knowledge sharing. It provides the possibility of course ontology reasoning and the necessary condition for personalized learning. In education domain, knowledge point is the basic elements and the foundation of the relationship between them. Hence, automatic extraction of knowledge is the key of ontology learning [7].

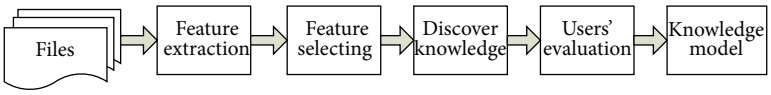


Figure 1: The general process of text mining.

Generally, there are three ways for automatic extraction of knowledge in the field of education: linguistics method, statistical method, and hybrid method [8]. There are the following several advantages of Linguistic method high accuracy, small amount of calculation, without relying on the corpus, ability to extract low frequency point of knowledge, but with poor portability, it is difficult to maintain the rules of language. Even not relying on syntactic and semantic knowledge base and ability to process incomplete sentences or phrases properly without the restriction of different language, statistical method bear the disadvantage of huge calculation and difficulty to extract multimeaning knowledge points and low frequency knowledge points. Hybrid method is combining Statistics knowledge with linguistic knowledge (syntactic and semantic information), taking the advantage of both methods [9]. Considering the particularity of online course, we use the hybrid method, using linguistic methods to process Chinese word segmentation and POS tagging, and using statistics method to handle score method for characteristics.

In order to construct the educational domain ontology automatically, automatic extraction of knowledge point is a very important job. First, it classifies the document and then makes Chinese word segmentation and POS tagging for each document, it uses vector space model (VSM) to calculate similarity and design the weight value to optimize the TF-IDF algorithm value as the score for each feature value, and then sequence these characteristics by rating sort. Finally, the higher scores are selected as knowledge points. The experiment results show that the automatic extraction for knowledge has high accuracy rate and high recalling rate, lay a solid foundation for future automatic construction of course ontology.

This paper is arranged as the following seven sections. Section 1 is the introduction of research background; Section 2 is related literature review; Section 3 expounds the methodology and technology as well as the TF-IDF algorithm, similarity calculation, and normalization method; Section 4 discusses the modeling and designing frameworks of automatic extracting course knowledge point; Section 5 illustrates the process and algorithm systematically; Section 6 is about the empirical analysis of "c programming language" course documents; and the conclusion and discussion are expressed in Section 7.

2. Related Researches

The sorting of information in text resource cannot be realized without the text mining technology. Figure 1 is the typical chart for the flow of text mining.

From Figure 1, it can be seen that the first step is to extract appropriate features from the text, which make the text into digital form that the computer can understand. According to the need for processing speed and accuracy, the features in text can be selected and optimized. Then, a variety of text mining methods will be used to discover the hidden knowledge patterns, the final output which meets the user's evaluation standard will be formed as useful knowledge to guide people's practice [10]. The essence of text mining is about text classification and feature extraction technology. The development of text classification has experienced two stages which are rule-based system and machine learning. Since 2000, the machine learning method has been widely used in text classification, when several training samples with manual annotation categories are designed, the system of machine-based learning can construct automatically text classification model, which improve the efficiency and performance of the classification. But no matter in which stage of text classification, expert's knowledge in the field plays a very important role; for example, the training samples should be labeled manually when using the classification method based on machine learning [11]. Therefore, in the design of text classification process, experts' knowledge in that field is taken as an important part of the system.

Generally, teaching document is semistructured or unstructured data; the knowledge point can be extracted automatically by using text mining. Research in other countries is mature and has proposed many fruitful methods, which is based on the study of English language. Missikoff's approach to ontology engineering uses an iterative process that involves automatic concept learning with OntoLearn [12]. Navigli et al. used it to automatically translate multiword terms from English to Italian [13]. Text mining produces a more structured analysis of textual knowledge than simple word searches and can provide powerful tools [14–16]. A personalized ontology model is proposed for knowledge representation and reasoning over user profiles [17]. As there is big difference between English language and Chinese language, there are fewer researches in the field of automatic extraction for Chinese language in China. Du et al. proposed a term extraction algorithm combining statistics-based method and rule-based method [18]. Zheng and Lu proposed a method that combined nonlinear function and "paired comparison method," considered the location and frequency of words, gave the weight of candidate word, and realized the automatic extraction of keywords [19]. Chen et al. proposed automatic acquisition of field words from a large unlabeled corpus by using Bootstrapping machine learning technology [20]. Liu proposed methods which extract automatically webontLearn in the web pages [21]. In his study, He studied the relationship between semantic concepts from the data in the web page and how to extract automatically web ontology through the analysis of the same application field of web page set.

In the concept extraction, statistical method is mainly adopted, which is also the current mainstream technology.

Rules-based approach is also applied to solve the key difficulty in field correlation of concept. By calculating the ratio between the frequency of the concept in the documents of particular field and frequency of the concept in the normal documents, correlation of the concept can be determined. That is, if the ratio is greater than the specified threshold, it means that the concept often appears in that particular field and is not often used in other fields.

3. Methodology and Technology

3.1. Concept Filters. Domain concept emerged in the field of corpus more frequently than it appeared in the General Corpus. If a concept appears in the field of corpus more frequently than it appears in the general corpus, it is considered related to the field [22, 23]. The concept of the area has the following two characteristics.

(1) The words appear in the field more frequently than in other areas.

(2) The concept in the field is commonly recognized, it is therefore widely used in the field.

These two characteristics can be measured, respectively, by the concept of Domain Relevant and Domain Consensus [24].

3.1.1. Domain Relevant. The domain relevance of a concept t in domain D_i is given as follows:

$$\mathrm{DR}(t, D_i) = \frac{p(t \mid D_i)}{\max p(t \mid D_j)}, \tag{1}$$

where DR is in [0, 1]. According to the large number theorem of probability theory that, under the premise that large sample has the same base, the sample's frequency is close to the probability value, so the maximum likelihood estimation value of the conditional probability $p(t, \mid D_i)$ is equal to the frequency of "t" appearing in the field of D_i, there is an equation that

$$p(t \mid D_i) = \frac{\mathrm{freq}(t \in D_i)}{\sum_{i=1}^{n} \mathrm{freq}(t \in D_i)}. \tag{2}$$

3.1.2. Domain Consensus. The domain consensus of a concept "t" in domain D_i is given as follows:

$$\mathrm{DC}(t, D_i) = H(p(t, d_j)) = \sum p(t, d_j) \times \log_2 \frac{1}{p(t, d_j)}, \tag{3}$$

where d_j is documents in D_i, and the probability $p(t, d_j)$ is estimated as follows:

$$\frac{\mathrm{freq}(t \in d_j)}{\sum_{d_j \in D_i} \mathrm{freq}(t \in d_j)}. \tag{4}$$

3.1.3. Concept Filters. After Qualify concept's Domain Relevant and Domain Consensus, the degree of importance for each candidate concepts "T" to domain D_i can be defined as follows:

$$\mathrm{CF}(T, D_i) = \alpha \times \mathrm{DR}(T, D_i) + \beta \times \mathrm{DC}(T, D_i). \tag{5}$$

In the above equation, $\alpha, \beta \in [0, 1]$.

3.2. TF-IDF. Term Frequency-Inverse Document Frequency is a numerical statistic that is intended to reflect how important a word is to a document in a collection. It is often used as a weighting factor in information retrieval and text mining. The importance of a word is highlighted with the increasing of the times of its appearing in a file, but the importance is decreased inversely as the frequency of its appearing in the corpus. If a word or phrase bears high frequency in an article while with very low frequency in other articles, the word or phrase is usually taken as keyword with ability for distinguishing.

3.2.1. Calculate TF. TF represents the number of a word appears in the document. Because documents have different lengths, the TF standardization is used to facilitate the comparison of different documents:

$$\mathrm{TF} = \frac{\text{The number of a word appears in the document}}{\text{The total number of words in the document}}. \tag{6}$$

3.2.2. Calculate IDF. IDF is a measure of the importance of a common word. IDF's main idea is as follows: if the document contains fewer entries, IDF becomes bigger; the entry bears the ability to distinguish between good categories.

$$\mathrm{IDF} = \log\left(\frac{\text{Total number of documents in corpus}}{\text{Total number of documents containing the term} + 1}\right). \tag{7}$$

3.2.3. Calculate TF-IDF. TF and IDF together can form TF-IDF measure:

$$\text{TF-IDF} = \mathrm{TF} \times \mathrm{IDF}. \tag{8}$$

As you can see, the value of TF-IDF is directly proportional to the frequency of a word's appearing in the file, but inversely proportional to the frequency of the word's appearing in the entire corpus.

3.3. Similarity Algorithm. Each word W is considered as a vector:

$$W = [w_1, w_2, w_3, \ldots, w_n]. \tag{9}$$

A lot of similarity algorithms have been proposed and widely applied on similarity calculation, such as cosine

similarity, Jaccard coefficient, and Pearson Correlation Coefficients. The details of different similarity measures are described as below.

(i) Cosine Similarity. Cosine similarity is a measure of similarity between two vectors, which measures the cosine of the angle between them [25]. The cosine of 0° is 1, and it is less than 1 for any other angle. Compared to the distance measure, the cosine similarity pays more attention to the differences between the two vectors in the direction, rather than the distance or length. The formula is as follows:

$$\operatorname{sim}\left(w_j, w_i\right) = \cos\theta = \frac{w_i \cdot w_j}{\left\|w_i\right\| \cdot \left\|w_j\right\|}. \tag{10}$$

(ii) Jaccard Coefficient. The Jaccard coefficient measures similarity as the intersection divided by the union of the objects. The Jaccard coefficient is mainly used for computing symbol metric or Boolean similarity between individual attributes, because the individual is symbol metric or a Boolean indicator therefore unable to measure the difference of specific value, can only get "is the same as" the results, the Jaccard coefficient is concerned only with the common features among individuals is consistent with this problem [26]. The formula is as follows:

$$\operatorname{sim}\left(w_i, w_j\right) = \frac{w_i \cap w_j}{w_i \cup w_j}. \tag{11}$$

The Jaccard Coefficient ranges between $[0, 1]$. The Cosine Similarity may be extended to yield Jaccard Coefficient in case of Binary attributes.

(iii) Pearson Correlation Coefficients. In statistics, Pearson correlation coefficient is used to measure the relationship between the two variables X and Y (linear), in the range $[-1, +1]$. Pearson correlation coefficient is widely used in academic research to measure the two variable linear correlations [27]. The formula is as follows:

$$\operatorname{sim}\left(w_i, w_j\right) = \frac{\operatorname{Cov}\left(w_i, w_j\right)}{\sqrt{\operatorname{Var}\left(w_i\right) \cdot \operatorname{Var}\left(w_j\right)}}. \tag{12}$$

$\operatorname{Cov}(w_i, w_j)$ represent the covariance of w_i and w_j, $\operatorname{Var}(W_i)$ represent the variance of w_i, and $\operatorname{Var}(W_j)$ represent the variance of w_j.

3.4. Normalization Method. Normalization method is a basic task of data mining; different evaluation index often have different dimension and dimensional units; this situation will affect the results of data analysis. In order to eliminate the dimensional effects between the indexes, normalization method is frequently used. After data standardization processing which is each index of the original data at the same level, suitable for evaluation of comprehensive comparison. The data is mapped to $[0, 1]$ interval method for data normalization includes: Min-Max normalization, log function, atan function, and zero-mean normalization. We use Min-Max normalization in this paper; the formula is as follows:

$$x^* = \frac{x - \min}{\max - \min}. \tag{13}$$

4. Frameworks and Processes

There is great difference between the extraction for course knowledge point and the extraction for general feature in common document. The extraction for general feature is to study and analyze mass documents and find out the feature value which can represent a field, commonly used in document classification, document clustering, information extraction, relation analysis, and so on. The following are the methods for feature extraction (evaluation): document frequency (referred to as DF), information gain (referred to as IG), mutual information (referred to as MI), expected cross entropy, the weight of evidence for document, odds ratio, and so on. The experimental results show that DF and IG result well [28]. There are a lot of researches on the feature selection. Yang et al. and Feng et al. pointed out that extraction of curriculum knowledge is to extract knowledge automatically from the curriculum teaching files, teaching content, database, and other documents by using Chinese segmentation and text mining techniques, that is to structure or semantic the unstructured documents for the follow-up research work of knowledge sharing and knowledge discovery [29, 30]. Because it is in a specific environment and there is a strong correlation between document and knowledge points in the online course, so using VSM model will greatly reduce the feature dimension. At the same time, by increasing the "knowledge-Document" matrix design weight algorithm and optimizing the document frequency method, improve the extraction effect for course knowledge point. Framework of automatic extraction for course knowledge points as shown in Figure 2.

The whole process consists of seven steps, as follows.

4.1. Documents Preprocessing. Curriculum resource of online course is rich; the content and style of the course are varied, they generally include teaching files, teaching content, exercises, case base, question library, video library and so on. The first step is to classify documents and taking the following three types of documents, which are very important in almost every course: the teaching files, teaching contents, and exercises. The teaching file is a programmatic document which has large and comprehensive contents; teaching contents include detail contents of each chapter; exercises is to measure teaching quality of this course. The above three documents contain all the knowledge points of a course. Secondly, considering the diversification of the types of the document which shows in PDF, HTML, XML, Excel, and other different formats, this document needs to be unified into a plain document file format (∗.txt) [31].

4.2. Chinese Word Segmentation and POS Tagging. Chinese language is read sentence by sentence, which is different from English word, so we need to perform segmentation on Chinese document. Chinese word segmentation is

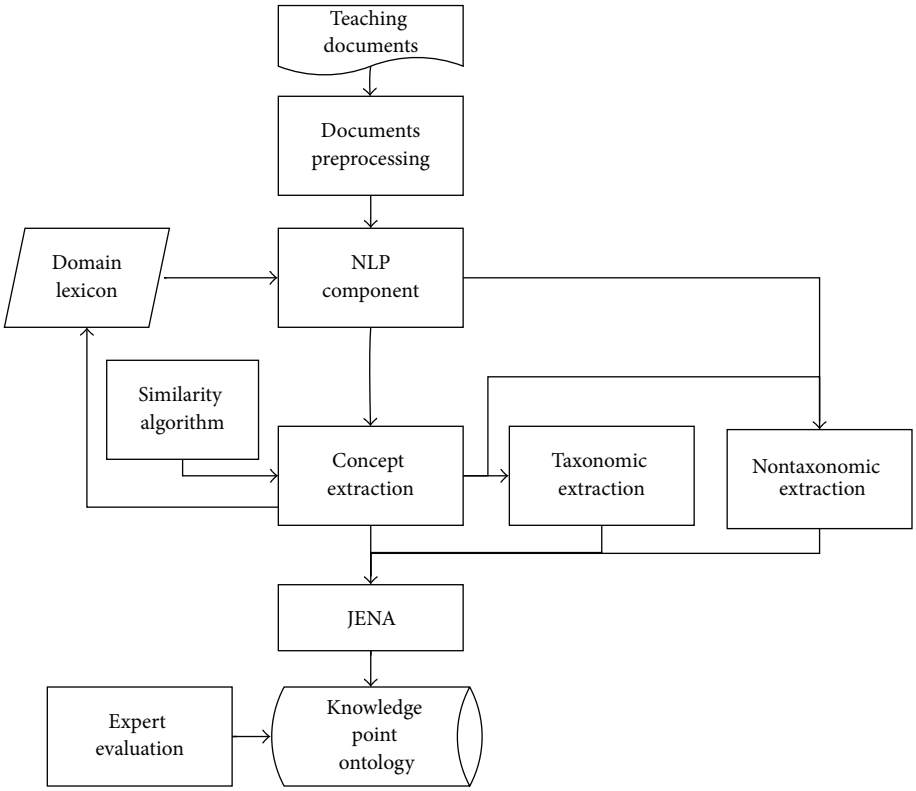


Figure 2: The frameworks of automatic knowledge points' extraction.

the process of dividing written text into meaningful units, such as Chinese words, Chinese sentences, or Chinese topics. Software ICTCLAS is used to divide sentences into words and tag words in this paper. Because dividing sentence into words belongs to the category of linguistics, different factors will lead to different results [32]. For example, "the foundation of program design" in Chinese idiom can be divided into "program," "design," and "the foundation of" or be divided into "program design" and "the foundation of" or be divided by other ways. Therefore, the dictionary should be referred to when the sentences were divided into words; a number of keywords in a field and corresponding frequency should be added into the dictionary. Considering the background of this study, the dictionary in education field, dictionary in computer science field, and dictionary in curriculum field should be composed.

4.3. Candidate Knowledge Point. To process the segmentation results, VSM model was used to calculate the characteristics of TF-IDF algorithm using the TF-IDF value, then candidate course knowledge points were obtained by sequencing. Because most of the knowledge points are names and verbs (a lot of knowledge is a verb, e.g., "cycle" is a very important knowledge, but in Chinese it refers to a verb), so to reduce the number of useless adjectives and adverbs, articles can greatly reduce the dimension and improve the time complexity degree for VSM model. Then, calculate their frequency and inverse-document frequency for each feature. Because the relations between knowledge points will be extracted, the property of each candidate course knowledge points should be contained, including the location of the document, the document size in bytes, the position of the paragraph, the sentence position and other candidate knowledge in the same sentence.

4.4. Similarity Calculation. Because there are couples of expression for a same knowledge point; for example, the "branch structure" in "C language program design" can also be called "conditional structure," "single branch," or "multi branch." So the similarity-value of knowledge points needs to be calculated. The knowledge points bearing similar similarity-value can be merged.

4.5. Weight Calculation and Normalization. Use "knowledge-document" matrix to calculate the weight of candidate knowledge points. Because all the documents are from the online course, there is strong relationship between knowledge and document. Considering the special nature of teaching content document and exercises for each chapter, "knowledge-document" matrix can be built to calculate the weight of each knowledge point, and then the weights are normalized.

4.6. Extraction for Knowledge Point. The frequency and correlation of candidate knowledge points are used to analyze weight and knowledge entropy weight and recalculate the frequency of candidate knowledge points. Then, course knowledge points are selected according to the sequence of the above calculating results.

4.7. Expert Evaluation. Experts determine knowledge point according to the characteristic of the curriculum field then compare to them by the knowledge points extracted automatically and analyze the reasons for the difference.

5. Algorithm Design

It is considered that online courses have distinctive feature; Automatically Extract Course Knowledge Points (AECKP) are designed in this paper to extract a certain course knowledge points automatically which includes the TF-IDF, similarity, weight algorithm, and the improved TF-IDF algorithms.

5.1. TF-IDF Calculation. The key point of TF-IDF (term frequency-inverse document frequency) is that if a knowledge point has high frequency in particular documents while seldom appears in other types of documents, this kind of knowledge point bears high capacity to distinguish category, thus has high degree of importance [33].

TF (Term Frequency) refers to the frequency a word appears in a document. Equation (14) means the frequency of kp (a knowledge point) in document d; $\mathrm{kp}_{\mathrm{all}}$ means all the candidate knowledge points:

$$\mathrm{tf}(\mathrm{kp}, d) = \frac{\mathrm{count}(\mathrm{kp}, d)}{\mathrm{count}(\mathrm{kp}_{\mathrm{all}}, d)}. \tag{14}$$

The main point of IDF Inverse Document includes the less the document which contains the knowledge point and the higher the IDF, which means the knowledge point is very important. Equation (15) represents the frequency of IDF in the whole documents collection, and N means the total number of documents in D_i document collection:

$$\mathrm{idf}(\mathrm{kp}) = \log\left(\frac{N}{\mathrm{docs}(\mathrm{kp}, D_i) + 1}\right). \tag{15}$$

Equation (16) is about TF-IDF model; it is to calculate the value of TF-IDF for each knowledge point according to tf and idf. D_{ij} means the document sequenced by j in D_i document collection, and N means the total numbers of documents in D_i document collection:

$$\text{tf-idf}(\mathrm{kp}, D_i) = \sum_{j=1}^{N} \mathrm{tf}(\mathrm{kp}, D_{ij}) * \mathrm{idf}(\mathrm{kp}). \tag{16}$$

While judging the importance of the documents, TF-IDF considers not only the frequency of a knowledge point in a document (word frequency) but also the IDF of the knowledge point in all kinds of documents.

5.2. Similarity Calculation. Extract the feature vector of two candidate knowledge points in any domain concept, respectively, and then calculates the semantic similarity between them using the cosine method. The equation can be as shown in

$$\cos(\mathrm{KP}_i, \mathrm{KP}_j) = \frac{\sum_{i=1}^{k} X_i Y_i}{\sqrt{\sum_{i=1}^{k} X_i}\sqrt{\sum_{i=1}^{k} Y_i}}. \tag{17}$$

Table 1: "Knowledge point-teaching content" matrix.

Knowledge point	Teaching content 1	Teaching content 2	...	Teaching content m1
Constant	2	1	···	0
Variable	8	3	···	1
Integer	3	2	···	1
Float	1	1	···	0
Array	0	0	···	0
Function	0	0	···	0
Style	0	6		0
⋮	⋮	⋮	⋮	⋮

In (17), KP_i and KP_j represent two knowledge points, X_i and Y_i represent the feature vector, and K represents the number of feature vector.

5.3. Weight Calculation and Normalization. The calculation of Document TF-IDF is for mass text mining; for this particular environment of online course, the effect is not ideal. This paper adopts "knowledge point-document" matrix to calculate the weight value of each knowledge point. According to the above classification, "knowledge point-teaching file," "knowledge point-teaching content," and "knowledge point-exercises" matrix were established. "Knowledge-teaching content" matrix is shown in Table 1.

Consider

$$\begin{aligned} a_1 &= \frac{\max_{i=1\cdots M_1}(\mathrm{count}(\mathrm{kp}, d_i))}{\mathrm{count}(\mathrm{kp}, D_1)}, \\ a_2 &= \frac{\max_{i=1\cdots M_2}(\mathrm{count}(\mathrm{kp}, d_i))}{\mathrm{count}(\mathrm{kp}, D_2)}, \\ a_3 &= \frac{\max_{i=1\cdots M_3}(\mathrm{count}(\mathrm{kp}, d_i))}{\mathrm{count}(\mathrm{kp}, D_3)}. \end{aligned} \tag{18}$$

In (18), a_1 represents the weights of knowledge point in the teaching file, D_1 represents the teaching file collection, M_1 represents the total number of teaching file collection, a_2 represents the weight of knowledge point in teaching contents collection, D_2 represents the teaching content collection, M_2 represents the total number of teaching content collection, a_3 represents the weight of knowledge point in exercise Library, D_3 represents the exercises in the document collection, and M_3 represents the total number of exercises in the document.

Min-Max normalization method is used to normalize the weight as shown in

$$a_i^* = \frac{a_i - \min(a_i)}{\max(a_i) - \min(a_i)}. \tag{19}$$

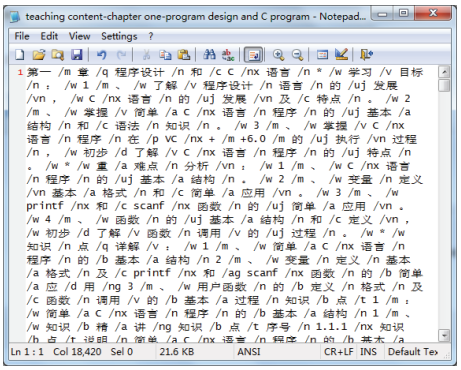


Figure 3: Result of word segmentation and POS tagging.

5.4. Improved TF-IDF Algorithm. In this paper, the TF-IDF algorithm is added with weight to form the improved TF-IDF, named I-TF-IDF:

$$\text{I-TF-IDF}(\text{kp}) = \sum_{i=1}^{3} \alpha_i^* \times \text{TF}(\text{kp}, D_i) * \text{IDF}(\text{kp}). \quad (20)$$

In (20), D_i represents the document collection numbered i, TF represents KP's frequency in D_i, IDF represents KP's inverse document frequency in D_i, and α_i^* indicates the normalization weight of document numbered i.

In this paper, the weighted word frequency values were calculated by I-TF-IDF algorithm, normalization, and sequencing. We choose 80 as the threshold value in 1st level of knowledge points and 200 as the threshold value in 2nd level of knowledge points; the knowledge point whose calculating results is greater than the threshold is taken as course knowledge point being extracted automatically.

6. Experiment

This experiment adopts C# language and SQL2005 to write program and uses SharpICTCLAS to make word segmentation and POS tagging. SharpICTCLAS is word segmentation system, which is provided by China Academy of Sciences.

In this paper, "C programming language" was selected as the experiment, the 68 study documents about "c language" were downloaded in the MOOCs platforms from 8 colleges and universities. The results of word segmentation and POS tagging about "c programming language" document as shown in Figure 3.

Course knowledge points were extracted automatically by using the AECKP algorithm; the precision rate, the recall rate, and F_{measures} were analyzed and were compared with the knowledge point marked by experts [34]:

$$\text{precision} = \frac{\text{correct}}{\text{ExpertsMark}} \times 100\%,$$

$$\text{recall} = \frac{\text{correct}}{\text{all}} \times 100\%, \quad (21)$$

$$F_{\text{measures}} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precison} + \text{recall}}.$$

Table 2: The results of accuracy rate of two level knowledge points.

Parameters	The 1st level of knowledge points	The 2nd level of knowledge points
The expert annotation number of knowledge points	66	258
Extract expert annotation number of knowledge points	48	193
Accuracy rate	72.7%	74.8%

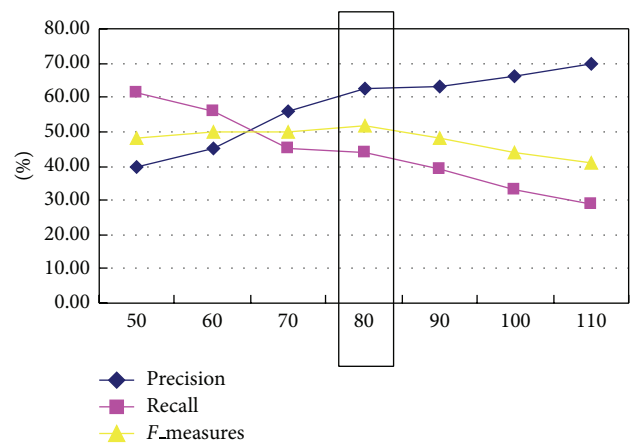


Figure 4: The best different threshold value in 1st level of knowledge points.

In (21), "correct" represents the number of correct knowledge points being extracted automatically, "all" represents the whole number of all knowledge points extracted automatically, "ExpertsMark" means the number of the knowledge point marked by experts.

Curriculum experts make hierarchical annotation for the knowledge in "C language program design," divided the knowledge points into two levels. There are 66 knowledge points for the 1st level; there are 258 knowledge points for the 2nd level. There are 1953 candidate knowledge points extracted through AECKP algorithm, including 48 knowledge points in the first level and 193 knowledge points in the second level. The results of accuracy rate of two level knowledge points are shown in Table 2.

From Table 2, we can see that there is no close relation between the number of the knowledge points extracted by experts and the accuracy of knowledge points' extraction.

In our experience, we choose different threshold in 1st level and 2nd of knowledge points, and the best different threshold values of them as shown in Figures 4 and 5.

From Figure 4, we can find 80 is the best threshold value. From Figure 5, it can be seen that 250 is the best threshold value. Then, the course knowledge points are greater than 80 in 1st knowledge points and greater than 250 in 2nd knowledge points are selected as the candidate knowledge points.

The study results are shown in Table 3.

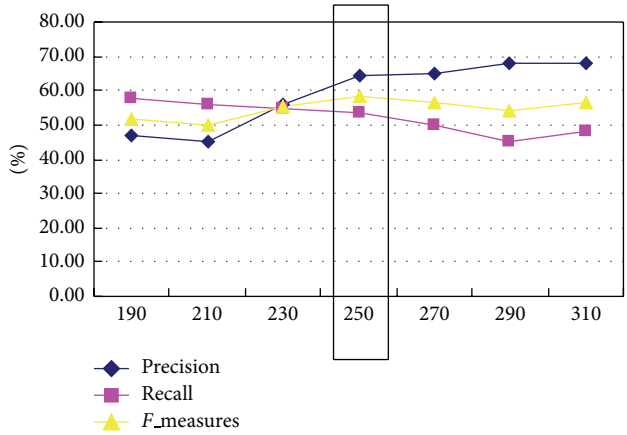


Figure 5: The best different threshold value in 2nd level of knowledge points.

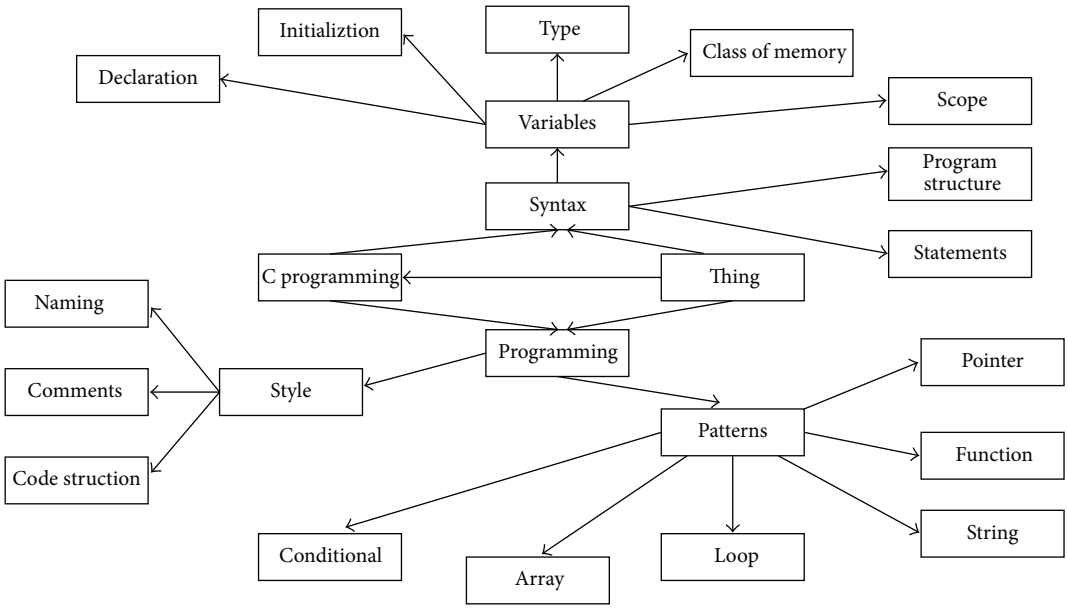


Figure 6: The partial educational ontology of C programming.

Table 3: The results contrast.

Index	The 1st level of knowledge points		The 2nd level of knowledge points	
	TF-IDF	AECKP	TF-IDF	AECKP
ExpertsMark	66	66	258	258
All	80	80	250	250
Correct	31	48	121	193
Precision	47.0%	72.7%	46.9%	74.8%
Recall	38.8%	60.0%	48.4%	77.2%
F_{measures}	42.5%	65.7%	47.6%	76.0%

From Table 3, it can be seen that once increase the number of expert annotation knowledge points, precision, recall and F test value will increase obviously. The main reason is that the number of candidate points did not change while the expert annotation knowledge increased in number, so the possibility of being relatively selected will increase. In addition, it can be seen from Table 3 that compared with TF-IDF algorithm, the accuracy and recall rate of AECKP algorithm on the course knowledge point extraction are improved to a certain extent, at the same time the extraction of low efficiency knowledge points is also improved.

In our studies, we use the AECKP algorithm to extract the C language curriculum knowledge points and then use Jena to generate ontology automatically, the partial educational ontology of C Programming as shown in Figure 6.

7. Discussion

The necessity of automatic extraction for course knowledge points in ontology learning is analyzed, and the weakness of

characteristics extraction algorithm which is usually used to extract common documents in online course is summarized in this paper.

Automatic ontology construction includes extracting ontological elements from input and building ontology from them [35]. It aims at building ontology from a given text corpus semiautomatically or automatically with a limited human exert. We usually define automatic ontology construction as a set of methods and techniques which are used to build ontology from scratch and use several sources in a semiautomatic fashion to enrich or to adapt to an existing ontology [36]. Automatic ontology construction uses methods from a diverse spectrum of fields, the field is varied from machine learning, knowledge acquisition, natural-language processing, information retrieval, artificial intelligence, and reasoning to database management [37, 38].

In addition, with the characteristics of education field considered, AECKP algorithm is proposed with details including algorithm frame, process, and algorithm design, and its performance is tested by experiment of which the results show high accuracy and recall rates. Due to the fact that the selected course "C language program design" contains both English and Chinese knowledge points, while the word segmentation module can only process Chinese words, therefore, English knowledge points are ignored during the statistical process.

Automatic extraction for course knowledge point is only a part of the course ontology learning. In future study, the relationship among knowledge points, including sequence relation and inclusion relation will be focused, extraction of relations among knowledge points automatically from the teaching document for automatic construction of course knowledge ontology will be studied to implement the ontology learning in a better way. Furthermore, the learners' interest as well as their possible emotional reactions may be considered as one of the features associated with the course knowledge points through the intelligent behavioral data-mining [39], speaker's recognition and affective computing on the vocal signals from learners' historical online study [40, 41].

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (no. 41174007) and Graduate Innovation Fund Program (no. CXJJ-2013-440 and no. CXJJ-2013-445) of Shanghai University of Finance and Economics, China.

References

[1] M. G. Moore, "Independent Learning, MOOCs, and the Open Badges Infrastructure," *The American Journal of Distance Education*, vol. 27, no. 2, pp. 75–76, 2013.

[2] M. Kim and S. Y. Choi, "An ontology-based adaptive learning system to enhance self-directed learning," in *Knowledge Management and Acquisition for Smart Systems and Services*, vol. 6232 of *Lecture Notes in Computer Science*, pp. 91–102, Springer, Berlin, Germany, 2010.

[3] S. K. Malik, N. Prakash, and A. M. Rizvi, "Developing a n university ontology in education domain using protégé for semantic web," *International Journal of Science and Technology*, vol. 2, no. 9, pp. 4673–4681, 2010.

[4] V. Jain and M. Singh, "Ontology development and query retrieval using protégé tool," *International Journal of Intelligent Systems and Applications*, vol. 9, pp. 67–75, 2013.

[5] X. L. Jia, *Document-Oriented Ontology Learning Concept Extraction and Relationship Extraction*, Central South University, 2007.

[6] C. Hu, C. Ouyang, J. Wu, X. Zhang, and C. Zhao, "NON-structured materials science data sharing based on semantic annotation," *Data Science Journal*, vol. 8, pp. 52–61, 2009.

[7] F. Hu, Z. Shao, and T. Ruan, "Self-supervised chinese ontology learning from online encyclopedias," *The Scientific World Journal*, vol. 2014, Article ID 848631, 13 pages, 2014.

[8] B. G. Liu, *A Study on Web-Based Domain Independent Ontology Learning*, Zhejiang University, 2007.

[9] H.-C. Keh, D.-A. Chiang, C.-C. Hsu, and H.-H. Huang, "The chinese text categorization system with category priorities," *Journal of Software*, vol. 5, no. 10, pp. 1137–1143, 2010.

[10] Y. H. Yan and L. G. Xiao, "EM algorithm's application in text mining," *Journal of Huzhou Vocational and Technical College*, vol. 3, pp. 13–15, 2008.

[11] B. Wang, *Research on the Characteristics of Technology in Text Classification*, University of National Defense Science and Technology, 2009.

[12] M. Missikoff, R. Navigli, and P. Velardi, "Integrated approach to web ontology learning and engineering," *Computer*, vol. 35, no. 11, pp. 60–63, 2002.

[13] R. Navigli, P. Velardi, and A. Gangemi, "Ontology learning and its application to automated terminology translation," *IEEE Intelligent Systems*, vol. 18, no. 1, pp. 22–31, 2003.

[14] S. Ananiadou, D. B. Kell, and J.-I. Tsujii, "Text mining and its potential applications in systems biology," *Trends in Biotechnology*, vol. 24, no. 12, pp. 571–579, 2006.

[15] R. Gopal, J. R. Marsden, and J. Vanthienen, "Information mining—reflections on recent advancements and the road ahead in data, text, and media mining," *Decision Support Systems*, vol. 51, no. 4, pp. 727–731, 2011.

[16] D. A. Thomas, "Searching for significance in unstructured data: text mining with leximancer," *European Educational Research Journal*, vol. 13, no. 2, pp. 235–256, 2014.

[17] X. Tao, Y. Li, and N. Zhong, "A personalized ontology model for web information gathering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 4, pp. 496–511, 2011.

[18] B. Du, H. Tian, L. Wang, and R. Lu, "Design of domain-specific term extractor based on multi-strategy," *Computer Engineering*, vol. 31, no. 14, pp. 159–160, 2005.

[19] J. Zheng and J. Lu, "Study of an improved keywords distillation method," *Computer Engineering*, vol. 31, no. 18, pp. 194–196, 2005.

[20] W. L. Chen, J. B. Zhu, T. S. Yao, and Y. X. Zhang, *Automatic Learning Field Wordsby Bootstrapping*, Tsinghua University Press, Beijing, China, 2003.

[21] B. G. Liu, "A study on ontology learning for the knowledge grid," *Computer Engineering and Applications*, vol. 20, pp. 1–5, 2005.

[22] R. Navigli and P. Velardi, "Learning domain ontologies from document warehouses and dedicated web sites," *Computational Linguistics*, vol. 30, no. 2, pp. 151–179, 2004.

[23] J. Kietz, A. Maedche, and R. Volz, "A method for semi-automatic ontology acquisition from a corporate intranet," in *Proceedings of the Procernational of Learning Language in Logic Workshop*, vol. 31, pp. 2–43, 2000.

[24] K. Fu, *The study of ontology learning from web pages [Ph.D. thesis]*, Wuhan University of Technology, Wuhan, China, 2007.

[25] N. Sandhya and A. Govardhan, "Analysis of similarity measures with wordnet based text document clustering," in *Proceedings of the International Conference on Information Systems Design and Intelligent Applications*, vol. 132, pp. 703–714, 2012.

[26] M. Kobayakawa, "Fast computation of similarity based on Jaccard coefficient for composition-based image retrieval," in *Advances in Multimedia Information Processing—PCM 2009*, pp. 949–955, 2009.

[27] P. Ahlgren, B. Jarneving, and R. Rousseau, "Requirements for a cocitation similarity measure, with special reference to Pearson's correlation coefficient," *Journal of the American Society for Information Science and Technology*, vol. 54, no. 6, pp. 550–560, 2003.

[28] Y. Xu, J.-T. Li, B. Wang, and C.-M. Sun, "Category resolve power-based feature selection method," *Journal of Software*, vol. 19, no. 1, pp. 82–89, 2008.

[29] L. H. Yang, Q. Dai, and Z. H. Yang, "Study on document categorization techniques," *Microcomputer Information*, vol. 22, pp. 209–211, 2006.

[30] J. Feng, G. Li, X. Sun, and H.-W. Feng, "Automatic generation method of teaching-oriented knowledge point library," *Computer Engineering*, vol. 38, no. 2, pp. 201–203, 2012.

[31] K. Shaban, "A semantic approach for document clustering," *Journal of Software*, vol. 4, no. 5, pp. 391–404, 2009.

[32] H. P. Zhang, H. K. Yu, and D. Y. Xiong, "HHMM-based Chinese lexical analyzer ICTCLAS," in *Proceedings of the 2nd SIGHAN Workshop on Chinese Language Processing (SIGHAN '03)*, pp. 184–187, Sapporo, Japan, 2003.

[33] C. Y. Shi, C. J. Xu, and X. J. Yang, "Study of TFIDF algorithm," *Journal of Computer Applications*, vol. 29, pp. 167–170, 2009.

[34] L. Fu, L. Q. Huang, and C. L. Fu, "Improved document-oriented algorithm for the domain-specific concept sieving," *Computer Science*, vol. 39, pp. 253–256, 2012.

[35] M. Shamsfard and A. A. Barforoush, "The state of the art in ontology learning: a framework for comparison," *The Knowledge Engineering Review*, vol. 18, no. 4, pp. 293–316, 2003.

[36] A. Gómez-Pérez and D. Manzano, *A Survey of Ontology Learning Methods and Techniques*, Universidad Politécnica de Madrid, 2003.

[37] M. Sabou, C. Wroe, C. Goble, and G. Mishne, "Learning domain ontologies for web service descriptions: an experiment in bioinformatics," in *Proceedings of the 14th International World Wide Web Conference (WWW '05)*, pp. 190–198, 2005.

[38] W. Dai, Y. You, W. Wang, Y. Sun, and T. Li, "Search engine system based on ontology of technological resources," *Journal of Software*, vol. 6, no. 9, pp. 1729–1736, 2011.

[39] X. Hu, T. Mu, W. Dai, H. Hu, and G. Dai, "Analysis of browsing behaviors with ant colony clustering algorithm," *Journal of Computers*, vol. 7, no. 12, pp. 3096–3102, 2012.

[40] Y. Wang, X. Hu, W. Dai, J. Zhou, and T. Kuo, "Vocal emotion of humanoid robots: a study from brain mechanism," *The Scientific World Journal*, vol. 2014, Article ID 216341, 7 pages, 2014.

[41] D. D. Li, Y. C. Yang, and W. H. Dai, "Cost-sensitive learning for emotion robust speaker recognition," *The Scientific World Journal*, vol. 2014, Article ID 628516, 9 pages, 2014.