

Construction of Affective Education in Mobile Learning: The Study Based on Learner's Interest and Emotion Recognition

Haijian Chen¹, Yonghui Dai^{2,*}, Yanjie Feng², Bo Jiang³, Jun Xiao¹, and Ben You²

¹ Shanghai Engineering Research Center of Open Distance Education, Shanghai Open University, Shanghai 200433, China
{xochj, xiao}@shtvu.edu.cn

² Management School, Shanghai University of International Business and Economics, Shanghai 201620, China
{daiyonghui, fengyanjie}@suibe.edu.cn, youben022@gmail.com

³ School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China
jiangbo@sui.edu.cn

Abstract. Affective education has been the new educational pattern under modern ubiquitous learning environment. Especially in mobile learning, how to effectively construct affective education to optimize and enhance the teaching effectiveness has attracted many scholars attention. This paper presents the framework of affective education based on learner's interest and emotion recognition. Learner's voice, text and behavior log data are firstly preprocessed, then association rules analysis, SO-PMI (Semantic Orientation-Pointwise Mutual Information) and ANN-DL (Artificial Neural Network with Deep Learning) methods are used to learner's interest mining and emotion recognition. The experimental results show that these methods can effectively recognize the emotion of learners in mobile learning and satisfy the requirements of affective education.

Keywords: Affective education, mobile learning, learner's interest, emotion recognition.

1. Introduction

In recent years, with the development of mobile communication technology and educational technology, profound changes have occurred in the way of learning. Especially, mobile learning model is quickly development. Generally speaking, mobile learning refers to learning facilitated by mobile devices such as mobile phones, tablet PCs or personal media players for distance learning [13][35]. Compared with traditional learning methods, there are two outstanding advantages of mobile learning, one is its learning flexibility, and the other is its abundant learning resources. As mobile learning is not limited to learning time and place, it can allow learner not only make full use of fragmentation time to learn but also easy to share learning content with others, which has brought great convenience to learners. Therefore, this way of learning is loved by more and more contemporary learners. However there are some problems in this way. Learners use mobile devices for learning and they are faced with a lack of emotional machine every day, which is easy to

make learners become indifferent, heartless and emotional imbalance. Therefore, how to improve the situation of the lack of emotion in mobile learning, and carry out the affective education to improve teaching effectiveness is an important research subject.

As we known, the key to the implementation of affective education is learners emotion recognition and interest acquisition. Taking into account the interaction of mobile learning is carried on by the learner's text and voice, we used text affective computing and speech emotion recognition in our study. Because the emotion recognition of mobile learners is closely related to the situational context, emotional state and so on, then it is always difficult to recognize the learners' emotion accurately. Especially, learners voice contains a large number of conversational continuous phrases, which makes the traditional speech emotion recognition method such as six discrete emotion categories cannot get good results. In order to solve the above problems, the three-dimensional PAD emotional model [1] and the neural network method [8] were used to calculate learner's emotion.

This paper is organized as follows. In section 2, related research of affective education, learner's interest and emotion recognition are introduced. In section 3, the framework of learner's emotion recognition and methodology are shown. In section 4, experiment is illustrated. Section 5 is the conclusion of this article.

2. Related works

As a hot research topic in the field of education, affective education has attracted the attention of many scholars. Overall, previous research related to construction of affective education in mobile learning can be summarized into three aspects, namely, affective education theory, interest mining and emotion recognition.

2.1. Affective education

Affective education is used as an educational concept and a part of the educational process, which is concerned with the findings, beliefs, attitudes, and emotions of students with their interpersonal relationships and social skills [29]. Since 1970s, the research of affective education has changed from the initial stage to the development stage, and related works research mainly focused on emotional education theory and affective education model, such as humanistic emotion theory [36], academic achievement emotion theory [41], scaffolding affective education theory [24] and affective education practice model [9] [33]. Among them, humanistic emotion theory emphasizes self-expression, emotion and subjectivity, and it not only pays attention to the development of cognition in teaching process, but also pays more attention to learners emotion, motivation and interest. For example, Connolly used this theory to study on the coaching process, and he thought that communication, self-concept, affect, personal values are the key emphases and strategies for humanistic coaching [4]. Academic achievement emotion theory mainly focuses on the students learning process and the emotion related to achievement [18], for example, learners anger to the homework, or learners disgust to the homework, and the results show that emotion has both positive and negative effects on learners academic achievement [22].

On the research of affective education practice model, Cheng put forward the implementation of affective education in a middle school in Chinas Guangzhou, and three

levels of affective education were described, namely, class-group level, manner-individual level and institutional-whole school level [3]. After that Ghasemaghahi et al introduced a framework for multimodal educational systems and human-computer interaction (HCI) emotion education [10].

2.2. Interest mining

Previous researches on interest mining have focused on log mining and interest modeling. For example, Stamou et al (2009) proposed to get users' preference by analyzing their clicked log (such as query word, browse pages and so on), as well as the semantic similarity from their query words and visited web pages [34]. Xu et al (2009) believed that the behavior of browsing the page contains the attention and interest of the content and it can be used to predict interest, and they pointed out that the user behavior of the relevant interests includes the residence time of browsing the learning page, web link of clicking, and click frequency of a page and so on [38]. Rao et al (2015) extracted user's interests from web log data, including the log of visit time and visit density, and they discussed the technological in data mining and its applications to personalization [28]. Maheswari et al (2015) studied on data preprocessing of web log files and how to predict user's interest, and their research is based on the mining of behavior logs [21].

On the research of learner's interest modeling, Nakatsuji et al (2012) proposed a collaborative filtering method based on time periods and classification for user's interest modeling, their used data were collected from the historical behaviors such as listen to music, users' tweets and visit restaurant, their study showed that their method can get good accuracy in the prediction of interest [25]. Sanchez et al (2013) constructed innovative consumption-modeling system to predict user's interest of TV contents, and their model was established based on Hidden Markov Model and Bayesian inference techniques, experimental results showed that their system was more reliability [30]. Li et al (2014) built users' interest model and offered personalized recommendation according to their reading preferences, their research suggested that the result are associated with long-term and short-term reading preferences [19].

2.3. Emotion recognition

The study of emotion recognition begins with the classification of emotion, and the psychological point of view thinks that the emotion is divided into basic emotion and dimensional space emotion [20]. Among, basic emotion refers to human emotions are divided into fixed categories, for example, the typical classification of six kinds of emotion, that is, anger, disgust, fear, joy, sadness and surprise[31]. Relatively speaking, dimensional space emotion theory holds that human emotion is different position in space. From their point of view, emotion can be divided from one dimension, two dimensions or three dimensions. For example, the widely used PAD three-dimensional emotional model (Pleasure-Displeasure, Arousal-Nonarousal, Dominance-Submissiveness) was divided from three dimensions [6], and it used dimension 'P' to represent someone's evaluation which is positive or negative, dimension 'A' represents the level of neural activation, and dimension 'D' represents the individual's ability to control situations and others people [11]. The three dimensions of continuous emotion classification are shown in Fig.1.

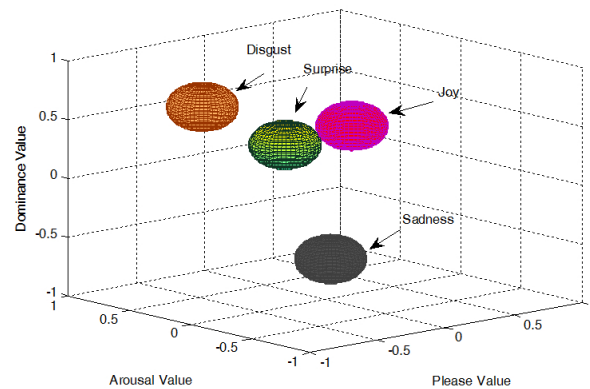


Fig. 1. Three dimensions of continuous emotion classification

Research on emotion recognition has made great progress since affective computing was proposed by Professor Picard in 1997 [27]. Overall, it is focus on text affective computing and speech emotion recognition. The realization of text affective computing is usually based on affective dictionary or machine learning. Because the voice volume, tone, sound speed and so on all contain the individual emotion, then the speech emotion recognition is relatively complex. At present, the main speech recognition methods include ANN (Artificial Neural Network) [39], HMM (Hidden Markov Model), SVM (Support Vector Machine), DBN (Deep Belief Nets) [42], GMM (Gaussian Mixture Model), DTW (Dynamic Time Warping), and mixed method [16]. For example, Schuller (2003) et al selected HMM model as their continuous speech emotion classification model, and their result showed that their method can get 86% recognition rate in recognition of seven discrete emotions [32]. At the same time, Nwe (2003) et al used discrete HMM model of vector quantization to classify six types of basic emotions, their method attained an average accuracy of 78% in the classification of six emotions [26]. Huang et al (2011) proposed a new algorithm that combined GMM and SVM to recognize speech emotion, the result showed that the average recognition rate of their method is 1.7%-3.7% higher than standard GMM method in accuracy [15]. Chavan et al (2012) used SVM to identify the anger, happiness, sadness, surprise and neutral state of the voice, and extracted MFCC as features, and their research obtained the 68% recognition rate [2].

3. Framework and Methodology

Because learning behavior is carried out through mobile devices in mobile learning, which makes the construction of affective education need to take full account of this characteristics. All the time, how to get the learners emotion timely and accurately has been a difficult problem to the implementation of affective education in mobile learning. In order to solve the above problem, the construction of affective education is built on our study, and it is based on learners interest and emotion recognition, which is shown in Fig.2.

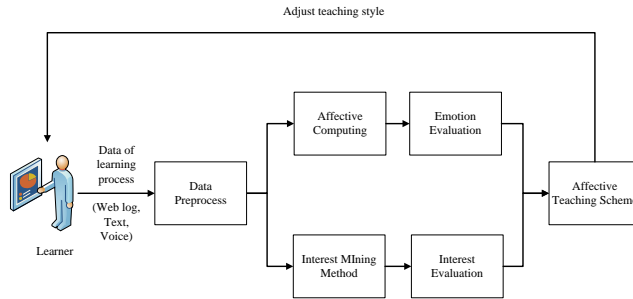


Fig. 2. Framework of affective education

It can be seen from Fig.2 that learner’s web log, text and voice data are firstly pre-processed, and then the method of interest mining and affective computing will be used to these data. Once learners interest and emotion are obtained, we can design the affective teaching scheme according to the above result and adjust the teaching style for affective education.

3.1. Interest mining for mobile learning

Generally speaking, if the learner is interested in some learning resource in mobile learning, he or she will usually carries out a series of online activities such as click resource link, add to favorites or post comment and so on. Therefore, if these data are used for mining and analysis, the learners’ behavior habits and their interest would be explored. It is well known that data mining method includes classification, clustering, regression analysis, association rules mining and so on. Among, association rules mining can discover the possible association or connection of objects from data, and it is especially good for mobile learners’ interest mining. As mobile learners’ interests are usually reflected in their learning behaviors, such as the length of their learning time, the times of being clicked of learning resources, the access order of hyperlink of learning resource. So if appropriate data mining methods are used to mine the above data, the interest of mobile learners can be obtained. In order to realize the mining of learners’ interest effectively, association rule analysis and ant colony clustering were applied in this study after referring to previous studies.

Association rules analysis The association rule is an implication of the form such as $R\{A\} \rightarrow R\{B\}$. It can be understood that if a transaction contains A , then the transaction is likely to contain B . Among them, A and B are named as the precursor and successor of association rules, AB is called association rule, which is support and confidence. The degree of support in association rules can be calculated as follows.

$$Support(A \rightarrow B) = \frac{R(A) \cap R(B)}{R_{all}} \tag{1}$$

The degree of confidence in association rules can be calculated as follows.

$$\text{Confidence}(A \rightarrow B) = \frac{R(A) \cap R(B)}{R(A)} \quad (2)$$

For example, when the value of support threshold is 0.06, a learner's preference for learning content is shown in Fig.3.

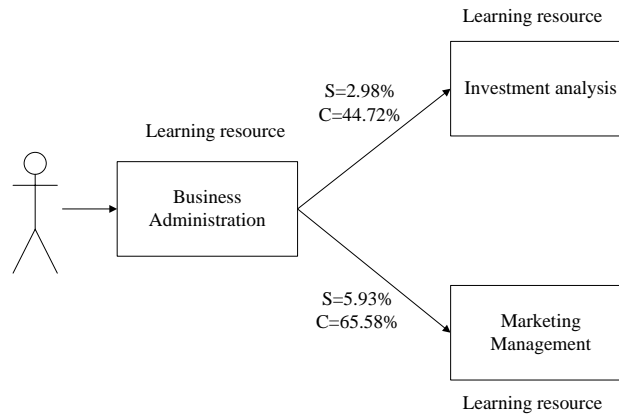


Fig. 3. Relation of learning content preference

It can be seen from Fig.3, the value of confidence of the course of Business administration and marketing management is 65.58%, which means learners are likely to have a preference for marketing management after learning the business administration.

Ant colony clustering algorithm In the course of mobile learning, learners often click on a number of learning resources, then how to effectively count and describe these resources for the purpose of interest mining is a very important task. Ant colony clustering algorithm is a clustering algorithm based on ant cleaning behavior, and the main idea is the process of transporting ants [14]. It can be assumed that the data objects to be clustered are randomly placed on a two-dimensional planar grid, and there are a number of artificial ants that allow them to move randomly in a two-dimensional plane. Each ant determines the probability of handling according to the similarity between the data object and the local environment, if the similarity is higher, the smaller probability of picking up, and the greater probability of dropping, After a certain number of iterations, the same kind of objects are clustered together in the same spatial region, and it realize the self-organizing clustering process. In this paper, an improved ant colony clustering algorithm was used, and it could be described as followings.

```

program AntCluster
  Init number of n Ants and place randomly;
  
```

```

begin
  repeat
    For all ants do
      Calculate the similarity object of each ant  $S_n$ ;
      Calculate the probability of Picking up  $P_p$ ;
      if  $P_p > \text{threshold}$ 
        Pick up object and remember current position;
        Add 1 to number of Load ant;
        Move randomly;
      else
        Dont move;
      end if
      Calculate the probability of Picking up  $P_d$ ;
      if  $P_d > \text{threshold}$ 
        Put down object and remember current position;
        Add 1 to number of unLoad ant;
        Move randomly;
      end if
    until repeat Maximum times
  end.

```

Definition 1 (Similarity) Similarity refers to the comprehensive similarity between an object and other objects in the environment. There are n objects in the data set D , and the similarity of objects is the arithmetic mean of the probability of each attribute of the object, and similarity of S_i is defined as follows.

$$f(S_i) = \frac{1}{n} \sum_{j=1}^n p_{ij} \quad (3)$$

Definition 2 (Pick up probability) The probability of picking up for ant is defined as follows.

$$P_p = 1 - \frac{1 - e^{-cf(S_i)}}{1 + e^{-cf(S_i)}} \quad (4)$$

Definition 3 (Dropping probability) The probability of dropping for ant is defined as follows.

$$P_d = 1 - \frac{1 - e^{-cf(S_i)}}{1 + e^{-cf(S_i)}} \quad (5)$$

Among them, the value of P_p and P_d belong to between 0 and 1. And the probability function of pick up and dropping is convex function, c is a constant that is used to adjust convergence speed, once the c value is different, and the function of convergence speed is different.

3.2. Text affective computing

There are usually two ways to implement the text affective computing, one is to rely on the emotion dictionary, and the other is based on machine learning. Considering that the text in mobile learning contains typical domain words, then the former way was used to our study. The process of text affective computing in our study is shown in Fig.4.

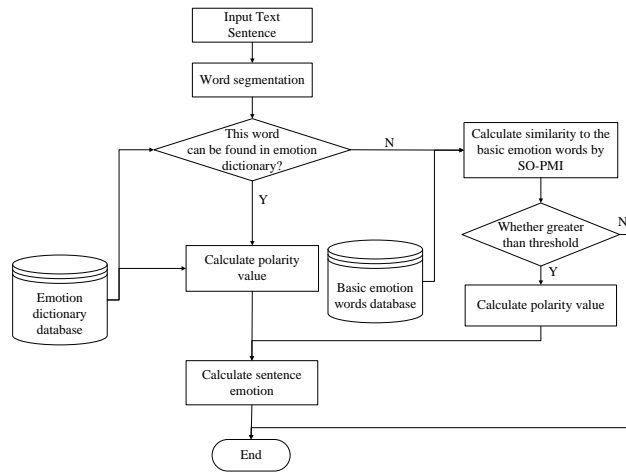


Fig. 4. Process of text affective computing

The process of text affective computing includes word segmentation, POS tagging, matching of emotional dictionary and calculation of emotional similarity and so on. For example, the statement of "Today's course is boring, I'm not interested at all", after the word segmentation and POS tagging, which is shown as follows. "Today /t 's /uj course /n is /v boring /a, I /r 'm /d not /v interested /a at all /d". Because adverbs, verbs and adjectives often contain emotions, then it was selected as candidate affective words. Therefore, 'boring /a', 'not /v', 'interested /a' and 'at all /d' were selected and were calculated according to emotional dictionary. Finally, the calculated results show that the result of this sentence reflects the negative emotions of the learners.

In addition, it can be seen from Fig.4, if the word is not in the emotion dictionary, then SO-PMI (Semantic Orientation-Pointwise Mutual Information) method will be used to calculate its similarity to the basic emotion words. The similarity of two words by SO-PMI is calculated as follows.

$$PMI(Word_1, Word_2) = \log_2 \left(\frac{P(Word_1 \& Word_2)}{P(Word_1)P(Word_2)} \right) \quad (6)$$

Among, $P(Word_1)$ is the probability of $Word_1$ appears independently in the corpus. $P(Word_2)$ refers to the probability of $Word_2$ appears independently in the corpus. And

$P(Word_1 \& Word_2)$ refers to the probability of $Word_1$ and $Word_2$ in the corpus at the same time.

3.3. Speech emotion recognition

In the process of mobile learning, some of learners interaction is carried out by their voice. If we can recognize the emotion by their voice, it will be helpful for the implementation of affective education. According to previous research, some feature of speech can represent the characteristics of human, such as volume of voice, short-term zero crossing rate, pitch frequency, formant parameters, and Mel Frequency Cepstrum Coefficient (MFCC) and so on [23]. Based on previous research and our experiments, twenty-four characteristic parameters including rhythm and tone quality were selected as speech emotion recognition, and an acoustic affective computing vector function was built based on the above acoustic characteristic parameters, namely $F(n)=[STE,SZR,PV,FF,NVB,VS,MFCC]$. Among them, STE means short-time energy (Maximum/Average/Minimum), SZR refers to short time average zero crossing rate (Maximum/Average/Minimum), PV means the value of pitch (Maximum/Average/Minimum), FF means the first formant of voice, NVB means number of voice break, VS means voice speed, and MFCC includes 12 order Mel frequency cepstrum coefficient.

Because the voice of learner in mobile learning often embodies the characteristics of fragmentation and context, traditional speech recognition methods are often difficult to get good effect. Recent studies have demonstrated that three-dimensional PAD emotion model and the ANN-DL method can get effective results [16]. Therefore, once the data of the above speech parameters are collected, the ANN-DL method will be used and the value of PAD will be calculated. For example, if a learner's pad value is [0.47 0.34 0.31], then we can get the learner's emotional state according to the PAD values for the typical emotion [6]. In order to identify the learner's emotions, some of the characteristic parameters need to be extracted, and some of the speech feature parameters are introduced as follows.

Volume of voice. It refers to the voice of the strength, and it can be regarded as the amplitude of the speech signal. When somebody is in happy or angry state, his volume of amplitude will be higher than that of calm state. On the contrary, if he is in grief and calm state, the volume will decrease the amplitude [37]. Generally, the volume of voice can be calculated by the sum of the voice signal amplitude.

Short-time energy. After the speech signal is divided into frames, the short-time energy of each frame can be calculated. The formula for calculating the short-time energy of frame $f_n(i)$ is as follows.

$$E_n = \sum_{i=0}^{N-1} f_n^2(i) \quad (7)$$

Pitch frequency. It refers to the frequency of the vocal cords. When people begin to speak, the sonant and airflow will pass through the glottis, which make the vocal cords

vibrate. At the same time it will generate excitation pulses and form the pitch frequency, which can be calculated by short-time autocorrelation function $R_n(k)$ or short-time average magnitude difference function $F_n(k)$. Among, $R_n(k)$ can be expressed as follows.

$$R_n(k) = \sum_{i=-\infty}^{+\infty} [x(i)w(n-i)][x(i+k)w(n-i-k)] \quad (8)$$

Where, k is called autocorrelation lag time, n is the N speech segment. In addition, $F_n(k)$ can be expressed as follows.

$$F_n(k) = \sum_{i=-\infty}^{+\infty} |x_n(i) - x_n(i+k)| \quad (9)$$

Mel Frequency Cepstrum Coefficient. It reflects the sensory judgments of the human ear on the short time amplitude spectrum of voice, and MFCC has been widely used in the field of speech recognition in recent years. The calculation of Mel frequency is expressed as follows.

$$f(Mel) = 2595 * \lg(1 + \frac{f}{700}) \quad (10)$$

The calculation process of MFCC coefficient includes steps as follows.

Step 1. Preprocessing of speech signal. It includes define the sampling length of each frame of the voice sequence (such as $N=256$), and pretreat each frame of speech signal $s(n)$.

Step 2. Calculation of discrete spectrum power. It gets the spectrum of each frame by the discrete FFT (Fast Fourier Transformation) and calculates the square of the value, and then gets the discrete power spectrum $s(n)$, which is the energy distribution on the spectrum.

Step 3. Power spectrum filtering. This step includes calculate $s(n)$ multiplied M triangular band-pass filter, and get M parameters P_m , where, $m=0, 1, \dots, M-1$.

Step 4. Logarithmic treatment. This step includes calculate the natural logarithm of P_m , and get L_m , where, $m=0, 1, \dots, M-1$.

Step 5. Discrete cosine transforms. This step includes calculate the discrete cosine transform of L_m , and get D_m , where, $m=0, 1, \dots, M-1$, then discard the DC component of D_0 , and take D_1, D_2, \dots, D_k as the MFCC coefficients.

Deep learning. A lot of fragmented voices bring great difficulty to emotion recognition in mobile learning. The success of deep learning algorithm in various industries has brought inspiration to our research. Overall, the concept of deep learning is based on artificial neural networks, and it is a multilayer perceptron with multiple hidden layers. It is combine low-level features to form a more abstract high-level representation of attribute

categories or features in order to discover the distributed feature representation of data. As a complex machine learning algorithm, deep learning can imitate human beings to solve realistic problems. Especially, it has developed rapidly and achieved great success in speech recognition. For example, it was reported that the accuracy of speech recognition of English in Google machine learning systems has reached 95%, and IFLYTEK Company has got more than 97% rate recognition in Chinese speech. Considering the advantages of deep learning, we used it to our speech recognition study. The structure of ANN-DL (Artificial Neural Network with Deep Learning) is shown in Fig. 5.

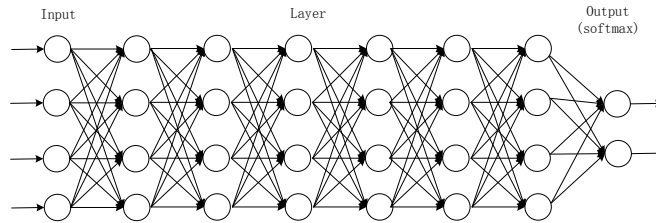


Fig. 5. The network structure of deep learning

Autoencoder is an important part of deep learning technology. As one of unsupervised learning algorithm, it uses the back propagation algorithm and tries to construct an identity function so that the output value is close to the target value. In addition, the gradient descent method is usually used to update the parameters in the Autoencoder algorithm, and the specific weights and bias of the update is shown as follows.

$$U_{ij} = U_{ij} - \alpha \frac{\partial J_{AE+wd}(\theta)}{\partial U_{ij}} \tag{11}$$

$$b_{ij} = b_{ij} - \alpha \frac{\partial J_{AE+wd}(\theta)}{\partial b_{ij}} \tag{12}$$

Where, α means learning rate.

4. Experiment

In this paper, our experimental data comes from a large online learning platform in China (<http://www.shlll.net>), which has more than 1 million 300 thousand registered learners. Thirty-two learners were randomly selected as subjects, and their text and voice in mobile learning was collected at the same time. In addition, basic speech emotional corpus is composed of CASIA (Chinese Academy of Sciences Institute of Automation) speech corpus [12] and three hundred marked historic sentences voice. Among, CASIA corpus includes happy, sad, angry, surprise, fear, and neutral six different emotional voices with the same semantic texts. On the basis of the establishment of text emotional database, Chinese Affective Dictionary of Information Retrieval Laboratory of Dalian University of Technology was selected as our study, which was created by Prof. Lin et al [40] and it

includes seven categories emotional words, that is, joy, love, anger, sorrow, fear, disgust and surprise.

4.1. Data processing

Generally speaking, learner behavior data needs to be pre-processed firstly in order to get better result, which includes processing of web log, speech signal de-noising and data transformation of interest preference.

Processing of web log. It includes data cleaning, user identification, session identification, formatting output and so on, and it is the key phase of learners interest mining, especially in the process of data cleaning and user identification. If the web log is not handled properly, it can greatly affect the efficiency or accuracy of interest mining. The processing of web log is shown in Fig. 6.

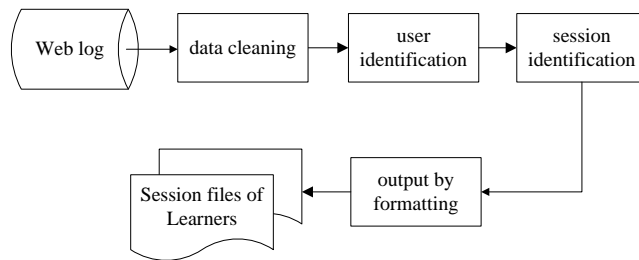


Fig. 6. Processing of web log

Speech signal de-noising. Because environmental noise is inevitably exists, the speech data need a series of pre-processing. Wavelet transform has been widely used to speech pre-processing, so it was used to our study. Firstly, the packet of db5 was selected as wavelet packet, and three level and 'Shannon' of entropy were used to the decomposing of speech signals. After above process, high frequency and low frequency coefficients were separated from initial signals. As the noise signal often exists in high frequency and it need to be discarded. Therefore, once the part of high frequency signals has more than a threshold, it will be discarded and the rest of the signals are recombined into new signals for analysis. Sample of initial signal and de-noising signal are shown in Fig. 7.

Data transformation of interest preference. On the data conversion processing of learning preferences, the threshold value will be set to deal with it. If the value is greater than the threshold value, then the content is the learner's preferences and its value is marked to

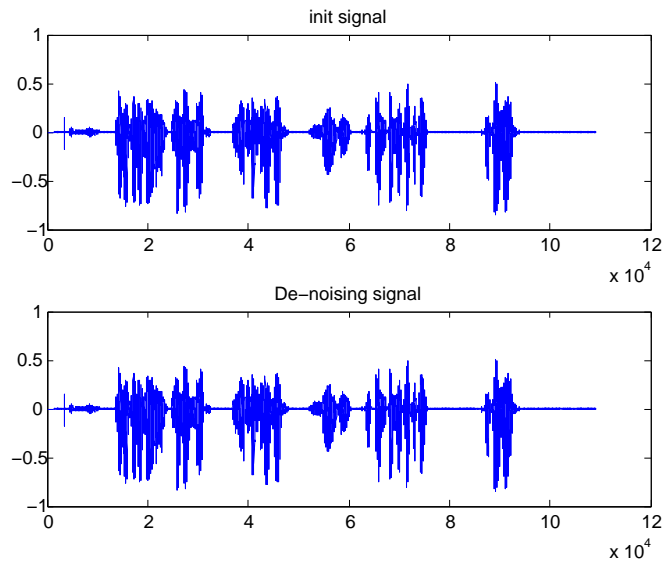


Fig. 7. Contrast of initial signal and de-noising signal

1. Otherwise, it is not the learners preference and its value is marked to 0. The expression is as follows.

$$preference = \begin{cases} 1, \text{ is preference} \\ 0, \text{ not preference} \end{cases} \quad (13)$$

After the data transformation, all data have become Boolean type and is very convenient to deal with. Taking the content of management course as an example, the preferences of learners are as follows.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \end{bmatrix} \quad (14)$$

In (14), n refers to the number of management course of learners, and the three columns of right of equation represent the learners interest preferences. For example, [1 0 1] indicates that the learner is interested in the first and third courses. However, he or she is not interested in the second course.

4.2. Experimental Results and Analysis

Learner’s interest evaluation. According to the association rule analysis and ant colony clustering method, we can get the interest of learners who participate in the experiment.

For example, the course of "management" includes five knowledge points, namely, planning, organizing, commanding, coordinating and commanding. Some learners exhibit different learning behaviors when they are learning the above knowledge points of the course. We found that a learner spend more time and much number of clicks on learning the knowledge of commanding than any other four knowledge points, especially when he was studying the knowledge of motivational theories and leadership behavior. Then we carried on data mining to this learner's learning behavior data by association rule analysis and ant colony clustering method. And some behavior rules of the learner were discovered such as he likes to browse the story of motivating employees, and he often clicked on the content of controlling chapter when he was studying motivational theories.

If a large number of learners' learning behavior rules are discovered, their interest would be easy to describe. Taking the management course as an example, we collected the learning data from thirty-two learners and described their interests according to five knowledge points of the course. And the interests of two of them are shown in Fig.8. It can be seen that the most interest of No. 1 learner is the knowledge of the commanding, and he has little interested in the knowledge of the planning. However, the most interest of No. 2 learner is the knowledge of the planning. Therefore, if we can get the learners' interest accurately, it will be a great help for the effective implementation of affective education.

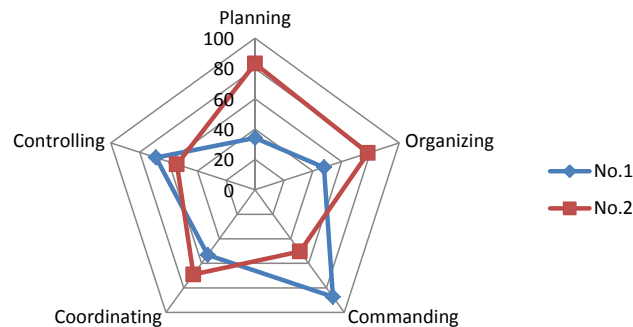


Fig. 8. The learner's interest of knowledge points

Learner's emotion evaluation. According to the result of affective computing, we can get learners' emotion who participates in the experiment. For example, on the recognition of speech emotion, the twenty-four speech feature parameters are extracted and deep neural network method is used for training and speech emotion recognition. By observed, we found learners is easy to show disgust, joy, sadness and surprise emotions in the learning process, then the above four kinds of emotion recognition rate is calculated and it is shown in Fig.9.

The recognition rates for emotion of disgust, joy, sadness and surprise are 86.32%, 91.79%, 83.72% and 90.42%, and the average recognition rate is 88.06%. Overall, this

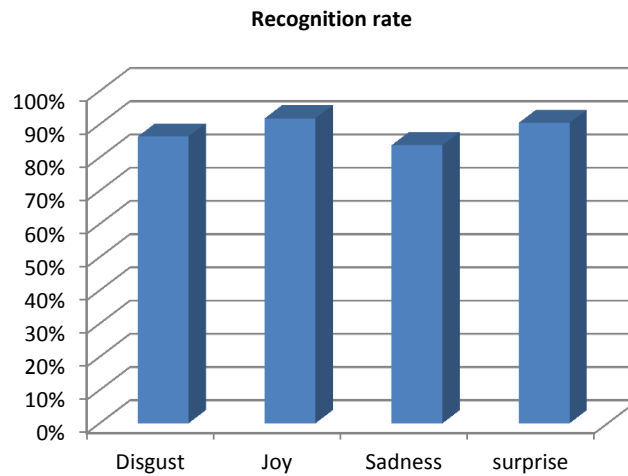


Fig. 9. The recognition rate of emotion states

method can help us to recognize the learner's emotion effectively. In addition, This proposed method has been applied satisfactorily to the affective education of foreign languages and emotional intelligence studies [5][7].

5. Conclusion

From the point of view of learning trend, mobile learning will become more and more popular. At the same time, how to implement affective education is a very meaningful research. Although emotion recognition has made great progress in recent years, however, many challenges still need to be faced and explored, such as the dynamic changes of learners' emotion, or the complex computing of mixed emotion, and so on.

This paper proposes the method of constructing the affective education based on learner's interest and emotion recognition, SO-PMI and ANN-DL methods are applied. Experimental results show that this method is effective and it can reach high recognition rates. From the perspective of future research, face recognition and other biometric verification of the human body can be used to affective education. Especially, combine emotional intelligence [17] with learner's behavior big data to study is the trend of the future research.

Acknowledgments. This work is supported by the Project of Shanghai Higher Education Society (No. GJSL1624), Program of Science and Technology Commission of Shanghai Municipality (No. 13DZ2252200), Project of Shanghai young university teachers training grant program project of Shanghai Municipal Education Commission (No. ZZSUIBE16026) and Project of Qtone Education of Ministry of Education of China (No. 2017YB115). Many thanks to Jinzhao Wang for her assistance to Yonghui Dai who is the corresponding author of this paper.

References

1. Cao, H.T., Li, M.C., Jiao, D., Feng, B.: A method of micro-blog sentiment analysis based on pad model. *Advanced Materials Research* 659, 186–190 (2013)
2. Chavan, V.M., Gohokar, V.V.: Speech emotion recognition by using svm-classifier. *International Journal of Engineering & Advanced Technology* (5), 11–15 (2012)
3. Cheng, K.Y.: Implementation of affective education: The case of a middle school in china's guangzhou. *Journal of Educational Research* 13(2), 47–63 (2010)
4. Connolly, G.J.: Applying humanistic learning theory: The "art" of coaching. *Strategies* 29(2), 39–41 (2016)
5. Dai, W.H., Duch, W., Abdullah, A.H., Xu, D.R., Chen, Y.S.: Recent advances in learning theory. *Computational Intelligence & Neuroscience* 2015(4), 1–4 (2015)
6. Dai, W.H., Han, D.M., Dai, Y.H., Xu, D.R.: Emotion recognition and affective computing on vocal social media. *Information & Management* 52(7), 777–788 (2015)
7. Dai, W.H., Huang, S., Zhou, X., Yu, X., Ivanović, M., Xu, D.R.: Emotional intelligence system for ubiquitous smart foreign language education based on neural mechanism. *Jitam* 21(3), 65–77 (2014)
8. Dai, Y.H., Han, D.M., Dai, W.H.: Modeling and computing of stock index forecasting based on neural network and markov chain. *ScientificWorldJournal*. 2014, 1–9 (2014)
9. Fatahi, S., Ghasemaghaee, N.: Design and implementation of an intelligent educational model based on personality and learner's emotion. *International Journal of Computer Science & Information Security* 7(3), 14–24 (2010)
10. Ghasemaghaei, R., Arya, A., Biddle, R.: The made framework: Multimodal software for affective education. *American Journal of Emergency Medicine* 8(5), 467–468 (2015)
11. Gong, S.P., Dai, Y.H., Ji, J., Wang, J.Z., Sun, H.: Emotion analysis of telephone complaints from customer based on affective computing. *Computational Intelligence & Neuroscience* 2015(5), 1–9 (2015)
12. Han, W.J., Li, H.F.: A brief review on emotional speech databases. *Intelligent Computer & Applications* 3(1), 5–7 (2013)
13. Herrington, A.: Authentic mobile learning in higher education. *Aare* 8(4), 489–496 (2008)
14. Hu, X.H., Mu, T., Dai, W.H., Hu, H.Z., Dai, G.H.: Analysis of browsing behaviors with ant colony clustering algorithm. *Journal of Computers* 7(12), 3096–3102 (2012)
15. Huang, Y.M., Zhang, G.B., Dong, F., Peng, D.F.: Speech emotion recognition based on two kind of gmm-ubm multidimensional likelihoods and svm. *Application Research of Computers* 28(1), 98–101 (2011)
16. Ijjina, E.P., Mohan, C.K.: Hybrid deep neural network model for human action recognition, vol. 46. Elsevier Science Publishers B. V. (2016)
17. Ivanovic, M., Budimac, Z., Radovanovic, M., Kurbalija, V., Dai, W., B?dic?, C., Colhon, M., Ninkovic, S., Mitrovic, D.: Emotional agents - state of the art and applications. *Computer Science & Information Systems* 12, 47–47 (2015)
18. Kim, C.M., Hodges, C.B.: Effects of an emotion control treatment on academic emotions, motivation and achievement in an online mathematics course. *Instructional Science* 40(1), 173–192 (2012)
19. Li, L., Zheng, L., Yang, F., Li, T.: Modeling and broadening temporal user interest in personalized news recommendation. *Expert Systems with Applications* 41(7), 3168–3177 (2014)
20. Luo, Y.J., Wu, J.H.: Emotional and mental control and cognitive strategy research. *Journal of Southwestern Normal University (Humanities and Social Sciences Edition)* 31(2), 26–29 (2005)
21. Maheswari, B.U., Sumathi, P.: Prediction of web users interest on buying behaviour for e-commerce solutions. *International Journal of Applied Engineering Research* 10(16), 37760–37764 (2015)

22. Mega, C., Ronconi, L., De, B.R.: What makes a good student? how emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology* 106(1), 121–131 (2014)
23. Milton, A., Ghosh, E.S.N., Mohan, S.S., Selvi, S.T.: Analysis towards optimum features and classifiers to recognize the emotion happiness from speech signals. *International Journal of Applied Engineering Research* 10(12), 32585–32600 (2015)
24. Morcom, V.: Scaffolding social and emotional learning in an elementary classroom community: A sociocultural perspective. *International Journal of Educational Research* 67, 18–29 (2014)
25. Nakatsuji, M., Fujiwara, Y., Uchiyama, T., Toda, H.: Collaborative filtering by analyzing dynamic user interests modeled by taxonomy. *Transactions of the Japanese Society for Artificial Intelligence* 28(6), 361–377 (2012)
26. Nwe, T.L., Foo, S.W., Silva, L.C.D.: Speech emotion recognition using hidden markov models. *Speech Communication* 41(4), 603–623 (2003)
27. Picard, R.W.: *Affective Computing*. MIT Press, Cambridge, Mass, USA (1997)
28. Rao, K.S., Krishnamurthy, M., Kannan, A.: Extracting the user's interests by using web log data based on web usage mining. *Journal of Computational & Theoretical Nanoscience* 12(12), 5031–5040 (2015)
29. Romi, S., Katz, Y.J.: Affective education: The nature and characteristics of teachers and students attitudes toward school in israel. *Educational Sciences Theory & Practice* 25(25), 35–47 (2003)
30. Sanchez, F., Barrilero, M., Alvarez, F., Cisneros, G.: User interest modeling for social tv-recommender systems based on audiovisual consumption. *Multimedia Systems* 19(6), 493–507 (2013)
31. Schröder, M.: The cognitive structure of emotion. *Pure & Applied Geophysics* 168(12), 2395–2425 (2011)
32. Schuller, B., Rigoll, G., Lang, M.: Hidden markov model-based speech emotion recognition. In: *Proceedings of the 2003 International Conference on Multimedia and Expo, IEEE Computer Society*. vol. 2, pp. 401–404 (2003)
33. Siu, K.W.M., Yi, L.W.: Fostering creativity from an emotional perspective: Do teachers recognise and handle students emotions? *International Journal of Technology & Design Education* 26(1), 105–121 (2016)
34. Stamou, S., Ntoulas, A.: *Search personalization through query and page topical analysis*. Kluwer Academic Publishers (2009)
35. Valk, J.H., Rashid, A.T., Elder, L.: Using mobile phones to improve educational outcomes: An analysis of evidence from asia. *International Review of Research in Open & Distance Learning* 11(1), 117–140 (2010)
36. Wang, N.: Highlighting the humanistic spirit in the age of globalization: Humanities education in china. *European Review* 23(2), 273–285 (2015)
37. Wu, J.D., Ye, S.H.: Driver identification based on voice signal using continuous wavelet transform and artificial neural network techniques. *Expert Systems with Applications* 36(2), 1061–1069 (2009)
38. Xu, B.: A user interest model based on the analysis of user behaviors. *Journal of Intelligence* 27(12), 76–78 (2009)
39. Xu, L., Ren, J.S.J., Liu, C., Jia, J.: Deep convolutional neural network for image deconvolution. *Advances in Neural Information Processing Systems* 2, 1790–1798 (2014)
40. Xu, L.H., Lin, H.F., Pan, Y., Ren, H., Chen, J.M.: Constructing the affective lexicon ontology. *Journal of the China society for scientific and technical information* 27(2), 180–185 (2008)
41. Zhao, X.: A review of progress and trend of foreign research on emotional education. *Comparative Education Review* (2013)
42. Zou, C.R., Zhang, X.R., Cheng, Z., Li, Z.: A novel dbn feature fusion model for cross-corpus speech emotion recognition. *Journal of Electrical and Computer Engineering* 2016(1), 1–11 (2016)

Haijian Chen is an associate professor at the Shanghai Engineering Research Center of Open Distance Education, Shanghai Open University, China. He received his Ph.D. in Management Science and Engineering from Shanghai University of Finance and Economics, China in 2015. His current research interests include Educational technology and cloud computing. Contact him at xochj@shtvu.edu.cn.

Yonghui Dai, corresponding author of this work, he is currently a lecturer at the Management School, Shanghai University of International Business and Economics, China. He received his Ph.D. in Management Science and Engineering from Shanghai University of Finance and Economics, China in 2016. His current research interests include Affective Computing, Intelligence Service and Big Data Analysis. His works have appeared in international journals with eighteen papers. Contact him at daiyonghui@suibe.edu.cn.

Yanjie Feng is an associate professor at the Management School, Shanghai University of International Business and Economics, China. She received her Ph.D. in Management Science and Engineering from Shanghai Jiao Tong University, China in 2001. Her research interests include Network information communication and financial risk management. Contact her at fengyanjie@suibe.edu.cn.

Bo Jiang is currently a Ph.D. candidate at the School of Information Management and Engineering, Shanghai University of Finance and Economics, China. His current research interests include Data Mining and Cloud Computing. Contact him at jiangbo@sui.edu.cn.

Jun Xiao is currently a professor at the Shanghai Engineering Research Center of Open Distance Education, Shanghai Open University, China. He received his Ph.D. in Educational technology from East China Normal University, China in 2006. His research interests include Educational technology and Management Information System. Contact him at xiaoj@shtvu.edu.cn.

Ben You is currently a Master candidate at the Management School, Shanghai University of International Business and Economics, China. He received his bachelor's degree in Traffic Engineering from Xiamen University of Technology, China in 2016. His current research interests include Data mining and Network behavior analysis. Contact him at youben022@gmail.com.

Received: January 10, 2017; Accepted: August 19, 2017.