



Research on Information Recommendation Optimization Mechanism Based on Emotional Expression and Cognition

Ke Zhong, Liqun Zhang^(✉), and Xiaolei Guan

Institute of Design Management, Shanghai Jiao Tong University, Shanghai, China
zhanglliqun@gmail.com

Abstract. The information revolution brought by the rapid spread of the Internet has made people be more dependent on the Internet to obtain information. The internet products based on information are increasing the construction of the information recommendation system. But at present, the information recommendation systems are all limited to the content of information itself, ignoring the differences in the emotional expression of information. This research tried to purpose an approach for the information emotional semantic classification and recommendation based on emotional cognitive model. This research used news Internet product as an example to construct a quantitative coordinates of emotional cognition system through variable control, samples association analysis and statistical data analysis to improve the present information recommendation system. The results of verification experiment indicated that it is practical and effective for information classify and recommend. It is foreseeable that the theory of this research can be applied to other internet products about business, social communication and so on. Meanwhile, the research methods and results can be applied to psychology, sociology research and other specific areas, playing a guiding and testing role.

Keywords: Quantitative coordinates of emotional cognition system
Recommendation system · Emotional cognitive · Emotional expression

1 Introduction

The rapid development of the Internet has led to the emergence of Web2.0, which has the main feature of advocating individualization, and UGC is coming with the development of Web2.0. In the context of this era, the data on the Internet are growing exponentially, and users are more dependent on the Internet to obtain information. And personalized recommendation system was born to solve the above problems. However, the information recommendation systems are all limited to the content of information itself, ignoring the differences in the emotional expression of information.

This research aims to purpose an approach for the information emotional semantic classification and recommendation based on emotional cognitive model. This research put news Internet product as an example, and used variable control to explore the differences between different emotional expressions of information under the same content.

In view of the differences found, professional researchers have proposed a quantitative coordinates of emotional cognition system.

The quantitative coordinates of emotional cognition system constructed in this research innovatively integrates the users' differences in emotion expression and cognition into the information recommendation mechanism, which is quantifiable.

This research creatively proposes the optimization mechanism based on the quantitative coordinates of emotional cognition system. The optimization mechanism has the following advantages,

- Integrating the users' subjective evaluation of information expression into the mechanism, which fills the blank and further enriches the information recommendation mechanism.
- Putting the differences in the emotional expression and cognition of users into the information recommendation system, and the related factors will make recommendation system more comprehensive and diversified, the similarity analysis between users will also be more diversified, eventually making personalized recommendations more precise and effective.
- Adding emotional factors to information recommendation system, which adds more data to the Internet from another dimension, expanding the value of data for the Internet.

The rest of the paper is organized as follows. An overview of emotion, UGC and intelligent recommendation system with current situation is presented in Sect. 2. Section 3 used News Internet Products as an example to explore the differences between different emotional expressions of information under the same content through variable control, samples association analysis and statistical data analysis. The information recommendation optimization mechanism with quantitative coordinates of emotional cognition system are presented in Sect. 4. The verification experiment in order to prove that it is practical and effective for information classify and recommend is presented in Sect. 5. Section 6 is the summary and prospect.

2 Desktop Research

Emotion. Emotion refers to the subjective feelings or experiences of the individual [1]. Emotional experience refers to the individual subjective experience of emotion [2]. Emotion is a part of attitude. It is in harmony with the introverted feelings and intentions in attitude. It is a more complex and stable physiological evaluation and experience of physiology [3, 4]. And for the same thing, each person always has a different emotional feedback and experience [2].

UGC. With the development of the Internet, the interaction of Internet users is embodied. The user is not only the browser of the content of the network, but also the creator of the content of the network [5].

UGC (User Generated Content) is arisen with the concept of Web2.0, which is the main feature of advocating personalization. It is not a specific business, but a new way

for users to use the Internet, from the original downloading to downloading and uploading [6, 7].

Recommendation System. Recommendation system is a subclass of information filtering system that seeks to predict the “rating” or “preference” that a user would give to an item [8, 9]. Recommendation system is the product of the development of Internet and e-commerce. It is a high-level business intelligence platform based on massive data mining, providing personalized information services and decision support to customers [10]. In recent years, many successful examples of large-scale recommender systems have emerged. For example, MovieLens recommends movies for users [11], Amazon recommends books, audio-visual resources and other products for users [12], and VERSIFI and TOUTIAO recommend news for users [13]. Meanwhile, personalized recommendation system has gradually become one of the research hotspots in academic circles [14–17] (Fig. 1).

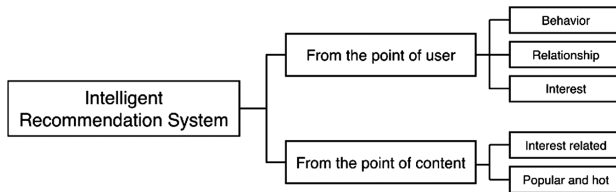


Fig. 1. Intelligent recommendation system

The current research of personalized intelligent recommendation system is limited to the research of user’s objective data and subjective behavior data, and the topic classification of content itself [18].

- From the point of users, personalized recommendation is made on the basis of user behavior (praise, comment and so on), user relationship (common friends, etc.), user interest, etc.
- From the point of content, personalized intelligent recommendation for users based on relevant content (keyword association, the same topic), popular content (hot content recommendation) and so on.

However, the research on user’s emotion cognition is still blank. Therefore, the main purpose of this study is to explore the relationship between the different expressions of the content and the different emotional cognition of the users.

3 Exploratory Research

There are many scenarios for the application of the intelligent recommendation system, such as the e-commerce industry: Amazon, Taobao, multimedia software: MovieLens, news app: Versifi, Toutiao and so on.

This paper aims to study the relationship between the content and the user’s emotional cognition, so we choose the news app as the research object, the reasons are:

- Personalized intelligent recommendation system is widely used in the news app. And news app takes text content as the main part, which is very representative.
- The main product of news app is based on text content, which facilitates the research of the relationship between the content and the user's emotional cognition.
- News app is now widely used in daily life, and the results are highly practical and extensible.
- Data in news app is rich, and easy to obtain.

3.1 Data Collection

The classification of news in news app can be divided into: sports, society, technology, military, health, education, women, estate, culture, automobile, finance and economics, international, entertainment, food, tourism, etc.

Professional researchers chose three categories in the above categories: sports, technology and society, which have a lot of news in each category and a low correlation between each other. And professional researchers collected 20 news with different expressions for the same content in each category. The source of the data is the common Chinese News app and internet.

3.2 Sample Process

Professional researchers processed the collected 60 news, extracting their text content (including the title), using a unified font size, word spacing, fonts to print them on the same white paper. And then researchers numbered them on the back of each white paper, A1 to A20 (Technology), B1 to B20 (Sports), C1 to C20 (Society), took them as the experimental samples of follow-up. Photo of the processed samples is shown in Fig. 2.

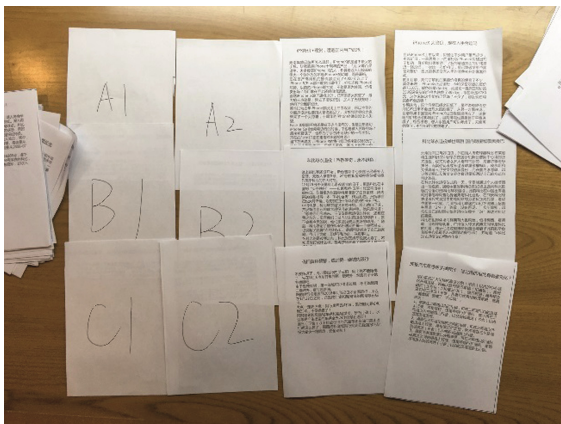


Fig. 2. Photo of the processed samples

3.3 Correlation Analysis

Subsequently, professional researchers respectively did correlation analysis to different kinds of news (20 samples of each kind of news), and observed their correlation results in the derived Euclidean distance model, trying to find out the key factors affecting the user cognitive and emotional differences on it.

Correlation Score

Under the hypothesis that samples in each kind of news search is extensive and representative, the experiment invited three professionals to grade the correlation between any two indicators of 20 samples in each kind of news. The score is on a 9-point scale ranging from ‘significant negative correlation’ to ‘significant positive correlation’ [19]. The scoring criteria are based on the professionals’ understanding and cognition of each news content. The relevant portion of the score result is shown in Fig. 3.

	A1	A2	A3	A4	A5	A6
A1	1	3	6	4	9	6
A2	3	1	7	2	9	4
A3	6	7	1	6	3	5
A4	4	2	6	1	9	3
A5	9	9	3	9	1	7
A6	6	4	5	3	7	1
A7	2	2	6	2	9	4
A8	3	2	8	3	9	5
A9	2	2	6	2	9	4

Fig. 3. Part screen shot of sample correlation score

Multidimensional Scaling Analysis

The correlation in 3 groups was scored by 3 professionals, and we got the results of 9 sets of correlation scores. Professional researchers then respectively numbered the data results to Group1-A, Group1-B, Group1-C, Group2-A, Group2-B, Group2-C, Group3-A, Group3-B, Group3-C.

SPSS software is used to analyze the samples correlation matrix through multidimensional scale analysis, and factor analysis. The reliability analysis results are as follows (Fig. 4).

According to the results of three sets of correlation scores for each sample, professional researchers selected two of the highest reliability as the follow-up experiment subjects: Group1-B, Group2-A, Group2-B, Group2-C, Group3-A, Group3-C. Meanwhile, we found that the 6 selected groups’ reliabilities were all above 0.7. According to the majority of scholars’ point of view on the SPSS reliability analysis, the reliability coefficient above 0.7 or more, means that the data needs to be modified, but it still has its value [20, 21]. On the basis that the test data is still valuable, researchers continue to analyze the data by multidimensional scaling analysis.

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.574	.677	20

a

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.754	.750	20

b

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.533	.535	20

c

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.745	.657	20

d

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.814	.769	20

e

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.753	.736	20

f

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.793	.815	20

g

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.717	.701	20

h

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.788	.778	20

i

Fig. 4. (a) Reliability statistics of professional Group1-A; (b) reliability statistics of professional Group1-B; (c) reliability statistics of professional Group1-C; (d) reliability statistics of professional Group2-A; (e) reliability statistics of professional Group2-B; (f) reliability statistics of professional Group2-C; (g) reliability statistics of professional Group3-A; (h) Reliability statistics of professional Group3-B; (i) reliability statistics of professional Group3-C.

Based on the relevance coefficient in the matrix to build N-dimensional space, the Euclidean distance formula (1) can be used to calculate the spatial distance of two samples. The closer, the more similar samples can be considered.

$$Euclid(1,2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{1}$$

Multidimensional scaling analysis can visually see the spatial distribution of all samples. The results of multidimensional scaling analysis are as follows (Fig. 5).

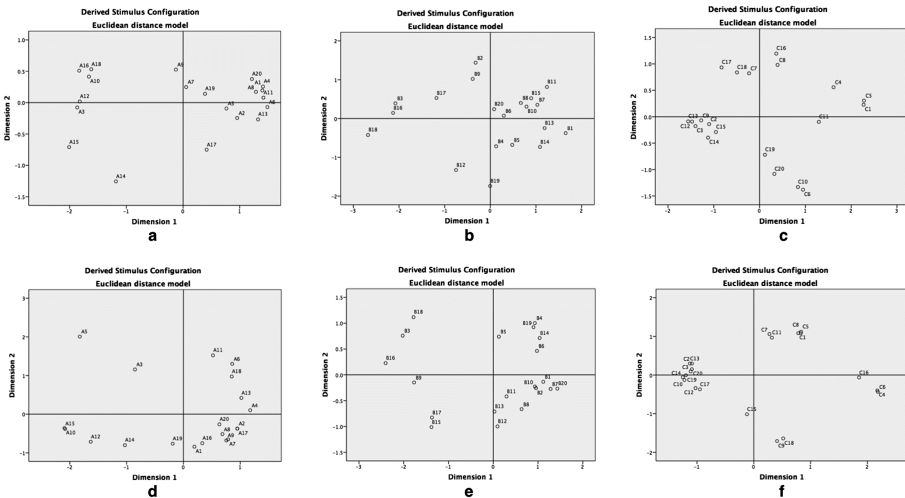


Fig. 5. (a) Multidimensional scaling analysis of Group2-A; (b) multidimensional scaling analysis of Group1-B; (c) multidimensional scaling analysis of Group2-C; (d) multidimensional scaling analysis of Group3-A; (e) multidimensional scaling analysis of Group2-B; (f) multidimensional scaling analysis of Group3-C;

Because the selected samples' difference is the expression of the news, combining the multidimensional scaling analysis results, professional researchers suspected that the main factors that affect the user's cognitive differences are following two dimensions: emotional (more popular and entertainment in expression) and rational (more rigorous and official in expression); simple and rich (the latter will use more relevant content to support expressing the same content).

4 Quantitative Coordinates of Emotional Cognition System

The current personalized recommendation system in news app is based on user's objective data and subjective behavioral data to judge and recommend the topics and columns that users interest in.

However, through the above exploratory experiments, professional researchers found that different expression of news content can cause great emotional and cognitive differences to users.

At the same time, after a certain text analysis, professional researchers suspected that the main factors that affect the user's cognitive differences are following two dimensions: emotional (more popular and entertainment in expression) and rational (more rigorous and official in expression); simple and rich (the latter will use more relevant content to support expressing the same content).

Therefore, quantitative coordinates of emotional cognition system is put forward to optimize personalized recommendation system in news app, make personalized recommendation system more targeted and accurate, to improve user's cognitive efficiency and experience when using the products. Quantitative coordinates of emotional cognition system is shown in Fig. 6.

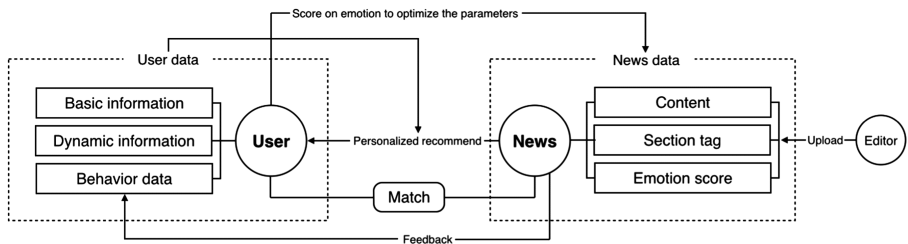


Fig. 6. Quantitative coordinates of emotional cognition system

4.1 News Data

Editor's Upload. When editor is uploading a news, he was traditionally uploading the content and choosing the section of the news. However, in this system, editor needs to score the emotional and rational, simple and rich of the news to get the coordinates of emotional cognition. Then, a news has three main parameters: Content, Section tag, Coordinates of emotional cognition. Given initial coordinates of emotional cognition of news C_{i0} .

User's Feedback to Optimize. After users read the news, some of them are randomly selected to score the emotional and rational, simple and rich of the news, coordinates of emotional cognition set $C = \{C_{i1}, C_{i2}, \dots, C_{in}\}$ can be gotten. The results of the score will be processed according to the relevance between the users' behavior and the news, as follows (Fig. 7).

BEHAVIOR	RELEVANCE
click the news to view when browsing	1
click the news to view after retrieval	1
stay on the news page longer than the average	1
put the news in the favorites after reading	1
comment the news after reading	0.5
forward the news to others after reading	0.5
like the news after reading	0.5
follow the editor of the news after reading	0.5
screenshot on the news page	0.25

Fig. 7. Relevance of users' behavior and the news

Then the processed results of the score will be attached to the original coordinates of emotional cognition.

As the users continue to read news and score, the coordinates of emotional cognition of the news will be constantly adjusted and optimized.

4.2 User Data

Basic Information. Users need to enter some basic information when they register in a news app, such as sex, age, hobbies, etc.

Dynamic Information. In addition to the basic information, some technologies accompanied by the development of the Internet like global positioning system, caching, cloud computing, big data and so on can also automatically obtain some objective dynamic information of the user. For example, based on global positioning system, we can obtain the dynamic location of the user, using this information, we can recommend some local news.

Behavior Data. When users use news app, they produce a lot of behavior data, which have a certain correlation with the news. It can be seen in Fig. 7.

According to the relevance between the users and the news, we can get the news users are interested in, which based on the three main parameters of the news mentioned above.

In addition, we can also get more information based on the user's behavior data. For example, based on user behavior data, we can get whether a user will be interested in different kinds of news in different periods of time and other details.

4.3 Personalized Recommend

User Similarity Computing. User similarity computing plays a very important role in collaborative filtering systems, user recommendation systems as well as social network services [22].

The input is a m user's score matrix for n news. We can find neighbor users who are similar to the current interest through user ratings. We can also recommend new news resources to the users according to the neighboring users. The method of finding neighbor users is the user similarity computing. The common methods of similarity measure are Pearson correlation coefficient, cosine similarity, etc.

Pearson correlation coefficient. If user a and u jointly evaluate excessive item set as I_{au} , Pearson correlation coefficient can be used to measure similarity between user a and u. User a and u similarity computation can be expressed as follows:

$$\text{sim}(a, u) = \frac{\sum_{i \in I_{au}} (r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i \in I_{au}} (r_{a,i} - \bar{r}_a)^2 \sum_{i \in I_{au}} (r_{u,i} - \bar{r}_u)^2}} \quad (2)$$

Among it, $r_{a,i}$ and $r_{u,i}$ respectively express user a and u's evaluation of item i. \bar{r}_a and \bar{r}_u respectively represent the average score for the items.

Cosine similarity. The user's score of n items are regarded as the scoring vectors on the n-dimensional projects. \vec{a} and \vec{u} denote the user a and u's scoring vectors respectively. Then the similarity between users can be measured by calculating the angle between different users' scoring vectors.

$$\text{sim}(a, u) = \cos(\vec{a}, \vec{u}) = \frac{\vec{a} * \vec{u}}{\|\vec{a}\| \|\vec{u}\|} \quad (3)$$

Computing user similarity can not only be helpful for recommending new news resources to the users, but also optimize user data and news data to enrich the data of the whole recommendation mechanism.

The Match Between the News Data and the User Data. Personalized recommend the news to the user based on the match of user data and news data, which is the most commonly recommendation system. But we can see that the recommendation system will be more accurate in the above recommendation mechanism, based on the section tag and emotion score.

5 Experimental Verification

5.1 Experiment Setting

Professional researchers had suspected that the main factors that affect the user's cognitive differences are following two dimensions: emotional (more popular and entertainment in expression) and rational (more rigorous and official in expression);

simple and rich (the latter will use more relevant content to support expressing the same content).

In this chapter, the same professionals used Likert scale to quantify the emotional coordinates of the samples, then the emotional coordinates of the samples would be gotten. Professional researchers classified the samples by the sample correlation score above. The classification results would be compared with emotional coordinates of the samples to see if there is a high degree of consistency to verify the conjecture described in the previous section.

5.2 Experiment Process

Likert Scale. Likert scale is a psychometric scale commonly involved in research that employs questionnaires. It is the most widely used approach to scaling responses in survey research, such that the term (or more accurately the Likert-type scale) is often used interchangeably with rating scale [23].

To verify the conjecture described in the previous section, the professional researchers designed the following 7 point Likert scale (Fig. 8).

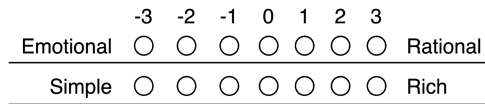


Fig. 8. 7 point Likert scale

Then the professionals scored the same samples in the Likert scale.

According to the scoring results of professionals, professional researchers draw the following coordinates of the 6 samples' scoring results (Fig. 9).

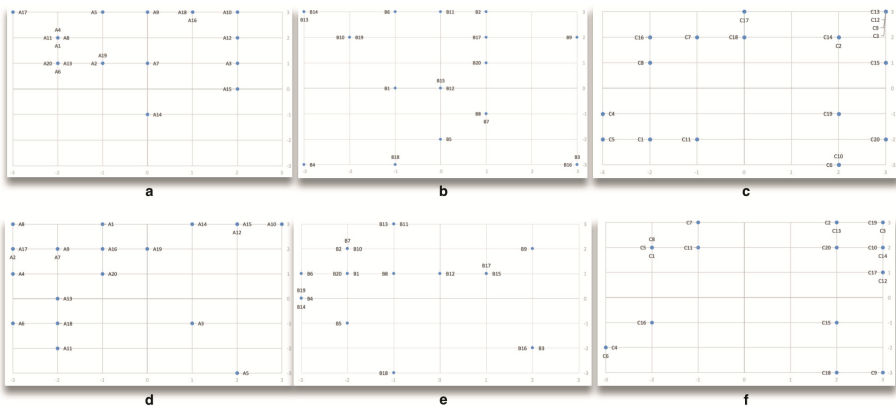


Fig. 9. (a) Emotional coordinates of Group2-A; (b) emotional coordinates of Group1-B; (c) emotional coordinates of Group2-C; (d) emotional coordinates of Group3-A; (e) emotional coordinates of Group2-B; (f) emotional coordinates of Group3-C;

With the dendrograms based on the sample correlation score, professional researchers clustered the samples, as follows (Fig. 10).

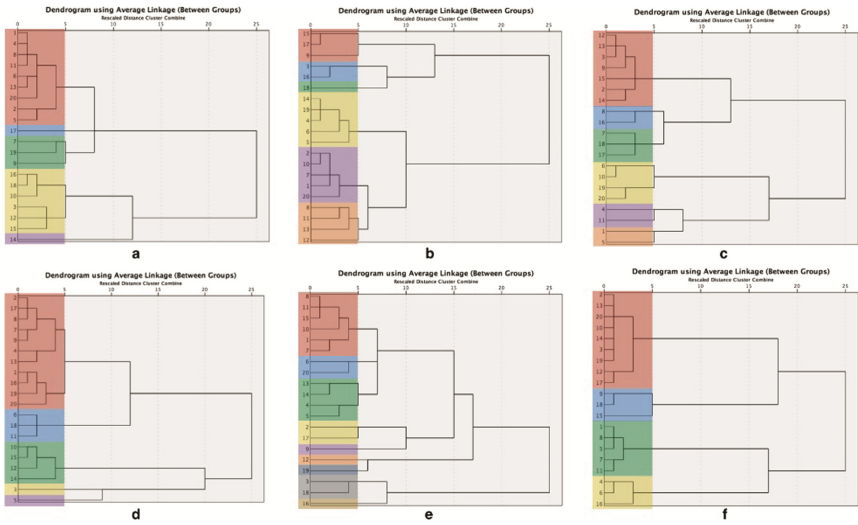


Fig. 10. (a) Dendrogram and sample clustering of Group2-A; (b) dendrogram and sample clustering of Group1-B; (c) dendrogram and sample clustering of Group2-C; (d) dendrogram and sample clustering of Group3-A; (e) dendrogram and sample clustering of Group2-B; (f) dendrogram and sample clustering of Group3-C;

The figure of their multidimensional scaling with clustering is shown in Fig. 11.

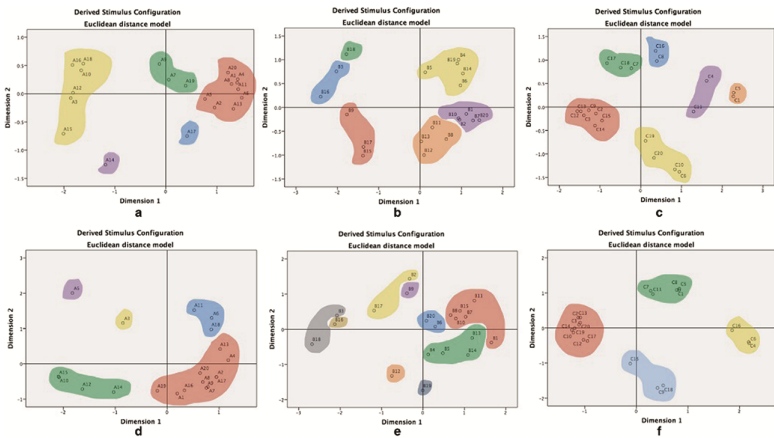


Fig. 11. (a) Multidimensional scaling with clustering of Group2-A; (b) multidimensional scaling with clustering of Group1-B; (c) multidimensional scaling with clustering of Group2-C; (d) multidimensional scaling with clustering of Group3-A; (e) multidimensional scaling with clustering of Group2-B; (f) multidimensional scaling with clustering of Group3-C;

Then the clustering results were put into the emotional coordinates obtained above, we could get the figure as follows (Fig. 12).

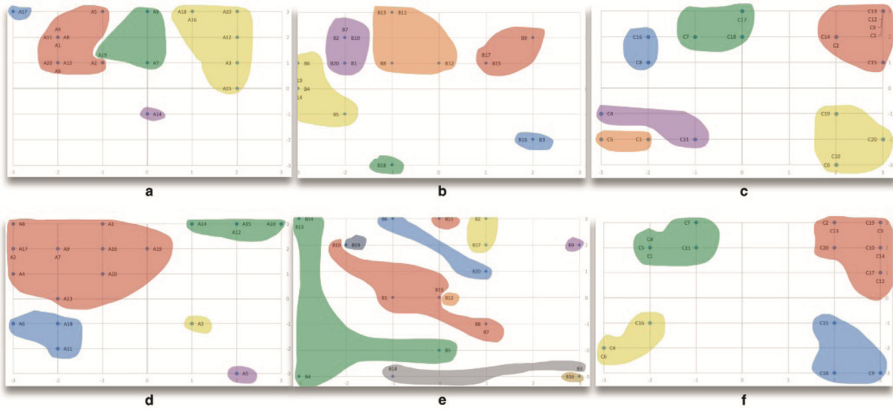


Fig. 12. (a) Emotional coordinates with clustering of Group2-A; (b) emotional coordinates with clustering of Group1-B; (c) emotional coordinates with clustering of Group2-C; (d) emotional coordinates with clustering of Group3-A; (e) emotional coordinates with clustering of Group2-B; (f) Emotional coordinates with clustering of Group3-C;

Professional researchers then compared the emotional coordinates with the clustering results based on the sample correlation score and the figure of multidimensional scaling with clustering.

We can clearly see that there is a high degree of consistency between them, proving that the system mentioned above is very practical and effective in information rank and recommend. It is reasonable to recommend information based on the emotional coordinates as optimization.

6 Conclusion

This paper initially envisages an information recommendation optimization mechanism based on the quantitative coordinates of emotional cognition system through variable control, samples association analysis and statistical data analysis. Its core idea is to integrate the factors of user's emotional expression and cognition to the original information recommendation mechanism to optimize the mechanism.

At the same time, based on behavior psychology, the users' behaviors provide much data for the modeling and make the model more representative. Besides, the users' subjective evaluation of information expression is integrated into the mechanism, which fills the blank and further enriches the information recommendation mechanism.

However, the factors that affect users' cognition and use efficiency should be multi-dimensional at the level of emotional expression and cognition, not just the two-dimensional displayed in the quantitative coordinates of emotional cognition system. Therefore, we actually need more samples which contain various fields and topics to analysis,

and more subjects, including professionals and ordinary users, to get more accurate multidimensional factors.

It is foreseeable that the theory of this research can be applied to other internet products about business, social communication and so on. Meanwhile, the research methods and results can be applied to psychology, sociology research and other specific areas, playing a guiding and testing role.

References

1. Fox, E.: *Emotion Science Cognitive and Neuroscientific Approaches to Understanding Human Emotions*. Palgrave Macmillan, Basingstoke (2008)
2. Carstensen, L.L., Pasupathi, M., Mayr, U., Nesselroade, J.R.: Emotional experience in everyday life across the adult life span. *J. Pers. Soc. Psychol.* **79**(4), 644 (2000)
3. Scherer, K.R.: What are emotions? And how can they be measured? *Soc. Sci. Inf.* **44**(4), 695–729 (2005)
4. Norman, D.A.: *Emotional Design: Why We Love (or Hate) Everyday Things* (2004)
5. 张森: 社会化媒体在市场营销中的应用研究. (Doctoral dissertation, 首都经济贸易大学) (2014)
6. 包传颖: 移动视频, 领跑 3 g 新应用. *信息技术* **3**(2) (2009)
7. Liang, T., Zhang, L., Xie, M.: Research on image emotional semantic retrieval mechanism based on cognitive quantification model. In: Marcus, A., Wang, W. (eds.) *DUXU 2017. LNCS*, vol. 10290, pp. 115–128. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58640-3_10
8. Ricci, F., Rokach, L., Shapira, B.: *Introduction to Recommender Systems Handbook*. ACM (2004)
9. Weaver, A.: Facebook and other pandora's boxes. *Access* **24**(4), 24 (2010)
10. 任明: 智能信息系统: 以关联知识优化数据建模的方法和实践. 浙江大学出版社 (2012)
11. Linden, G., Smith, B., York, J.: Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Comput.* **7**(1), 76–80 (2003)
12. Miller, B.N., Albert, I., Lam, S.K., Konstan, J.A., Riedl, J.: MovieLens unplugged: experiences with an occasionally connected recommender system. In: *International Conference on Intelligent User Interfaces*, pp. 263–266. ACM (2003)
13. Billsus, D.: Adaptive interfaces for ubiquitous web access. *Commun. ACM* **45**(5), 34–38 (2002)
14. Hill, W., Stead, L., Rosenstein, M., Furnas, G.: Recommending and evaluating choices in a virtual community of use. In: *SIGCHI Conference on Human Factors in Computing Systems*, vol. 1, pp. 194–201. ACM Press/Addison-Wesley Publishing Co. (1995)
15. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews. In: *ACM Conference on Computer Supported Cooperative Work*, pp. 175–186. ACM (1994)
16. Shardanand, U., Maes, P.: Social information filtering: algorithms for automating “word of mouth”. In: *SIGCHI Conference on Human Factors in Computing Systems*, vol. 110, pp. 210–217. ACM Press/Addison-Wesley Publishing Co. (1995)
17. Plume, M.L.: SPSS (statistical package for the social sciences). *Encyclopedia Inf. Syst.* **38**(4), 187–196 (2003)
18. 邓以克: 基于改进的 SVM-KNN 算法的中文网页层次式分类. (Doctoral dissertation, 浙江大学计算机科学与技术学院 浙江大学) (2010)

19. Xie, M., Zhang, L., Liang, T.: A quantitative study of emotional experience of *Daqi* based on cognitive integration. In: Marcus, A., Wang, W. (eds.) DUXU 2017. LNCS, vol. 10288, pp. 306–323. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58634-2_24
20. Nie, N.H., Bent, D.H., Hull, C.H.: SPSS: Statistical Package for the Social Sciences. McGraw-Hill, New York (1970)
21. Green, S.B., Salkind, N.J., Jones, T.M.: Using SPSS for Windows; Analyzing and Understanding Data. Prentice Hall PTR, Upper Saddle River (1996)
22. Fang, A., Ya-Zi, L.I.: Computing algorithm of user similarity in virtual medical communities. *J. Med. Inform.* (2011)
23. Wuensch, K.L.: What is a Likert Scale? And How Do You Pronounce “Likert?”. East Carolina University, 4 October 2005. Accessed 30 Apr 2009