Information Relationship Identification in Team Innovation

Xinmiao Li[1], Xinhui Li[2], and Pengzhu Zhang[3]

[1] School of Information Management and Engineering,
Shanghai University of Finance and Economics,
200433, Shanghai, China
xinmiaolisunny@hotmail.com
[2] DongBeiWang, Haidian,
100094, Beijing, China
miaoxinhuili@gmail.com
[3] Antai Management School, Shanghai Jiaotong University,
200052, Shanghai, China

Abstract. In the 1990s, innovation has become the main resource of competitive advantage for firms. Computer-mediated team innovative activities provide much more information, thus resulting in information overload. In this paper, the method for identifying the affirmative and negative relationship between the solution and its comment is proposed. The identification model is set up and applied in team innovation. In the model, we improve the traditional feature extraction method from three aspects: classifying feature words into six levels according to their affirmative and negative tone, introducing the new feature of the prior comment's attitude and the new feature of the number of sub-sentences. The results show that the method realizes the identification of the affirmative and negative relationship effectively. It can benefit organizing the large amount of team comments effectively and increase the efficiency of information organization.

1 Introduction

From the 1990s, innovation, replacing efficiency and quality, is regarded as the main resource of competitive advantage for firms [1]. Computer-mediated team innovative activities generate much more redundant information [2][3], thus results in information overload [4][5]. In this case, it cannot satisfy the needs of team innovation by using manual method. Therefore, it is necessary and imperious to study the organization of team information effectively by use of the advanced information technology. However, the process of team innovation is unstructured and complex, which results in the complicated information relationship and difficult information organization. On the other hand, due to the limitation of Chinese text mining technology, complete automation of Chinese information organization is difficult to realize. The essence of information organization is the classification of information. This paper particularly studies the identification of the affirmative and negative relationship between the solution and its comments during team innovation.

2   Literature Review

Traditional team activities usually depend on the secretary to record and pack up the team information, which costs time and effort. With the development of information technology, more and more scholars pay attention to the research on effective methods and tools to support team information organization during team activities. Some scholars study information classification produced in e-meeting by mining the comparability of information topics [6][7]. Most researches about information classification usually make the concepts involved in information as classes [8]. In team innovation, aiming at a solution, some comments bear a clearly and definitely affirmative or negative relationship with the solution. If the comments are picked up and identified, the team can recognize and master the tendentiousness of the comments, which favors the solution selection and decision. Therefore, the identification of the affirmative and negative relationship between the solution and its comments deserves the research. Moreover, most researches deal with English text. Because of the particularity of Chinese, those researches cannot treat Chinese text effectively. On the basis of text mining and Neural Network (NN), this paper studies the identification of the affirmative and negative relationship between a solution and its comments.

3   Model of Team Information Relationship Identification

In the real process of team innovation, there are various relationships between a solution and its comments. As for the solution selection and decision, the affirmative or negative relationship is what the team is concerned about most. In this paper, the model of information relationship identification is proposed. The model is used to identify the affirmative or negative relationship between the solution and its comment. Firstly, team comment is picked up from the team information base as the target information for identification. Then, the model extracts the features of comments. The next step is to put those features into the NN. The relationship between the comments and solution is identified by the NN. Finally, the results identified are put back into the team information base and used to the team information organization.

4   Application of the Identification Method to Team Innovation

The method above has been applied to an innovative project. During the project, a design solution is selected randomly. 100 comments on the solution are selected randomly as training samples. Another 100 comments on the solution are selected randomly as testing samples.

4.1   Identification Model

Feature Extraction of Team Information
The characteristic vector of comment should be extracted first during the identification. The algorithm of feature extraction has a direct impact on information relationship identification. Vector Space Model (VSM) is mainly used. However, the traditional

method is not appropriate to address the problem mentioned in this paper. We improve the traditional method of feature extraction.

Firstly, words and phrases representing affirmative or negative meaning will be selected as characteristic items.

Secondly, because the comments studied by this paper are short, many dimensions of the characteristic vectors deriving from the traditional extracting method are zero. In order to reduce the dimensions of characteristic vectors greatly and avoid serious losing of features, based on the result of experiments, the feature words are sorted into six levels representing different degrees of affirmation or negation. The six levels separately express the meanings of feature words: extremely affirmation, affirmation, somehow affirmation, extremely negation, negation, and somehow negation.

Thirdly, in order to further improve the precision of the comment attitude, the following two features are particularly added.

(1) *Number of sub-sentences*: This feature describes the number of sub-sentences included by a comment. The sub-sentences are separated by punctuations. This feature can represent the distribution of feature words in a text, which can usually reflect the attitude of the current text.

(2) *Attitude of the prior text*: This feature reflects the implied syntax relationship between neighbored texts in continuous comments. It is mainly used to improve the identification effect on the condition of continuous comments.

BPNN Classifier

A three-layer Back Propagation Neural Network (BPNN) is used to identify the relationship of team information. The input layer has eight neurons. They are the weight of extremely affirmative feature words, the weight of affirmative feature words, the weight of somehow affirmative feature words, the weight of extremely negative feature words, the weight of negative feature words, the weight of somehow negative feature words, the number of sub-sentences, and the attitude of the prior text. Among which, the values of the first seven neurons are $\{0, 1, 2, 3, 4, 5\}$. The value of the last one is $\{-1, 0, 1\}$. There is one neuron in the output layer. The value of the output node is $\{-1, 1\}$. There are four neurons in the hidden layer. *tansig* is the action function of the hidden layer. *purelin* is the action function of the output layer.

4.2   Results of the Identification Model

The experimental results show that the correct ratio of the training set is 100%, and the correct ratio of the testing set is 73.33%. The applied effect can be further described by Recall and Precision. The Precision and the Recall of the affirmative relationship is separately 76.47% and 72.22%. The Precision and the Recall of the negative relationship is separately 75% and 78.95%. The average Recall is 75.74%. The average Precision is 75.59%. The results show that the model can support the identification of the affirmative or negative relationship between the solution and its comments. It increases the efficiency of information organization.

5   Conclusions

In this research, we explore the identification method of the affirmative and negative relationship between the innovation solution and its comments during team innovation. In the identification model, we improve the traditional feature extraction method from three aspects. The results show that the method realizes the identification of the affirmative and negative relationship between the solution and its comment effectively. It can increase the efficiency of information organization.

References

1. Bolwijn, P.T., Kumpe: T. Manufacturing in the 1990s-Productivity, Flexibility and Innovation. Long Range Planning (1990) 44-57
2. Michael Parent, R. Brent Gallupe. Knowledge Creation in Focus Groups: can Group Technologies Help? Information & Management (2000) 47-58
3. Schilling, M.A., Hill, C. Managing the New Product Development Process: Strategic Imperatives. Academy of Management Executive (1998) 67-81
4. Trauth E M, Jessup L M.: Understanding Computer-Mediated Discussions Positives and Interpretive Analyses of Group Support System Use. MIS Quarterly (2000) 43-79
5. Grise M, Gallupe R B.: Information Overload Addressing the Productivity Paradox in Face-to-Face Electronic Meetings. Journal of Management Information Systems (1999-2000) 157-185
6. H. Chen, K.J. Lynch,: Automatic Construction of Networks of Concepts Characterizing Document Databases. IEEE Transactions on Systems, Man, and Cybernetics (1992) 885- 902
7. H. Chen, P. Hsu, R. Orwig L. Hoopes, and J. F. Nunamaker.: Automatic Concept Classification of Text from Electronic Meetings. Communications of the ACM (1994) 56-73
8. Dmitri Roussinov, J. Leon Zhao.: Automatic Discovery of Similarity Relationships through Web Mining. Decision Support Systems (2002) 1-18