Topic map construction using bee colony algorithm

Student：Wei-Jhong Ji　　　　　　　Advisor：Dr. Chieh-Yuan Tsai

Institute of Industrial Engineering and Management
Yuan-Ze University

ABSTRACT

With the rapid development of the information technology, information overload is becoming a serious problem during the information acquisition process. Information overload leads users spend more time to find necessary knowledge. To relieve this difficulty, knowledge map is a systematic approach to reveal the underlying relationships between abundant knowledge sources. However, few studies focused on optimizing the coordinates of objects in the map. In addition, too many parameters should be set which lead them complicated and not intuition. To solve the above problems, this thesis presents a novel knowledge map approach to transform high-dimensional objects into a 2-dimensional space to help understand complicated relatedness among high-dimensional important topics. First, the papers related to certain domains are collected from the knowledge database and papers as the knowledge items that contains many keywords. Second, the collected knowledge items are presented as the vector space model (VSM). In VSM, keywords can be represented as a term vector in m-dimensional space where the term frequency-inverse document frequency (TF-IDF) approach is used for term weighting so that the tf-idf value increases proportionally to the number of times a keyword appears in the knowledge item. Third, hierarchical clustering is used find important topics. Additionally, high-dimensional relationships among objects are transformed into a 2-dimensional space using the multi-dimension scaling method. The optimal transformation coordinate matrix is also determined by using the artificial bee colony (ABC) algorithm. Then, this transformation coordinate matrix is used to construct a two-dimensional knowledge map so that the relationship among all important topics can be visualized easily. According to experiments, it is found that setting appropriate number of clusters is important for visual perception in the knowledge map. In addition, population size and iteration number in ABC algorithm can affect the results. This paper also shows the example of using the proposed topic knowledge map for research trend analysis in IOT during years 2011 to 2016.

Keyword: Green supply chain、Game theory、dual-channel、integrated channel、

decentralized channel