

DISSERTATION INFORMATION

Title: **DEVELOPING METHODS FOR IDENTIFYING AND ANALYZING THE RISK BASED ON THE FUZZY CLUSTERING**

Major: Computer Science

Major Code: 62.48.01.01

PhD Student: Doan Huan

Advisors: Assoc. Prof. Dr Nguyen Dinh Thuan, Assoc. Prof. Dr Ho Bao Quoc

University: University of Information Technology – Vietnam National University - Ho Chi Minh City.

1. ABSTRACT

Risk management is one of the most important problems in organizations and enterprises. Solving this problem is very challenging because it depends on many factors. Two important steps of the risk management problem are risk identification and analysis. In practice, traditional methods such as surveys and interviews are often used to identify and analyze risks. These ways take a lot of time and effort, but they have not taken advantage of existing data sources. Recently, a number of studies in the world have used computer science techniques for supporting risk identification and analysis. However, most of these works only use existing techniques but have not developed or improved them to make it more suitable for the problem, so the efficiency is not high. The strength of knowledge discovery techniques of computer science is the search for information and knowledge potentially hidden in big data. In order to take advantage of that strength, this dissertation has researched, developed, and applied knowledge discovery techniques in risk identification and analysis. Firstly, this dissertation proposes a framework based on knowledge discovery techniques to make a base as a general orientation for building risk identification and analysis methods. Secondly, based on this framework, two methods of identifying and analyzing the clear and fuzzy risks are developed. Finally, the dissertation has also studied and developed the FCM-E and FCM-R algorithms to integrate into these methods in order to support building the risk identification and analysis methods in such a way as to make them is better. In particular, the FCM-R algorithm not only has the ability to automatically determine the appropriate number of clusters of the data set but also has the ability to rank the clusters obtained according to the risk level measurement, therefore

it will be a very good tool for risk analysis. Experimental implementation of the methods which are developed in the dissertation on published standard data sets and enterprise's actual data sets show that these methods are effective.

2. THE MAIN CONTRIBUTIONS OF THE DISSERTATION

Based on the research results achieved, the main contributions of the dissertation are summarized as follows:

- 1) Propose a two-stage framework based on knowledge discovery techniques to make a base as a general orientation for building the methods for identifying and analyzing latent risk.
- 2) Build a method for identifying and analyzing clear risk by using classification algorithms at the identification stage and FCM-R fuzzy clustering algorithm at the risk analysis stage.
- 3) Build a method for identifying and analyzing fuzzy risk by integrating fuzzy logic and knowledge discovery techniques, especially the FCM-R fuzzy clustering algorithm.
- 4) Develop FCM-E fuzzy clustering algorithm by building the index that automatically determines the appropriate number of clusters of the data set. After that, integrate this index into the FCM algorithm.
- 5) Develop FCM-R fuzzy clustering algorithm by building a measure of the risk level of clusters obtained after clustering process and integrating it into the FCM-E algorithm.

The results of the research and experiments of the dissertation are published in eight scientific articles. Among them, there are three international journal articles, one of them is indexed by ISI (IAJIT, IF: 0.724, SCImago: Q2), four articles in specialized international conferences which have proceedings were published by the prestigious publishers: Springer, IEEE, and one article in the specialized national conference.

Advisors

Phd Student

Assoc. Prof. Dr Nguyen Dinh Thuan, Assoc. Prof. Dr Ho Bao Quoc

Doan Huan