Award Details

Mining Interesting Useful Patterns

Research Details			
Competition Year:	2012	**Fiscal Year:**	2015-2016
Project Lead Name:	Leung, CarsonKaiSang	**Institution:**	University of Manitoba
Department:	Computer Science	**Province:**	Manitoba
Award Amount:	22,000	**Installment:**	4 - 5
Program:	Discovery Grants Program - Individual	**Selection Committee:**	Computer Science
Research Subject:	Database management	**Area of Application:**	Information systems and technology
Co-Researchers:	No Co-Researcher	**Partners:**	No Partners

Award Summary

Frequent pattern mining is an important data mining task that finds sets of frequently co-occurring items. Many existing algorithms find frequent patterns from precise data (e.g., supermarket transactions), in which the contents of datasets are precisely known. However, there are real-life situations in which data are imprecise or uncertain (e.g., sensor data, medical test results) due to factors like inherited measurement inaccuracies or sampling frequency. Despite their uncertainty, these data contain a rich set of useful knowledge. Over the past few years, I have developed algorithms that use probabilistic approaches to find frequent patterns from uncertain data, in which items in each transaction are assumed to be independent. However, this assumption may not hold in many real-life situations. Hence, I propose a research program with an objective to build an exploratory, efficient, user-friendly, and powerful mining framework--which consists of systems that mine useful patterns that are interesting to users from data streams and/or uncertain data. Specifically, I plan to (i) explore non-probabilistic approaches in finding frequent patterns, (ii) relax the above assumption so as to handle more realistic situations where items in each uncertain transaction may be related, (iii) incorporate user preferences in the mining process so as to allow users to find other useful (frequent or infrequent) patterns that are interesting to users, (iv) further improve performance so as to provide users with real-time responses, (v) develop visual analytics tools so as to enable users to visualize and analyze static (or dynamic) datasets of precise (or uncertain) data. Consequently, for this proposed research program, I and my HQP would develop new data mining technology for mining interesting useful patterns. This, in turn, advances knowledge of researchers in the field. Moreover, I also plan to apply the proposed system to various real-life applications (e.g., mining Web data, telecommunication data, agro-meteorological data, and tweets from social networks) so as to demonstrate the effectiveness of the proposed systems in addressing scientific/business needs of the application users when mining interesting useful patterns.