

M.Sc. Thesis Defense
Frequent Pattern Mining of
Uncertain Data Streams
by
Fan Jiang

Date: March 22, 2012 at 9 am
Place: E2-229 EITC

Abstract

In the data mining field, data can be categorized into two types: precise data and uncertain data. When dealing with uncertain data, users may not be certain about the presence of an item in the database. For example, due to inherent instrumental imprecision or errors, data collected by sensors are usually uncertain. In various real-life applications, uncertain databases are not necessarily static, new data may come continuously and at a rapid rate. For instance, uncertain data can come in batches, which forms a data stream. To discover useful knowledge in the form of frequent patterns from streams of uncertain data, algorithms have been developed to use the sliding window model for processing and mining data streams. However, for some applications, other stream processing models such as the landmark window model and the time-fading model are more appropriate. In this M.Sc. thesis, I propose algorithms for uncertain data streams frequent pattern mining. Specifically, I develop tree-based algorithms that use the landmark window model or the time-fading model to mine frequent patterns from streams of uncertain data.