

# Mining Evolving Data Streams for Real-time Monitoring Applications

Philip S. Yu  
University of Illinois at Chicago  
psyu@cs.uic.edu

September 29, 2009

## **Abstract**

With the advance of data gathering and communication technologies, it becomes increasingly possible to support real-time monitoring of large amount of information from diverse information sources. This can have wide spread applications. For example, to reduce the carbon footprints of buildings, vehicles, and airplanes would require continuous monitoring of sensors and detecting deviation from desired behavior. In this kind of applications, data is viewed as a continuous stream. Problems such as data mining which have been widely studied for traditional data sets cannot be easily solved for the data stream domain. This is because the large volume of data arriving in a stream renders most algorithms to inefficient as most mining algorithms require multiple scans of data which is unrealistic for stream data. More importantly, the characteristics of the data stream can change over time and the evolving pattern needs to be captured. In this talk, I'll provide an overview, discuss the issues and focus on how to mine evolving data streams for real-time monitoring applications.