

Computer Science Seminar Series

National Capital Region

Grammar-Based Time Series Pattern Mining and Visualization

Speaker: Prof. Jessica Lin
George Mason University
Friday, April 10, 2015
1:00PM- 2:00PM, NVC 325

Abstract

Massive amounts of data are generated daily at a rapid rate. As a result, the world is faced with unprecedented challenges and opportunities on managing the ever-growing data, and much of the world's supply of data is in the form of time series. Time series data mining has thus attracted an enormous amount of attention in the past two decades. This talk will focus on the discovery of novel and non-trivial patterns in time series data, including frequently encountered patterns (motifs) and rare patterns (anomalies). The ability to efficiently detect frequent and anomalous patterns in time series allows for the exploration, summation, and compression of data. In addition, such information is crucial to a variety of application domains where these patterns convey critical and actionable information. In recent work, we demonstrate that grammar induction, the process of learning rules of a formal language from a set of observations, allows the discovery of hierarchical structures and regularities from input time series. We proposed several algorithms based on grammar for efficient discovery of co-existing variable-length approximate motifs and anomalies without any prior knowledge about their length, shape, or minimal occurrence frequency. We present GrammarViz, an interactive tool for grammar-driven mining and visualization of variable-length time series patterns.

Biography



Dr. Jessica Lin is an Associate Professor in the Department of Computer Science at George Mason University. She received her PhD from University of California, Riverside in 2005. Her research interests encompass broad areas of data mining and machine learning, on large-scale multimedia datasets including time series, spatiotemporal databases, text, and images. Dr. Lin teaches advanced topics on data mining at GMU, concentrating on mining multimedia and high-dimensional data. Her research work is partially funded by the National Science Foundation, U.S. Army, Naval Research Lab, and Intel Corporation.