Fast Algorithms for Mining Co-evolving Time Series

Lei Li

Monday, September 12, 2011
1:00 p.m.
Gates 8102

Time series data arise in numerous applications, such as motion capture, computer network monitoring, data center monitoring, environmental monitoring and many more. Finding patterns and learning features in such collections of sequences are crucial to solve real-world, domain specific problems, for example, to build humanoid robots, to detect pollution in drinking water, and to identify intrusion in computer networks.

In this talk, we focus on fast algorithms on mining co-evolving time series, with or without missing values. We will present a series of our effort in analyzing those data: (a) time series mining and summarization with missing values, and (b) learning features from multiple sequences. Algorithms proposed in the first work allow us to obtain meaningful patterns effectively and efficiently. Thus they enable vital mining tasks including forecast, compression, and segmentation for co-evolving time series, even with missing values. We also propose "PLiF" and Complex Linear Dynamical System (CLDS), novel algorithms to extract features from multiple sequences. Such features will serve as a corner stone of many applications for time series such clustering and similarity search. Our algorithms scale linearly with respect to the length of sequences, and outperform the competitors often by large factors. In addition, we will briefly mention several other time series mining problems and algorithms, including natural motion stitching, bone constrained occlusion filling, a parallelization of our algorithms for multi-core systems, and an forecasting algorithm for thermal conditions in data centers.

Thesis Committee:
Christos Faloutsos, Chair
Nancy Pollard
Eric Xing
Jiawei Han, University of Illinois at Urbana-Champaign