DISSERTATION
DEFENSE

Preserving Individual Privacy in Spatio-Temporal Data Analytics

Liyue Fan
Emory University

Abstract: We live in the age of big data. With an increasing number of people, devices, and sensors connected with digital networks, individual data now can be largely collected and analyzed by data mining applications for social good as well as for commercial interests. However, the data generated by individual users exhibit unique behavioral patterns and sensitive information, and therefore must be transformed prior to the release for analysis. The AOL search log release in 2006 is an example of privacy catastrophe, where the searches of an innocent citizen were quickly re-identified by a newspaper journalist. In this talk, I will present a novel framework to release continuous aggregation of private data for an important class of real-time data mining tasks, such as disease outbreak detection and web mining, to name a few. The key innovation is that the proposed framework captures the underlying dynamics of the continual aggregate statistics with time series state-space models, and simultaneously adopts filtering techniques to correct the observed, noisy data. I will show that the new framework provides a rigorous, provable privacy guarantee to individual data contributors without compromising the output analysis results. Extensive empirical studies confirm that it will enable privacy-preserving data analytical tasks in a broad range of application domains.

Friday, June 27, 2014, 2:00 pm
Mathematics and Science Center: W301

Advisor: Li Xiong

MATHEMATICS AND COMPUTER SCIENCE
EMORY UNIVERSITY