

Monday, July 23, 10:10

Room: C3 - Main Auditorium

PLENARY SPEAKER

BIG DATA, GOOGLE AND DISEASE DETECTION: A STATISTICAL ADVENTURE

SAMUEL KOU

kou@stat.harvard.edu

Department of Statistics, Harvard University, USA

Keywords: influenza, dengue fever, Internet search, forecasting.

ABSTRACT

Big data collected from the Internet have generated significant interest in not only the academic community but also industry and government agencies. They bring great potential in tracking and predicting massive social activities. We focus on tracking disease epidemics in this talk. We will discuss the applications, in particular, Google Flu Trends, some of the fallacy and the statistical implications. We will propose a new model that utilizes publicly available online data to estimate disease epidemics. Our model outperforms all previous real-time tracking models for influenza epidemics at the national level of the US. An extended version of the model gives accurate tracking of Dengue fever in Asian and South American countries. We will also draw some lessons for big data applications.