You are cordially invited to a talk in the Edmond J. Safra Center for Bioinformatics Distinguished Speaker Series.
The speaker is Prof. Roni Rosenfeld, School of Computer Science, Carnegie Mellon University.

**Title:** "Forecasting Epidemics"

**Time**: Thursday, 6 November 2014, at 13:00 sharp (refreshments from 12:50)

**Place:** Schreiber building, room 008

**Host:** Prof. Saharon Rosset, saharon@post.tau.ac.il, Department of Statistics and Operations Research, School of Mathematical Sciences, Tel Aviv University

**Abstract:** Epidemics of infectious diseases such as Influenza and Dengue cause consistent, considerable and widespread morbidity, mortality and economic burden. With access to accurate and reliable forecasts of a current or upcoming epidemic, policy makers can design and implement more effective countermeasures, and the general public can make more informed decisions. Our long term vision is to make epidemiological forecasting as universally accepted and useful as weather forecasting is today.

This past year, the U.S. CDC hosted the 'Influenza Challenge', to predict key epidemiological measures for the 2013/2014 US influenza season with the help of digital surveillance data. We developed a framework for in-season forecasts of epidemics using a semi-parametric Empirical Bayes methodology, and applied it to predict the weekly percentage of outpatient doctors visits for influenza-like illness, as well as the season onset, duration, peak time, and peak intensity, with and without using data from Google Flu Trends. Previous work on epidemic modeling used mechanistic models of disease behavior or applied time series tools to explain historical data. However, these models do not always capture the range of possible epidemic behaviors. Our approach instead uses importance sampling to produce posterior trajectories for the epidemic curve using modified versions of previous seasons, allowing for reasonable variations in timing, pace, and intensity of epidemics, and observational noise. We report prospective forecasts and compare retrospective cross-validated prediction error on historical data to a variety of simpler baseline methods.

Since the framework does not make domain-specific assumptions, it can easily be applied to other diseases. In June-July 2014, the 2014 FIFA Soccer World Cup saw some 600,000 foreign visitors descending on Brazil, a country with the world's highest number of annual dengue transmissions. This raised great concern about the risk of returning travelers seeding epidemics in other countries. In Late April we applied a slightly modified version of our method to weekly dengue surveillance data from previous years (2001–2013) and the first 19 weeks in 2014 to estimate the risk of dengue during the World Cup in each of the 12 game cities, and found that the expected incidence rates were relatively low in most locations with a few notable exceptions.

Finally, in August 2014 DARPA announced a $500,000 competition to predict the spread of Chikungunya in the Western Hemisphere in the coming 6 months. This time, we fielded a 'committee

of experts' forecasting system, which we expanded into a 'wisdom of crowds' system for the upcoming flu season, with results still pending.

Joint work with Ryan Tibshirani, Rob Tibshirani, Logan Brooks, David Farrow, and Sangwon Hyun.

Roni Rosenfeld (TAU, Math and Physics, 1985) is Professor of Machine Learning, Language Technologies, Computer Science and Computational Biology in the School of Computer Science at Carnegie Mellon University. His current research interests are in epidemic forecasting, and the use of speech and language technologies to aid international development. He has also performed research in statistical language modeling, machine learning, speech recognition and viral evolution. He has published over 100 scientific articles in academic journals and conferences.
http://www.cs.cmu.edu/~roni.