

## BIOEN 509 – DEPARTMENTAL SEMINAR SERIES

Thursday, Jan. 12<sup>th</sup> 2012, 12:30-1:20 PM

Foege Bioengineering Building N130

# Integrative Analysis and Interactive Exploration of Data from the Cancer Genome Atlas

*Prof. Ilya Shmulevich*

*Professor at The Institute for Systems Biology & Affiliate Prof of UW BioE & EE*

The data collected within the TCGA project is exceptionally heterogeneous. Molecular profiling data generated by different measurement modalities as well as clinical information collected on each patient give rise to continuous, discrete, and categorical data with different distributional properties. Additional data is generated by a variety of analyses carried out on the individual data sets. Examples include functional or structural annotations of mutations, assignment of an expression subtype of the tumor, and enrichment or activity of molecular pathways, for each patient sample. The data also include missing values as well as interdependencies among the features that undoubtedly extend beyond pairwise correlations. I will describe our efforts towards identifying strong multivariate associations in the TCGA data using a framework based on random forest regression as well as development of web-based tools to interactively explore such associations.

*Dr. Ilya Shmulevich received his Ph.D. in Electrical and Computer Engineering from Purdue University, West Lafayette, IN, in 1997. From 1997-1998, he was a postdoctoral researcher at the Nijmegen Institute for Cognition and Information at the University of Nijmegen and National Research Institute for Mathematics and Computer Science at the University of Amsterdam in The Netherlands, where he studied computational models of music perception and recognition. In 1998-2000, he worked as a senior researcher at the Tampere International Center for Signal*



*Processing at the Signal Processing Laboratory in Tampere University of Technology, Tampere, Finland. From 2001-2005, he was an Assistant Professor at the Cancer Genomics Laboratory in the Department of Pathology at The University of Texas M. D. Anderson Cancer Center and an Adjunct Professor in the Department of Statistics in Rice University. Presently, he is a Professor at The Institute for Systems Biology, where he directs a Genome Data Analysis Center that is part of The Cancer Genome Atlas (TCGA) project. He is an Affiliate Professor in the Departments of Bioengineering and Electrical Engineering at the University of Washington, Department of Signal Processing in Tampere University of Technology, Finland, and Department of Electronic and Electrical Engineering in Strathclyde University, Glasgow, UK. He is an Associate Editor of EURASIP Journal on Bioinformatics and Systems Biology and a Senior Member of the IEEE. His research interests include systems biology, nonlinear signal and image processing, and computational learning theory.*

*For more information please visit <http://courses.washington.edu/bioetalk>*