UNIVERSITY OF RHODE ISLAND
Department of Mathematics

Applied Mathematics and
Scientific Computing Seminar

Location: Lippitt Hall 204
Time: Monday, October 24, 2016, 1:00pm
(refreshments at 12:50 p.m.)

Entropy-Scaling Search of Massive Data

by Dr. Noah Daniels

Dept. of Computer Science & Statistics, URI

Abstract: Recent technological improvements have increased the scale and content of data produced across a variety of fields, from astronomy and biology to global trade and social networks. In many cases, the scale, richness, and noise of the data have limited our computational ability to make significant discoveries. For example, the advent of shotgun sequencing in the field of genomics has led to growth of next-generation sequencing data that outpaces Moore's and Kryder's laws for computation and storage. Although computers are getting faster and cheaper, they are not doing so at a rate fast enough to keep pace with this exponential explosion of data.

I introduce a novel approach to approximate search, which scales in both time and space with the entropy of the underlying data (albeit with two distinct measures of entropy). This approach exploits the structural properties—low metric entropy and low fractal dimension—that occur in high-dimensional data sets when the data generation is constrained, for example by biological evolution.

I briefly discuss applications to bioinformatics, chemistry, and astronomy.