

**When:** Friday 12:40 – 13:40

**Where:** ETB 1020

**Speaker:** Shahin Boluki

Department of Electrical & Computer Engineering  
Texas A&M University

**Title:** Knowledge-Driven Prior Construction for Bayesian Learning

**Date:** 2-22-2019

**Abstract:** In applications where small sample sizes are ubiquitous, leveraging prior knowledge in addition to the data may be helpful. Small samples are commonplace in phenotypic classification based on genomic data and, for these, use of prior knowledge is critical. Bayesian approaches that treat uncertainty directly on the feature-label distribution provide a candidate framework for incorporating prior knowledge through specifying the corresponding prior distributions. For example, if knowledge regarding the feature-label distribution in the form of genetic pathways or molecular interaction networks is available, then it can potentially be used in learning. However, in many applications including genomics, the given prior knowledge is uncertain, incomplete, and may be inconsistent. The question is how we can effectively incorporate the state of prior knowledge in the modeling and learning process. I present a prior construction methodology based on a general framework of constraints in the form of conditional probability statements. The formal structure involves a constrained optimization in which the constraints incorporate existing (scientific) prior knowledge, while at the same time avoiding inadvertent over restriction of the prior, an important consideration with small samples. This new framework is rather flexible and can naturally handle the potential inconsistency in archived regulatory relationships and conditioning can be augmented by other knowledge, such as population statistics. I also present the extension to mixture models, which are applicable in cases with inter- and intra-heterogeneity in samples.

**Bio:** Shahin Boluki is pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. His current research interests include Bayesian learning, computational biology, and uncertainty quantification in complex systems.