# Auburn University Montgomery

# Department of Mathematics & Computer Science

# Mathematics and Computer Science Colloquium

Time: Friday, 11/04/2016, Noon-1:00PM
Place: Goodwyn Hall 204 (AUM)
Speaker: Dr. Imren Dinc (Troy University Montgomery)
Title: Associative Data Analytics and its Application to Protein Crystallization Analysis

Data analytics focuses on processing raw data to extract important information and involves 3 main stages: data acquisition, data analysis, and visualization. In data acquisition, an efficient experimental design methodology is necessary to maximize the amount of information obtained. It is quite difficult to employ supervision in this domain due to limited data size and skewed data distribution. Once the data is collected, proper analysis methods are necessary to make valuable decisions regarding the type of data. Finally, the decisions and the data should be visualized effectively to increase the ease of understanding.

This research presents a data analytics framework called ``*Associative Data Analytics* (*ADA*)'', which targets to improve the reliability and efficiency of data analytics applications. We evaluated the performance of the *ADA* on the protein crystallization problem. *Protein crystallization* (*PC*) is the process of growing a protein crystal in a solution to visualize the protein at the atomic level. The process begins with setting up experiments in a wet-lab, then continues using spatio-temporal analysis of the solution using a microscope system, and ends with optimizations.



Our results show that the *ADA* yielded successful results and made a significant impact on this domain. Under the *ADA* framework, the *AED* generated novel conditions yielding protein crystals that are difficult to crystallize. We analyzed the crystal regions in the protein images using our novel image thresholding method, Super-thresholding. Super-thresholding improved accuracy around 10% compared to the other best single thresholding method. Final results are displayed using our visualization tool, Visual-*X2*.
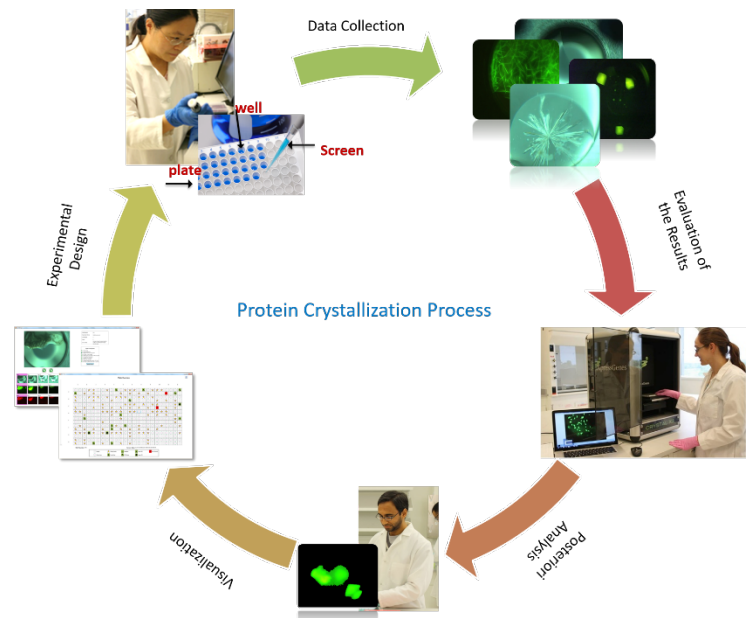
**About Imren Dinc:** Imren Dinc is an adjunct assistant professor at computer science department in Troy University. She received her Ph.D. degree at University of Alabama in Huntsville in 2016. She received her B.Sc. degree from computer engineering at Dokuz Eylul University, Turkey in 2012 with first rank in the department and an honor degree. She has been working on experimental design & analysis for protein crystallization domain and visualization of these experimental results. Her research areas are data mining, bioinformatics, image processing, machine learning, and biochemistry.