

COMPUTER SCIENCE  
SEMINAR

*Reading Between the Lines of Datacenter Logs*

Dr. Nosayba El-Sayed  
MIT

**Abstract:** Designing datacenters that are reliable, energy-efficient, and capable of delivering high performance and high utilization is a nontrivial problem facing scientists, businesses, and governments alike. In this talk, I will demonstrate how analyzing large datasets from different organizations helped us uncover interesting (and often surprising) patterns in the behavior of systems and applications in these large-scale platforms. I will show how real-world data helped us tackle critical questions such as how does temperature impact server reliability in places like Google, or how well do users configure the computing jobs they submit to shared clusters (spoiler alert: not very well!). Finally, I will demonstrate how simple machine learning techniques can be leveraged to accurately predict job failures in datacenters, while using data that is easily collected in current platforms.

Tuesday, February 20, 2018, 1:00 pm  
Atwood Chemistry Building Room 240

MATHEMATICS AND COMPUTER SCIENCE  
EMORY UNIVERSITY