Striking the balance between cyber security and convenience has been a long-standing challenge: while a strict security policy prevents intrusion, it may severely disrupt critical services of an organization through automated blocking. Traditionally, static whitelists are manually maintained to capture such critical services which should not be interfered with, or which are trusted out of necessity. However, it is extremely difficult to be accurate and exhaustive as in reality they change over time. In this talk we discuss a machine learning based approach that aims to model the relationship between the target organization and the external resources, and dynamically suggest changes to the whitelists. We extract features from network flow summaries, bipartite graph analysis, and the contents of web pages crawled from the hostname of the resources. The resulting model could also be used as part of intrusion alert prioritization, response selection, as well as exfiltration discovery.

Dr. Harris Lin is a postdoctoral researcher in Ames Laboratory, operated for U.S. Department of Energy by Iowa State University. He is part of the Cyber Security research program, currently working on the research, design, and development of a dynamic impact modeling tool called CyDIME which aims to complement existing network intrusion detection and response systems. He obtained his Ph. D. from Iowa State University where he focused on machine learning for distributed and semantically rich data also known as Linked Open Data. His current research interests are to apply large-scale and/or distributed machine learning methods to better analyze and understand cyber activity data, with applications to intrusion detection, anomaly detection, interactive visualization, and knowledge discovery in general.