Big Data in the Era of the Internet of Things

Participant: Yaman Sharaf-Dabbagh (PhD student)

The Internet of things (IoT) is a promising technology in which physical objects will be endowed with cyber and computation capabilities via smart sensors, RFID tags, smartphones, and wearable devices. The IoT environment consists of a large number of connected devices with heterogeneous capabilities. For instance, Cisco Internet Business Solutions Group (IBSG) estimates that the IoT will consist of almost 50 billion objects by 2020. This large scale and heterogeneous mix of objects raise challenging concerns in the fields of data analytics, security, and privacy.

In this research, we focus on developing novel data-oriented optimization, security, and privacy techniques tailored to the unique nature cyber-physical nature of the IoT. For example, conventional IoT security solutions typically focus on cryptographic constructs such as secure networking protocols and key exchange mechanisms. However, these existing solutions may be difficult to implement on the computationally limited and portable IoT objects. The first stage in this research is to develop novel machine learning tools that enable the IoT to dynamically identify, classify, and authenticate devices based on their cyber-physical environment and with limited available prior data. The proposed framework take advantage of device fingerprinting techniques to identify each IoT object based on unique cyber and physical features. For example, the figure below shows an insider attack scenario on an IoT system. Fingerprinting based techniques are particularly suitable for the IoT due to the low overhead compared to sophisticated cryptographic techniques. The second stage of our work is to validate the data coming from IoT objects to ensure that those objects are behaving as they should be. To achieve this, we extract relationships between objects, and then use these relationships to identify any odd behavior of an object. Anomalous behavior in an object reflects either an attack or a malfunctioning object. Transfer learning and deep learning tools are used to extract relationships and analyze behavior across time. Our work received the **best paper award** in the proceedings of 26th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Hong Kong, September 2015.