A practical approach for applying Machine Learning in the detection and classification of network devices used in building management

Maroun Touma[1], Shalisha Witherspoon[1], Shonda Witherspoon[1], and Isabelle Crawford-Eng[2]

[1]IBM
[2]University of Pennsylvania

December 2, 2020

Abstract

With the increasing deployment of smart buildings and infrastructure, Supervisory Control and Data Acquisition (SCADA) devices and the underlying IT network have become essential elements for the proper operations of these highly complex systems. Of course, with the increase in automation and the proliferation of SCADA devices, a corresponding increase in surface area of attack on critical infrastructure has increased. Understanding device behaviors in terms of known and understood or potentially qualified activities versus unknown and potentially nefarious activities in near-real time is a key component of any security solution. In this paper, we investigate the challenges with building robust machine learning models to identify unknowns purely from network traffic both inside and outside firewalls, starting with missing or inconsistent labels across sites, feature engineering and learning, temporal dependencies and analysis, and training data quality (including small sample sizes) for both shallow and deep learning methods. To demonstrate these challenges and the capabilities we have developed, we focus on Building Automation and Control networks (BACnet) from a private commercial building system. Our results show that "Model Zoo" built from binary classifiers based on each device or behavior combined with an ensemble classifier integrating information from all classifiers provides a reliable methodology to identify unknown devices as well as determining specific known devices when the device type is in the training set. The capability of the Model Zoo framework is shown to be directly linked to feature engineering and learning, and the dependency of the feature selection varies depending on both the binary and ensemble classifiers as well.

Hosted file

`MADI_AppliedAI_Letters.pdf` available at https://authorea.com/users/380704/articles/496579-a-practical-approach-for-applying-machine-learning-in-the-detection-and-classification-of-network-devices-used-in-building-management