Automatic Generation of High Quality Indicators of Compromise

**Indicators of Compromise** or IOCs are a set of expressions that can be used by security products to describe and detect specific malware to help block and mitigate their effects. IOCs are usually delivered in a standardized format such as Yara rules or TAXII/STIX. The figure below shows an example IOC. This project will involve automating the generation of IOCs using big data sets and machine learning. In a recent meeting with the Chief Technology Officer of SWIFT, he mentioned that large banks have well-established sharing networks, but small banks would be big beneficiaries of better quality generation and sharing of IOCs.

![IOC Diagram](image)

**Figure: An example Indicator of Compromise.**

**Project Idea**

This main goal of this project would be to develop a prototype system on AWS that automatically generates IOCs. This would involve 1) integrating an existing Yara rule generator into our platform as a micro-service using an AWS Lambda, 2) investigate decision trees and rule induction as a method of automatic IOC generation, 3) displaying IOCs as output using a web-based interface (static webpage, only HTML+JS). If time permits, the group can investigate feeding these IOCs to an Endpoint we will provide to facilitate sharing of the IOCs. All data and code to serve as a starting point will be provided.