Using Machine Learning to solve network security problem

Wenqing Yan, Wenjun Xiong, Hao Chen, Yibei Li, Boris Petkovic

Group 5

An Introduction to the Problem

Network security problems:

1. Increased attack surface, e.g. Bluetooth
2. Increased attack types, e.g. network intruders and criminals

Network security solutions and their characteristics:

1. Signatures, Packet-filtering firewalls - Recognize known threats, Difficult to implement, Cannot prevent application-layer attacks
2. Heuristic scanning, Sandbox protection - Recognize malicious indicators, Rely on known indicators
3. Machine Learning - Robust

Goal: Prediction, Prevention, Detection, Response, Monitoring

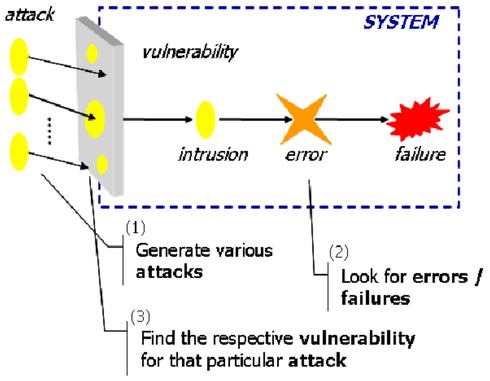

Figure 1: Intrusion Attack

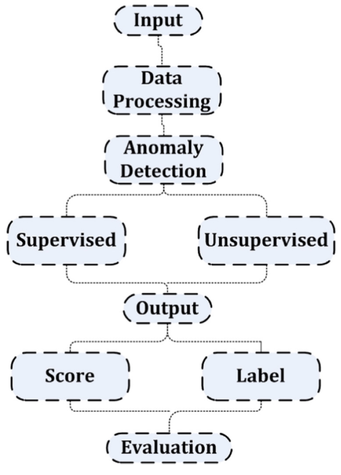

Figure 2: Framework for Network Detection Anomaly

Network Anomaly Detection Methods

Classification of Operation Modes:

1. Supervised classification - Pre-labeled data, Task: train predictive models for classification
2. Semi-supervised recognition - Only model normality
3. Unsupervised clustering - Does not require training data, Build statistic model for data

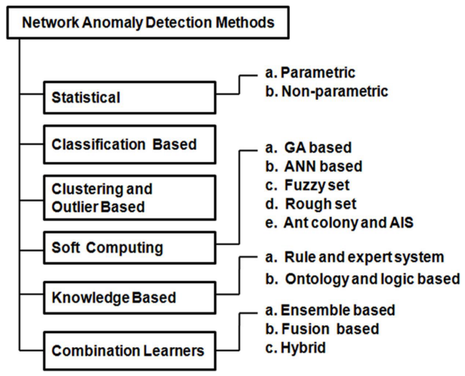

Figure 3: Existing methods

Statistical methods:

1. Method: processes the data as a static distribution
2. Outliers: the most remote points
3. Task: estimate statistic distribution (pdf)

Classification-based methods

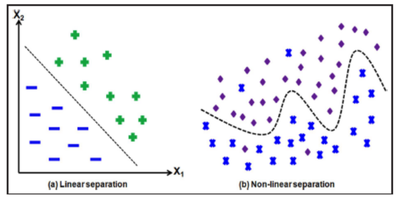

Figure 4: Classification Methods

Clustering and Outlier-based methods

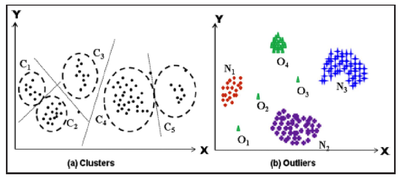

Figure 5: Clustering Methods

Study Case

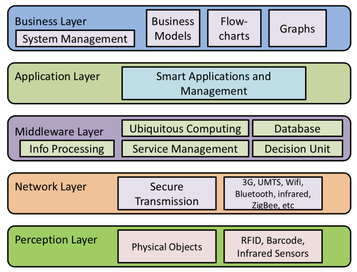

Figure 3: The IoT Architecture.

Figure 6: Different Attacks in IoTs

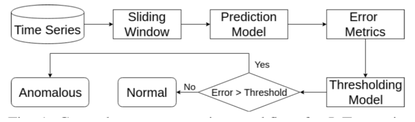

Figure 7: General System Workflow

$$\tau(data) = \begin{cases} \text{anomalous} & \text{if } E > \text{ treshold} \\ \text{normal} & \text{if } E < \text{ treshold} \end{cases}$$

Algorithm 1 The proposed AOT algorithm.

Input: Prediction error
Output: Detected anomalies
Initialize:
Anomalies, expanding mean, threshold as EmptyArray, $\mu = 0$, $\sigma = 0$
for i = 1 to Error.length **do**
  if $i < 100$ and Error[i] > threshold **then**
    expanding mean[i] = μ
    threshold[i] = $\mu + (\lambda \times \sigma)$
    Anomalies[i] = 1
    Update λ with small forgetting coefficients
  else
    μ, σ = Compute running mean, standard deviation.
    Update λ with large forgetting coefficients
  end if
end for
return Anomalies

Figure 8: Tresholding Method

References

1. Monowar H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita - Network Anomaly Detection: Methods, Systems and Tools
2. Fangyu Li, Aditya Shinde, Yang Shi,Jin Ye, Xiang-Yang Li and WenZhan Song - System Statistics Learning-Based IoT Security: Feasibility and Suitability
3. Victoria J. Hodge, Jim Austin - A Survey of Outlier Detection Methodologies