Unsupervised Real-Time Network Intrusion and Anomaly Detection by Memristor Based Autoencoder

Md Shahanur Alam
Adviser: Tarek M Taha

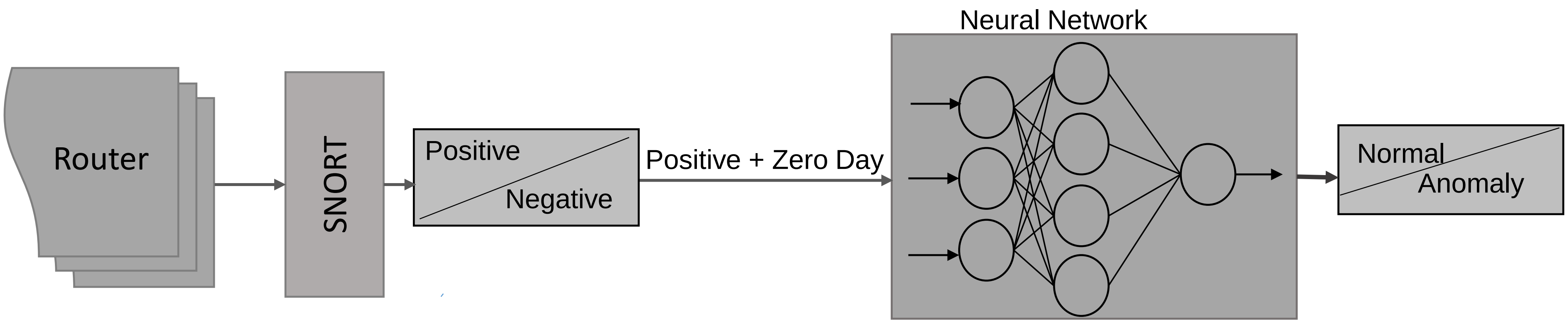
***Dept. Of Electrical and Computer Engineering, University of Dayton,* Dayton, OH, USA**

Outline

- Introduction
- Anomaly Detection Methods and Applications
- Motivation and Challenges
- Proposed Anomaly Detection System
- Results of Intrusion and Anomaly Detection System
- Summary
- Future work

Introduction

- Network Intrusion
- Intrusion Detection system
- SNORT
- What if new unknown packet comes?
  E.g. ‘Zero Day’



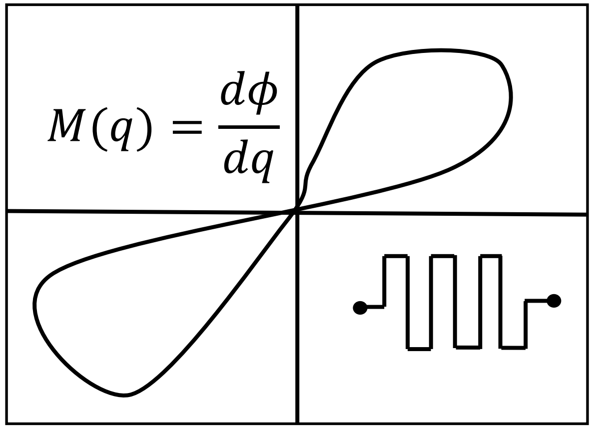
Block diagram of the neural network based intrusion detection system

Introduction (Contd.)

Deep Learning Vs Power Consumption

≈200W

Deep Learning for IoTs and Edge Devices



Memristor

- Memristive system could be a solution

Anomaly Detection Methods and Applications

What are the anomalies?

- Abnormalities/outliers

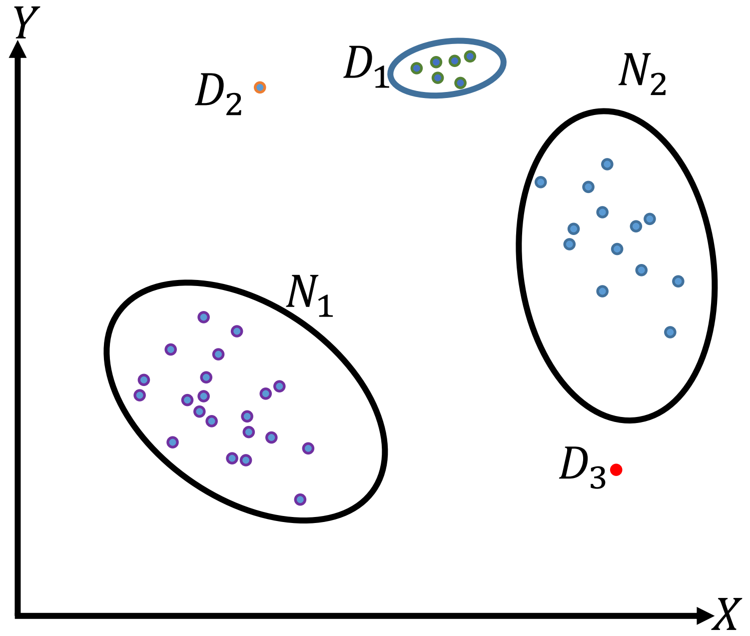


Illustration of anomalies in two-dimensional data set

Anomaly detection Methods:

- Unsupervised (AE, GAN, RNN, LSTM etc)
- Supervised (DNN, CNN)
- Hybrid model (AE+SVM)
- One-Class Neural Network

Applications:

- Cyber-Intrusion Detection
- Malware Detection
- Internet of Things (IoTs) Big Data Anomaly Detection
- Fraud Detection
- Medical Anomaly Detection
- Industrial Damage Detection

Motivation and Challenges

Motivation:

- Deep learning implementation for IoTs and edge devices
- Detection and learning of anomalies in real-time

Challenges:

- Boundary between normal and malicious is not explicitly defined
- Continual learning and the catastrophic forgetting

Our Contribution

- Design and implementation of unsupervised autoencoder in memristor crossbar devices
- Develop autoencoder training in memristor device
- Proposed an online learning system for network anomaly detection

Dataset Preprocessing

- NSL-KDD network dataset← KDD Cup'99 dataset
- Training data has125,973 packets, 23 different data types
- 43 attributes, consists numerical and alphanumeric data
- Preprocessed and sorted out the packets
- Network is pretrained with 90% of Normal
- Tested with 10% normal and 10% of total malicious data

Normal Packet

0,tcp,ftp_data,SF,491,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,0,0,
0,0,1,0,0,150,25,0.17,0.03,0.17,0,0,0,0.05,0,normal,20

Malicious Packet

0,tcp,ftp_data,SF,334,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,2,2,0,0,
0,0,1,0,0,2,20,1,0,1,0.20,0,0,0,0, warezclient,15

Preprocessed Malicious Packet

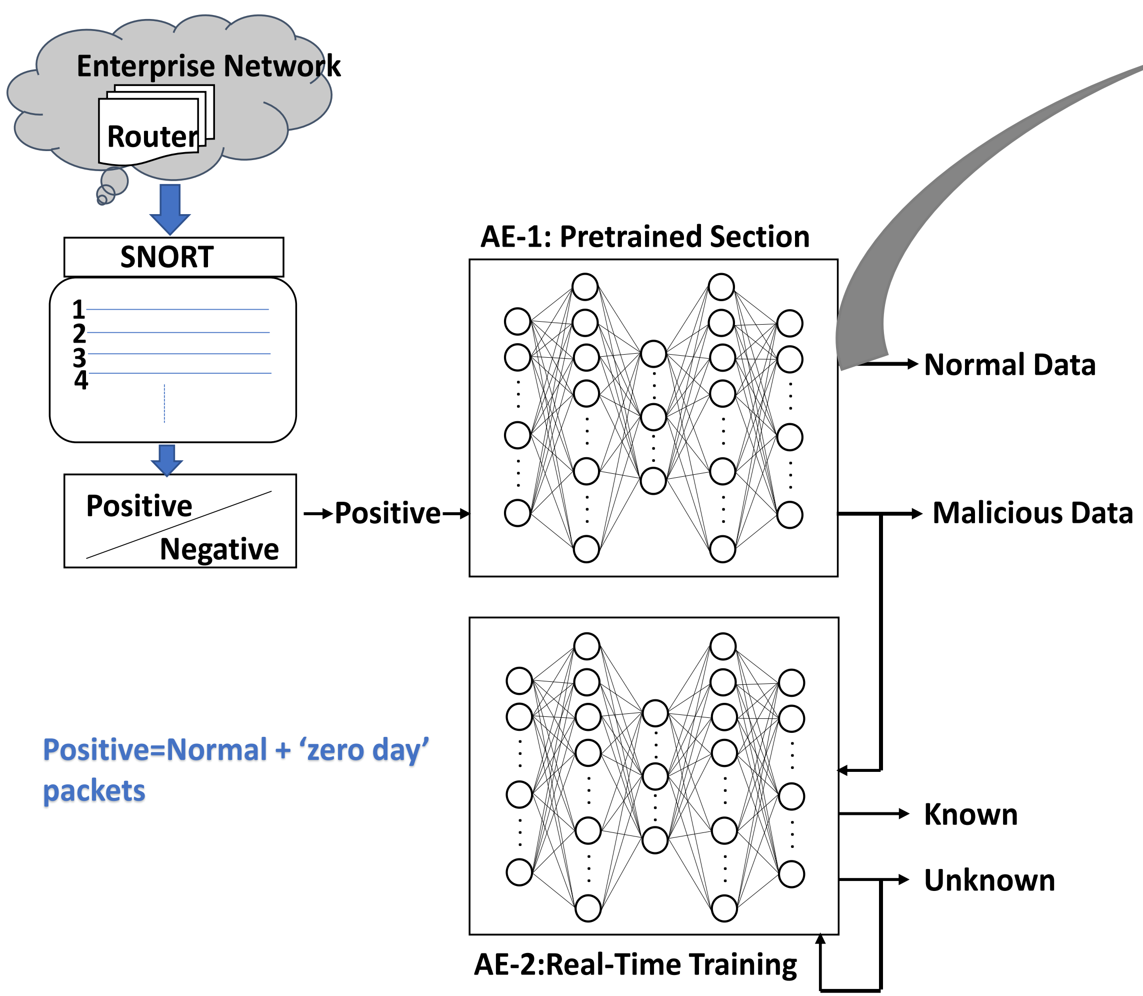
0,0.5,0.188,0.629,$3.55e^{-7}$,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0.003
91,0.00391,0,0,0,0,1,0,0,0.588,0.098,0.17,0.03,0.17,0,0,0,0.05
,0,0,0.9523

Preprocessed Malicious Packet

0,0.5,0.188,0.629,$2.42e^{-7}$,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0.003
91,0.0039,0,0,0,0,1,0,0,0.0078,0.078,1,0,1,0.2,0,0,0,0,1,0.714

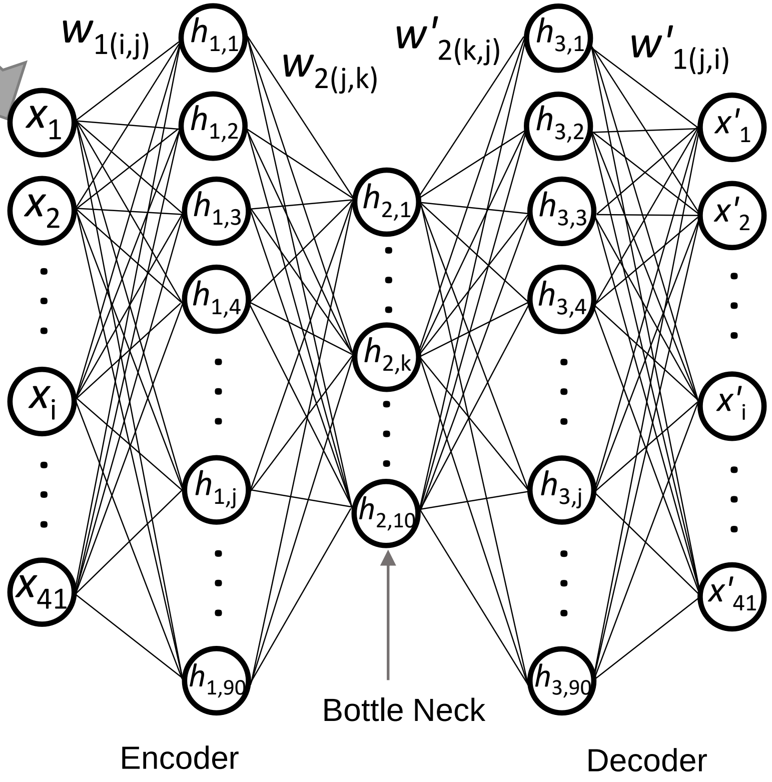
Proposed Anomaly Detection System

System Prototype Model



Intrusion And Anomaly Detection System with AE neural Network

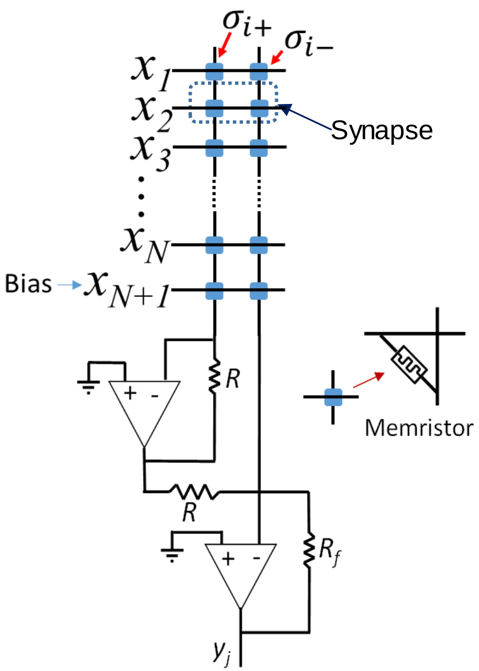
Autoencoder (AE) Neural Network



41→90→10→90→41

- AE learns to regenerate the input data at output
- AE can reduce the dimension of input data

Memristive Neural Network and Crossbar Circuit



Single Neuron

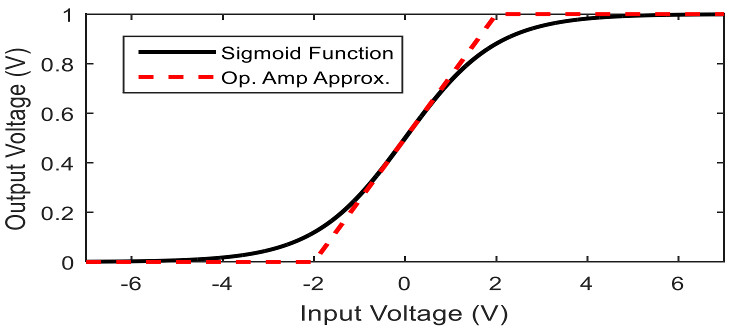
DOT Product:

$$DP_j = \sum_{i=1}^{N+1} x_i \times \left(\sigma_{ij}^{+} - \sigma_{ij}^{-}\right) \quad (1)$$

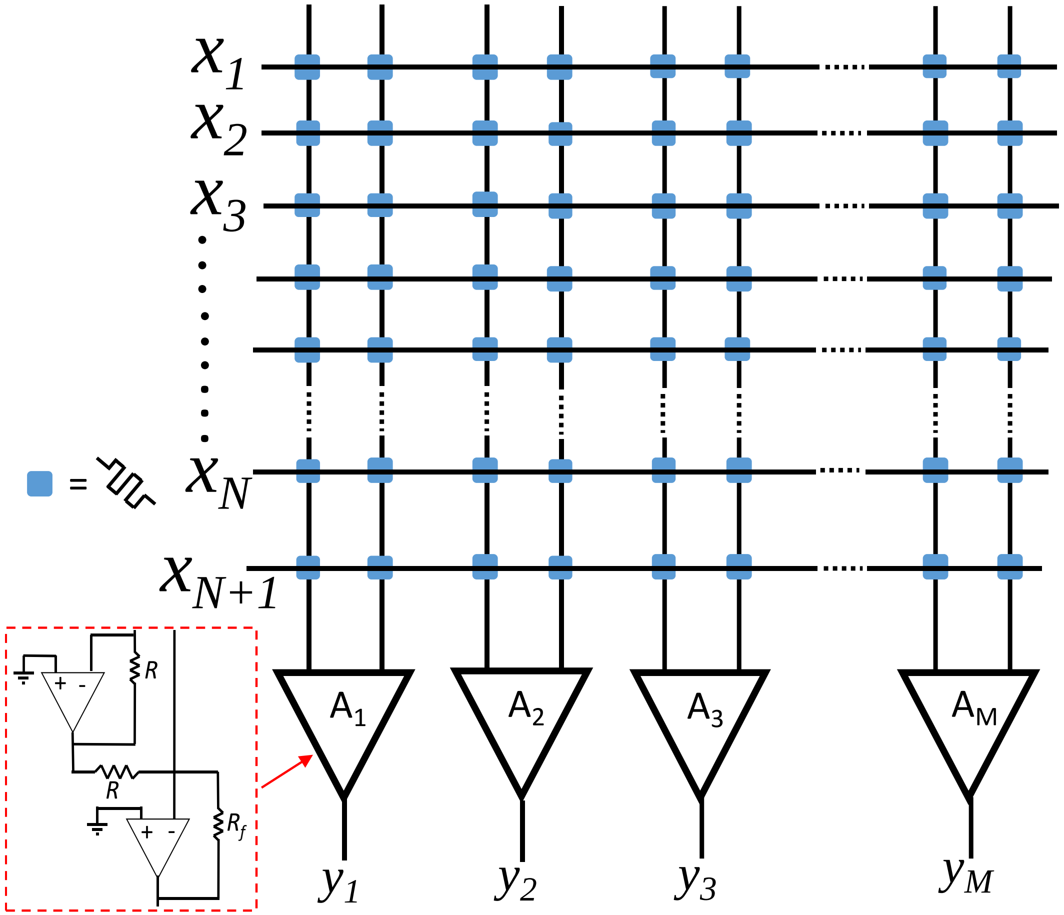
Sigmoid Approximation:

$$f(x) = \frac{1}{1+e^{-x}} \quad (2)$$

$$g(x) = \begin{cases} 1, & x > 2 \\ 0.25x + 0.5, & |x| \leq 2 \\ 0, & x < 2 \end{cases} \quad (3)$$



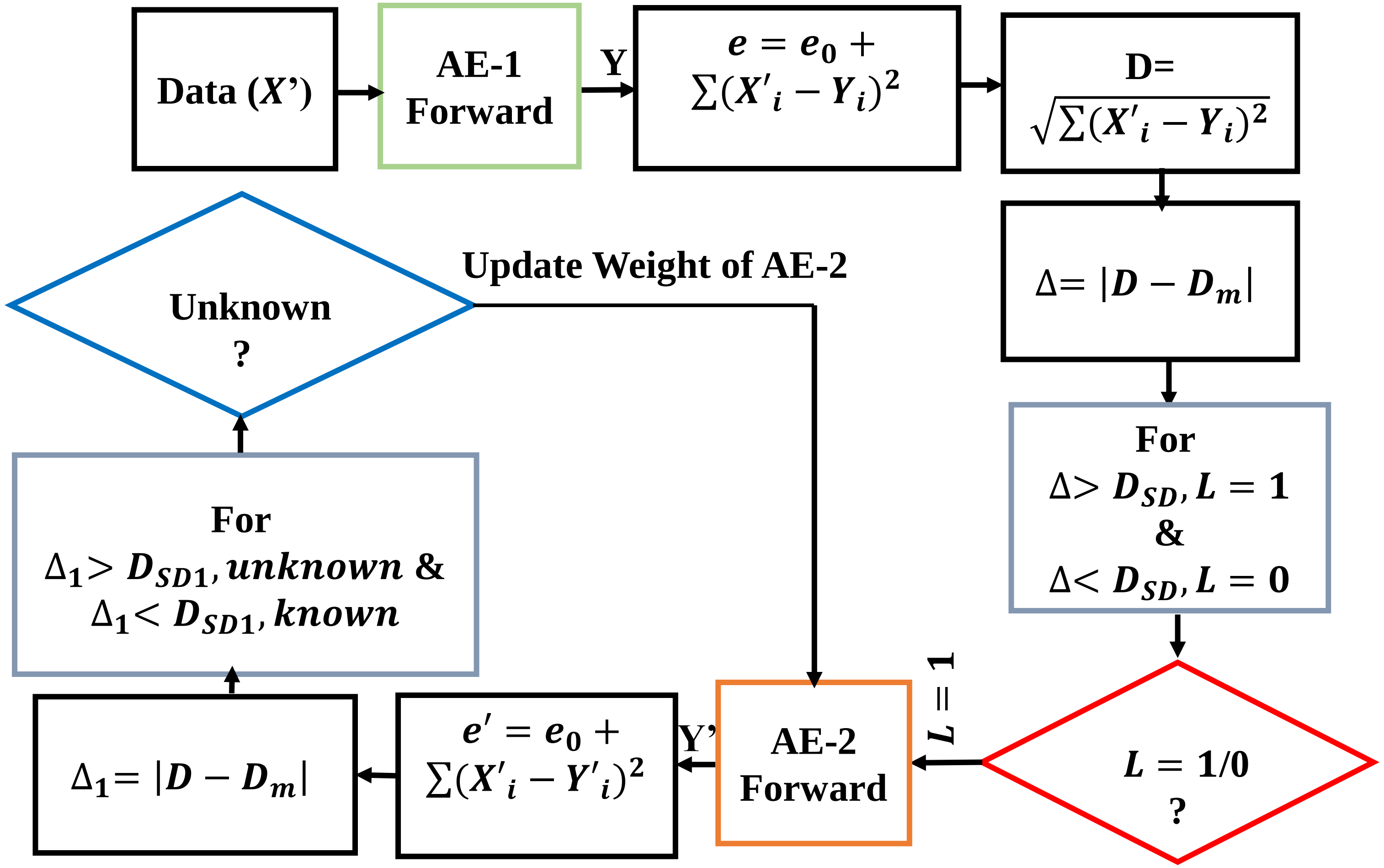
Ideal and approximate Sigmoid Function



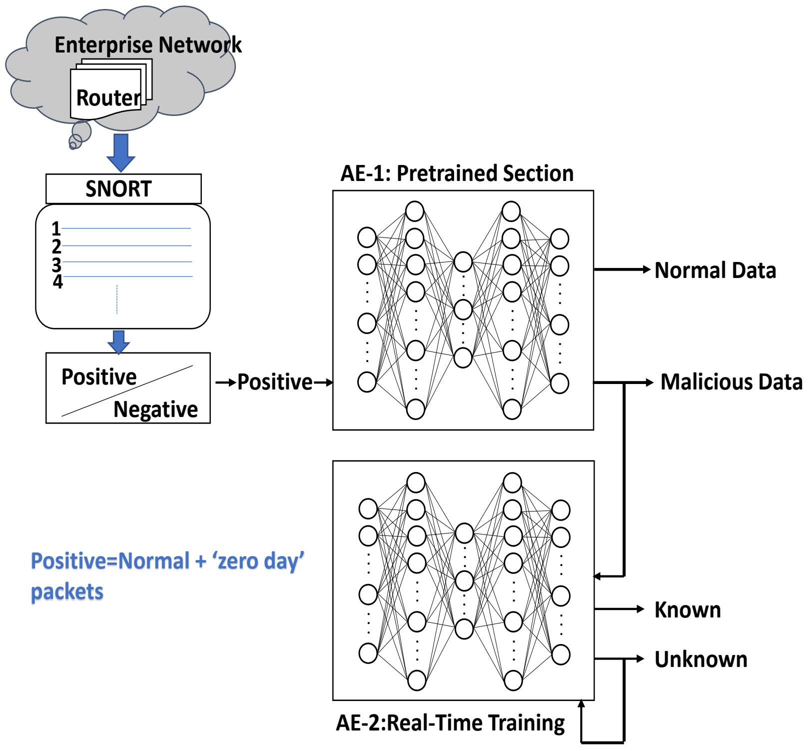
Training of the Network

- apply x_i
- crossbar computes the dot product DP_j
- output signal y_j
- error : $\delta_j = (x_i - y_j) f'(DP_j)$
- backpropagate the error $\delta_j = \sum_k \delta_k w_{k,j} f'(DP_j)$ in each hidden layer
- update the weights according δ_j as $\Delta w_j = \eta \delta_j x$
- calculate D= $\sqrt{\sum (X_i - Y_j)^2}$ and $D_{SD} = \sqrt{\frac{\sqrt{\sum (D - D_m)^2}}{N}}$

System Flowchart of Anomaly Detection System



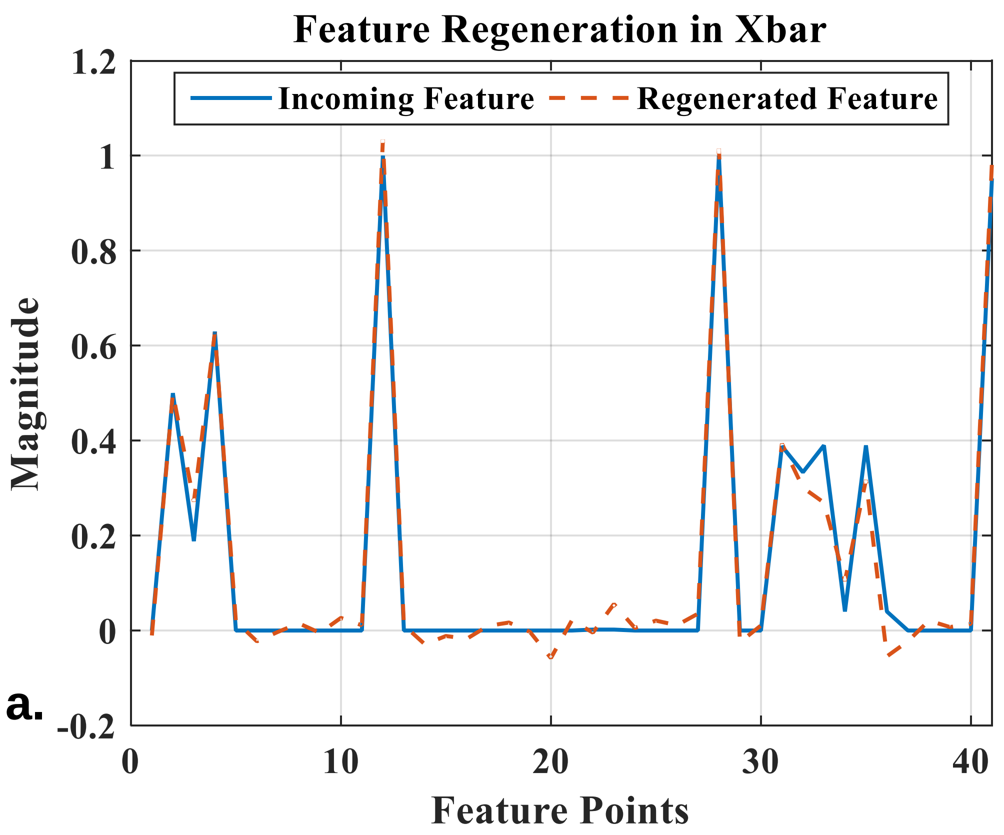
Flowchart of Real-time Anomaly detection System



Anomaly Detection System

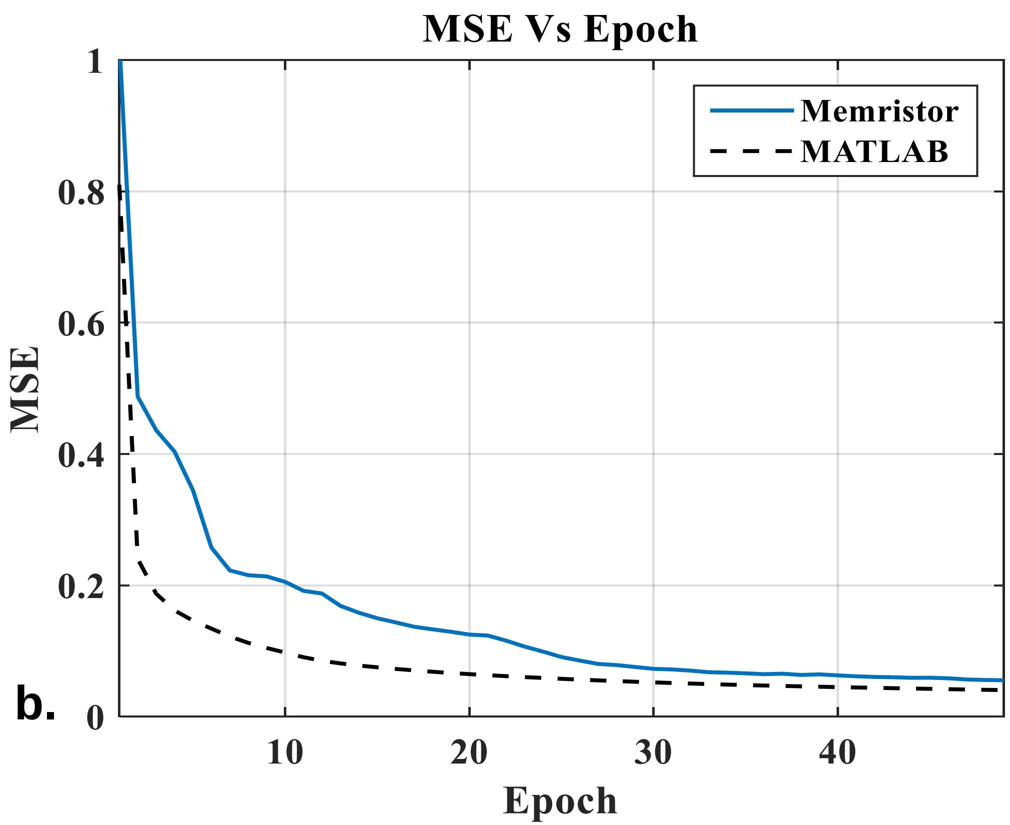
Pretraining of Autoencoder-1 (AE-1)

a.



Input feature and regenerated feature of a sample through (AE-1)

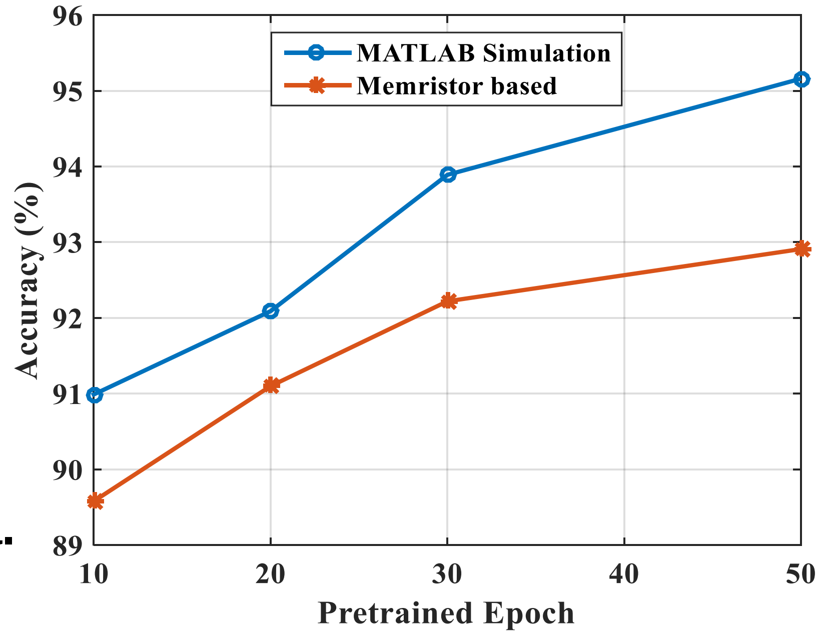
b.



Training Error (MSE) in software and memristor X-bar

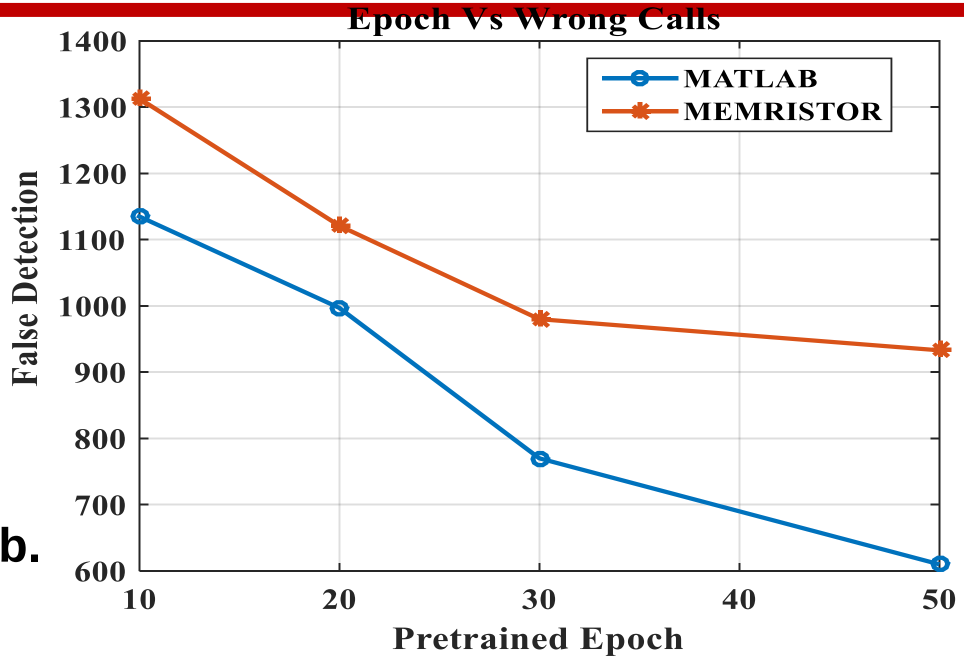
Intrusion Detection Accuracy

a.



Intrusion detection Accuracy (AE-1)

b.



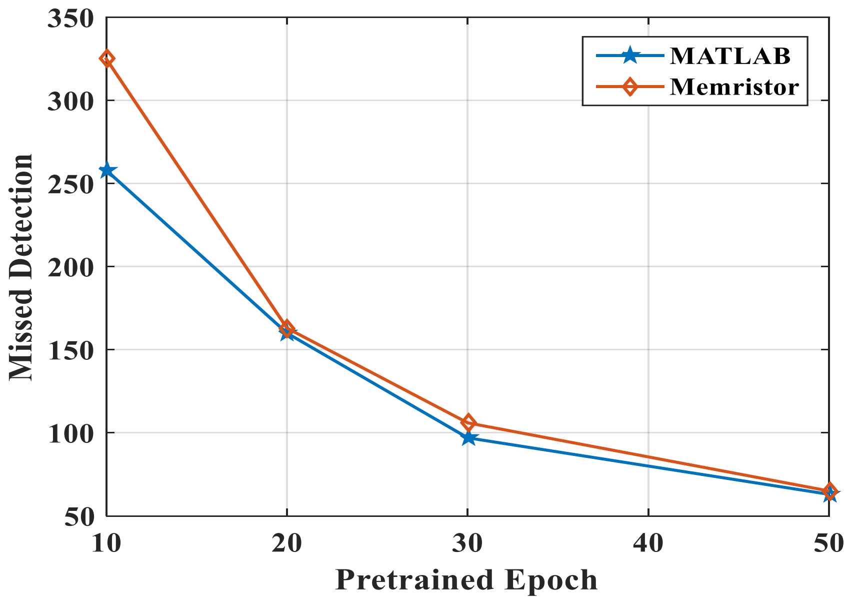
False Detection (Malicious + Normal)

$$Accuracy = \frac{N_S - N_F}{N_S} \times 100\%$$

Pretraining Epochs	***Global Accuracy***	N_{MN}	N_{NM}	N_F	***Case***
50	95.22%	56	546	602	Software
50	92.91%	65	868	933	Memristor

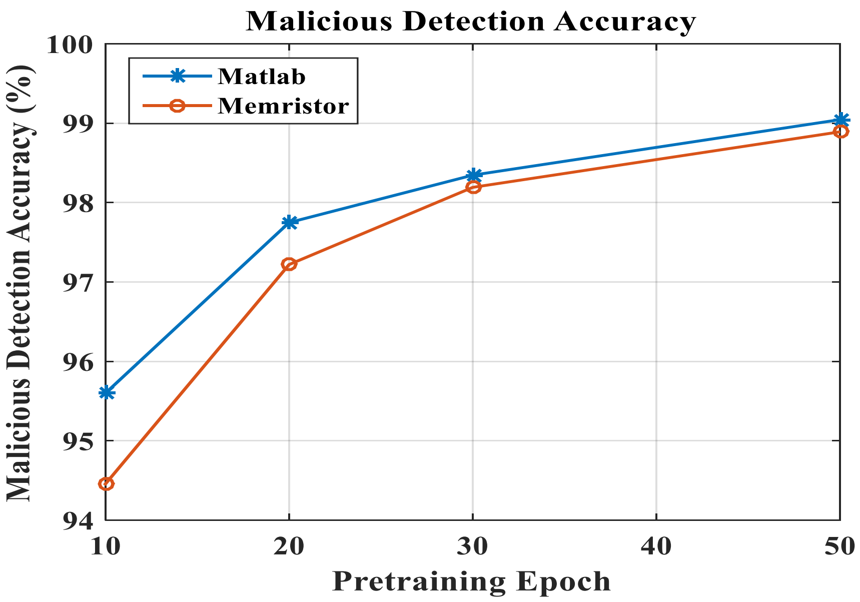
Intrusion Detection Accuracy (contd.)

a.



Malicious Packet Vs Epochs

b.



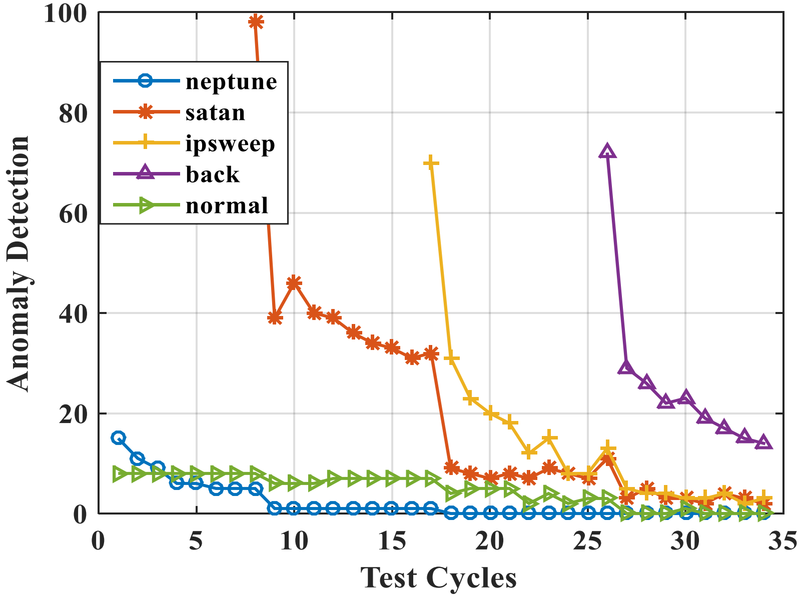
Malicious Packet Detection Accuracy Vs Epochs

Real-time learning and anomaly detection

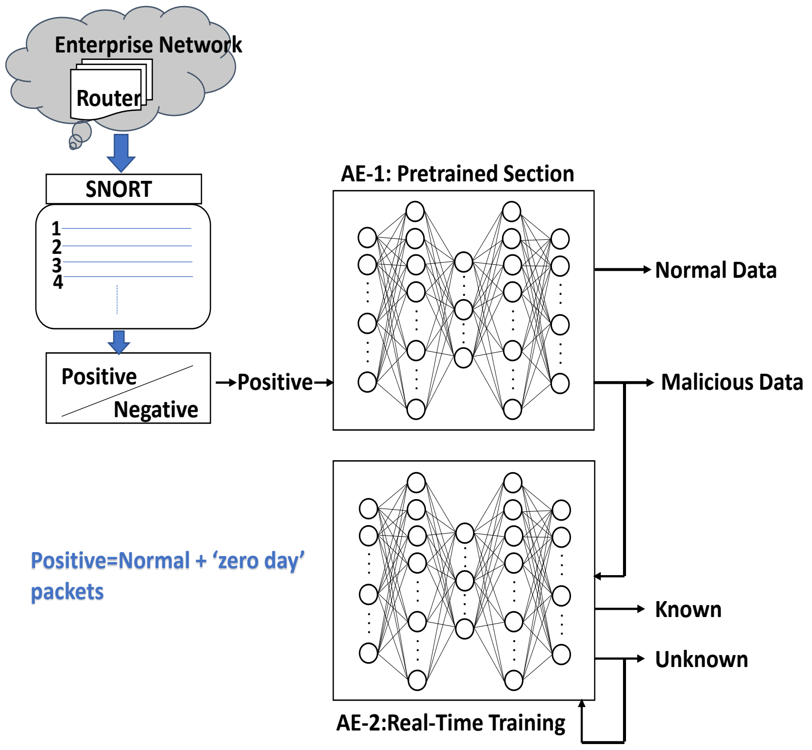
Real-Time Anomaly Detection:

$$T_1 = x_1^1, x_2^1, x_1^2, x_2^2, x_1^3, x_2^3, \ldots$$
$$T_2 = x_1^1, x_2^1, x_3^1, x_1^2, x_2^2, x_3^2, \ldots$$
$$T_3 = x_1^1, x_2^1, x_3^1, x_4^1, x_1^2, x_2^2, x_3^2, x_4^2, \ldots$$
$$T_4 = x_1^1, x_2^1, x_3^1, x_4^1, x_5^1, x_1^2, x_2^2, x_3^2, x_4^2, x_5^2, \ldots$$

$x_1 = normal, x_2 = neptune, x_3 = satan, x_4 = ipsweep, x_5 = back$



Anomaly Detection in real-time



Anomaly Detection System

Power, Area and Timing Analysis

- $R_{off} = 1 \times 10^7 \Omega, R_{on} = 5 \times 10^4\ \Omega$
- Wire Resistance =5 Ω, $V_{mem} = 1.3volt$
- Transistor Feature Size : F= 45nm
- Op-amp power = $3 \times 10^{-6}\ watt$
- Transistor Size= $50F^2$
- Memristor area= $1 \times 10^4\ nm^2$

Parameter	Training Data	Recognition Data
Area (mm^2)	*0.00271*	*0.00271*
Power (mW)	*20.6*	*7.56*
Time (µs)	*4.02*	*0.384*
Energy/One Sample (nJ)	*82*	*2.90*

Summary

- Introduced the problem and proposed a possible solution
- Presented the Autoencoder with memristor X-bar and the functionalities
- Over all accuracy 92.91% with malicious packet detection accuracy 98.89%
- Presented real-time anomaly detection system
- Explained the power and energy requirement

Current and future work

- Incremental learning algorithm & unseen class detection
- Adaptive learning system for battery power devices

THANK YOU

Questions?