The IoT Security Concept of IoTAC

Botnet Attacks on IoT Networks: Malicious Traffic to Compromised Devices

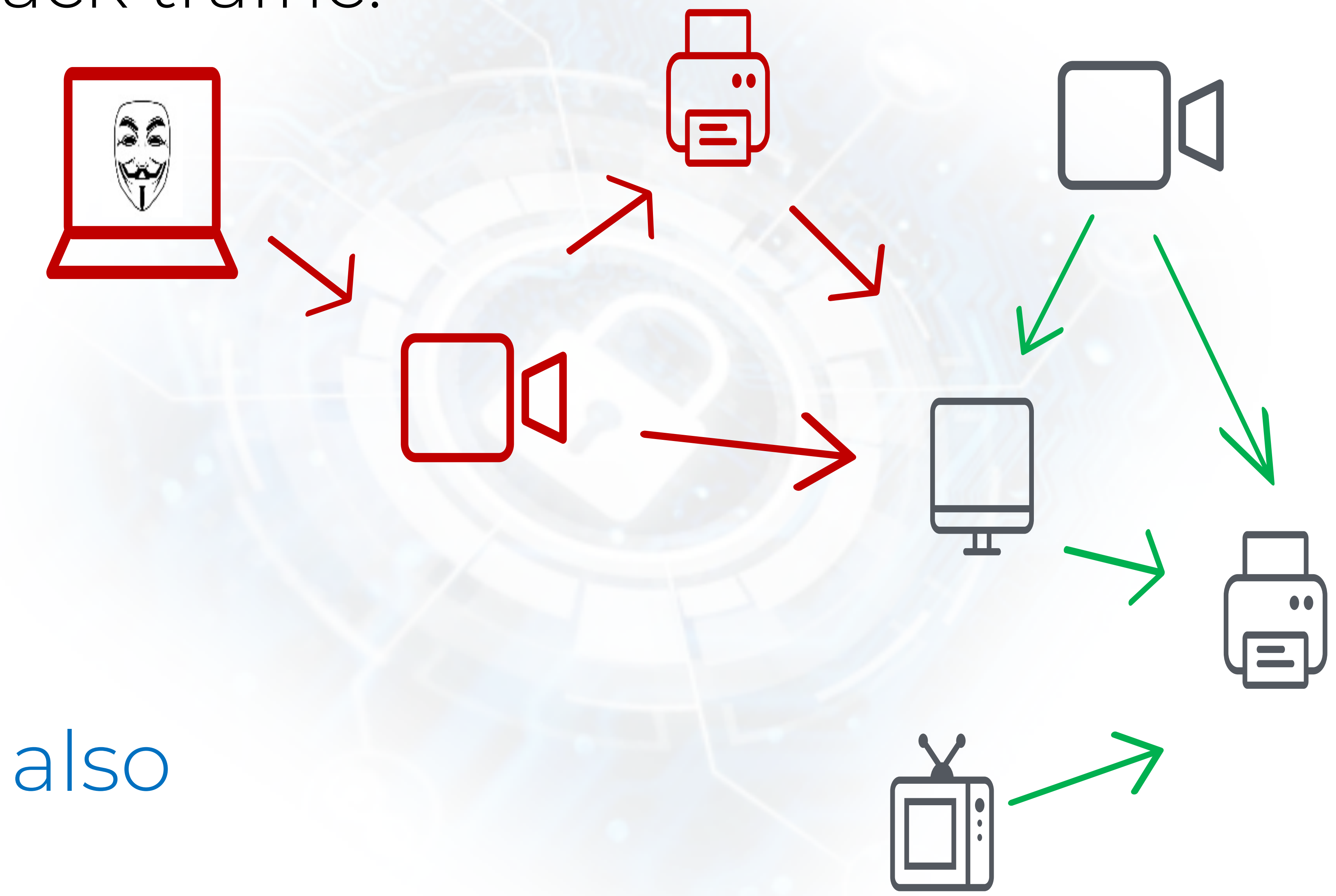
Mert NAKIP
Institute of Theoretical and Applied Informatics,
Polish Academy of Sciences

IoT Week 2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Mirai Botnet Attacks

- ➢ Mirai ("future" in Japanese) Botnet is a form of DDoS attack
  - Sends TCP SYN requests to a large number of IP addresses
  - If the victim responds these requests then attacker uses the weak login credentials.
  - The infected victim becomes a bot which generates attack traffic.
- ➢ Mirai attack spreads to IoT devices over the network.
- ➢ Every device infected by Mirai turns into a bot and generates more traffic than usual, causing a DDoS in the network.



- ➢ It is crucial to identify not only malicious packets but also compromised IoT devices for massive IoT networks.
- ➢ Successful identification of compromised devices may pave the ways to prevent the attack from growing with the spread of malware.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Detecting Malicious Traffic and Compromised Devices via Machine Learning on the Traffic Statistics

Known Basics:

- ➢ Compromised devices will try to increase the total traffic to overload the network by sending more packets.
- ➢ Thus, the attack packets that are generated by compromised devices will certainly have some traceable effects.

Proposed Detection Technique:

- ➢ Various statistics are defined to capture the effects of the attack on the network traffic.
- ➢ Machine Learning algorithm, called Dense Random Neural Network (RNN), is used to create Auto-Associative Memory for the statistics
    - "Off-line **training"** or "on-line incremental training" for malicious traffic detection
    - "On-line sequential training" for compromised device identification

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Why auto-associative memory?

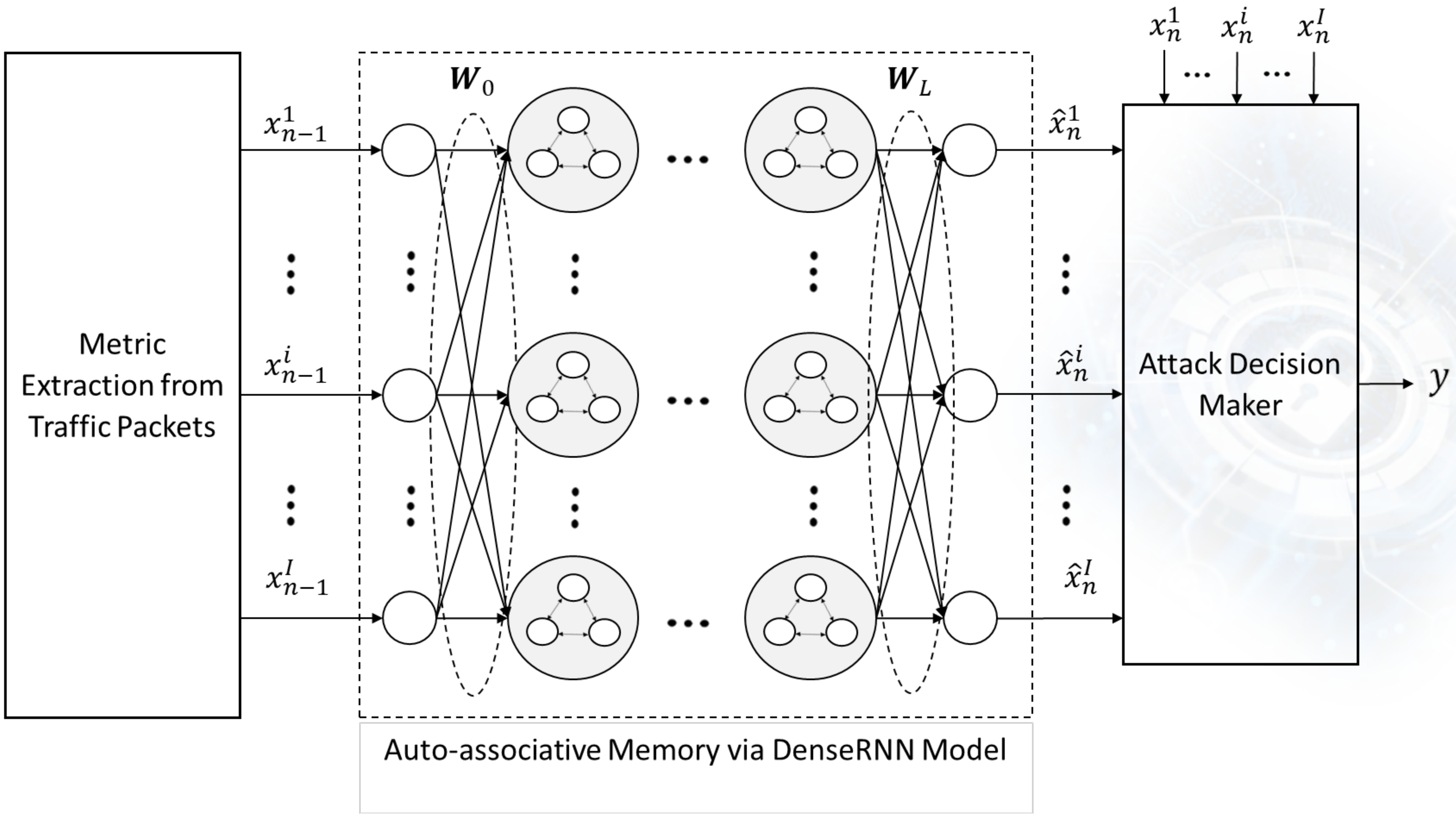
- Ability to react to anomalies / rare events by learning only the normal operation of the system
  - Does not require data on "attack traffic" for training
    - Eliminates data collection via simulations which may be misleading and computationally intensive
    - Enables real-time (online) training of attack detector on the normal traffic
  - High generalization ability
    - Towards the changes of the footprints of attacks on the considered statistics
    - For the detection of various types of attacks via a single detector

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Detecting Malicious Traffic

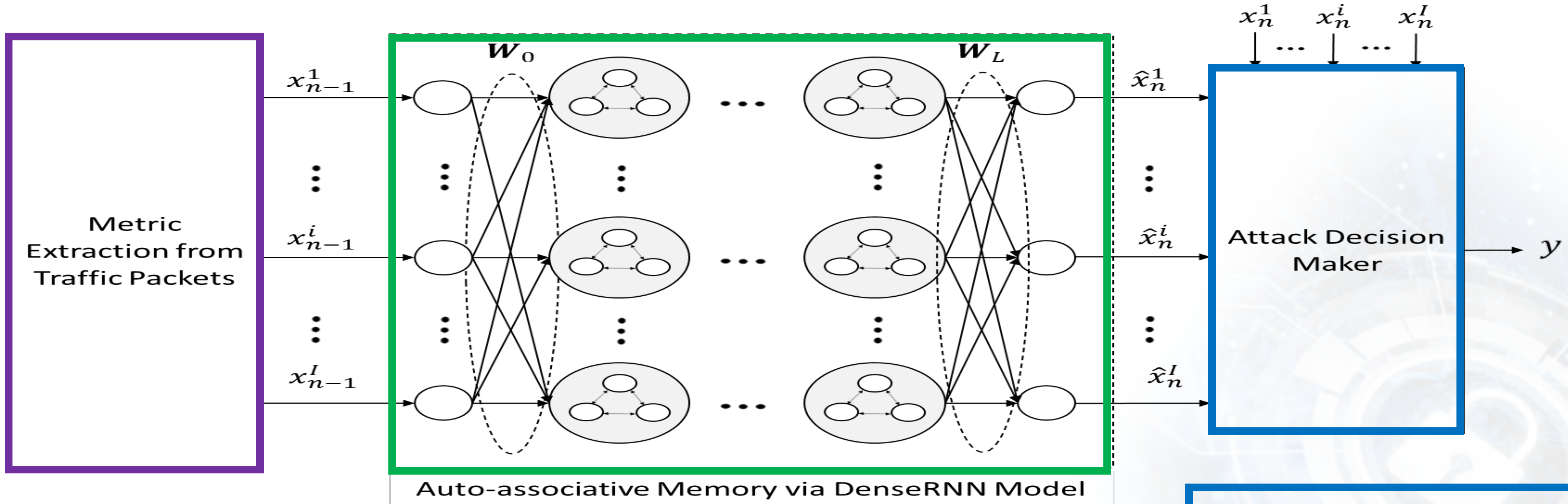
The content was partially published and presented in GLOBECOM'21

Auto-Associative Dense Random Neural Network for Attack Detection



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Auto-Associative Dense Random Neural Network for Attack Detection



1) Total size of the last N packets
2) Average inter-transmission times of the last N packets
3) Total number of packets that are transmitted in a time window with a duration of T

$$O_0^i = \min(X^i, 1),$$
$$O_l^i = q(O_{l-1}^i \mathbf{W}_{l-1}^i) \ \forall l \in \{1, \ldots, L\},$$
$$\hat{X}^i = O_L^i \mathbf{W}_L^i,$$

$$d_n^i = |x_n^i - \hat{x}_n^i|$$

1) Absolute difference between the expected and actual statistics

$$y = \mathbf{1}[\sum_{i \in \{1,\ldots,I\}} \alpha_i d_n^i \geq \Theta]$$

2) Thresholding on the weighted average of absolute differences

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Performance Evaluation

TABLE I
COMPARISON OF ATTACK DETECTION METHODS WITH RESPECT TO ACCURACY AS WELL AS EACH OF THE TRUE POSITIVE, FALSE NEGATIVE, TRUE NEGATIVE AND FALSE POSITIVE PERCENTAGES

Attack Detection Methods	Accuracy	True Positive	False Negative	True Negative	False Positive
AA-Dense RNN	99.84	99.82	0.18	99.98	0.02
KNN	99.79	99.79	0.21	99.75	0.25
Lasso	99.78	99.75	0.25	99.95	0.05
Simple Thresholding	93.18	93.09	6.94	93.63	6.37

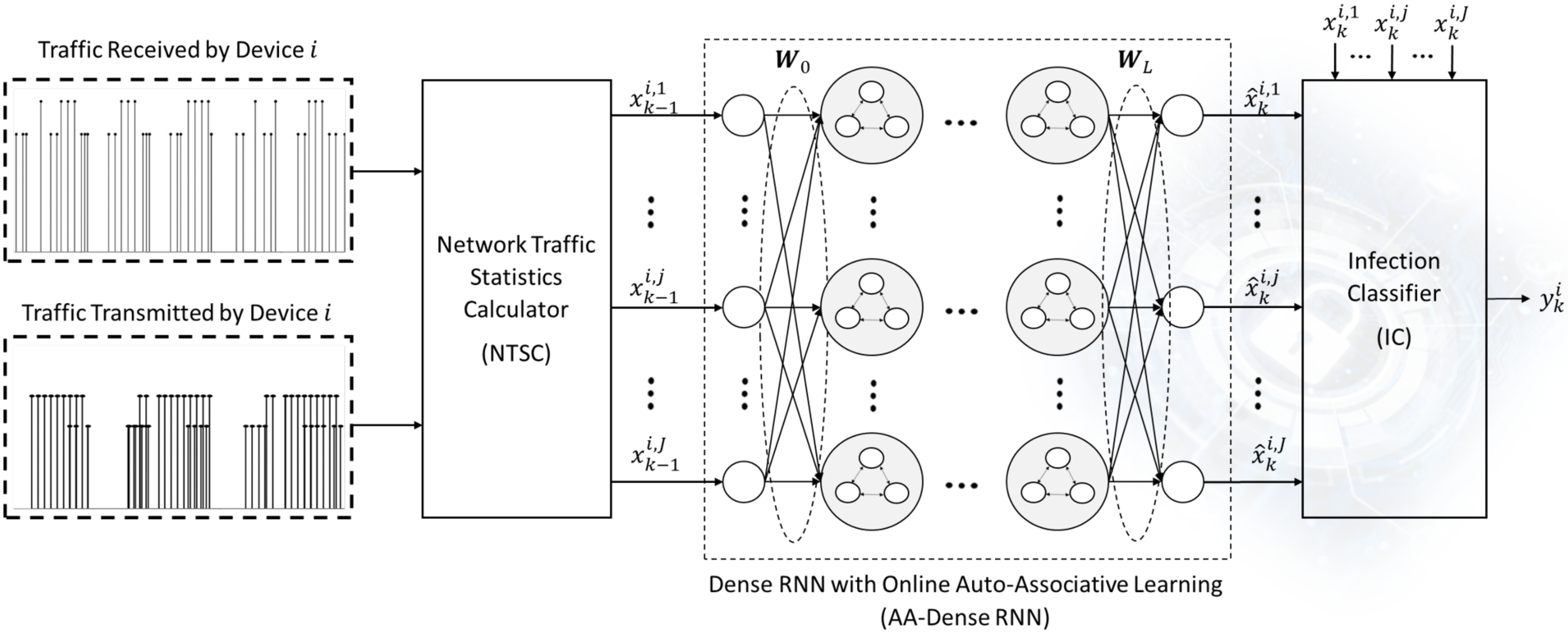
➢ AA-Dense RNN outperforms all compared methods in terms of accuracy as well as true positive and true negative percentages.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Identifying Compromised Devices

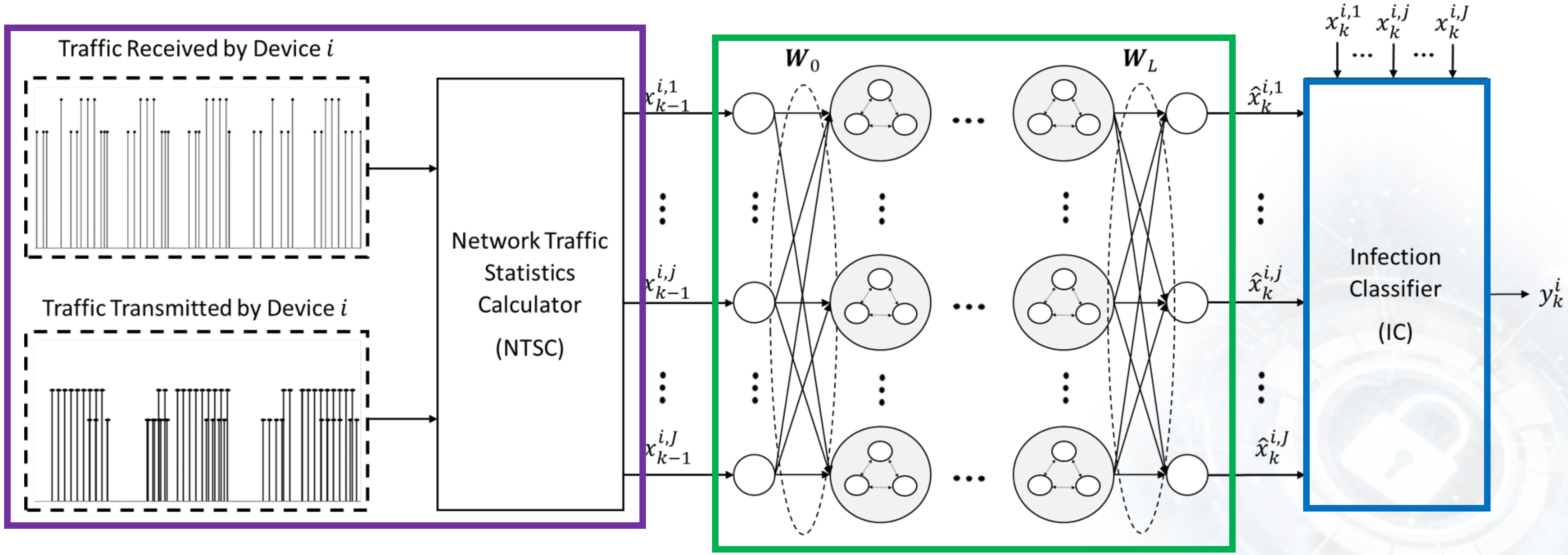
The content was partially submitted for possible publication in IEEE Access

Determine whether a Device is Compromised



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Determine whether a Device is Compromised



Received Traffic

1) The average size of traffic received from different sources
2) The maximum size of traffic received from a single source
3) The average number of packets received from different sources
4) The maximum number of packets received from a single source

Transmitted Traffic

5) Total size of the traffic transmitted
6) Total number of packets that are transmitted

$$\psi_k^i = \max_{j \in \{1,\ldots,J\}} |x_k^{i,j} - \hat{x}_k^{i,j}|,$$

1) Absolute difference between the expected and actual statistics

$$y_k^i = \mathbf{1}[\psi_k^i \geq \gamma_i],$$

2) Thresholding on the maximum of absolute differences

Sequential (online) Auto-Associative Learning

- ➢ Works in conjunction with the execution of the AADRNN system
- ➢ (Offline) Collected data for normal or attack situations is not required
- ➢ Only the benign network traffic is used
  - No labeling is needed

1) For each layer $l \in \{0, \ldots, L-2\}$, by using Fast Iterative Shrinkage-Thresholding (FISTA) [46] algorithm, we first solve

$$\min_{\mathbf{W}_l^i} ||\mathbf{X}_k^i - adj(\zeta(\mathbf{X}_k^i \mathbf{W}_{\text{rand}}^i))||^2 + ||\mathbf{W}_l^i||_{l1} \text{ s.t. } \mathbf{W}_l^i \geq 0,$$

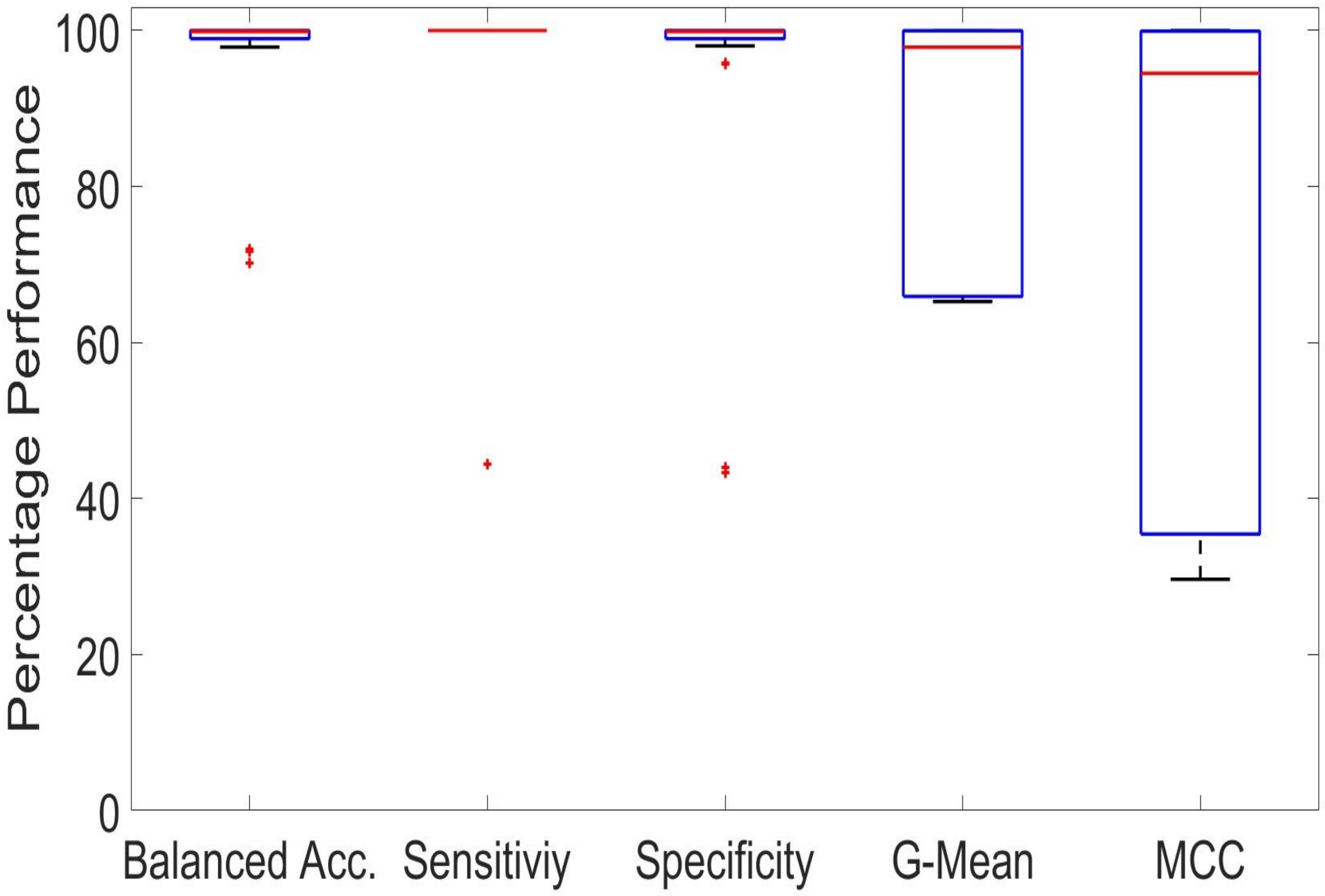
where the matrix of weights $\mathbf{W}_{\text{rand}}^i$ has randomly generated elements in the range $[0, 1]$. In addition, adj($\mathbf{A}$) is a linear mapping of elements of the matrix $\mathbf{A}$ into the range $[0, 1]$, applies z-score (standard score), and adds a positive constant to remove negativity. Then, $\mathbf{W}_l^i \leftarrow 0.1(\mathbf{W}_l^i / \max(\zeta(\mathbf{X}_k^i \mathbf{W}_l^i)))$, and $\mathbf{X}_k^i \leftarrow \zeta(\mathbf{X}_k^i \mathbf{W}_l^i)$.

2) $\mathbf{W}_{L-1}^i$ is randomly generated from uniform distribution in range $[0, 1]$.

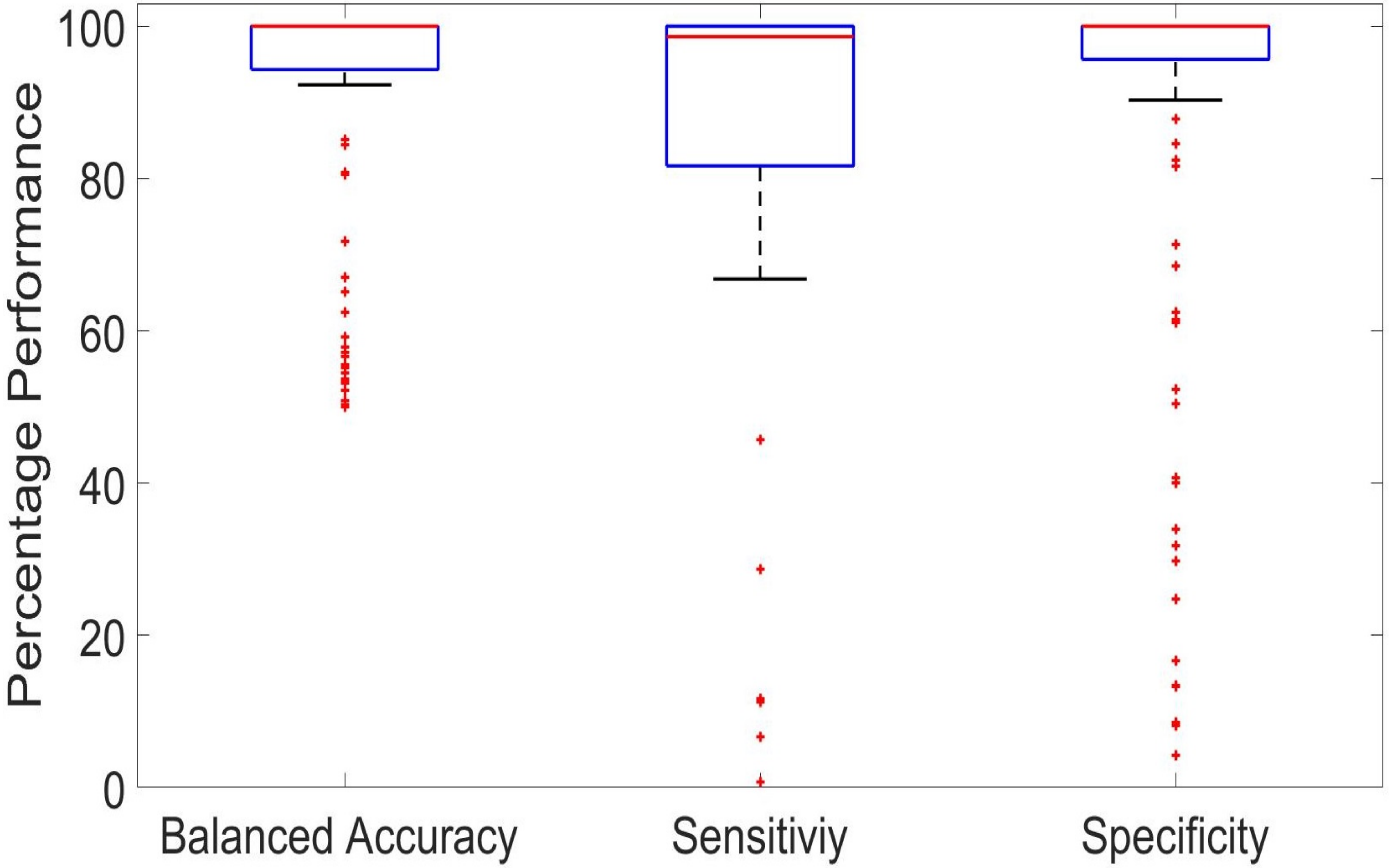
3) $\mathbf{W}_L^i \leftarrow \zeta(\mathbf{X}_k^i \mathbf{W}_{L-1}^i)^+ \mathbf{Y}_k^i$, where $\mathbf{A}^+$ denotes the pseudo inverse of matrix $\mathbf{A}$.

Performance Evaluation

Network with 24 IP Addresses



Network with 107 IP Addresses



- ➢ Almost 100% median for Balanced Accuracy, Sensitivity and Specificity
- ➢ Only 2 outlier IP addresses for which the Balanced Accuracy is 72%

- ➢ Performance decreases for a larger network set-up
- ➢ More outlier IP addresses with low identification performance

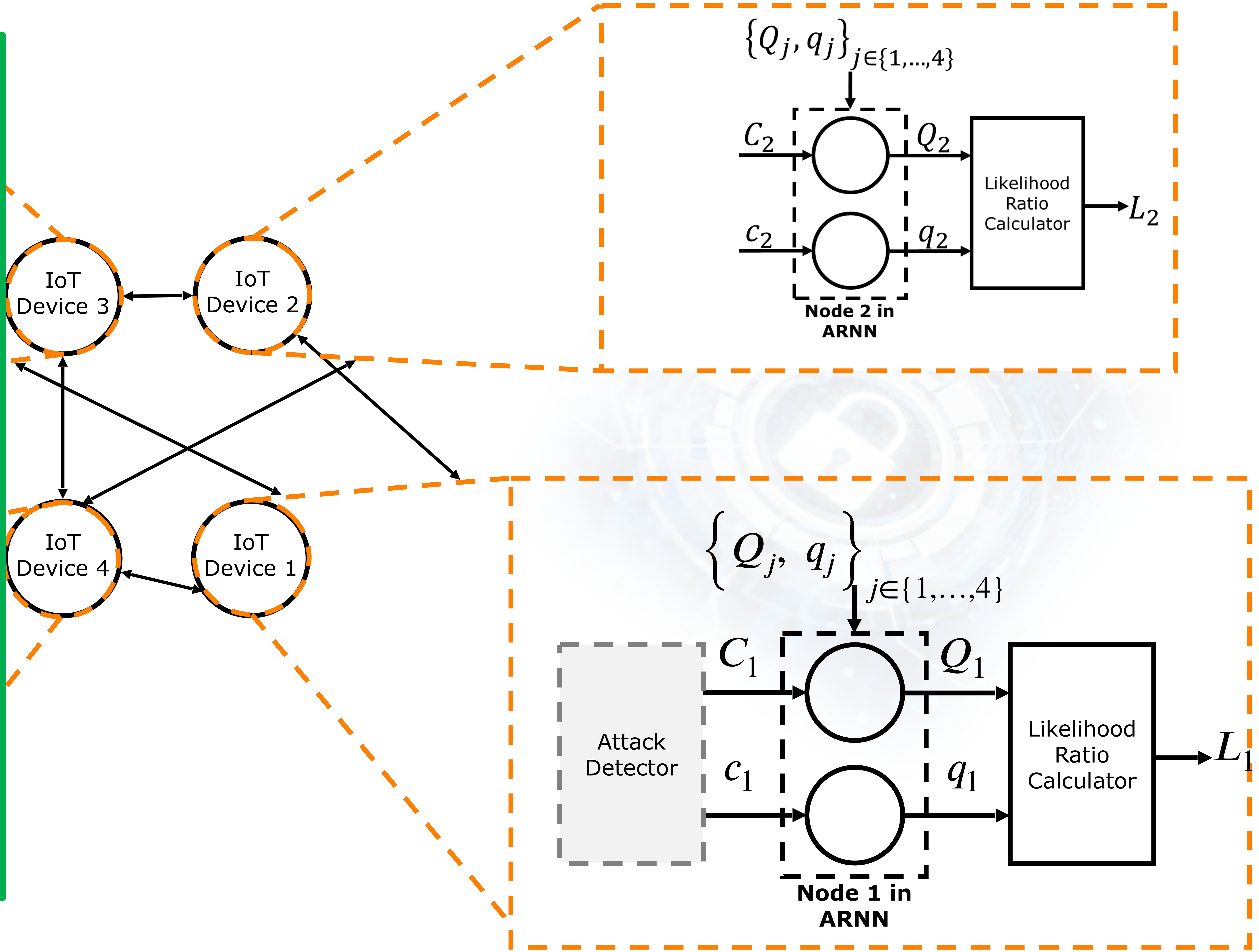
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

Adversarial RNN for Connected Devices

Ongoing research

Adversarial RNN for Network-Wide Attack Assessment

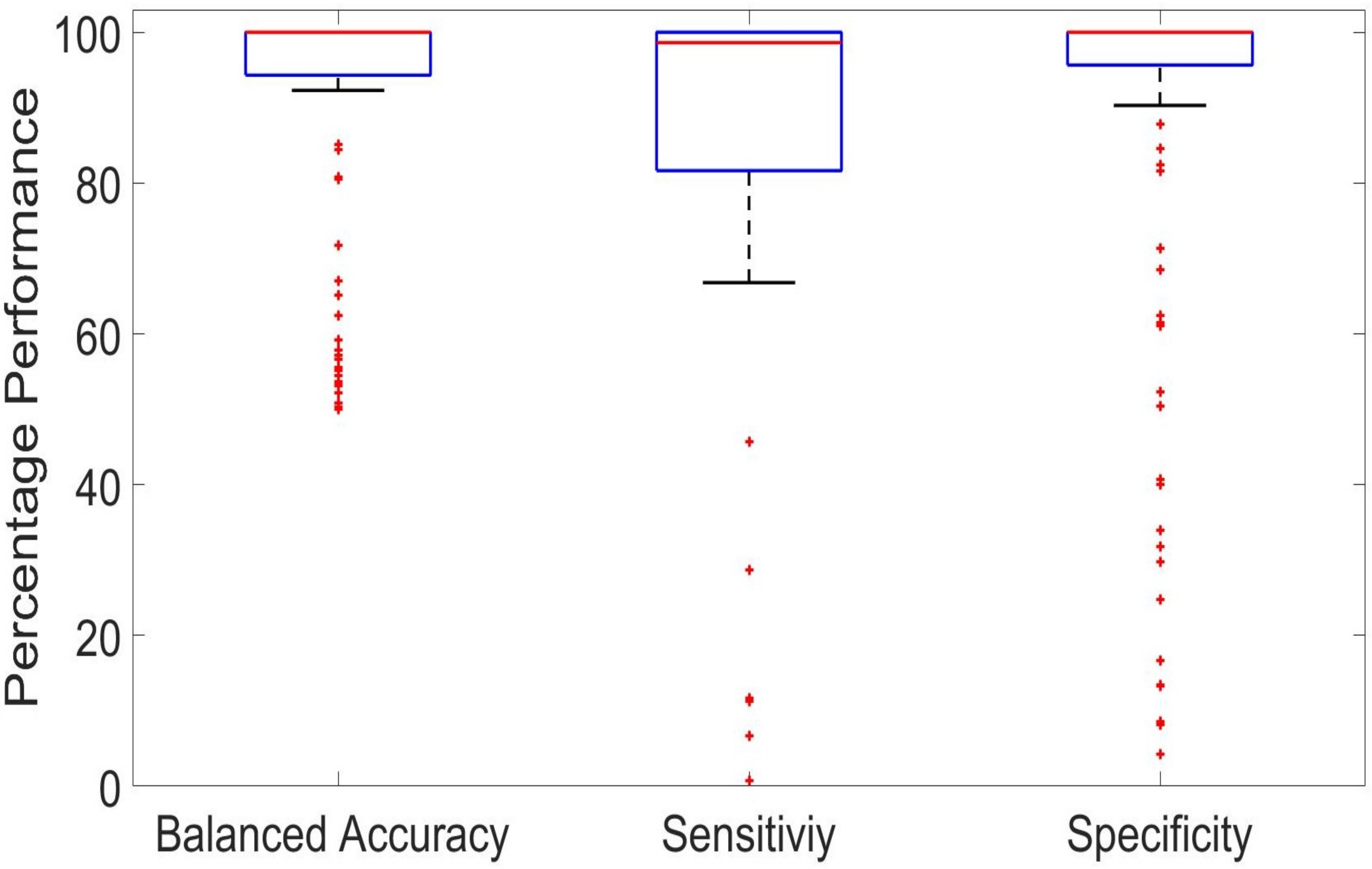
- ➢ From the decisions of “Local Detectors” to network-wide decision
- ➢ Learns the spread of the attack over the network
- ➢ No necessity to have local attack detector at all IoT nodes
- ➢ Easy to scale-up with adding two neurons per each new node



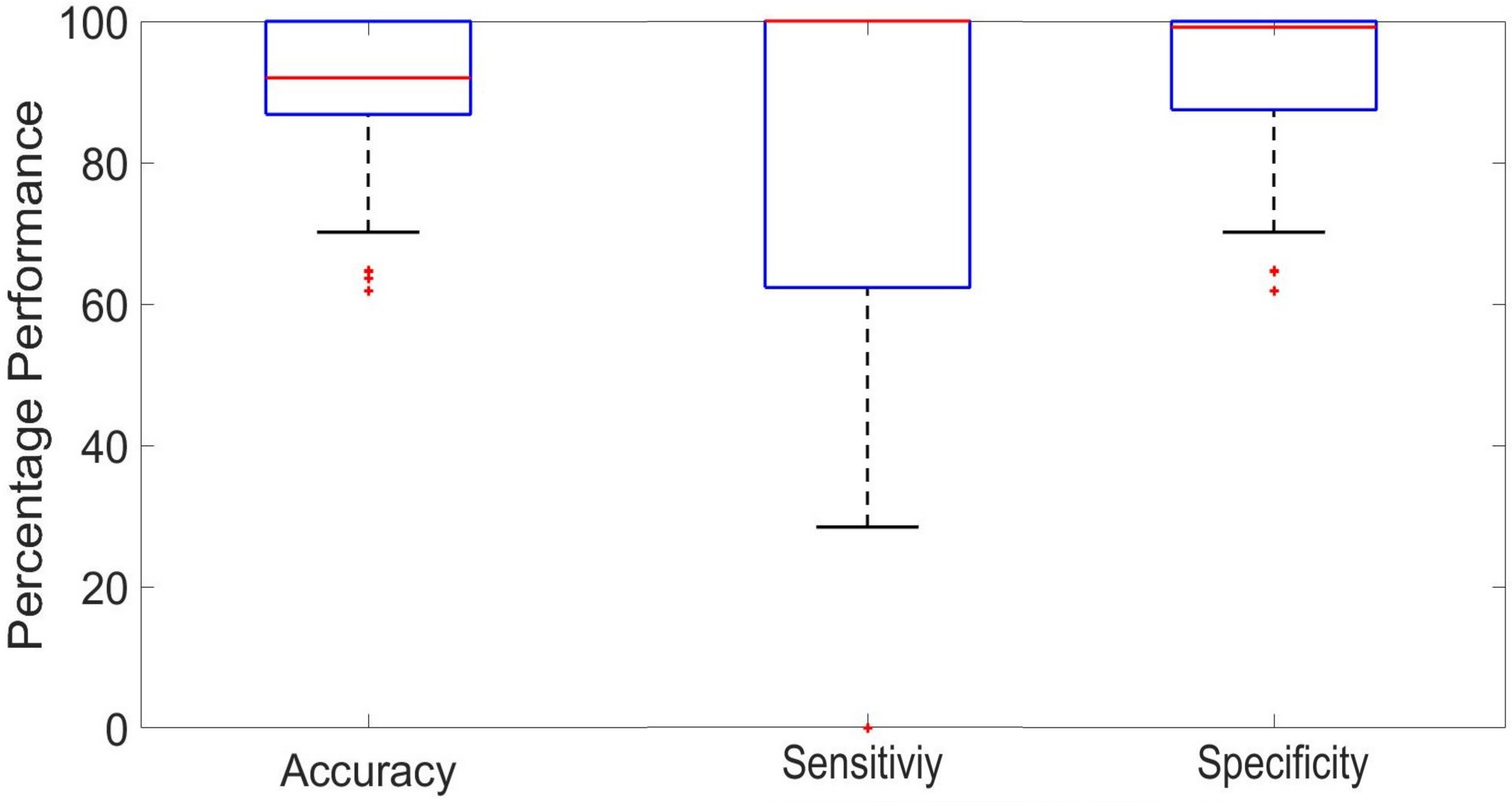
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 952684.

Performance Evaluation and Discussions
on 107 IP Addresses

Identification for Single Node Decision



Adversarial RNN for Network-Wide Attack Assessment



- ➢ Performance is significantly increased by Network-Wide Assessment via Adversarial RNN

  *Only one outlier IP with zero Sensitivity *More than 60% accuracy for all IPs

- ➢ HIGH SCALABILITY

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.

THANK YOU!

IoTAC Web Page

https://Iotac.eu

CONTAC

mnakip@iitis.pl

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952684.