IEEE ICC

14-23 June 2021 // Virtual / Montreal

"A Self-Learning Approach for Detecting Intrusions in Healthcare Systems

P. Radoglou-Grammatikis et al.

University of Western Macedonia

pradoglou@uowm.gr

Under SPEAR Project

Authors

University of Western Macedonia

https://ithaca.ece.uowm.gr/

Panagiotis Radoglou Grammatikis

Panagiotis Sarigiannidis

0Infinity Limited

http://0infinity.net/

Georgios Efstathopoulos

International Hellenic University

https://www.cs.ihu.gr/

Thomas Lagkas

SIDROCO HOLDINGS LIMITED

https://sidroco.com/

Antonios Sarigiannidis

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 787011 (SPEAR).

https://www.spear2020.eu/

Internet of Medical Things (IoMT) Cybersecurity

Among other critical infrastructures, the healthcare sector is the most vulnerable due to the vast amount of personal and administrative data

Vulnerabilities of Smart Medical Devices and EHRs

Smart medical devices use insecure communication protocols like Modbus/TCP, while Electronic Health Records (EHRs) are characterized by severe vulnerabilities

Countermeasures

Need for efficient Intrusion Detection and Prevention Systems (IDPS) using Machine Learning (ML) and Deep Learning (DL) techniques.

Challenge

Lack of available datasets in order to feed the ML and DL models

Goal

Providing an IDPS using an active learning approach for detecting intrusions in a healthcare ecosystem. The proposed IDPS is re-trained continuously, thus optimizing the detection efficacy by itself.

Introduction

A Self-Learning Approach for Detecting Intrusions in Healthcare Systems

Providing an IDPS utilizing active learning in order to detect and mitigate Modbus/TCP and HTTP cyberattacks against a healthcare ecosystem.

Related Work

Previous works related to detecting intrusions against healthcare ecosystems

2020
T. Yaqoob et al.
Security vulnerabilities, attacks, countermeasures, and regulations of networked medical devices—a review

2020
G. Thamilarasu et al.
An intrusion detection system for internet of medical things

2015
R. Mitchell and I. Chen
Behavior rule specification-based intrusion detection for safety critical medical cyber physical systems

2020
A. I. Newaz et al.
Heka: A novel intrusion detection system for attacks to personal medical devices

2019
M. B. Mohamed et al.
Intrusion cancellation for anomaly detection in healthcare applications

2008
J. Stokes et al.
Aladin: Active learning of anomalies to detect intrusions

Contributions

Three main contributions

C1

IDPS for Healthcare Ecosystem

Detecting a plethora of HTTP (5) and Modbus/TCP (14) cyberattacks.

C2

Active Learning Approach

The proposed IDPS is re-trained continuously, thus optimizing the detection efficacy by itself

C3

ML/DL Methods Evaluation

Evaluating a plethora of ML/DL methods

Threat Identification

HTTP and Modbus/TCP Cyberattacks

Electronic Health Records

EHRs use typical ICT protocols like HTTP

Smart Medical Devices

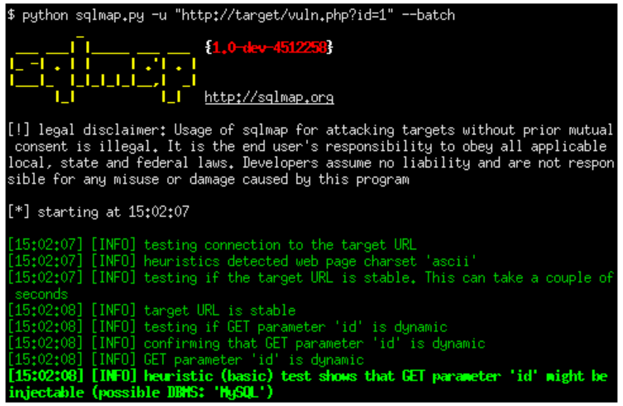
Modbus is an industrial protocol widely adopted by both legacy and smart medical devices.

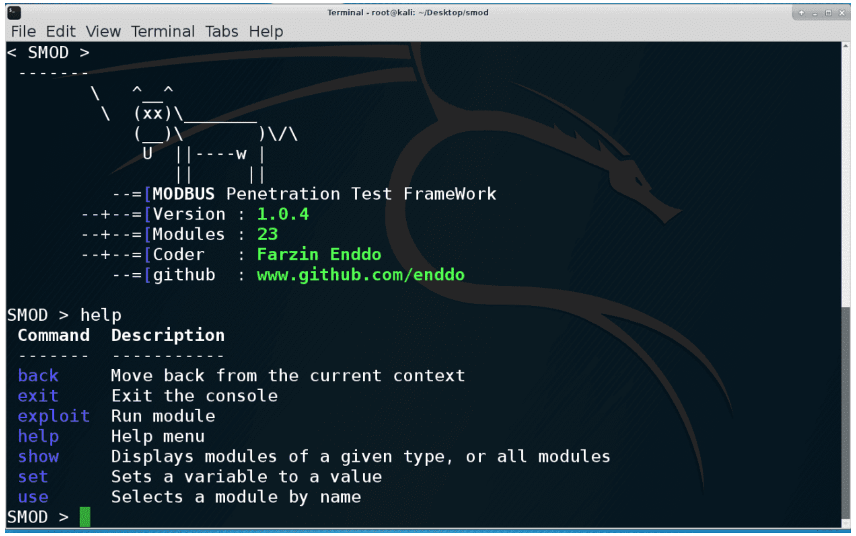
Four HTTP Cyberattacks

(a) Dos, (b) SQL Injection, (c) Bruteforce and (d) XSS

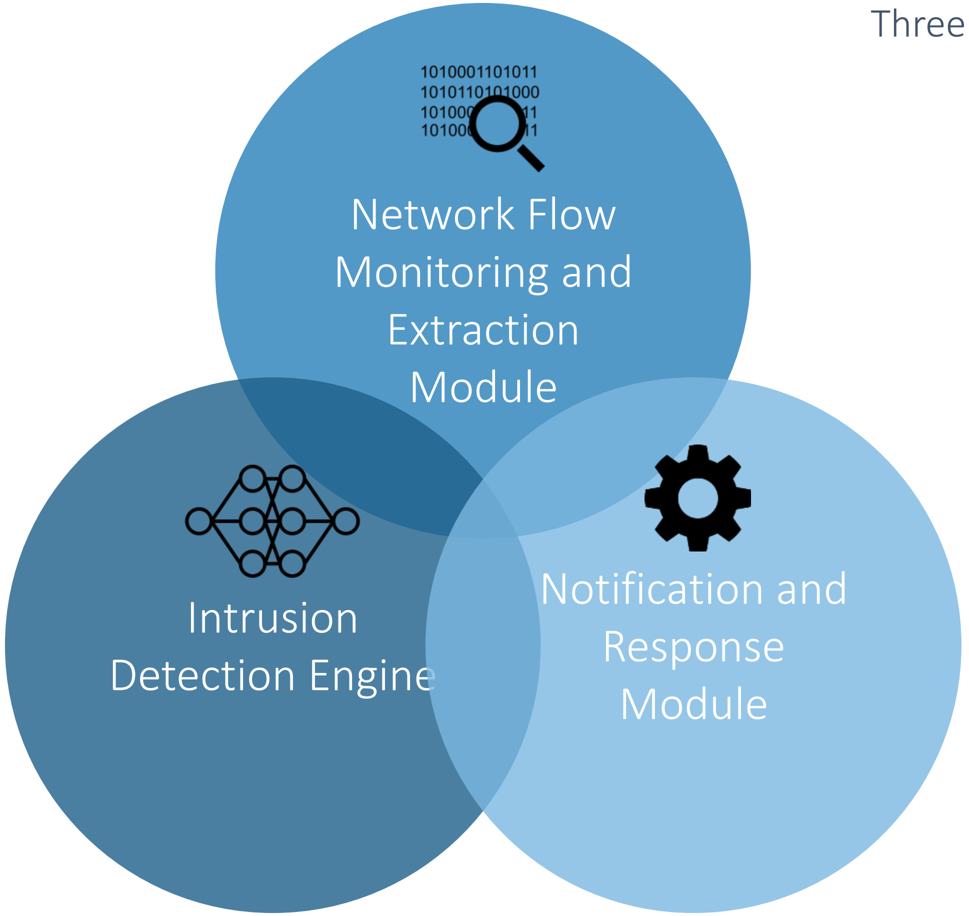
13 Modbus/TCP Cyberattacks

(a) modbus/function/readHoldingRegister
(b) modbus/scanner/uid
(c) modbus/function/readDiscreteInput,
(d) modbus/dos/writeSingleCoils,
(e) modbus/function/writeSingleRegister,
(f) modbus/function/readInputRegister,
(g) modbus/function/readCoils (DoS
(h) modbus/function/readHoldingRegister (DoS),
(i) (modbus/function/readDiscreteInputs (DoS)),
(j) modbus/dos/writeSingleRegister,
(k) modbus/scanner/getfunc,
(l) modbus/function/writeSingleCoils and
(m) modbus/function/readInputRegister (DoS)





Proposed IDPS Architecture



Three main components

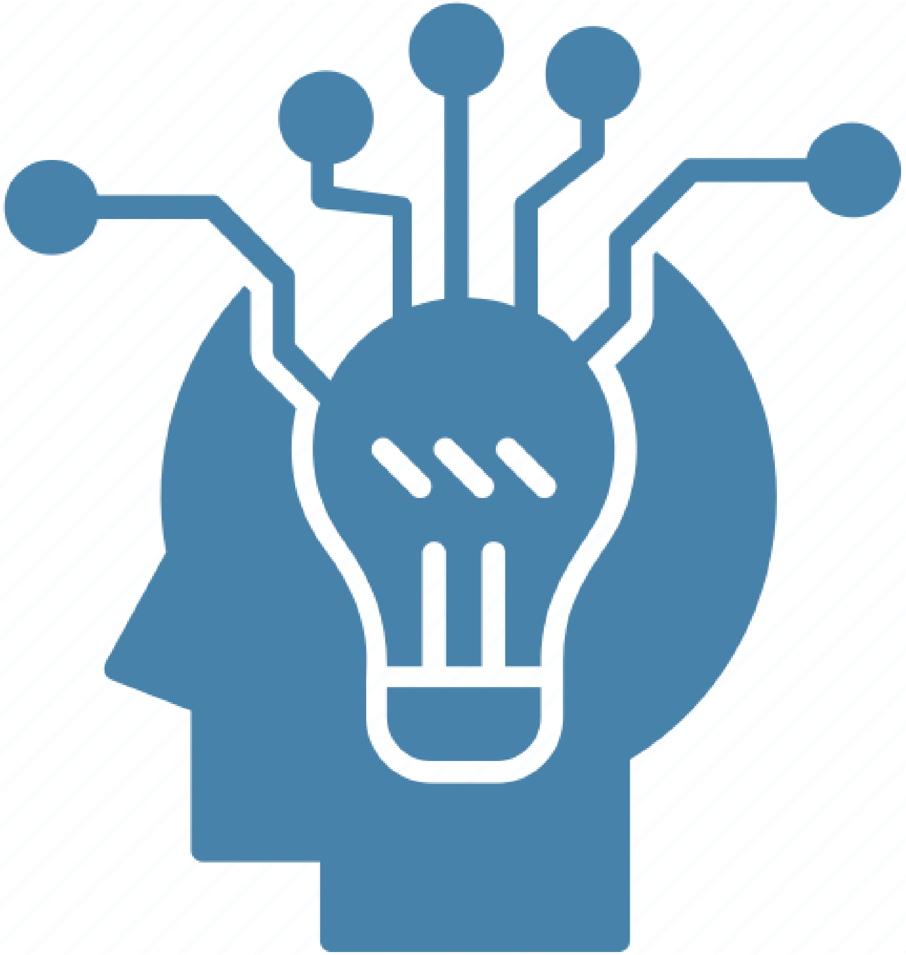
- **Network Flow Monitoring and Extraction Module**
  Through SPAN and Tcpdump, it monitors the examined healthcare infrastructure. It generates bi-directional Modbus/TCP and HTTP flow statistics.

- **Intrusion Detection Engine**
  Responsible for detecting the aforementioned Modbus/TCP and HTTP cyberattacks. Decision Tree → HTTP cyberattacks, Random Forest → Modbus/TCP cyberattacks

- **Notification and Response Module**
  Notifies the security administrator about possible security events. Indicates and applies firewall (iptables) rules.

Active Learning

Overview & Methods



When?

When there are no available labelled training datasets as in our case since CIs cannot label and disclose their sensitive data.

What?

Selects the most informative data from a set of unlabelled data in order to optimize and construct a training dataset

Oracle

Usually, there is an external factor that annotates the data investigated

Query Synthesis

Synthesizes the data samples de novo, producing never observed data samples.

Stream-based Selective Sampling

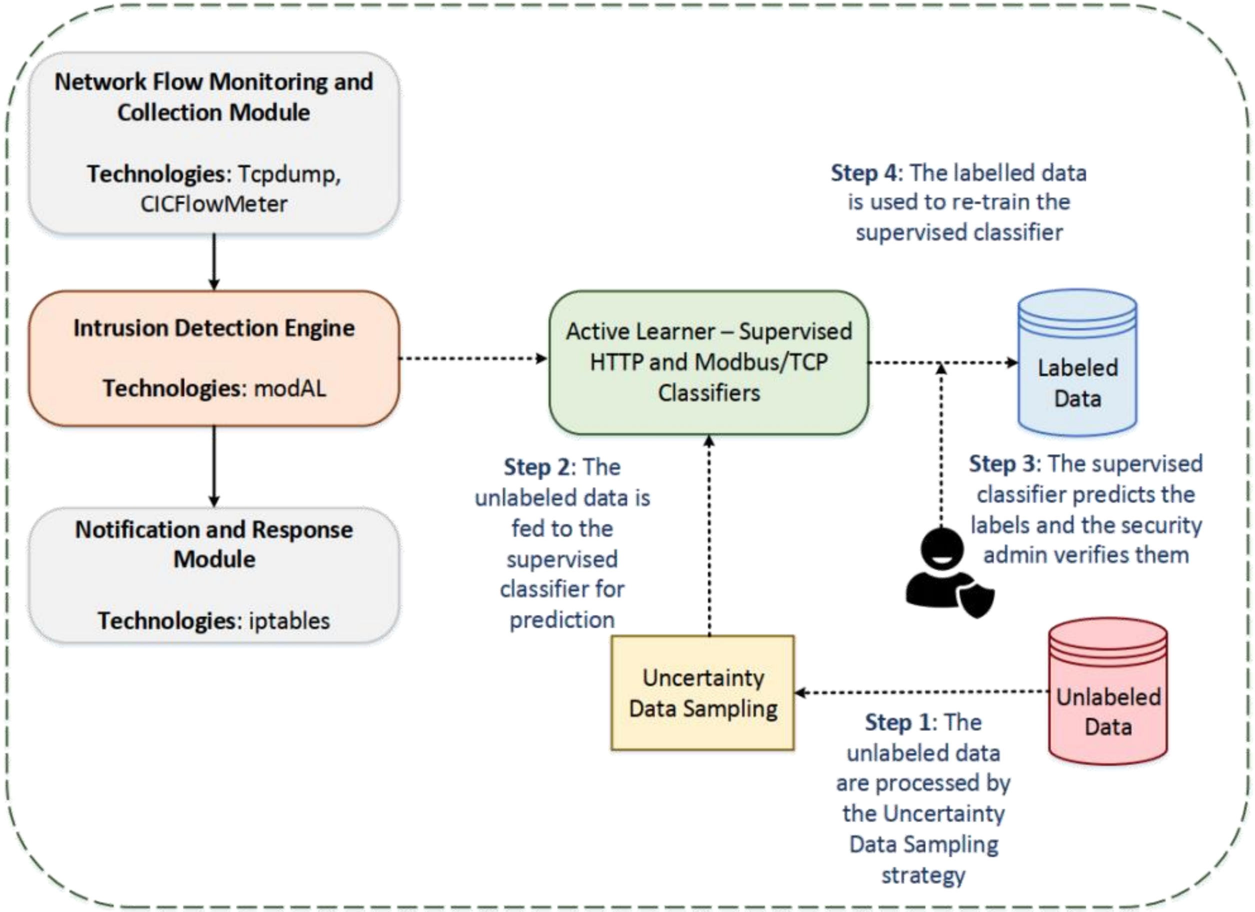
Receives data samples as streams continuously and decides based on a **query strategy** which data samples should be labelled or not.

Pool-based Sampling

Creates first a pool with unlabelled data samples and sequentially decides based on a **query strategy** which of them will be labelled.

Active Learning

How it is adopted by the proposed IDPS



Step #1: Data Investigation & Assessment

The unlabelled data is assessed by the **query strategy** named Uncertainty Data Sampling.

Step #2: Feeding supervised classifiers

The data approved by the Uncertainty Data Sampling is fed to the supervised classifiers

Step #3: Data Prediction and Verification

The supervised classifiers predict the labels that also are verified by a security expert.

Step #4: Training Dataset Update

The new labelled data is introduced to the new training dataset

Active Learning

Implementation

Definitions

- Let x be an unlabelled network flow from the input space X and y the respective label
- Let U be a set of unlabelled TCP/IP network flows within a pool.
- Let L be the training dataset consisting of the labelled TCP/IP network flows
- $f(x) = y$: the target function, which absolutely classifies the unlabelled TCP/IP network flows in the correct classes.
- $h(x) = y$: supervised classifier (**hypothesis**) predicting the label of an unlabelled TCP/IP network flow after the training process

Generalization Error *E*

$$E[l(h)] \int_{-\infty}^{\infty} l(h(x), f(x))\, dx$$

$$l(h(x), f(x)) = (h(x) - f(x))^2$$

Uncertainty Sampling Strategy

- Identify and label those unlabelled network flows in the pool that will be used to re-train the supervised classifiers (**hypothesis**)
- Ask the external factor about those unlabelled network flows for which the hypothesis is less confident.
- The external factor is the same hypothesis since IDPS should be re-trained by itself. A security expert confirm the labels

- The uncertainty criterion is the entropy: H

$$H = -\sum_{i=1}^{m} p_\theta(y_i|x) \log_2(p_\theta(y_i|x))$$

$$x^* = argmax(x) + H > \delta$$

Algorithm 1: Active Learner: Pooling-based Sampling and Uncertainty Sampling Strategy

Data: U, L, h
Result: Re-train h
initialization;
while *size(U)* > 0 **do**
  if *classifier_uncertainty(U(i))* $> \delta$ **then**
    Predict y(i) using h;
    Verify or change the prediction of h through the security expert ;
    Add U(i) and y(i) in L;
    Re-train h
  end
  Remove U(i) from U;
end

Evaluation

Evaluation Methodology

A

Step One
HTTP Dataset Preparation

CICIDS2017 Dataset

I. Sharafaldin et al.

Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization

B

Step One
Modbus/TCP Dataset Preparation

Emulating Modbus/TCP Cyberattacks

P. Radoglou-Grammatikis et al.

Implementation and Detection of Modbus Cyberattacks

C

Step Three
Data Preprocessing

D

Step Four
Feature Selection: Flow Duration, TotLen Fwd Pkts, Fwd Pkt Len Mean. Fwd Pkt Len Mean, Bwd Pkt Len Std, Flow IAT Std, Bwd Pkts/s, Subflow Bwd Pkts, Init Bwd Win Bytes, Active Mean

E

Step Five
Evaluation

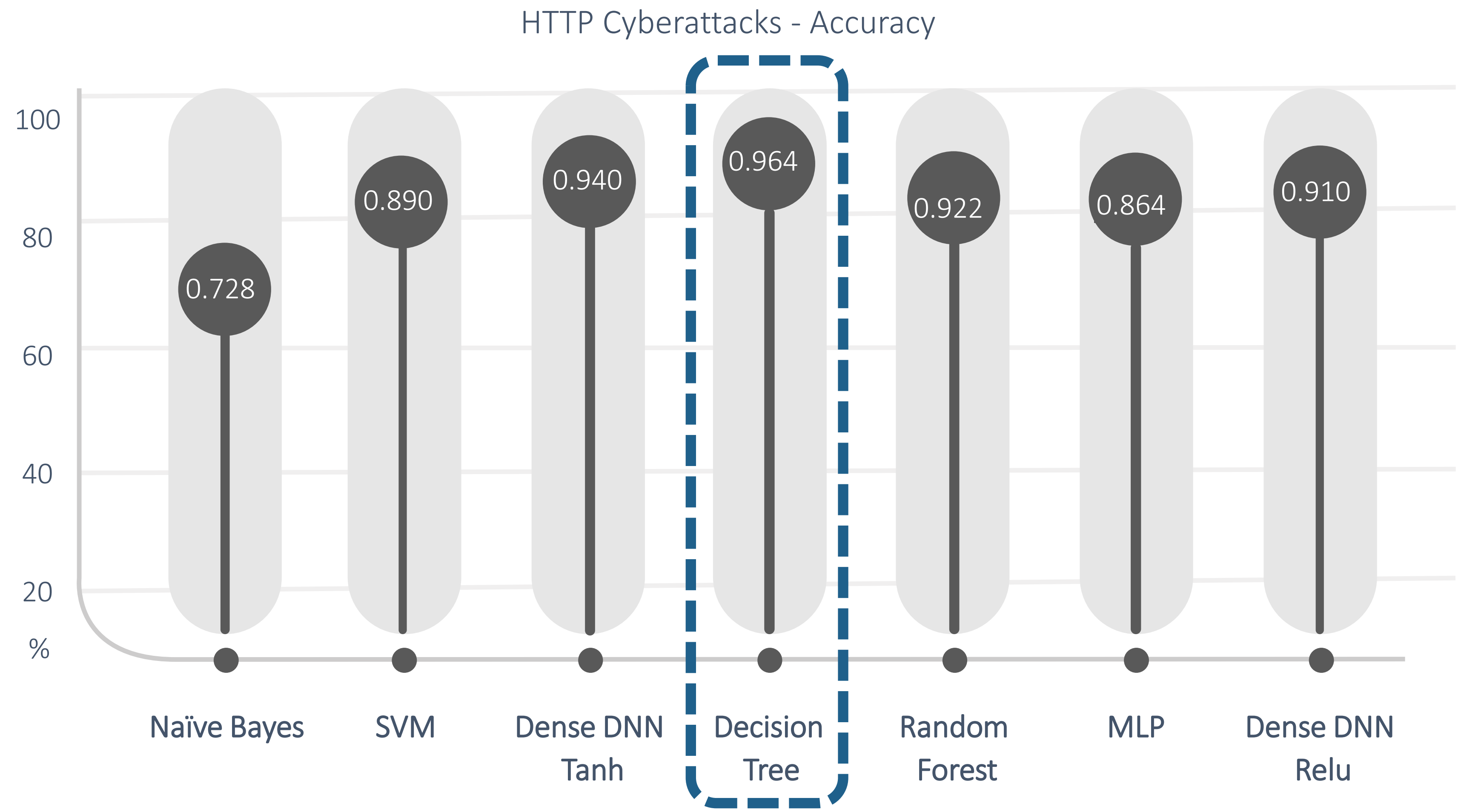
$$Accuraccy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$TPR = \frac{TP}{TP + FN}$$

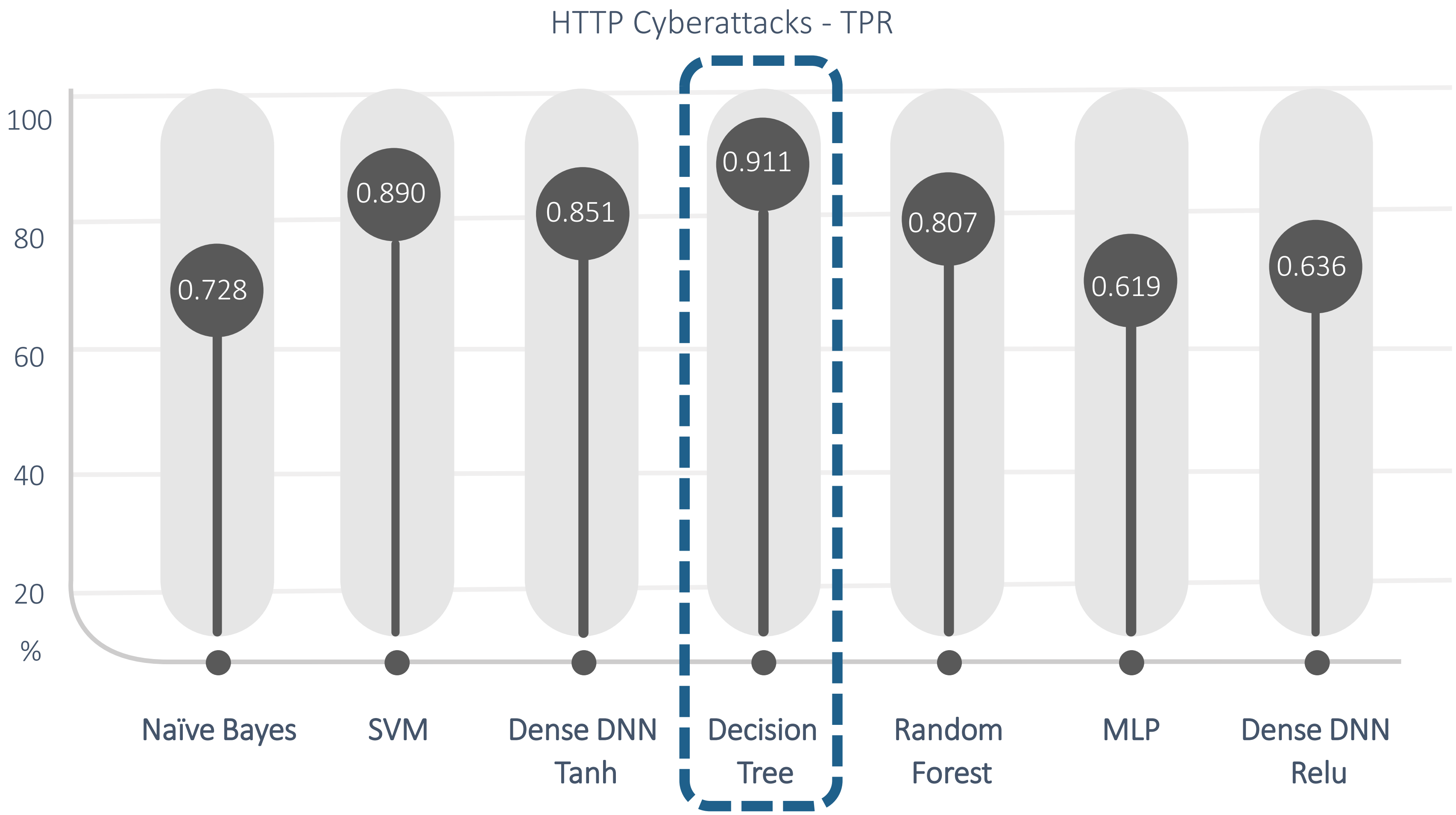
$$FPR = \frac{FP}{FP + TN}$$

$F1 = \frac{2 \times Precision \times TPR}{Precision + TPR}$ where $Precision = \frac{TP}{TP+FP}$

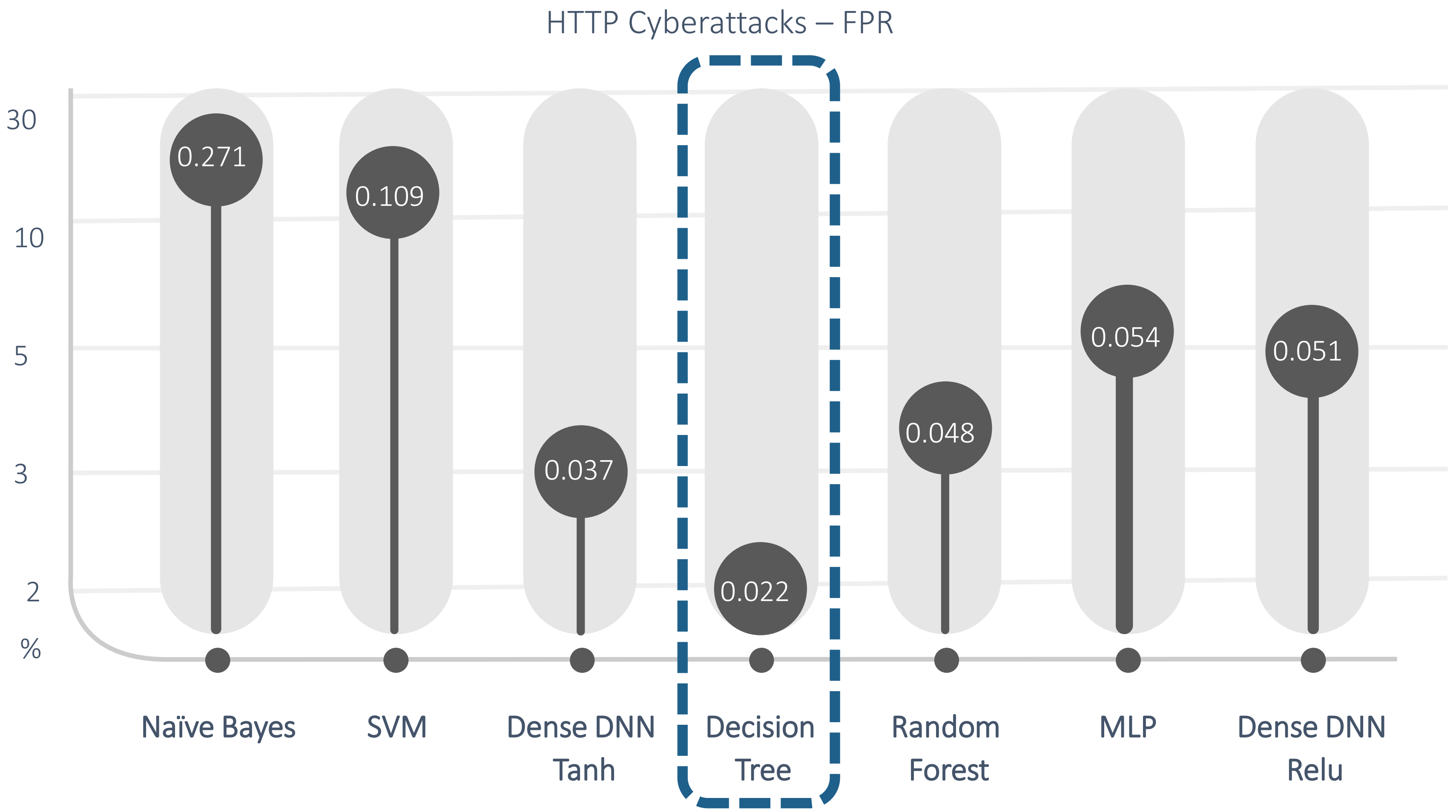
Evaluation


HTTP Cyberattacks - Accuracy
100
80
60
40
20
%
0.728
0.890
0.940
0.964
0.922
0.864
0.910
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

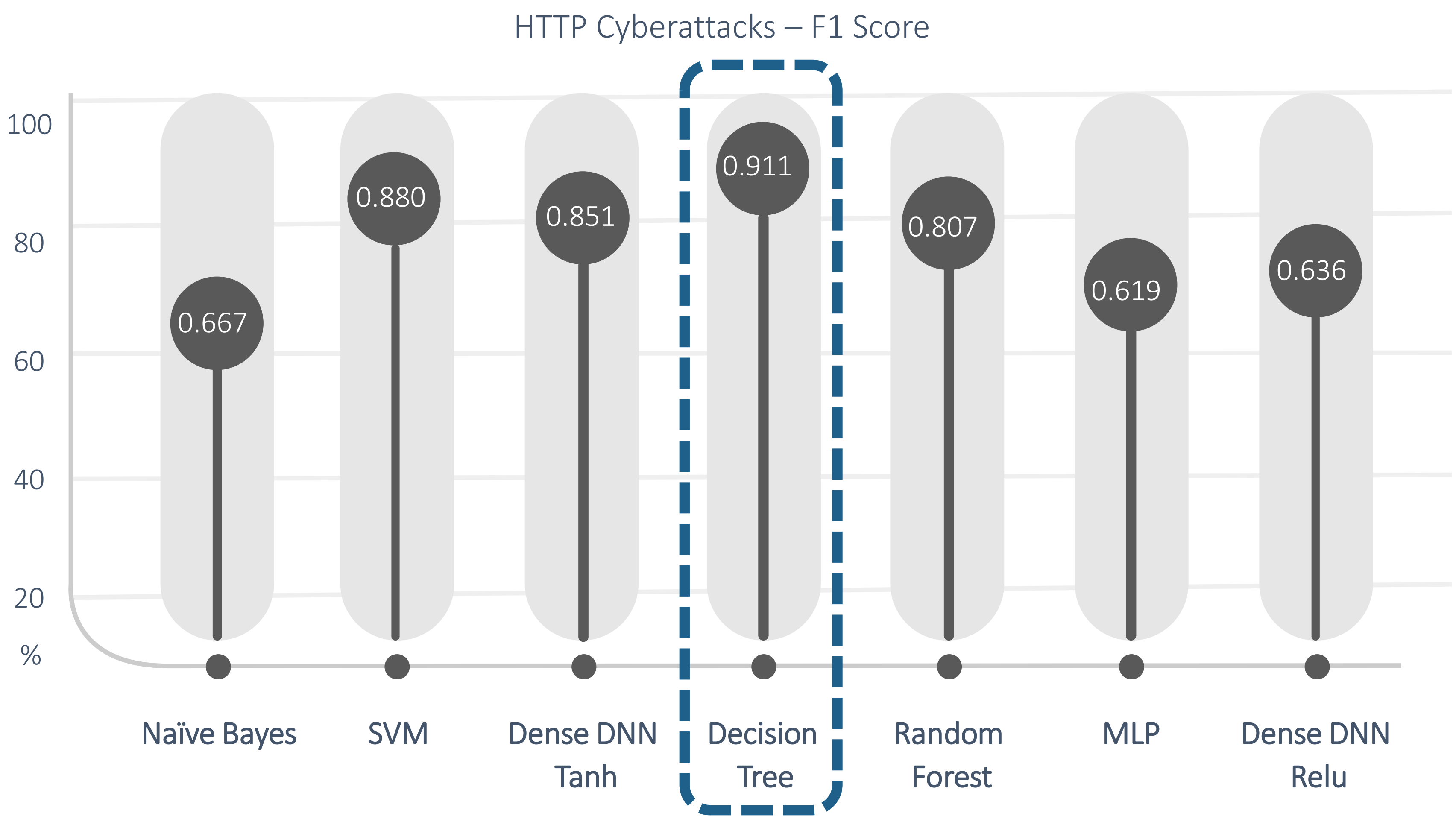
Evaluation


HTTP Cyberattacks - TPR
100
80
60
40
20
%
0.728
0.890
0.851
0.911
0.807
0.619
0.636
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation


HTTP Cyberattacks – FPR
30
10
5
3
2
%
0.271
0.109
0.037
0.022
0.048
0.054
0.051
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation


HTTP Cyberattacks – F1 Score
100
80
60
40
20
%
0.667
0.880
0.851
0.911
0.807
0.619
0.636
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation

HTTP Cyberattacks – Aggregative Results

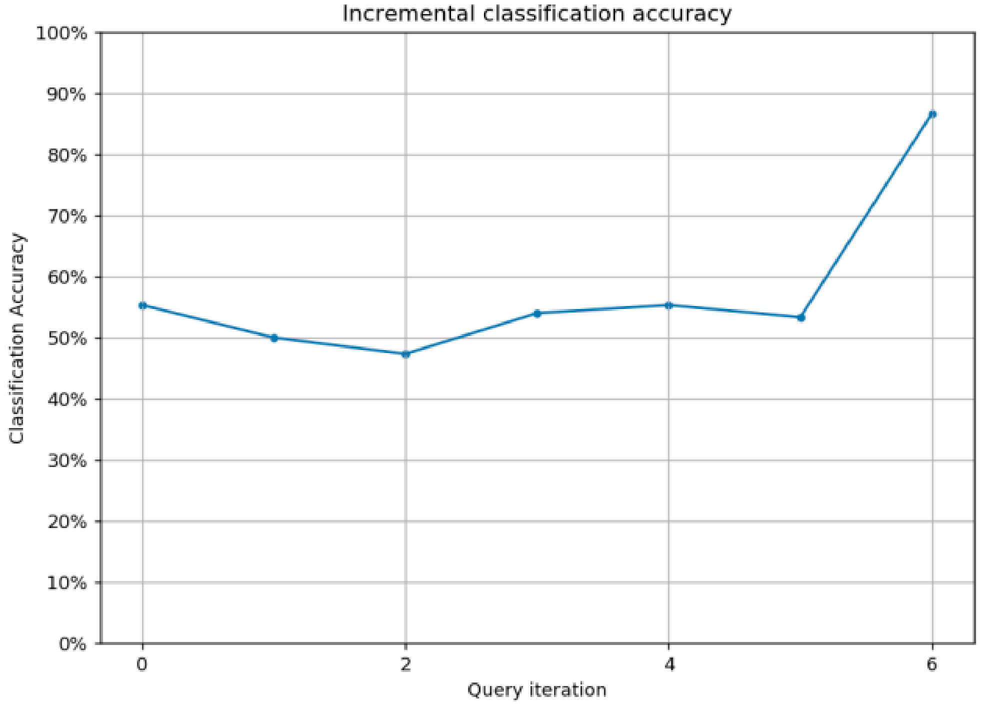
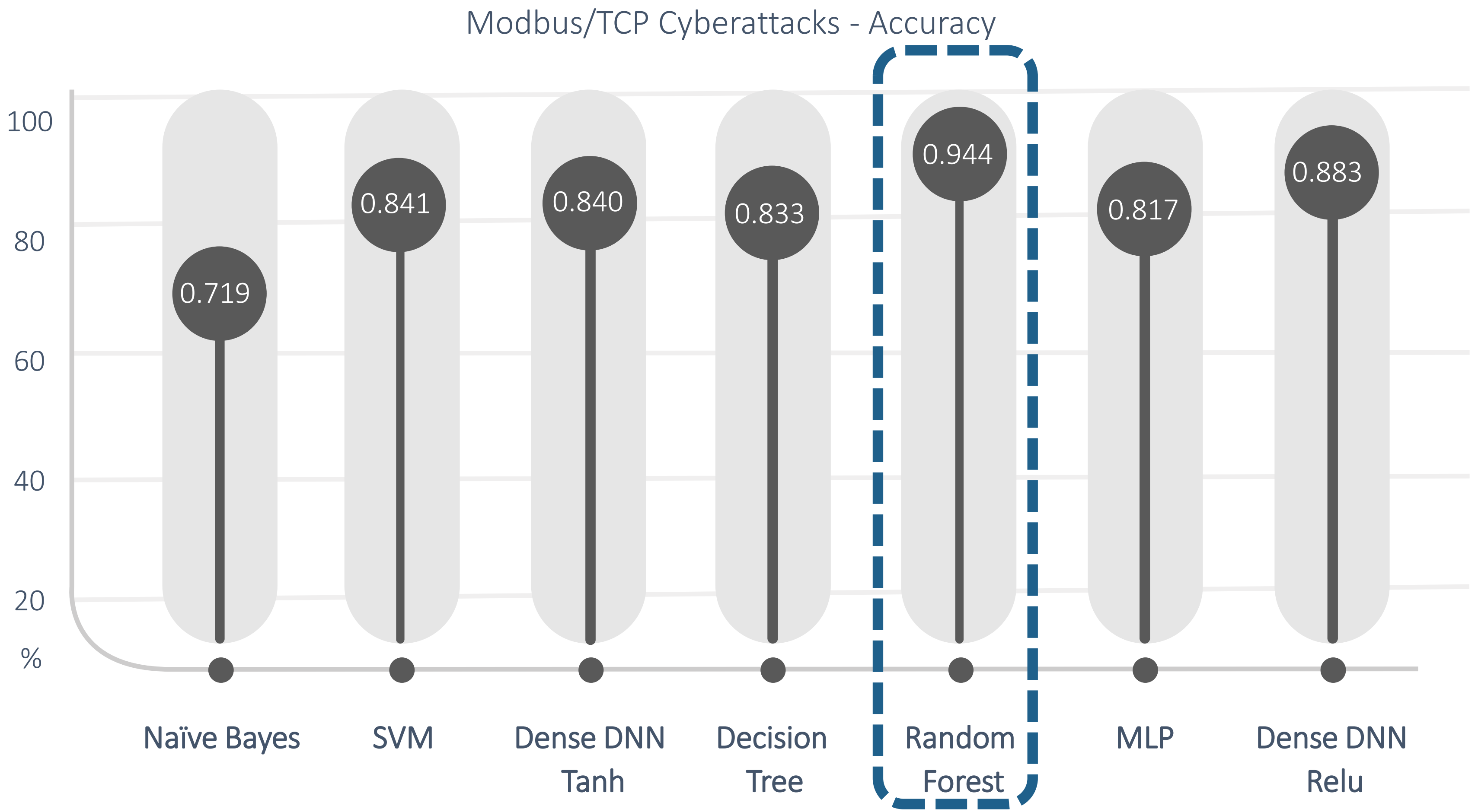


Fig. 2: Decision Tree - Accuracy Increment during the re-training phases

TABLE I: Evaluation Results related to the cyberattacks against HTTP

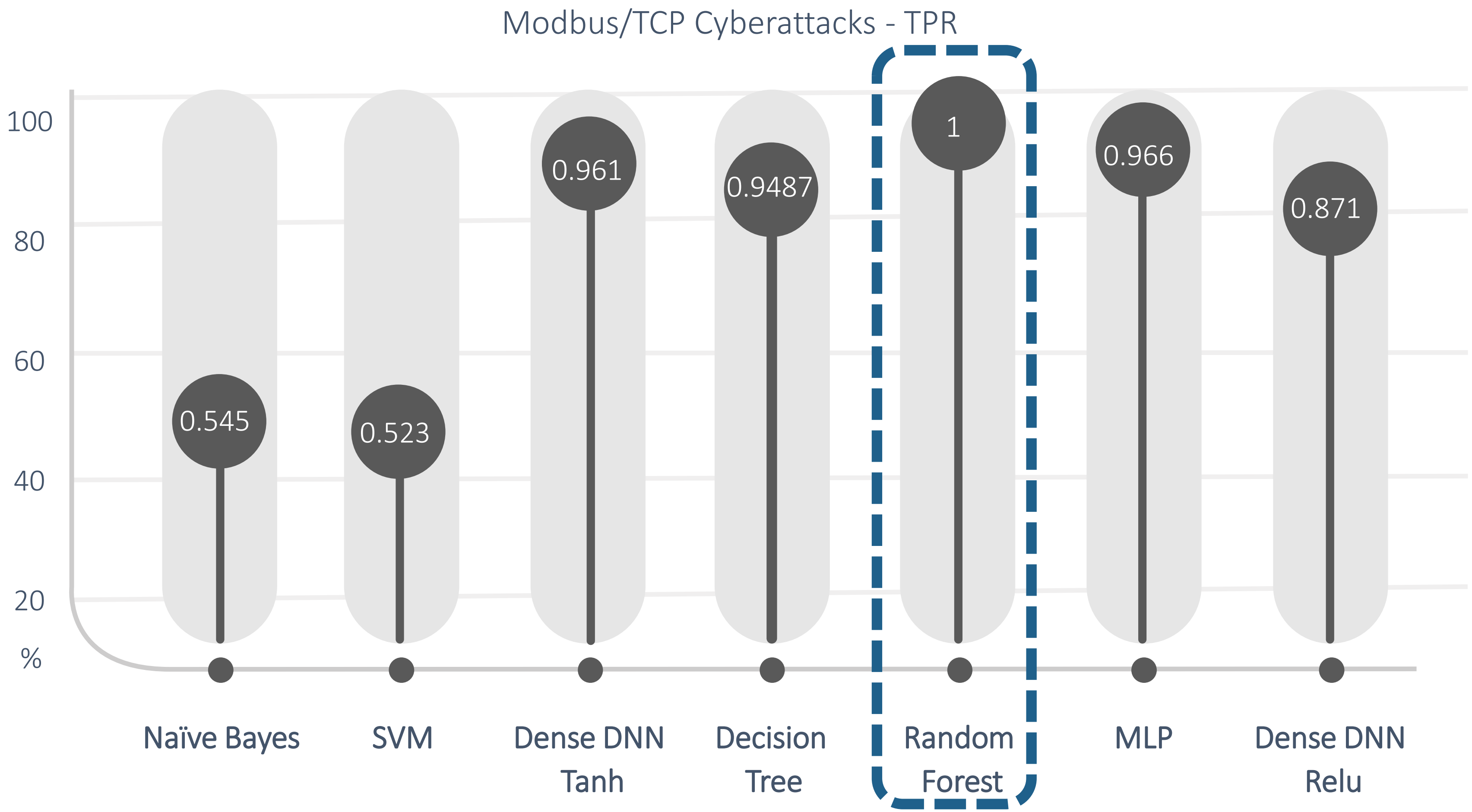
ML Method	Accuracy	TPR	FPR	F1
Decision Tree Classifier	**0.9644**	**0.9111**	**0.0222**	**0.9111**
Naive Bayes	0.7288	0.72888	0.27111	0.66744
SVM	0.89075	0.89075	0.10924	0.88027
Random Forest	0.92296	0.80740	0.04814	0.80740
MLP	0.90478	0.61915	0.05440	0.61915
Dense DNN Relu	0.90908	0.63633	0.05195	0.63633
Dense DNN Tanh	0.94074	0.85185	0.03703	0.85185

Evaluation



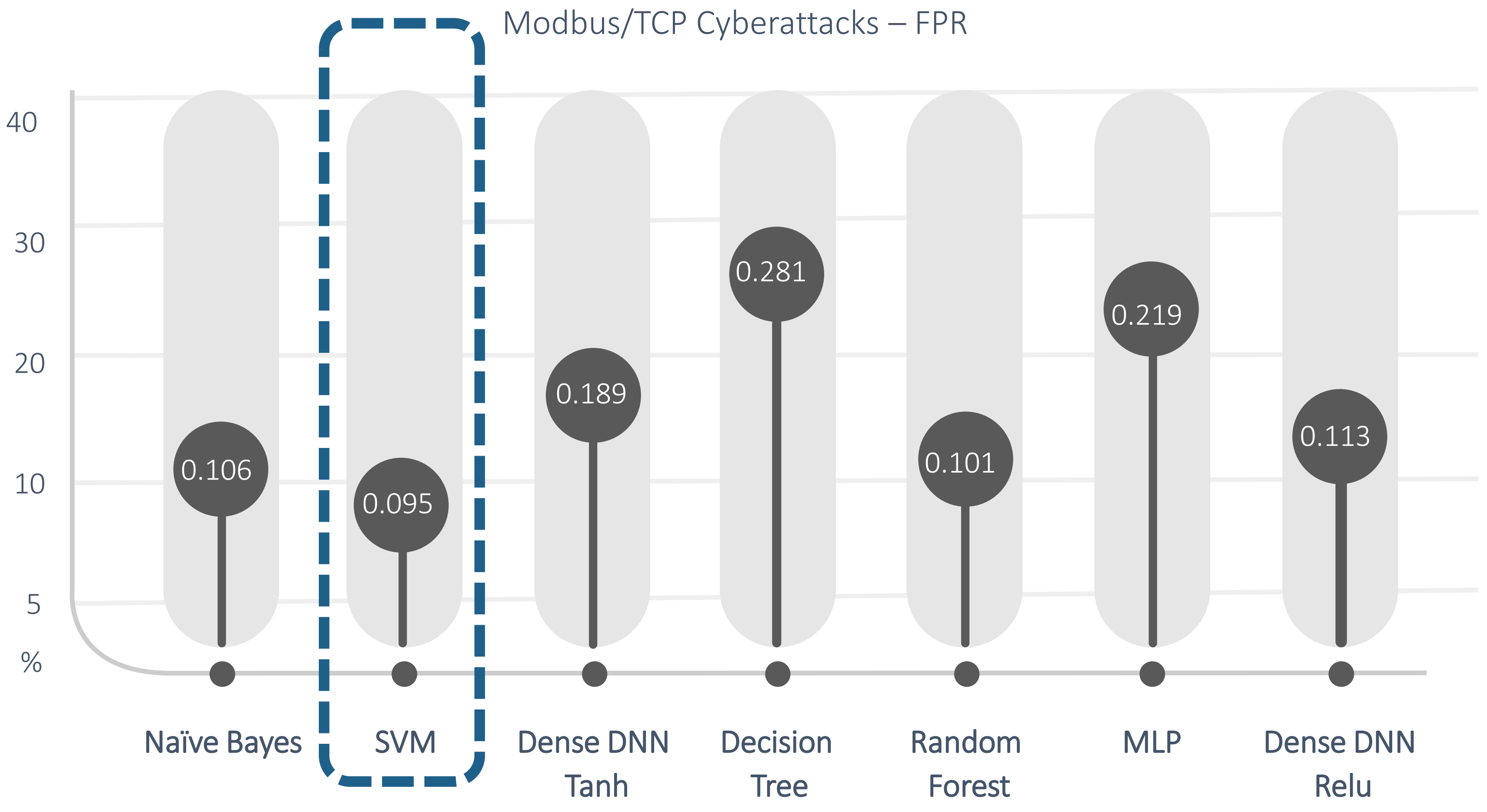
Modbus/TCP Cyberattacks - Accuracy
100
80
60
40
20
%
0.719
0.841
0.840
0.833
0.944
0.817
0.883
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation

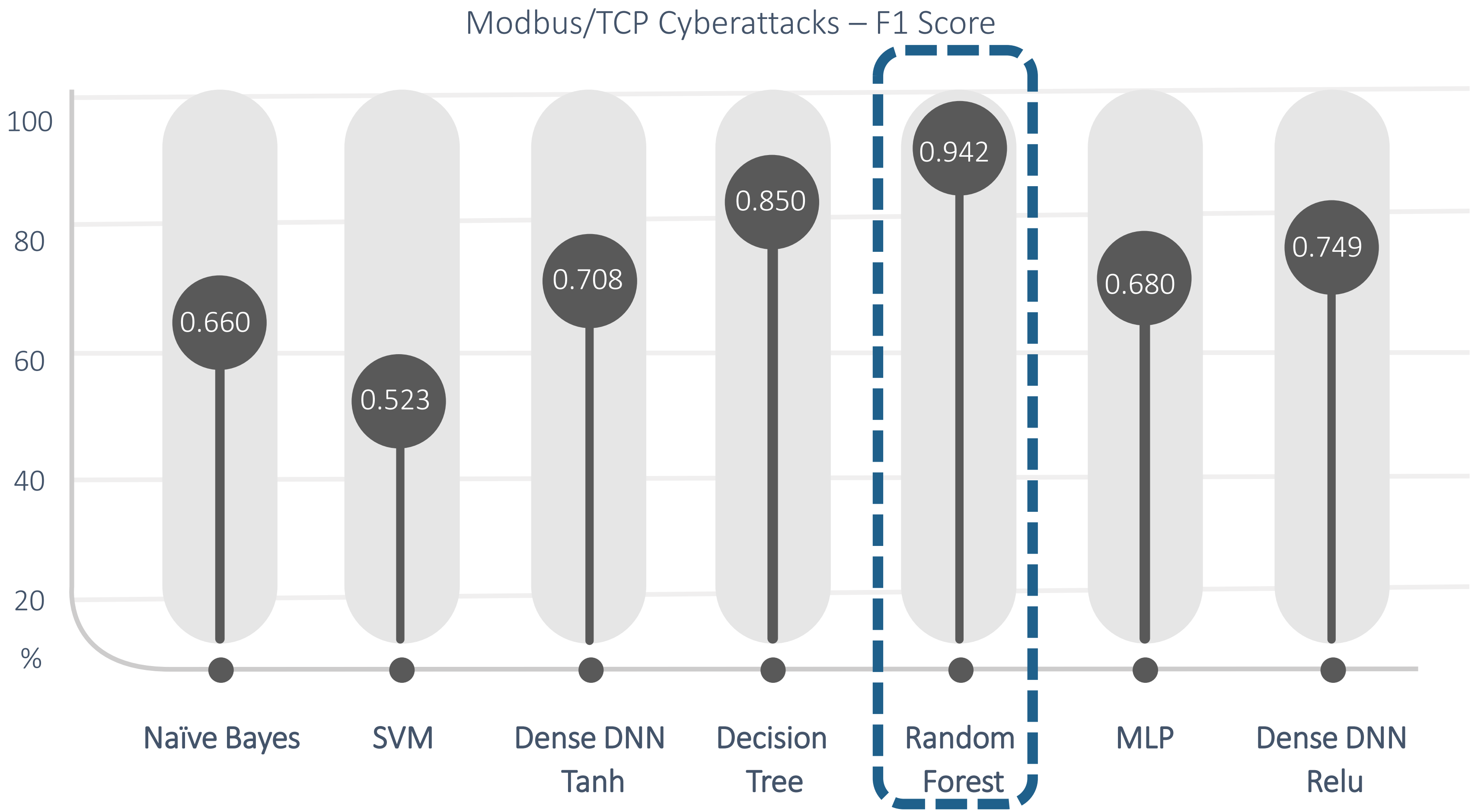


Modbus/TCP Cyberattacks - TPR
100
80
60
40
20
%
0.545
0.523
0.961
0.9487
1
0.966
0.871
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation


Modbus/TCP Cyberattacks – FPR
40
30
20
10
5
%
0.106
0.095
0.189
0.281
0.101
0.219
0.113
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation



Modbus/TCP Cyberattacks – F1 Score
100
80
60
40
20
%
0.660
0.523
0.708
0.850
0.942
0.680
0.749
Naïve Bayes
SVM
Dense DNN Tanh
Decision Tree
Random Forest
MLP
Dense DNN Relu

Evaluation

Modbus/TCP Cyberattacks – Aggregative Results

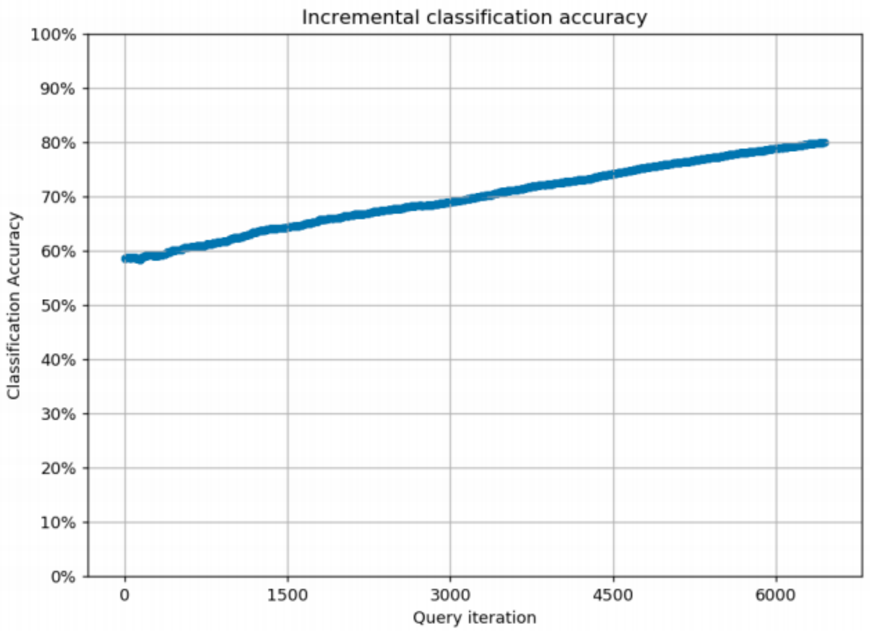


Fig. 3: Random Forest - Accuracy Increment during the re-training phases

TABLE II: Evaluation Results related to the cyberattacks against Modbus/TCP

ML Method	Accuracy	TPR	FPR	F1
Decision Tree Classifier	0.83333	0.94827	0.28160	0.85051
Naive Bayes	0.71982	0.54597	0.10632	0.66086
SVM	0.841	0.523	0.095	0.523
Random Forest	**0.94454**	**1**	**0.10166**	**0.94250**
MLP	0.81797	0.96663	0.21924	0.68018
Dense DNN Relu	0.88357	0.87158	0.11341	0.74989
Dense DNN Tanh	0.84078	0.96122	0.18952	0.70827

Conclusions

Final Remarks

IoMT and Legacy Healthcare Systems

IoMT and legacy helatcare systems are characterised by severe security issues

EHR - HTTP

EHRs are threatened by HTTP cyberattacks

IoMT - Modbus

Smart medical devices are threatened by Modbus/TCP cyberattacks

Accurate Intrusion Detection

Necessity of intrusion detection mechanisms – lack of available datasets

Active Leering Approach

The proposed IDPS is re-trained by itself. It uses the pool-based zampling method and the uncertainty sampling strategy with the entropy criterion.

Average Accuracy

95%

The evaluation analysis demonstrates the efficiency of the proposed IDPS against HTTP and Modbus/TCP cyberattacks, showing additionally how the overall accuracy is increased during the re-training phases.

Future Plans

Optimization of the proposed active learning approach with reinforcement learning techniques, thus eliminating the presence of the cybersecurity expert.

Thank You & Q/A

Contact us

pradoglou@uowm.gr

https://www.spear2020.eu/

https://www.linkedin.com/company/spear2020/

https://www.youtube.com/channel/UCw6-d5G01ToBhCmaUnHIcpw

Thank You

Q/A ?

This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No. 787011 (SPEAR).