“An Anomaly Detection Mechanism for IEC 60870-5-104

Panagiotis Sarigiannidis

University of Western Macedonia

psarigiannidis@uowm.gr

Authors

Under the H2020 SPEAR Project

Panagiotis Radoglou Grammatikis, Panagiotis Sarigiannidis

University of Western Macedonia

Antonios Sarigiannidis, Dimitrios Margounakis, Apostolos Tsiakalos

SIDROCO Holdings

Georgios Efstathopoulos

0INFINITY LIMITED

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

Introduction

An IDS system for the IEC 60870-5-104 protocol

Summary

After the study of the IEC 60870-5-104 (IEC-104) protocol, an Intrusion Detection System (IDS) for the IEC-104 protocol is provided. The efficiency of the proposed IDS is demonstrated by the Accuracy and the F1 score metrics that reach 98% and 87%, respectively.

Smart Grid Security Status

The critical infrastructures and especially the electrical grid suffers from severe cybersecurity and privacy issues due to its insecure legacy and IoT assets.

IEC 60870-5-104 Security Issues

IEC 60870-5-104 does not include essential security mechanisms, such as authentication and authorization, thus enabling various cyberattacks.

IEC 60870-5-104 IDS

It is based on access control and outlier detection mechanisms.

Related Work

Previous Research Works related to IEC 60870-5-104

01 — 2014

- P. Maynard, K. McLaughlin, and B. Haberler, “Towards understandingman-in-the-middle attacks on iec 60870-5-104 scada networks,” in2ndInternational Symposium for ICS & SCADA Cyber Security Research2014 (ICS-CSR 2014) 2, 2014, pp. 30–42

02 — 2017

- E. Hodo, S. Grebeniuk, H. Ruotsalainen, and P. Tavolato, “Anomalydetection for simulated iec-60870-5-104 trafiic,” inProceedings of the12th International Conference on Availability, Reliability and Security,2017, pp. 1–7.

03 — 2018

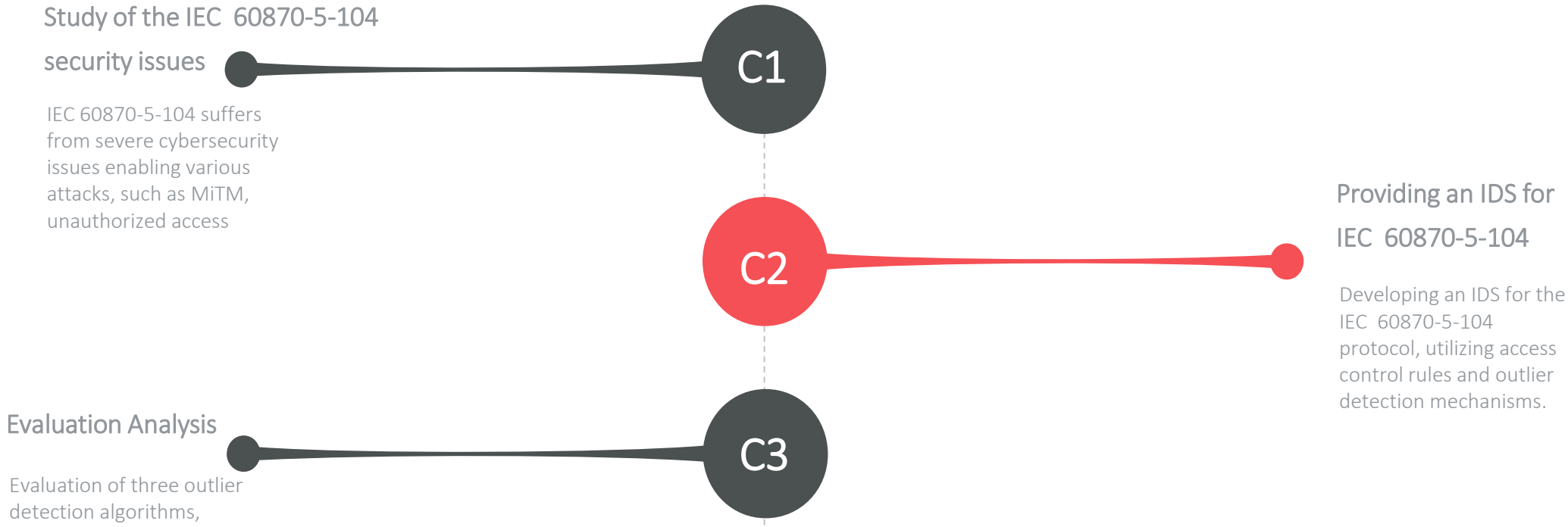
- C.-Y. Lin and S. Nadjm-Tehrani, “Understanding iec-60870-5-104 trafficpatterns in scada networks,” inProceedings of the 4th ACM Workshopon Cyber-Physical System Security, 2018, pp. 51–60

04 — 2019

- P. Radoglou-Grammatikis, P. Sarigiannidis, I. Giannoulakis, E. Kafetzakis, and E. Panaousis, “Attacking iec-60870-5-104 scada systems,” in2019 IEEE World Congress on Services (SERVICES), vol. 2642-939X,July 2019, pp. 41–46.

Contributions

3 Main Contributions



Background

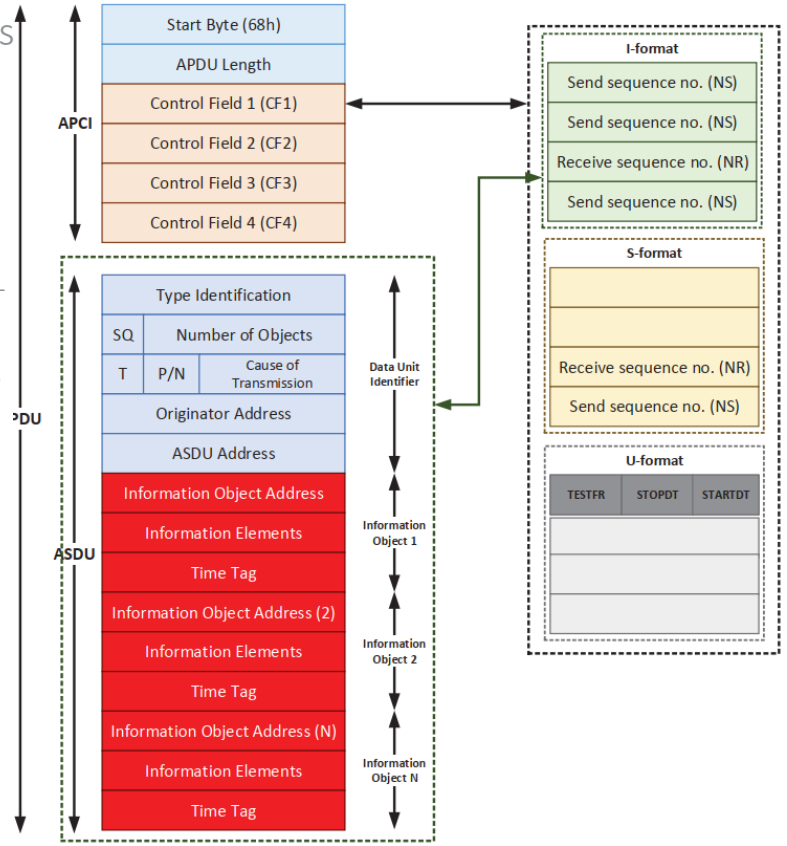
IEC 60870-5-104 security, Typical IDS Architecture, Intrusion Detection Techniques, ML-based Detection and Outlier Detection Algorithms

IEC 60870-5-104 Security

Lack of Authentication and Authorisation

IEC 60870-5-104 Security Issues

- ✓ A severe security issue of IEC-104is the transmission of data without any encryption mechanism, thus making it possible to execute traffic analysis and MiTM attacks. In addition, many IEC-104 commands, such as reset commands, interrogation commands, read commands do not integrate authentication and authorisation procedures, thereb yallowing the unauthorised access.
- ✓ This vulnerability is crucial since a cyberattacker is capable of controlling the field devicesand possibly, the overall operation of the infrastructure.



Risk Assessment based on

P. Radoglou-Grammatikis, P. Sarigiannidis, I. Giannoulakis, E. Kafetzakis, and E. Panaousis, “Attacking iec-60870-5-104 scada systems,” in2019 IEEE World Congress on Services (SERVICES), vol. 2642-939X,July 2019, pp. 41–46.

Traffic Analysis

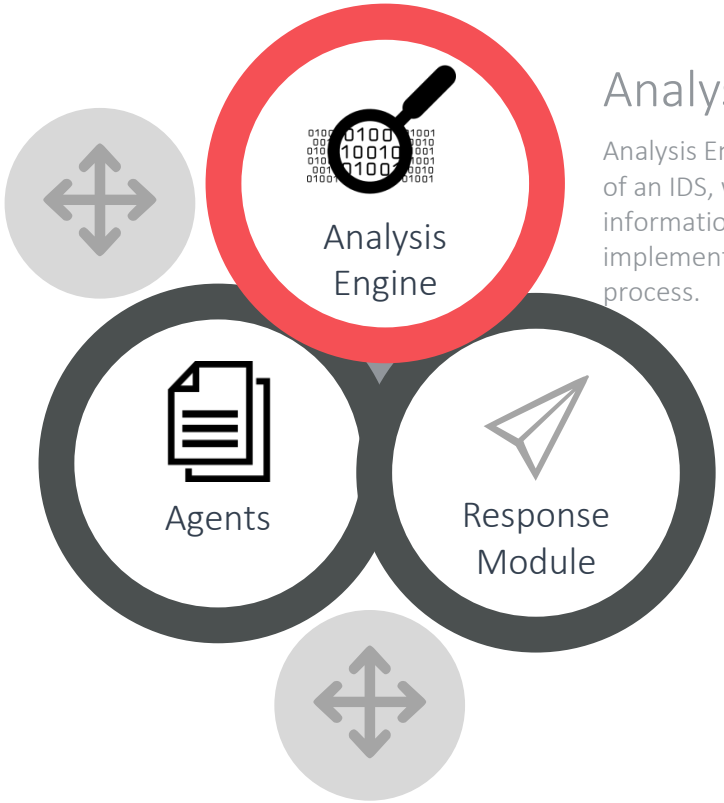
DoS

Unauthorized Access

MITM

Typical IDPS Architecture

3 Main Components



Analysis Engine

Analysis Engine is the core component of an IDS, which receives the information of the various Agents and implements the intrusion detection process.

Agents

Agents undertake to monitor the examined infrastructure, thus collecting and sometimes pre-processing the necessary data for the detection process.

Response Module

The Response Module notifies the responsible operator. It can perform automate mitigation processes.

Intrusion Detection Techniques

3 Main Intrusion
Detection Techniques

Signature-based

specific rules called signatures that reflect malicious patterns. If the characteristics of the monitoring data match with those of the signatures, then a possible security violation takes place.

Anomaly-based

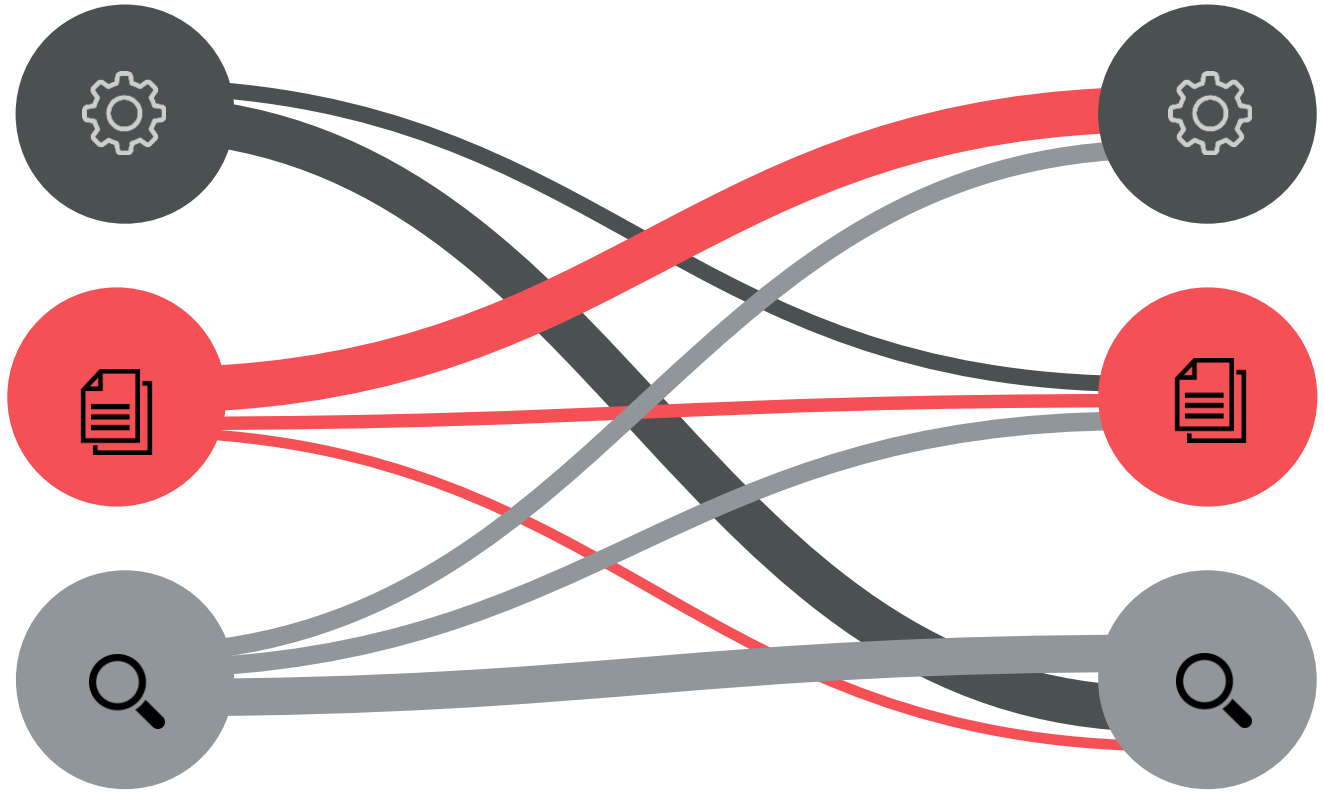
The anomaly-based detection applies statistical analysis and Artificial Intelligence (AI) methods.

Specification-based

Set of rules called now specifications that define the normal operation of the monitored system/infrastructure. If the characteristics of the monitored data do not agree with those of the specifications, then a security violation is carried out.

ML-Base Detection

Three Main Steps



Preprocessing

Processes appropriately the input data so that it will be in accordance with the corresponding ML model. Usually, data-preprocessing methods are applied, such as min-max scaling, normalisation, standardisation, robust scaler and max abs scaler

Training

Supervised detection methods, unsupervised/oulier detection methods and semi-supervised/novelty detection methods

Prediction

The ML model can be deployed in order to predict unknown data after the execution of the same pre-processing tasks of the first phase

Outlier Detection Algorithms

Three Algorithms

One-Class SVM

Bernhard Schölkopf, Robert C Williamson, Alex J Smola, John Shawe-Taylor, John C Platt in 2000

One-Clas Support Vector Machine (SVM) aims to find a hyperplane that can separate the vast majority of data from the origin in the projected high dimensional space without making any assumptions about their distribution. In particular, One-Class SVM separates all the data points from the origin (in feature space) and maximises the distance from this hyperplane to the origin. This results in a binary function, which captures regions in the input space where the probability density of the data lives. The idea of One-class SVM for anomaly detection is to find a function that is positive for regions with a high density of points, and negative for small densities.

Local Oilier Factor (LOF)

Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng and Jörg Sander in 2000

LOF relies on the concept of a local density, where locality is given by k nearest neighbors, whose distance is utilised to estimate the density. By comparing the local density of an object to the local densities of its neighbors, one can identify regions of similar density, and points that have a substantially lower density than their neighbors. These are considered to be outliers.

Isolation Forest

Fei Tony Liu, Kai Ming Ting and Zhi-Hua Zhou in 2008.

The Isolation Forest algorithm finds anomalies by deliberately “overfitting” models that memorize each data point. Since outliers have more empty space around them, they take fewer steps to memorize. The algorithm is using full decision trees (every leaf is a single data point) and we measure the path length between the root and each leaf (data point). The final measure for each data point would be the average path length. Abnormal data points should be classified easily thus the average path should be relatively short.

Proposed IEC 60870-5-104 IDS

Architecture & Evaluation

Proposed IEC 60870-5-104 IDS

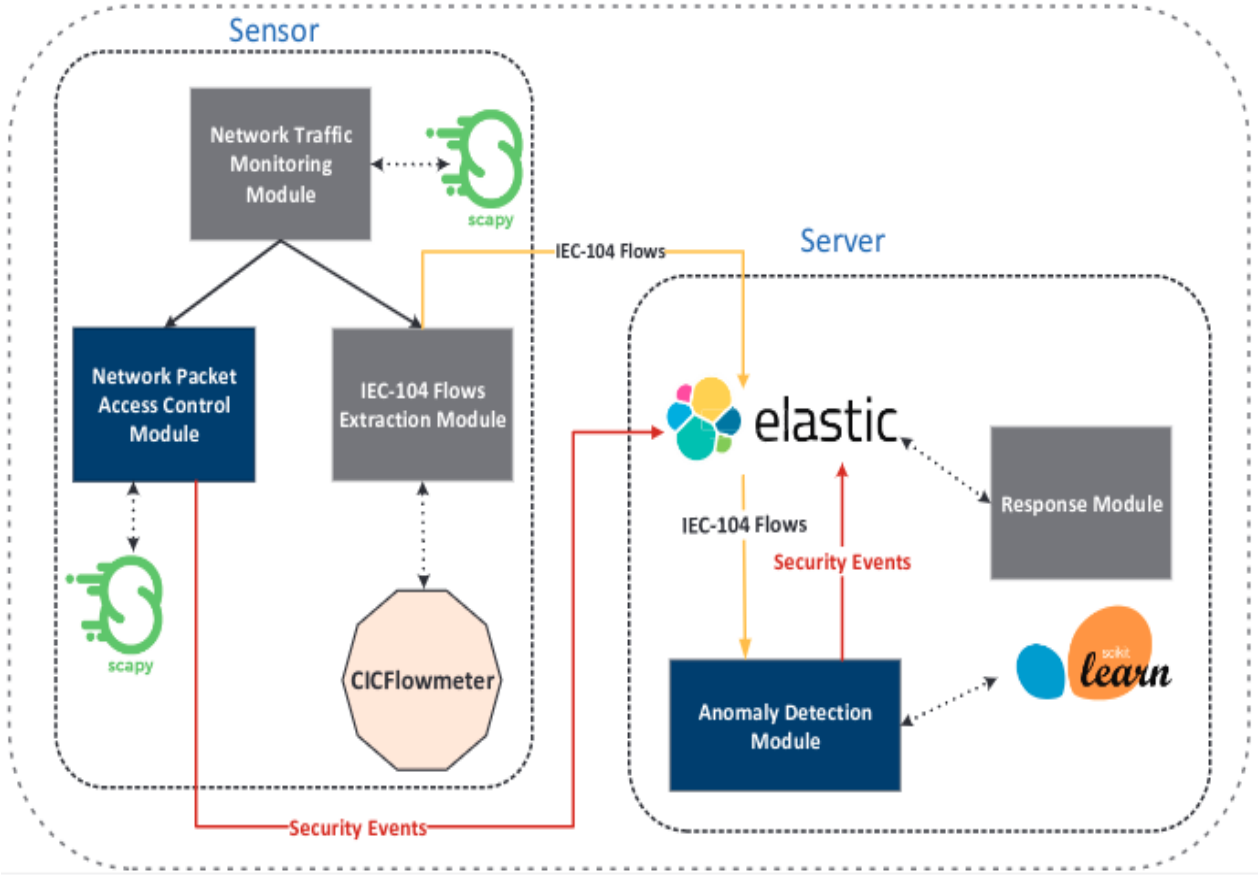
Architecture

Sensor

It consist of three modules, namelya)Network Traffic Monitoring Module,b) Network Packet AccessControlandc) IEC-104 Flows Extraction Moduleresponsiblerespectively for monitoring and analysing the entire networktraffic generated in the infrastructure.

Server

It is a centralized point where the anomaly detection processes take place, and the security events are stored. In particular, it is composed of an Elasticsearch database, the Anomaly Detection Module and the Response Module.



Network Traffic Monitoring Module

It relies on Scapy and is responsible for monitoring and capturing the overall network traffic based on a predefined frequency.

Network Packet Access Control Module

Based on a whitelist, it applies access control rules regarding the IP and MAC addresses as well as the TCP and UDP ports. Thus, it generates security events that are stored in Elasticsearch.

IEC-14 Flows Extraction Module

It extracts IEC 60870-5-104 network flows based on CICFlowMeter. Different flow timeouts can be used, thus adapting the network flow statistics.

Anomaly Detection Module

Applies the Outlier Detection Algorithms in order to recognize IEC 60870-5-104 anomalies. Thus, it generates security events that are stored in Elasticsearch.

Response Module

It informs the user regarding the various security events via Kibana. Moreover, through Kibana, it also generates statistic charts that assist the user in understanding better the security status of the monitored infrastructure.

Evaluation

Evaluation Methodology

Step 1

Normal IEC 60870-5-104 Network Flows

Emulated substation environment equipped with real industrial devices such as RTUs, IEDs. Via SPAN, the Data Monitoring Module can receive the overall DNP3 traffic, extracting the normal network flow statistics.

Step 2

Malicious IEC 60870-5-104 Network Flows

P. Maynard, K. McLaughlin, and S. Sezer, "An open framework fordeploying experimental scada testbed networks," in5th InternationalSymposium for ICS & SCADA Cyber Security Research 2018 5, 2018,pp. 92–101.

Step 3

Feature Selection & Training

- Total packets in the forward direction
- Total size of the packets in the backward direction
- Standard deviation size of the packets in the forward direction
- Number of the flow bytes per second
- Maximum time between two packets sent in the flow
- Minimum length of a packet
- Average number of bytes in a sub-flow in the backward direction
- Maximum time where a flow was active before becoming idle

Step 4

Evaluation

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

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

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

$$Precision = \frac{TP}{TP + FP}$$

Evaluation

The Outlier Detection Evaluation Results for flow-timeout 15s

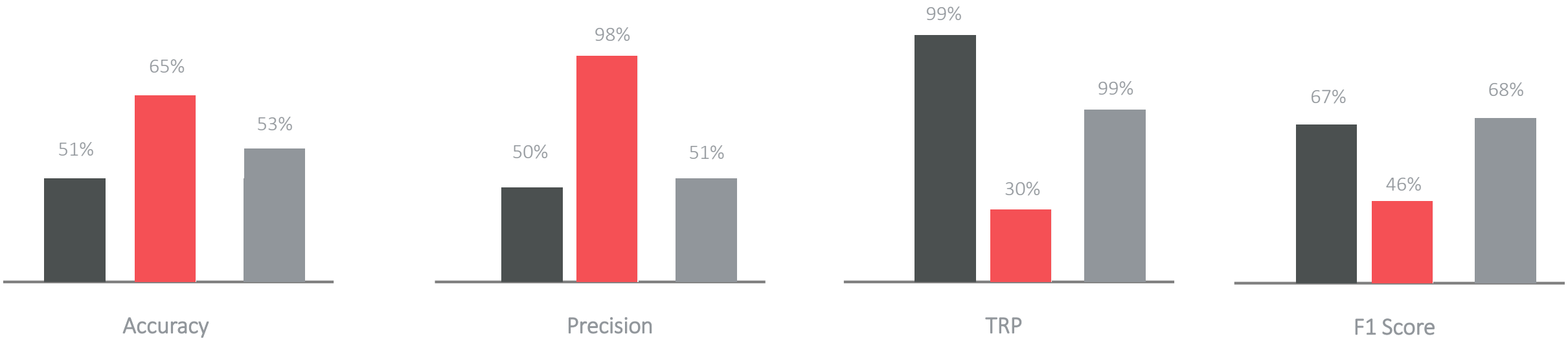


TABLE I: The Outlier Detection Evaluation Results for flow-timeout 15s.

Model	Accuracy	Precision	TPR	F1
OS-SVM	0.519	0.509	0.993	0.673
LOF	0.65	0.98	0.30	0.46
Isolation Forest	0.536	0.519	0.992	0.6817

OC-SVM LOF Isolation Forest

Evaluation

The Outlier Detection Evaluation Results for flow-timeout 30s

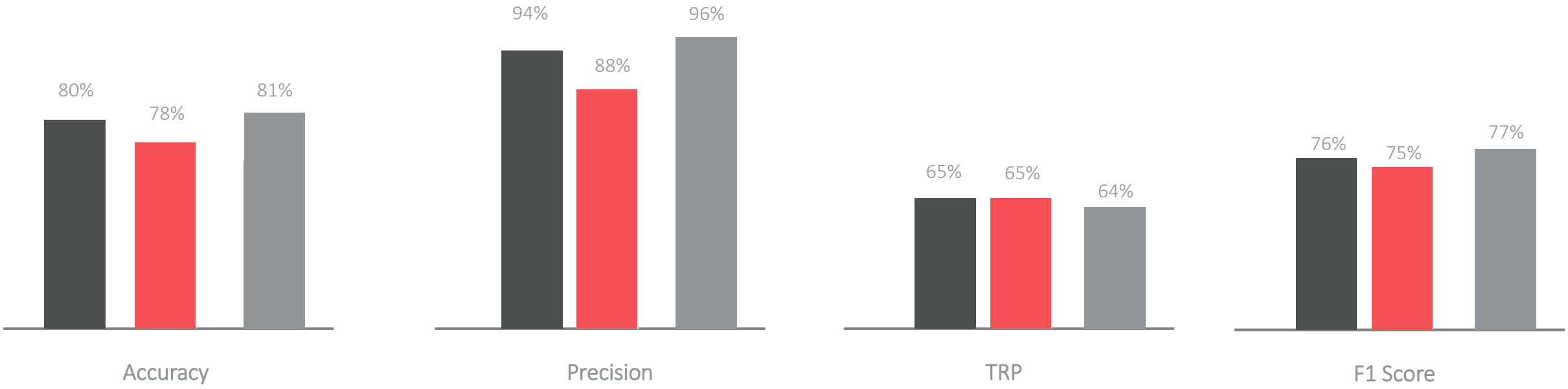


TABLE II: The Outlier Detection Evaluation results for flow-timeout 30s.

Model	Accuracy	Precision	TPR	F1
OS-SVM	0.805	0.943	0.650	0.769
LOF	0.783	0.886	0.650	0.750
Isolation Forest	0.811	0.964	0.647	0.774

Evaluation

The Outlier Detection Evaluation Results for flow-timeout 60s

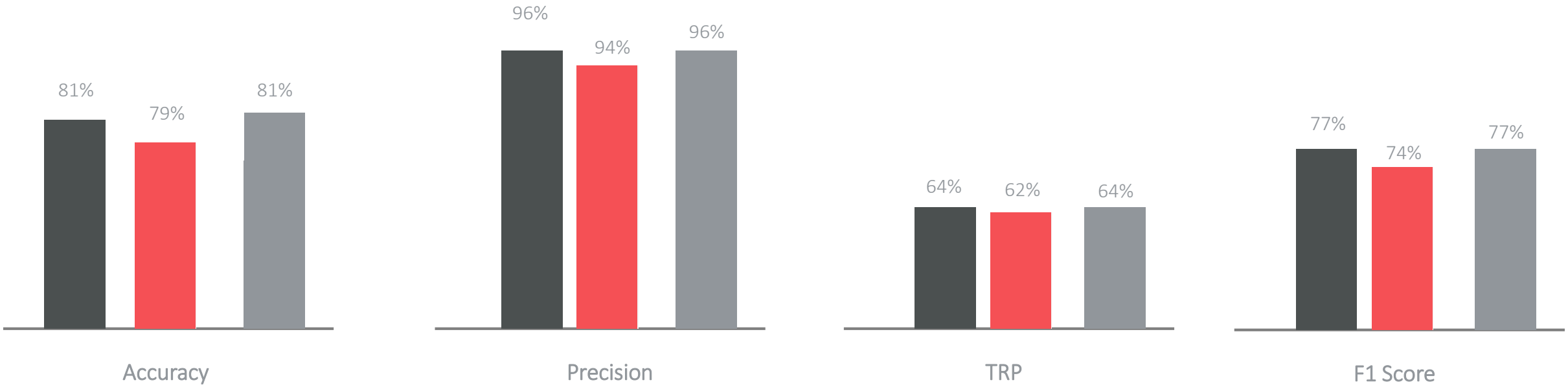


TABLE III: The Outlier Detection Evaluation for flow-timeout 60s.

Model	Accuracy	Precision	TPR	F1
OS-SVM	0.811	0.964	0.647	0.774
LOF	0.790	0.941	0.620	0.747
Isolation Forest	0.812	0.964	0.647	0.775

Evaluation

The Outlier Detection Evaluation Results for flow-timeout 120s

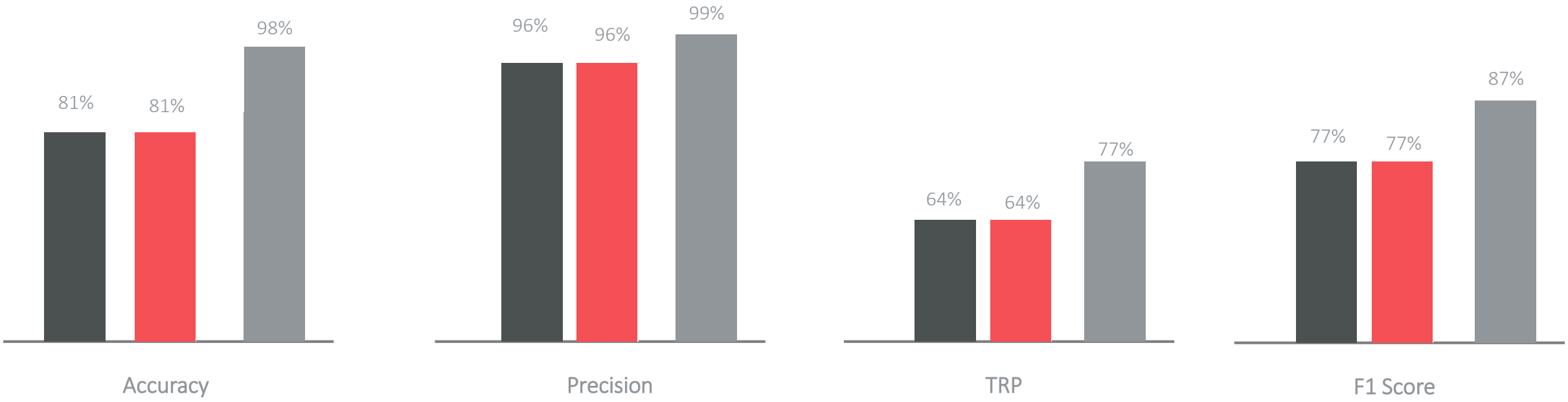


TABLE IV: The Outlier Detection Evaluation results for flow-timeout 120s.

Model	Accuracy	Precision	TPR	F1
OS-SVM-Linear	0.812	0.962	0.647	0.774
LOF	0.812	0.964	0.647	0.775
Isolation Forest	0.982	0.990	0.777	0.875

OC-SVM LOF Isolation Forest

Conclusions

Study of the IEC 60870-5-104 security issues

IEC 60870-5-104 suffers from severe cybersecurity issues enabling various attacks, such as MiTM, unauthorized access

C1

C2

Providing an IDS for IEC 60870-5-104

After investigating IEC-104 security issues, we provided a relevant IDS, which applies access control and outlier detection mechanisms in order to detectIEC-104 anomalies.

Evaluation Analysis

C3

According to the evaluationresults, when the flow-timeout value is equal to 120s, theIsolation Forest method achieves the highest Accuracy, Precision, TPR and F1 that reach0.982,0.990,0.777and0.875respectively

Thank You & Q/A

Contact us

psarigiannidis@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).