5GCIDS An Intrusion Detection System for 5G Core with AI and Explainability Mechanisms

P. Radoglou-Grammatikis, G. Nakas, G. Amponis, S. Giannakidou, T. Lagkas, V. Argyriou, S.Goudos,
P. Sarigiannidis

Authors

George Nakas, Panagiotis Radoglou-Grammatikis, George Amponis, Sofia Giannakidou
Research and Development Department
K3Y Ltd., Bulgaria
{gnakas, pradoglou, gamponis}@k3y.bg

Thomas Lagkas
Department of Computer Science, International Hellenic University, Greece
tlagkas@cs.ihu.gr

Vasileios Argyriou
Department of Networks and Digital Media
Kingston University, United Kingdom
vasileios.argyriou@kingston.ac.uk

Sotirios Goudos
School of Physics
Aristotle University of Thessaloniki, Greece
sgoudo@physics.auth.gr

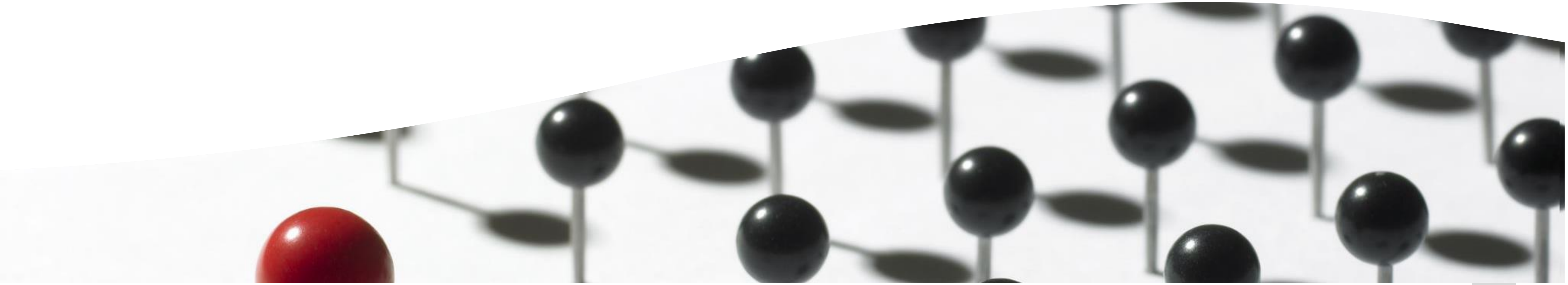
Panagiotis Radoglou-Grammatikis, Panagiotis Sarigiannidis
Department of Electrical and Computer Engineering
University of Western Macedonia, Greece
{pradoglou, psarigiannidis}@uowm.gr

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101097122 (ACROSS).

Outline

- Motivation
- Threat Modelling
- Design & Implementation
- Evaluation
- Conclusion

Motivation



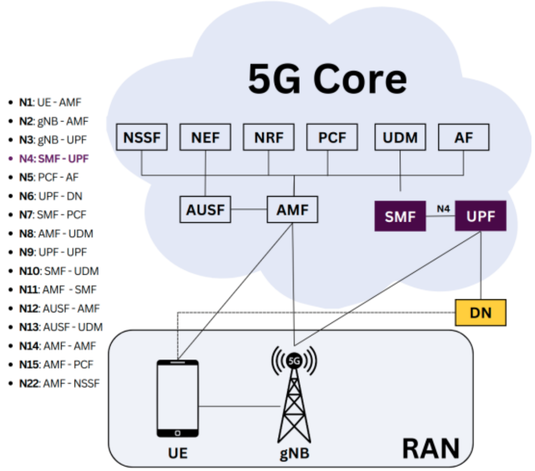
Motivation[1/2] – Introduction

- The progression of fifth-generation (5G) networks provides multiple advantages, such as faster speed, reduced latency and increased capacity.

- However, despite the various benefits, 5GC is prone to several cyberthreats that can result in catastrophic effects.

- Signature/specification-based IDS uses pre-defined rules that can match malicious patterns, while anomaly-based IDS use statistical analysis and AI methods in order to discriminate potential abnormalities.

- AI-based IDS suffer from explainability issues, given that the security administrators cannot comprehend and therefore trust the decision-making process of the AI models.



Motivation[2/2] – Aim of this work

The development of an Intrusion Detection System (IDS) called **5GCIDS**, focusing on the Packet Forwarding Control Protocol (PFCP), which is utilised for the communication between SMF and UPF.

The main contributions include:

- **Implementation of PFCPFlowMeter**: A new flow statistics generator is implemented, generating bidirectional flow statistics related to the application-layer attributes of PFCP.
- **Detection of PFCP-related Attacks:** Through ML methods, namely eXtreme Gradient Boosting (XGBoost) and decision tree, the proposed IDS can successfully detect four cyberattacks against PFCP.
- **Development of AI-powered IDS with Explainability Mechanisms**: A new AI-powered IDS with an explainability mechanism is implemented. For this purpose, the SHAP (SHapley Additive exPlanations) method is used.

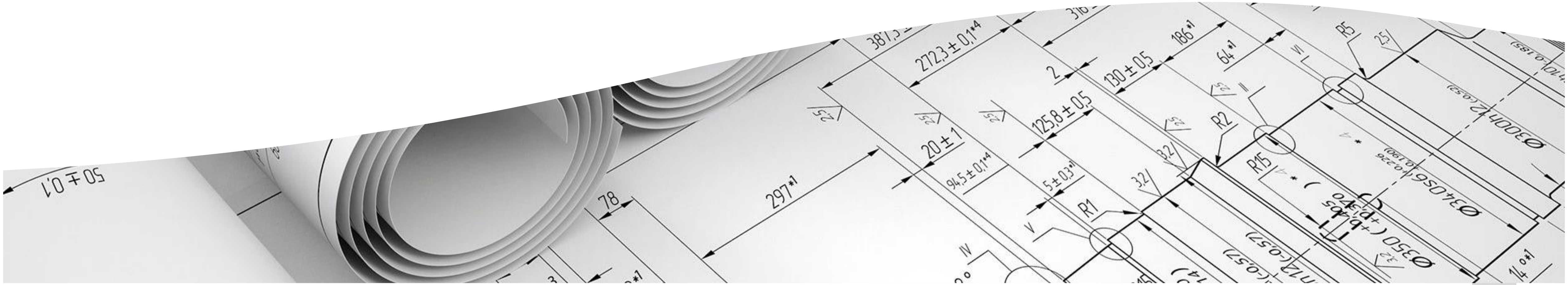
Threat Modelling

Threat Modelling – PFCP Attacks

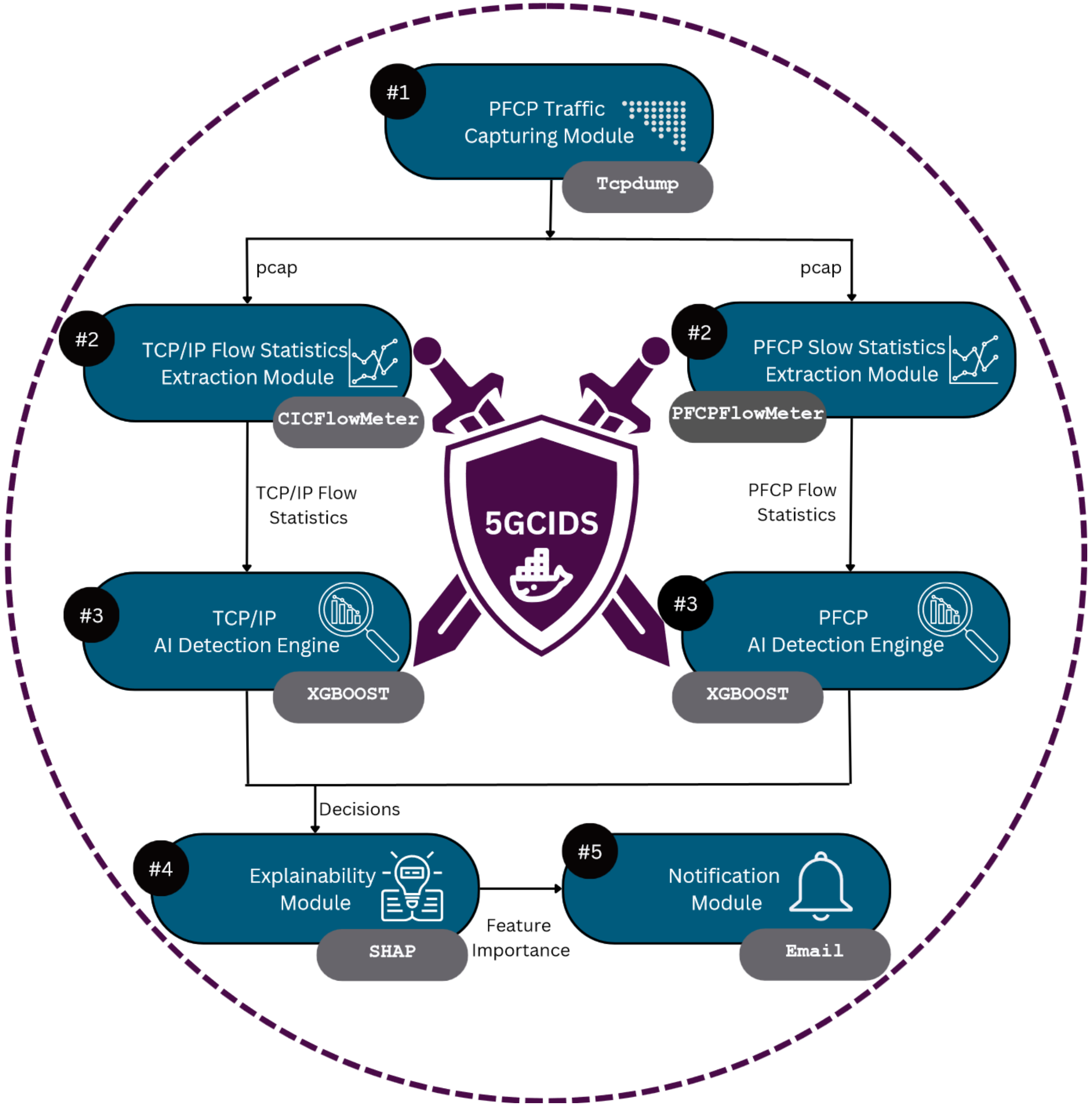
Our threat model focuses mainly on the N4 interface between SMF and UPF, where PFCP is used. Therefore, we investigate four PFCP-related cyberattacks:

- **PFCP Session Establishment DoS Attack:** This attack floods the UPF with valid Session Establishment and Heartbeat Requests, aiming to drain its resources.
- **PFCP Session Deletion DoS Attack:** This attack aims to disconnect a UE from the DN. The attack script disrupts client-DN PDU sessions without disconnecting the UE from the 5G RAN or 5GC network.
- **PFCP Session Modification DoS Attack (DROP):** This attack aims to invalidate session-specific packet handling rules in order to disconnect a targeted UE from the DN. During rules update, the UPF removes Forwarding Action Rule (FAR) entries associated with the Tunnel Endpoint Identifier (TEID) and base station IP address.
- **PFCP Session Modification DoS Attack (DUPL):** This attack leverages the DUPL flag in the Apply Action field to compel the UPF to duplicate session rules, resulting in the creation of multiple paths for data originating from a single source.

Design & Implementation



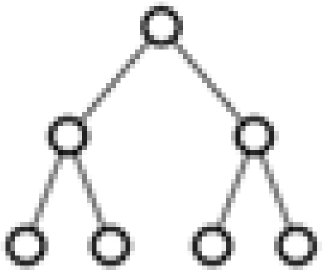
Methodology – 5GCIDS Architecture



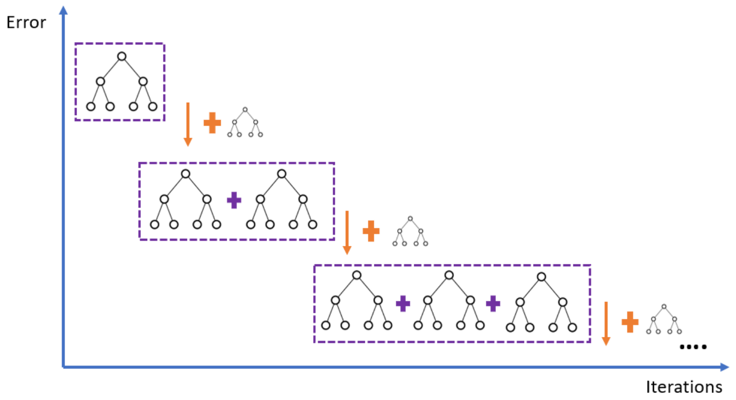
Methodology – Detecting PFCP Cyberattacks

For the detection process two modules are used:

- **PFCP AI Detection Engine** uses a decision tree model, which can be described as a set of if-else statements, categorizing the various instances into particular classes based on the various features.

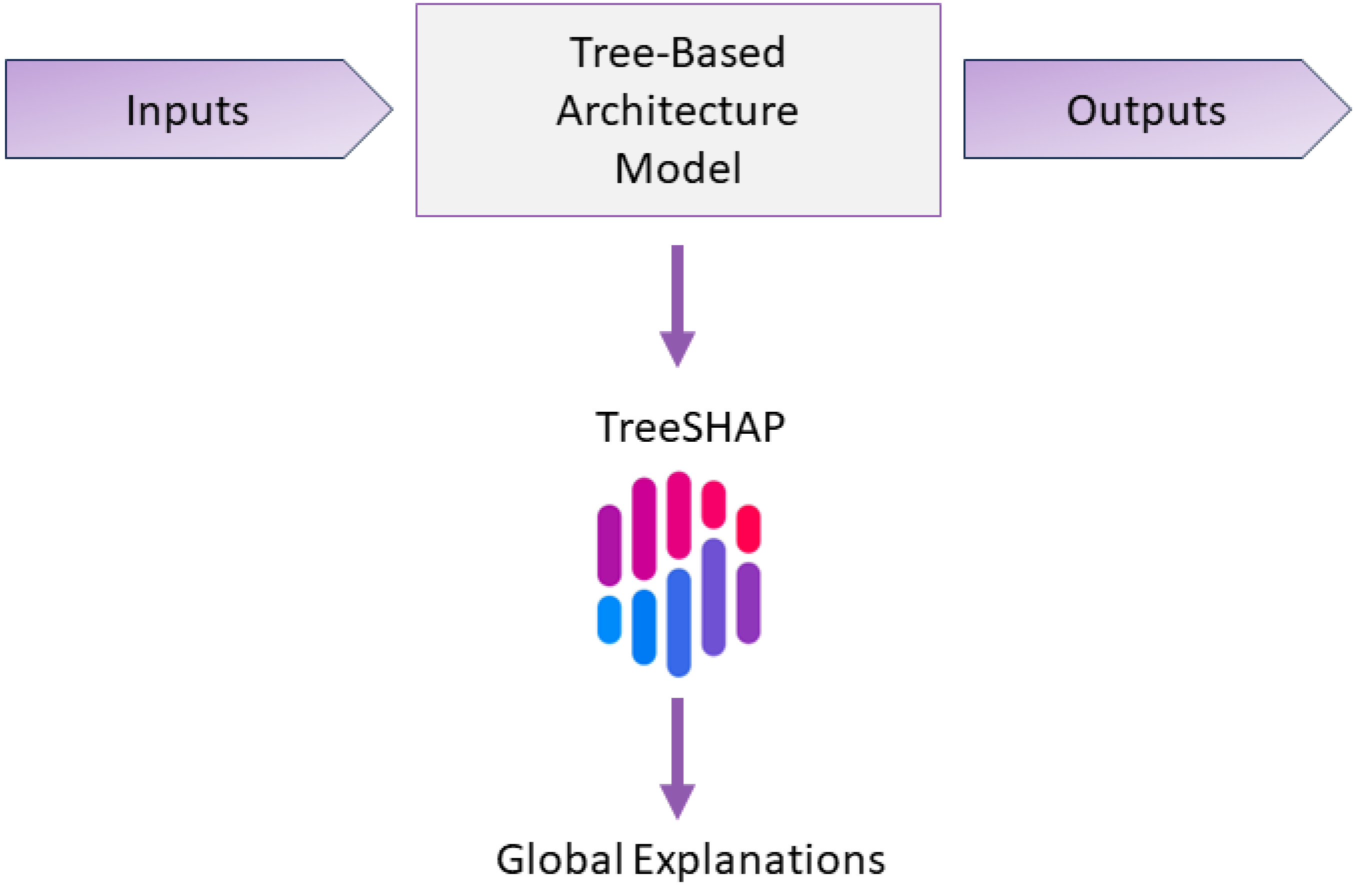


- **TCP/IP AI Detection Engine** uses a model based on XGBoost, which is an ensemble ML method that is able to produce an ensemble based on weak learners. In this case, short decision trees are used. In particular, the ensemble is sequentially constructed by training each weak learner with the data samples that are previously classified incorrectly.



Methodology – Explainability of the Predictions

5GCIDS Explainability Module



Evaluation

Evaluation[1/4] – Evaluation Data

- Multi-class classification metrics:

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

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

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

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN}$$

- Based on the tree-based methods used for the detection process, the TreeSHAP method is chosen in order to provide global explanations. TreeSHAP is a model-specific explanation method designed for tree-based models, based on Shapley values.

Evaluation[2/4] – ML/DL Models comparison

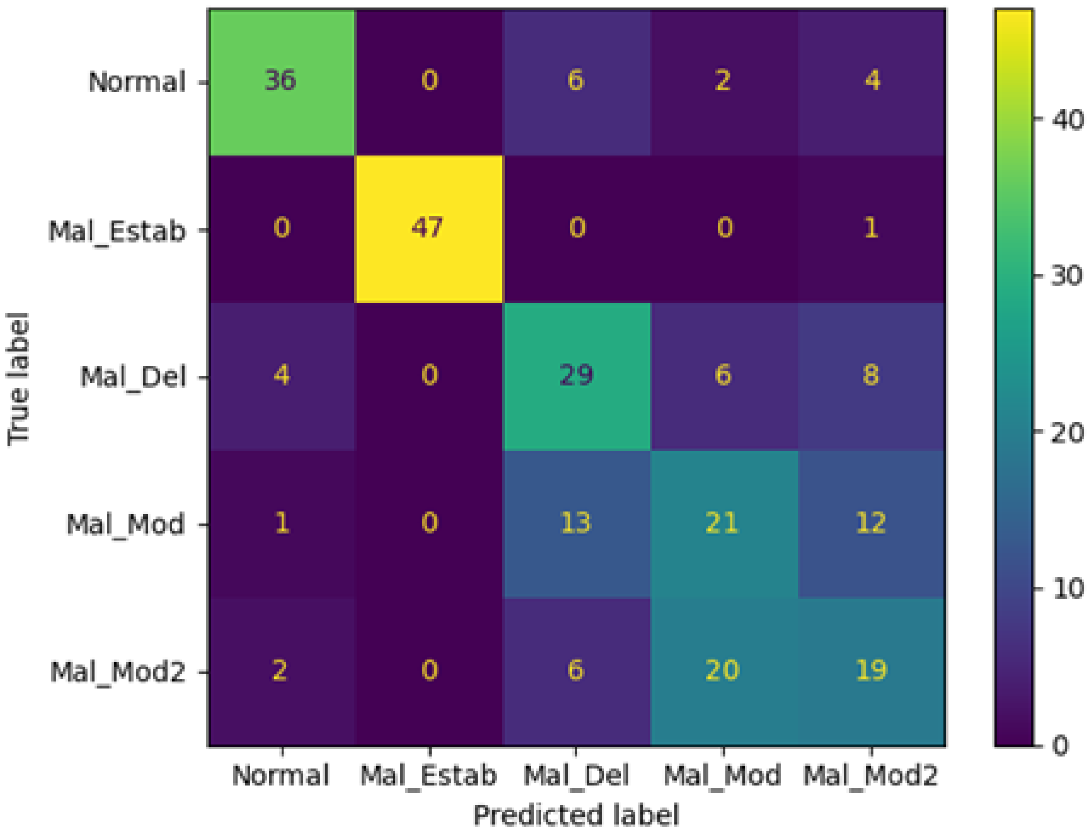
ML/DL Method	ACC	TPR	FPR	F1
Logistic Regression	0.536	0.533	0.116	0.486
LDA	0.536	0.533	0.116	0.492
QDA	0.54	0.538	0.115	0.463
KNN	0.612	0.61	0.097	0.614
Decision Tree	**0.641**	**0.639**	**0.09**	**0.642**
Naive Bayes	0.54	0.538	0.115	0.492
SVM	0.532	0.529	0.117	0.482
AdaBoost	0.62	0.618	0.095	0.627
Gradient Boosting	0.62	0.618	0.095	0.629
Random Forest	0.629	0.627	0.093	0.633
Extra Trees	0.633	0.631	0.092	0.641
XGBoost	0.603	0.601	0.099	0.612
LightGBM	0.565	0.563	0.108	0.579
MLP	0.536	0.533	0.116	0.486

- Comparative Study of ML/DL Models trained with Application Layer PFCP Flow Statistics/Features.

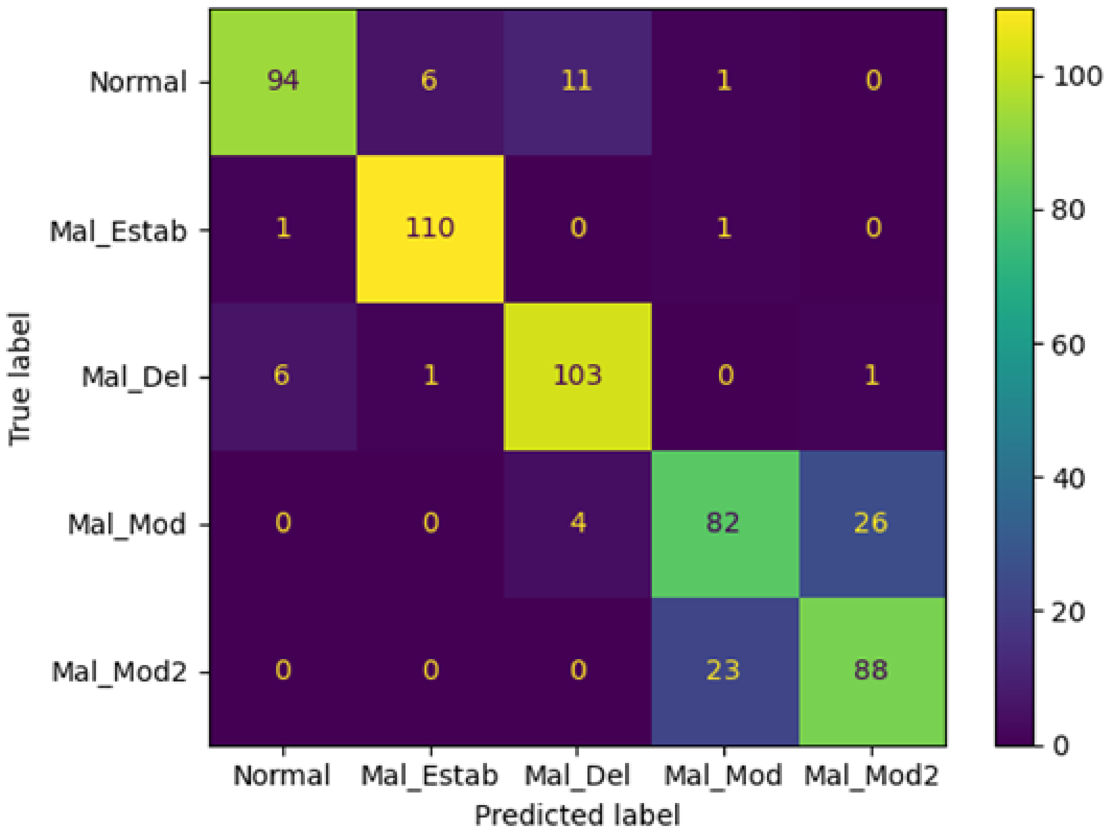
ML/DL Method	ACC	TPR	FPR	F1
Logistic Regression	0.559	0.559	0.11	0.535
LDA	0.566	0.566	0.108	0.565
QDA	0.547	0.546	0.113	0.508
KNN	0.769	0.769	0.058	0.769
Decision Tree	0.817	0.817	0.046	0.817
Naive Bayes	0.487	0.487	0.128	0.425
SVM	0.489	0.489	0.128	0.432
AdaBoost	0.681	0.68	0.08	0.667
Gradient Boosting	0.844	0.844	0.039	0.844
Random Forest	0.83	0.83	0.043	0.83
Extra Trees	0.832	0.832	0.042	0.831
XGBoost	**0.855**	**0.855**	**0.036**	**0.854**
LightGBM	0.841	0.84	0.04	0.84
MLP	0.559	0.559	0.11	0.536

- Comparative Study of ML/DL Models trained with TCP/IP Flow Statistics/Features.

Evaluation[3/4] – Confusion Matrix

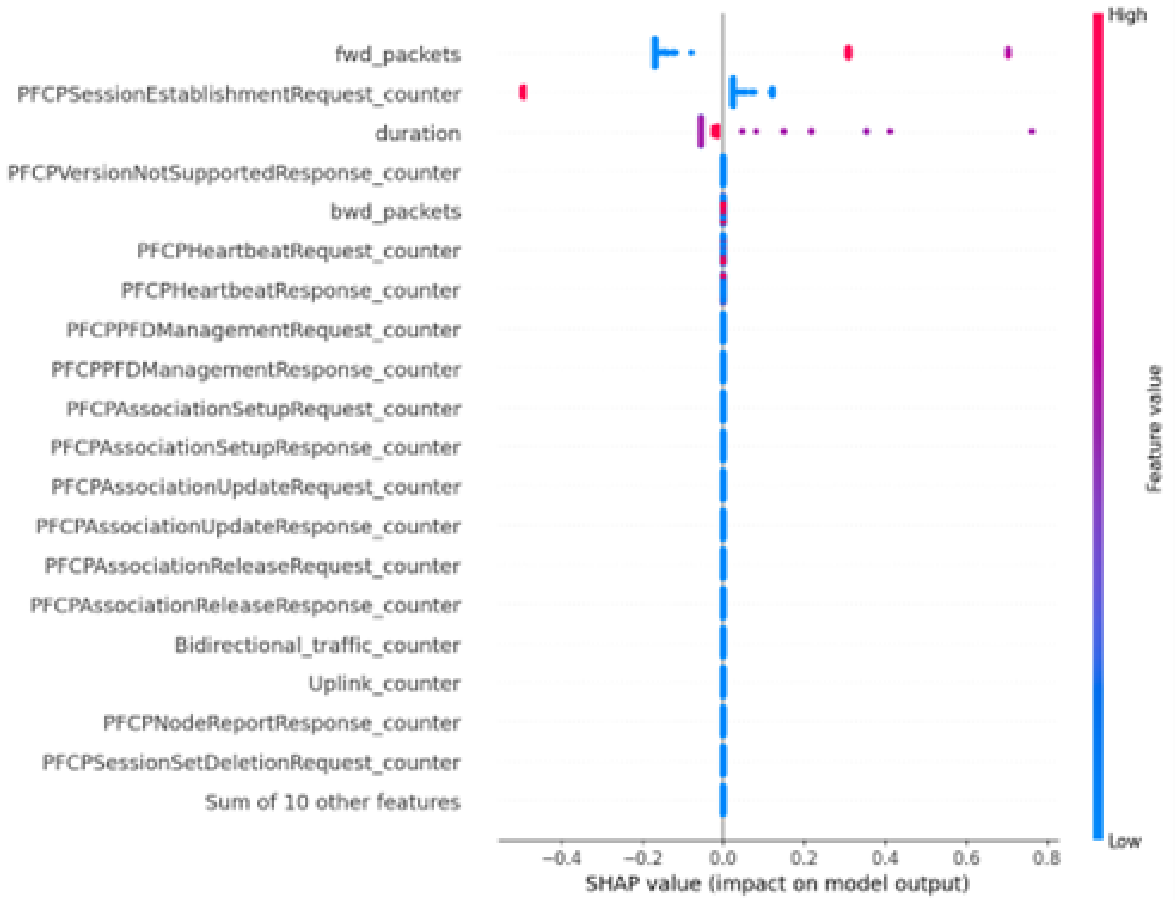


- Confusion Matrix Explanations for the Decision Tree Model trained with Application Layer PFCP Flow Statistics/Features.

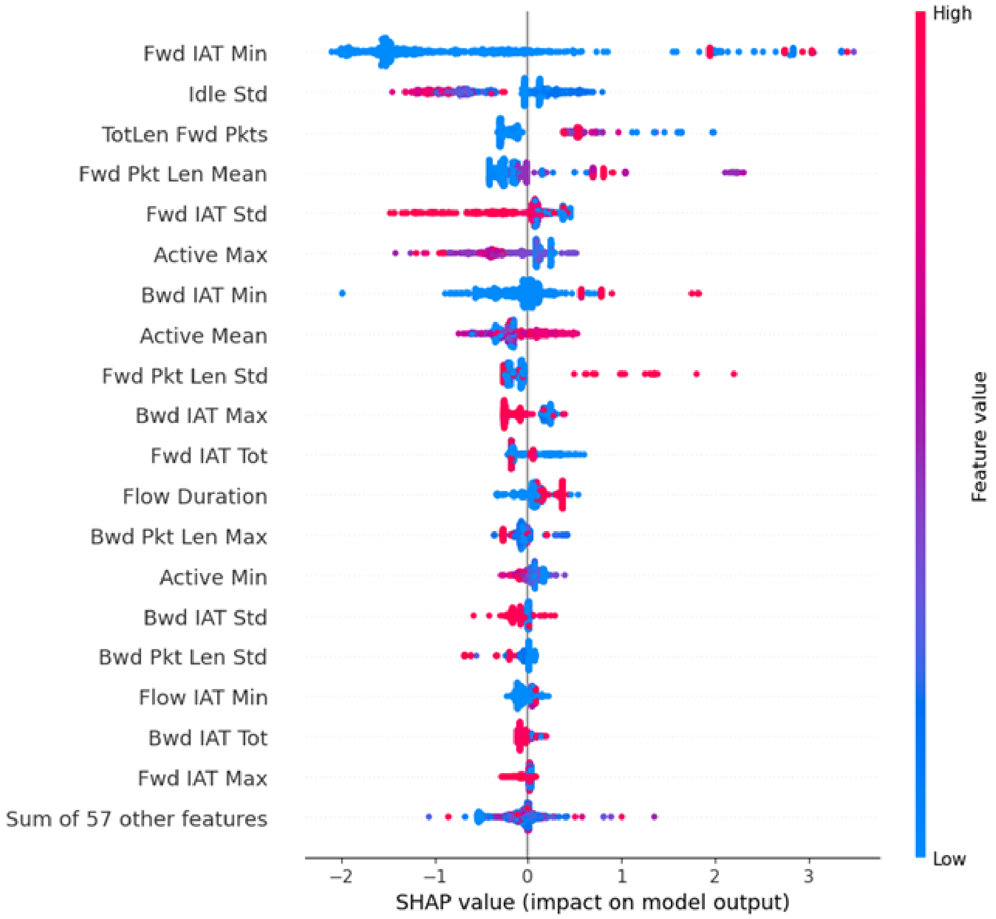


- Confusion Matrix for the XGBoost Model trained with TCP/IP Flow Statistics/Features.

Evaluation[4/4] – Global SHAP explanations



- Global Shap Explanations for the Decision Tree Model trained with Application Layer PFCP Flow Statistics/Features.



- Global Explanations for the XGBoost Model trained with TCP/IP Flow Statistics/Features.

Conclusions

Conclusions

- An **AI-powered IDS for 5GC** with an **explainability** service is introduced, focusing on the **N4 interface** of 5GC between SMF and UPF.

- The proposed IDS called **5GCIDS** emphasises on four cyberattacks against PFCP, which is used for the N4 interface. 5GCIDS takes into account both **TCP/IP flow statistics** and application-layer PFCP flow statistics. PFCPFlowMeter was implemented in the context of this work.

- For the detection process, **XGBoost** is used with the **TCP/IP flow statistics**, while the **decision tree** is used with the application-layer **PFCP flow statistics**.

- Finally, the explainability module relies on the **TreeSHAP** methods.

- The performance and the global explanations of the proposed IDS are demonstrated by the evaluation results.

Thank you for your attention!

https://across-he.eu/

gamponis@k3y.bg