ENI ISG PoC Report Template

1 General

The following normative disclaimer shall be included on the front page of a PoC report:

Submission of this ENI ISG PoC Report as a contribution to the ENI ISG does not imply any endorsement by the ENI ISG of the contents of this report, or of any aspect of the PoC activity to which it refers.

2 ENI ISG PoC Report

2.1 PoC Project Completion Status

- Overall PoC Project Completion Status:
  Completed

2.2 ENI PoC Project Participants

Specify PoC Team; indicate any changes from the ENI ISG PoC Proposal:

- PoC Project Name: **Intelligent Traffic Profiling**
- Network Operator/Service Provider: China Mobile Research Institute Contact: Weiyuan Li
- Manufacturer A: Huawei _____ Contact: Yali Wang, Shucheng Liu
- Manufacturer B: Intel _____ Contact: Tong Zhang
- Additional Members: Tsinghua University Contact: Dan Li
- Additional Members: Vodafone Contact: Mostafa Essa

2.3 Confirmation of PoC Event Occurrence

- PoC Demonstration Event Details: The PoC was showcased in three events that took place in Beijing (China) on September 27th, 2019 during Network Intelligent Forum, Guangzhou (China) during the during the China Mobile Global Partners Conference in the week of November 14th-16th, and France during the ENI #12 conference in the week of December 9th-12th. Photos of the 3 events follow.

Figure 2-1: The PoC was shown in Beijing, China, Sep 27, 2019

Figure 2-2: The PoC was shown in Guangzhou, China, Nov 14-16, 2019

Figure 2-3: The PoC was shown in Sophia Antipolis France, Dec 9-12, 2019

2.4 PoC Goals Status Report

According to the proposal, these were the specific projects goals:
This PoC contributes to demonstrate the use case [#2-8: AI enabled network traffic classification]. In particular, the proposed mechanism is compliant with its initial context configuration, triggering conditions, operational flow, and post-conditions, as defined in GS ENI 001 [1].
This PoC will demonstrate various requirements that are identified in GS ENI 002 [2], including General requirements, Service orchestration and management, Data collection and analysis, Data learning, model training and iterative optimization, etc.
This PoC intends to test and validate functional blocks of ENI Reference Architecture that are identified in GS ENI 005 [3] and report on the suitability of ENI Reference Architecture.
This PoC aims to verify the feasibility and the benefits of the use of AI/ML for network traffic classification, including the encrypted traffic, and demonstrate in a testbed environment that how ENI system can support intelligent traffic profiling and mechanism generalization.

- PoC Project Goal #1: **Traffic Categorization**. Demonstrate the use of ML algorithms to be able to categorize the network traffic into a number of application classes, e.g. video, games, VoIP and so on

- Goal Status (Demonstrated/Met?) Demonstrated

The first scenario demonstrates demonstrated the PoC Project Goal #1.
As shown in the following figure, the entire solution consists of two procedures: training and inference procedure.

1) Training Procedure：
- Offline training: Statistics-based Technique Module using statistical packet features (e.g. packet length, inter-packet arrival time, and session time).
- Online training: Data stream-based Technique Module using data stream features (e.g. destination IP address, destination port number and protocol type) and historical classification results stored in Knowledge Depository.

2) Inference Procedure :
- Statistical packet features and data stream features are extracted from real-time traffic flow.
- Classification results are determined based on both inferences in Statistics-based Technique Module and Data stream-based Technique Module.

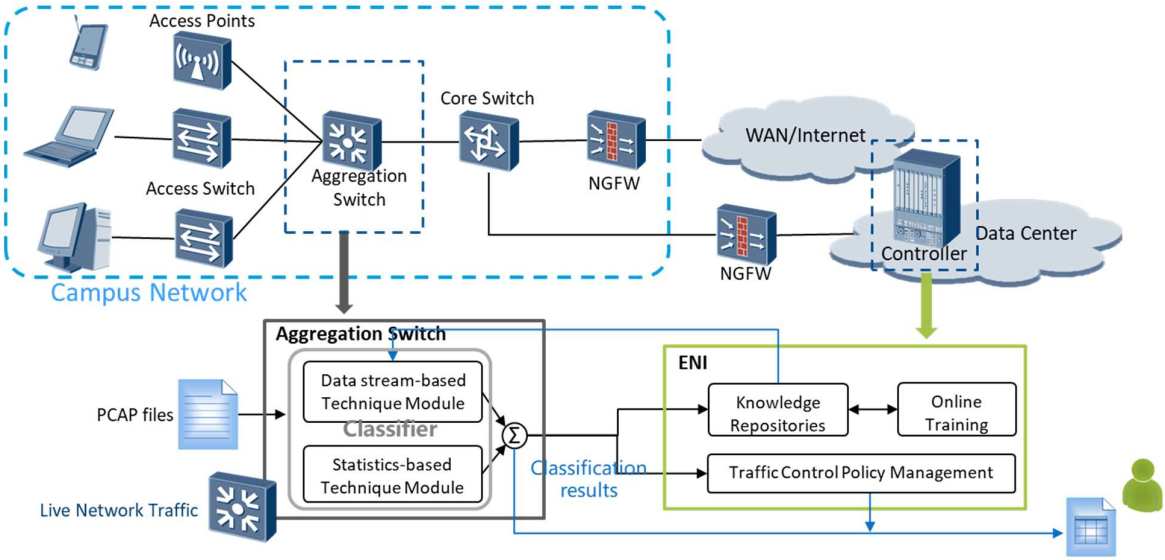

Figure 2-4: Training and inference procedures of Scenario #1

Following is an example for identifying Cloud Desktop Application.

1) Initial parameters:

- Confidence weight vector [w1 w2] = [0.4 0.6]
- Classification threshold θ1 = 0.5, Model update threshold θ2 = 0.7
- Identify two streams: stream a has atypical behaviour and stream b has typical behaviour as shown in the following figure.

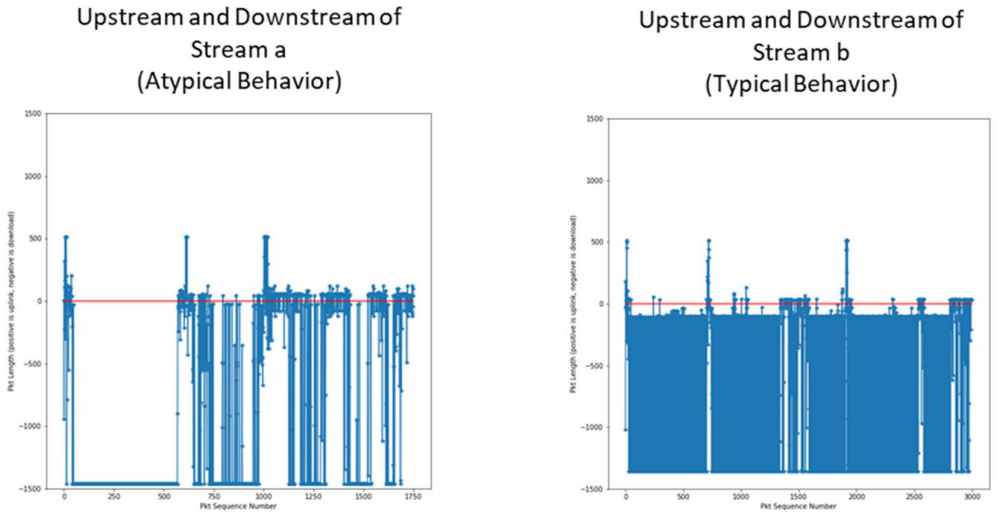

Figure 2-5: Examples of upstream and downstream

2) Data stream a and b are ingested into Classifier:

#	Dst. IP	Dst. Port	Label & Confidence	Synthesis Confidence
a	10.129.74.5	8443	Data stream-based Technique Module: desktop cloud & 0	0*0.4+0.5*0.6=0.3<0.5 Stream a is not Cloud Desktop.
			Statistics-based Technique Module: desktop cloud & 0.5	
b	10.129.56.39	443	Data stream-based Technique Module: desktop cloud & 0	0*0.4+0.9*0.6=0.54>0.5 Stream b is Cloud Desktop.
			Statistics-based Technique Module: desktop cloud & 0.9	

3) The classification result of stream b is transmitted to controller:

- Report to the controller to request high-priority QoS.

- As synthesis confidence is lower than 0.7, classification result is transmitted to Knowledge Repository in Knowledge Management FB defined in ENI Reference Architecture.

4) Knowledge Repository is edited, and online training procedure is triggered:

- Data stream-based Technique Module is updated.

#	Dst. IP	Dst. Port	Label & Confidence	Synthesis Confidence
a	10.129.74.5	8443	Data stream-based Technique Module: desktop cloud & 0.6 Statistics-based Technique Module: desktop cloud & 0.5	0.6*0.4+0.5*0.6=0.54>0.5 Stream a is Cloud Desktop.
b	10.129.56.39	443	Data stream-based Technique Module: desktop cloud & 1 Statistics-based Technique Module: desktop cloud & **0.9**	1*0.4+0.9*0.6=0.94>0.5 Stream b is Cloud Desktop.

5) Finally, the identify results are that both of stream a and b are Cloud Desktop application.

- PoC Project Goal #2: **Identification of Application Subactions**. Demonstrate the use of ML algorithms to be able to identify various subactions in a specific application, e.g. picture, voice, red packet, etc. in WeChat
- Goal Status (Demonstrated/Met?) Demonstrated

The second scenario demonstrated the PoC Project Goal #2.
The scenario includes two procedures, one is training and the other is inferring. As shown in Figure 2-6, the training phase includes three functional blocks.
Automatic clicking tool: is responsible for acquiring traffic data with labels. By using the automatic clicking tool, 5000 different subactions, including sending red packets, sending pictures and calling, are performed. Then 48166 TCP flows are captured, corresponding to these subactions.
Feature extraction: is responsible for extracting features (e.g. port, packet length, inter arrival packet time). N-dimension features are extracted from each TCP flow and then an N*48166 matrix which serves as a training data set is formed.
Model training: is responsible for feeding training data into machine learning models (e.g. Random Forests, Convolutional Neural Network, and Recurrent Neural Network). After iterations, a well-trained traffic classifier is acquired.

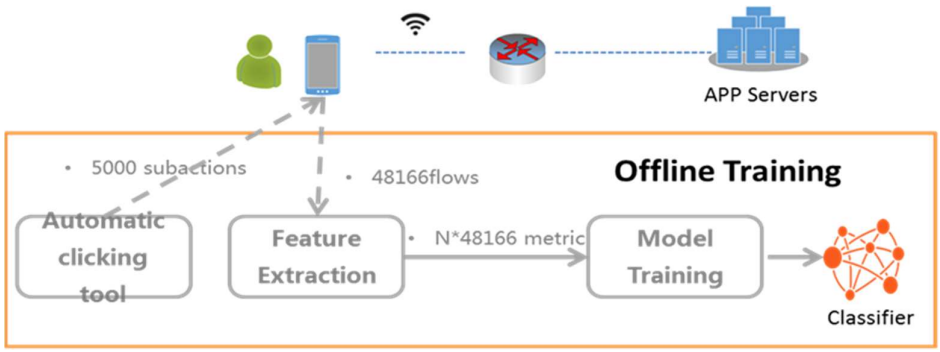

Figure 2-6: Training procedure of Scenario #2

The figure 2-7 shows the accuracy with the model iterates.

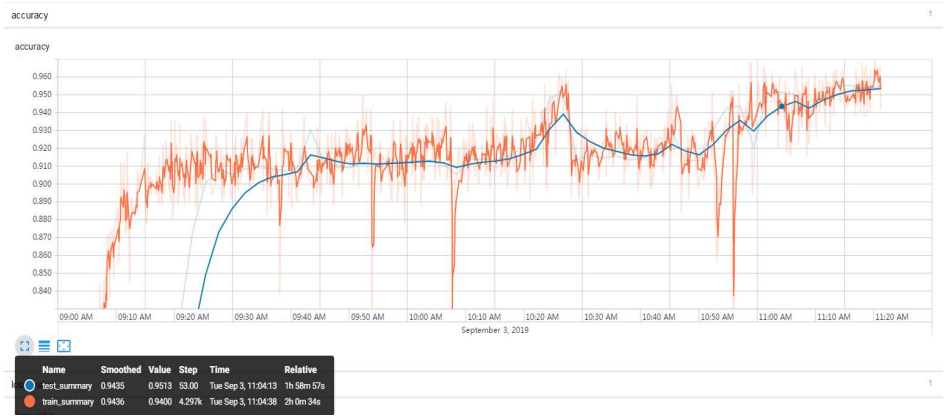

Figure 2-7: Training accuracy of classification model

As shown in Figure 2-8, in the inference phase, the well-trained modules are implemented in ENI system and intended to perform network traffic classification.

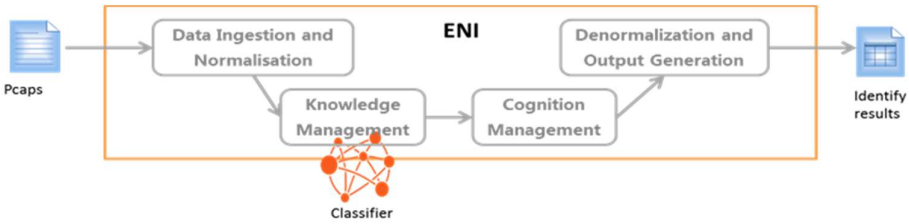

Figure 2-8: Inferring procedure of Scenario #2

The figure 2-9 shows how the traffic classification results are displayed. Based on the traffic classification results, traffic statistical analysis is realized, including traffic proportion of different types, the trend of different traffic over time, etc.

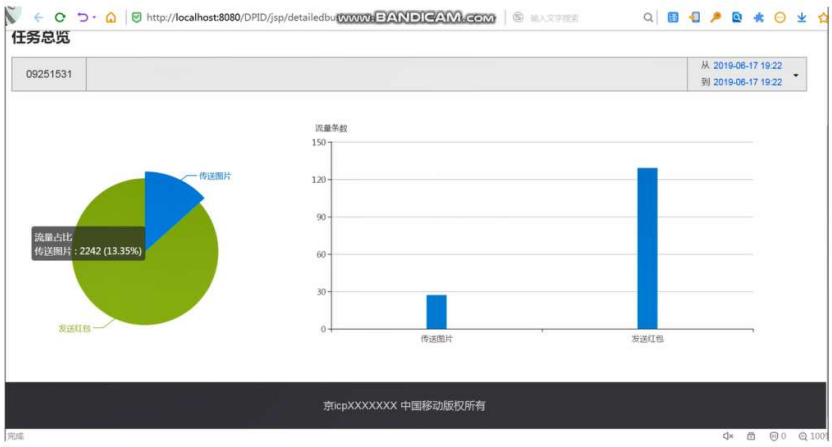

Figure 2-9: Result display interface of Scenario #2

Besides, we also developed the third scenario, which shows that the malware traffic and normal traffic can be identified based on the AI/ML algorithms.

As shown in figure 2-10, the main routine of the scenario includes PCAP Parser, FFEL, DAAL, OpenVINO.

PCAP Parser: parse PCAP files and get TCP/UDP flows.

FFEL: Flow feature extraction library

DAAL: Data analytics acceleration library, aiming to improve the speed of model training.

OpenVINO: Deep learning deployment toolkit

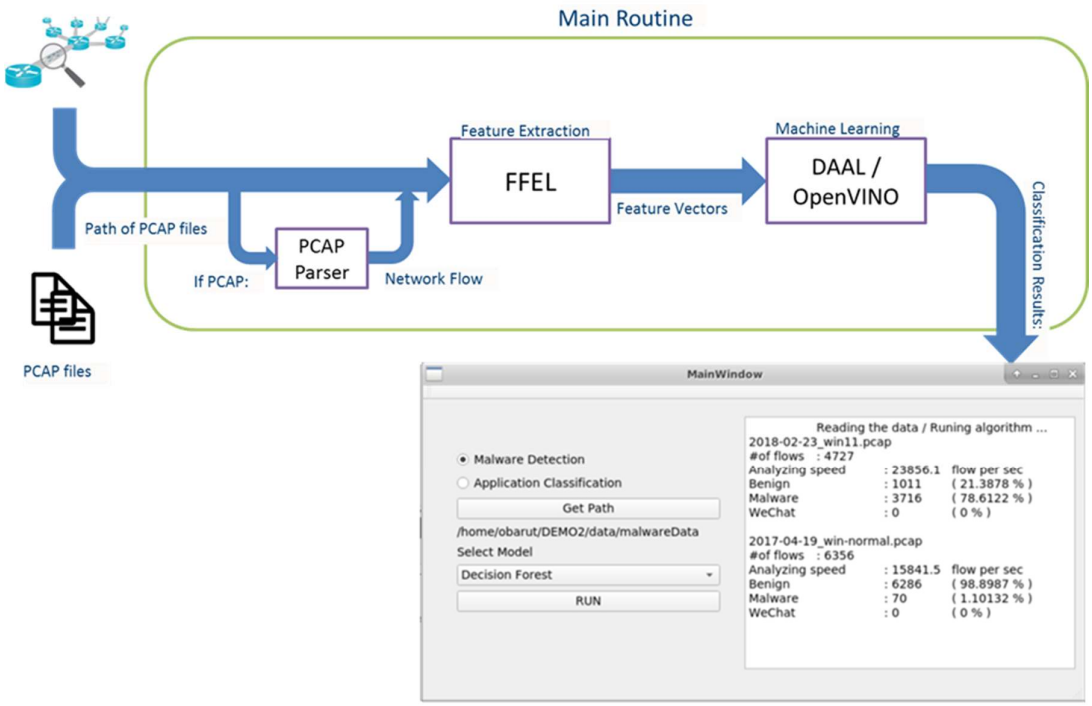

Figure 2-10: Main Routine of Scenario #3

The figure 2-11 shows traffic classification results based on SVM and Decision Tree

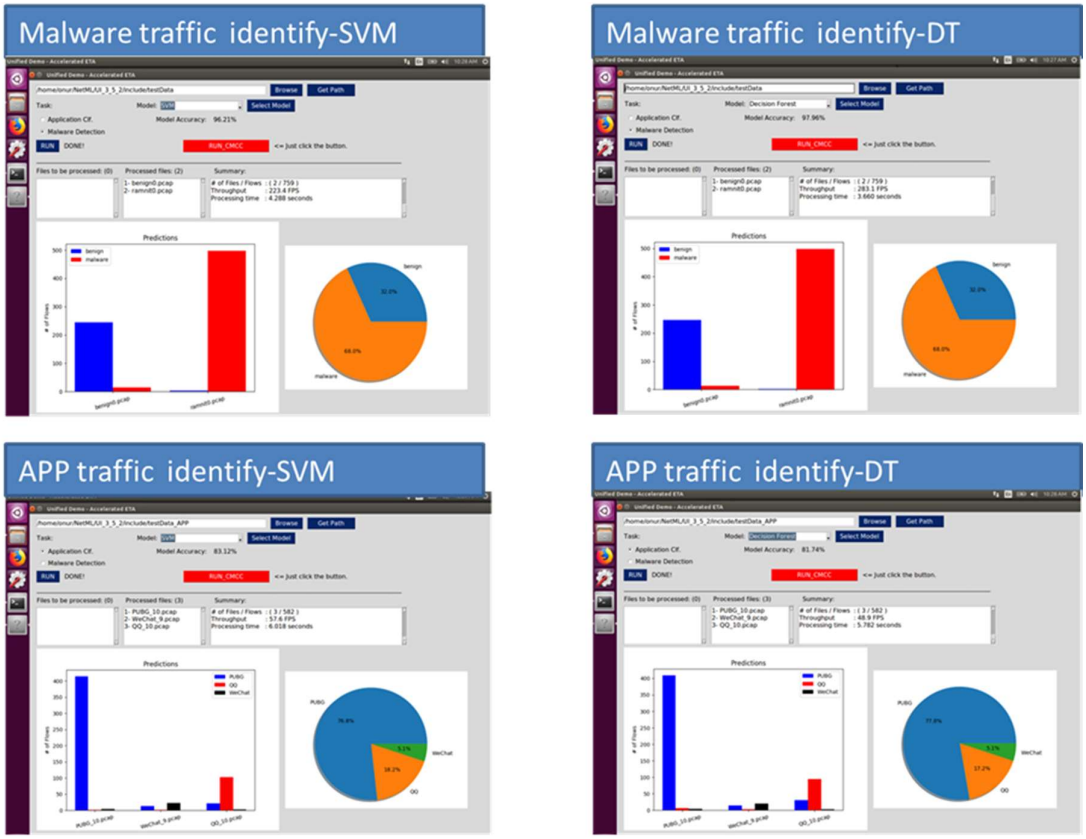

Figure 2-11: Result display interface of Scenario #3

2.5 PoC Feedback Received from Third Parties (Optional)

- Where applicable, provide in a free text, feedback received from potential customers, Ecosystem partners, event audience and/or general public.

3 ENI PoC Technical Report (Optional)

3.1 General

PoC Teams are encouraged to provide technical details on the results of their PoC using the PoC Scenario Report template below.

3.2 PoC Contribution to ENI ISG

Use table B.1 to list any contributions to the ENI ISG resulting from this PoC Project.

Table 1

Contribution	WG/EG	Work Item (WI)	Comments
Feasibility of traffic data pre-processing		ETSI GS ENI 005	We could successfully test traffic data pre-processing, corresponding to data cleaning and data labeling in Data Ingestion and Data Normalization functional block
Feasibility of online learning and offline learning		ETSI GS ENI 001, 005	We tested a pre-trained traffic classification model and an online learning model.
Feasibility of traffic types identification		ETSI GS ENI 002	We successfully tested traffic type identification in different granularity levels, corresponding to new requirements in Network optimization.

3.3 Gaps identified in ENI standardization

Use table B.2 to indicate Gaps in standardization identified by this PoC Team including which forum(s) would be most relevant to work on closing the gap(s).Where applicable, outline any action(s) the ENI ISG should take.

Table B.2

Gap Identified	Forum (ENI ISG, Other)	Affected WG/EG	WI/Document Ref	Gap details and Status

3.4 PoC Suggested Action Items

3.5 Additional messages to ENI

3.6 Additional messages to Network Operators and Service Providers?

References

[1] RGS/ENI-008 (GS ENI 001), “Experiential Networked Intelligence (ENI); ENI use cases”, v2.0.8 (early draft), Sec 5.3.8.
[2] RGS/ENI-007 (GS ENI 002), “Experiential Networked Intelligence (ENI); ENI requirements”, v2.0.4 (early draft).
[3] DGS/ENI-005 (GS ENI 005), “Experiential Networked Intelligence (ENI); System Architecture”, v0.0.22 (early draft).