DUST Project - Identification and Obfuscation of Security and Behavioral Vulnerabilities in IoT

Fernando Nakayama PhD

Center for Computational Security sCience (CCSC)

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

UFMG

Belo Horizonte – Brazil

July 17, 2024

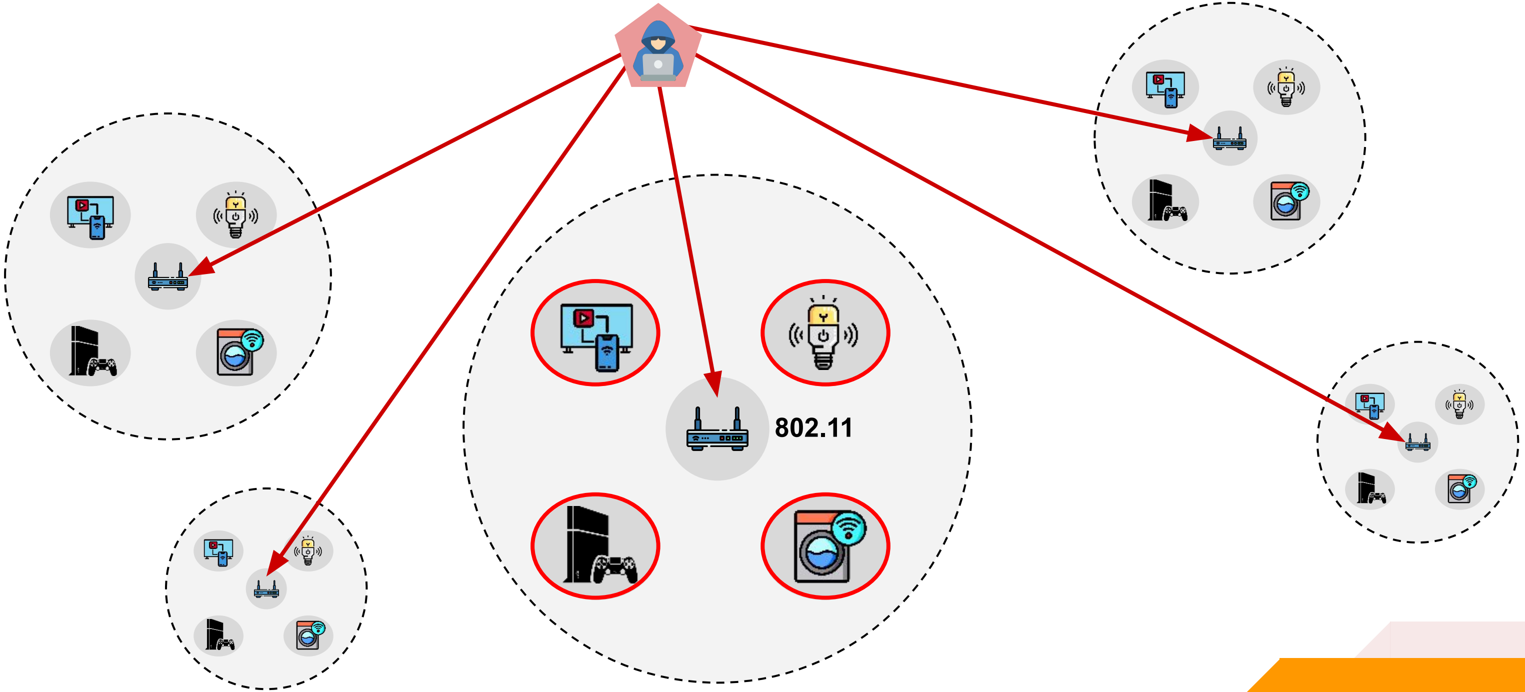
FAPESP

www.ccsc-research.org

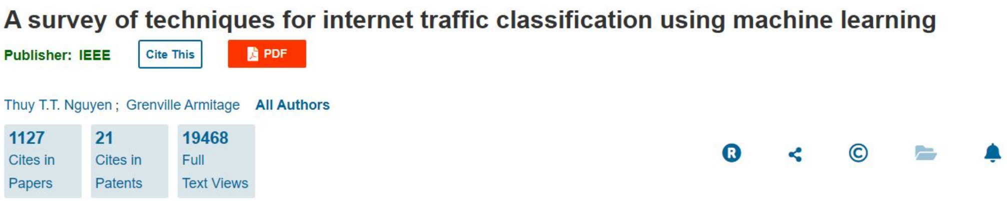
DUST Project
Objectives

- Identify security threats in IoT environments
  - Research line #1
- Obfuscate security and behavioral vulnerabilities in IoT
  - Research line #2

DUST Project
Information leakage (side-channel)



DUST Project
Information leakage (side-channel) - traffic classification

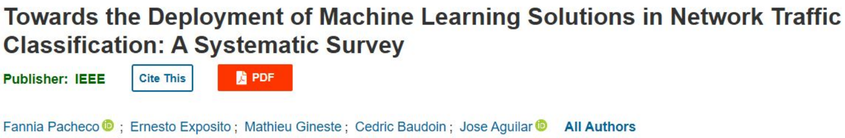


A survey of techniques for internet traffic classification using machine learning

Publisher: IEEE | Cite This | PDF

Thuy T.T. Nguyen ; Grenville Armitage All Authors

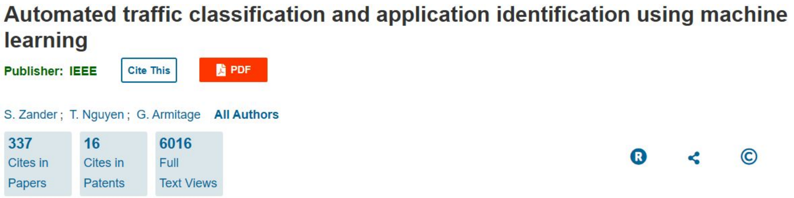
1127	21	19468
Cites in Papers	Cites in Patents	Full Text Views



Towards the Deployment of Machine Learning Solutions in Network Traffic Classification: A Systematic Survey

Publisher: IEEE | Cite This | PDF

Fannia Pacheco ; Ernesto Exposito ; Mathieu Gineste ; Cedric Baudoin ; Jose Aguilar All Authors

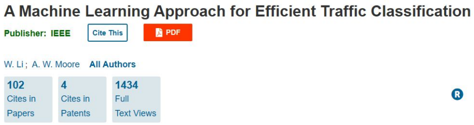


Automated traffic classification and application identification using machine learning

Publisher: IEEE | Cite This | PDF

S. Zander ; T. Nguyen ; G. Armitage All Authors

337	16	6016
Cites in Papers	Cites in Patents	Full Text Views

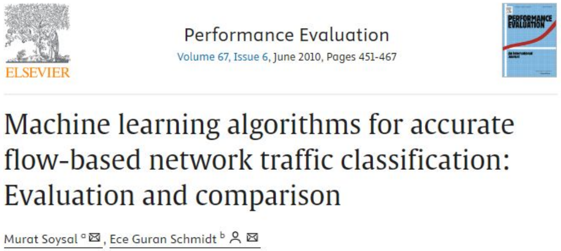


A Machine Learning Approach for Efficient Traffic Classification

Publisher: IEEE | Cite This | PDF

W. Li ; A. W. Moore All Authors

102	4	1434
Cites in Papers	Cites in Patents	Full Text Views



Performance Evaluation
Volume 67, Issue 6, June 2010, Pages 451-467

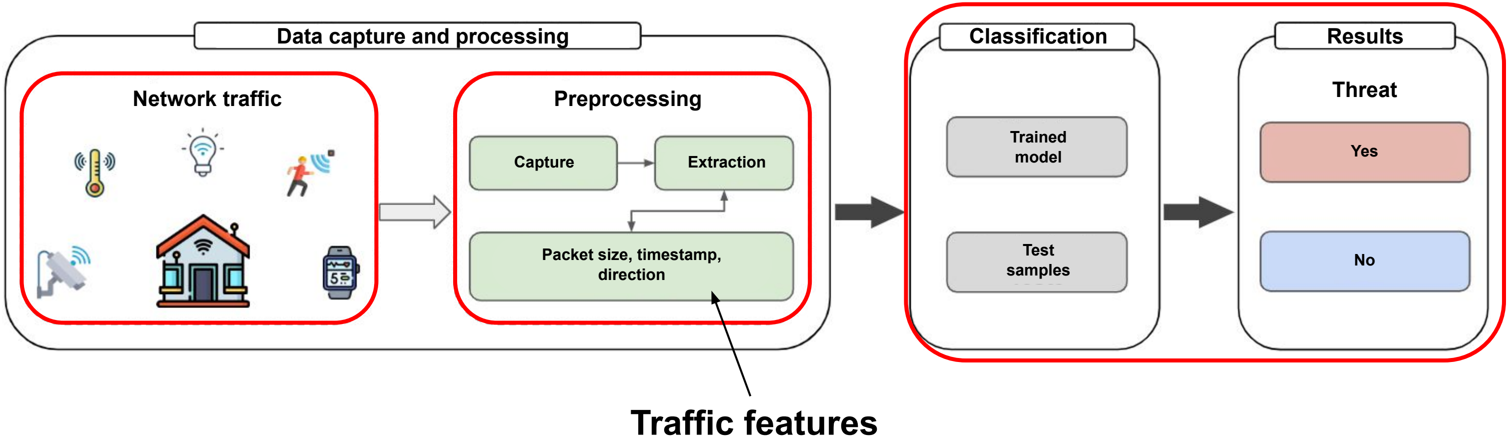
Machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison

Murat Soysal [a], Ece Guran Schmidt [b]

DUST Project

Identification of threats - profiling and classification

- Traffic features to build a behavior profile



Data capture and processing

Network traffic

Preprocessing

Capture

Extraction

Packet size, timestamp, direction

Traffic features

Classification

Trained model

Test samples

Results

Threat

Yes

No

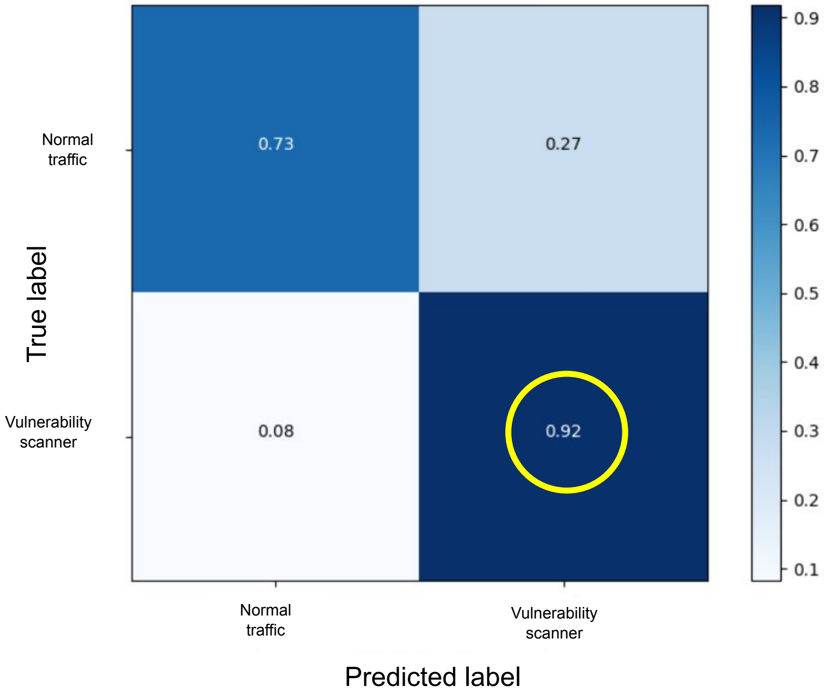
DUST Project
Identification of vulnerabilities - profiling and classification

- Selected vulnerability - Portscanner
  - Profile built with multiple attack examples (data fusion)
  - Tested with different datasets (topologies)

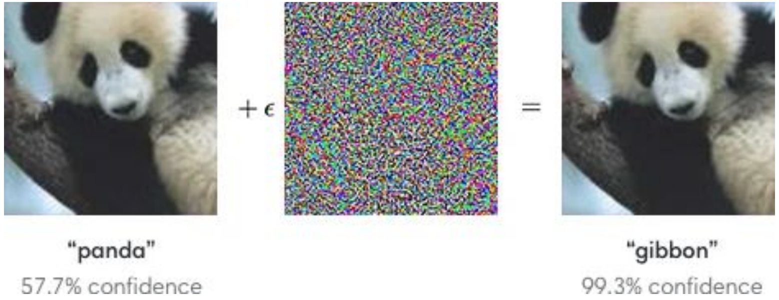
Port Scanner

UFPR UFMG

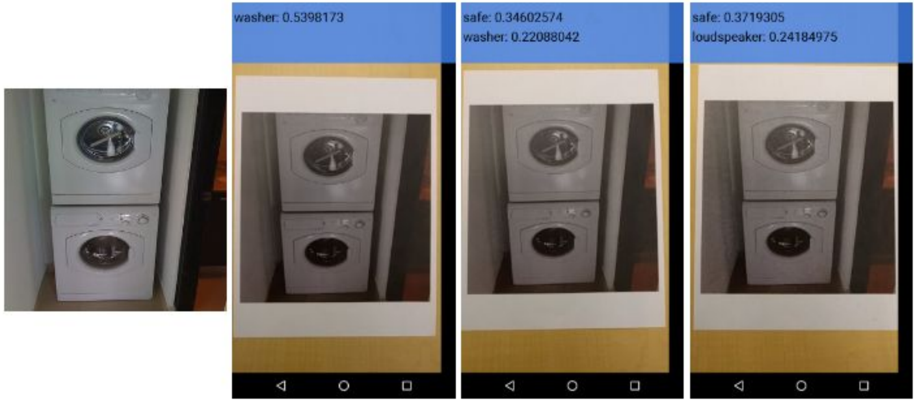
DUST Project
Identification of vulnerabilities - selected result



DUST Project
Obfuscation against information leakage - AML



Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. "Explaining and harnessing adversarial examples." *arXiv preprint arXiv:1412.6572* (2014).



Kurakin, Alexey, Ian J. Goodfellow, and Samy Bengio. "Adversarial examples in the physical world." Artificial intelligence safety and security. Chapman and Hall/CRC, 2018. 99-112.

DUST Project
Obfuscation against information leakage

- Obfuscation against traffic classification models
- Different data structures (features, size, etc)
- Reverse AML

DUST Project
Obfuscation against information leakage - AML

Replication

Adaptation

Evaluation

DUST Project

Obfuscation against information leakage - AML

Adversarial network configuration

$$\begin{cases} \text{LeakyReLU(Linear(10,64))} \\ \text{LeakyReLU(BN(Linear(64,128)))} \\ \text{LeakyReLU(BN(Linear(128,256)))} \\ \text{Tahn(BN(Linear(256,10)))} \\ \text{nn.Linear(10,2)} \end{cases}$$

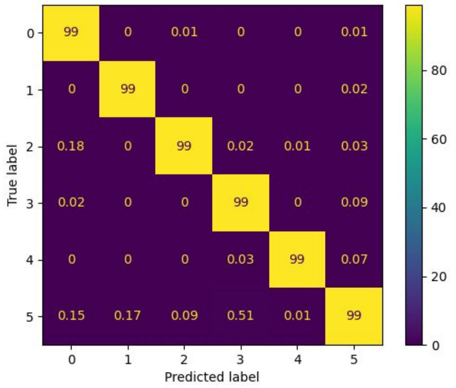
Adversarial attacks techniques

Carlini-Wagner (CW2)

Fast gradient sign method (FGSM)

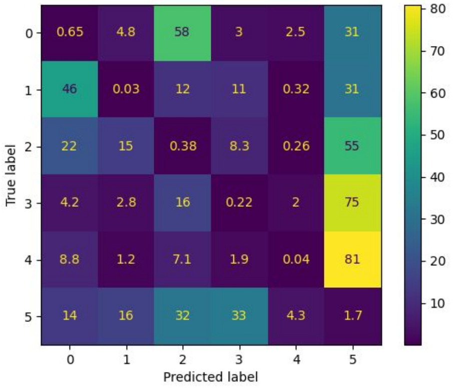
DUST Project
Obfuscation against information leakage - results

Original data



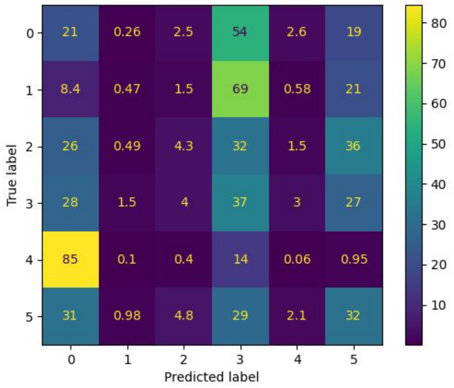
True label \ Predicted label	0	1	2	3	4	5
0	99	0	0.01	0	0	0.01
1	0	99	0	0	0	0.02
2	0.18	0	99	0.02	0.01	0.03
3	0.02	0	0	99	0	0.09
4	0	0	0	0.03	99	0.07
5	0.15	0.17	0.09	0.51	0.01	99

Adversarial (CW2)



True label \ Predicted label	0	1	2	3	4	5
0	0.65	4.8	58	3	2.5	31
1	46	0.03	12	11	0.32	31
2	22	15	0.38	8.3	0.26	55
3	4.2	2.8	16	0.22	2	75
4	8.8	1.2	7.1	1.9	0.04	81
5	14	16	32	33	4.3	1.7

Adversarial (FGSM)



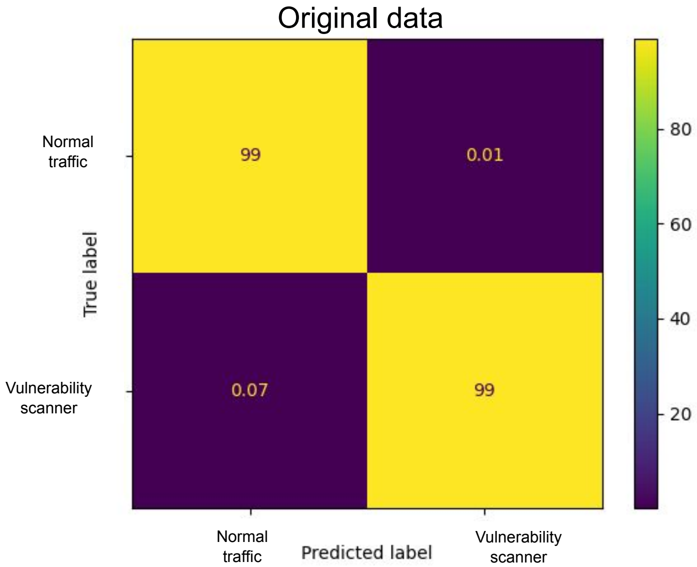
True label \ Predicted label	0	1	2	3	4	5
0	21	0.26	2.5	54	2.6	19
1	8.4	0.47	1.5	69	0.58	21
2	26	0.49	4.3	32	1.5	36
3	28	1.5	4	37	3	27
4	85	0.1	0.4	14	0.06	0.95
5	31	0.98	4.8	29	2.1	32

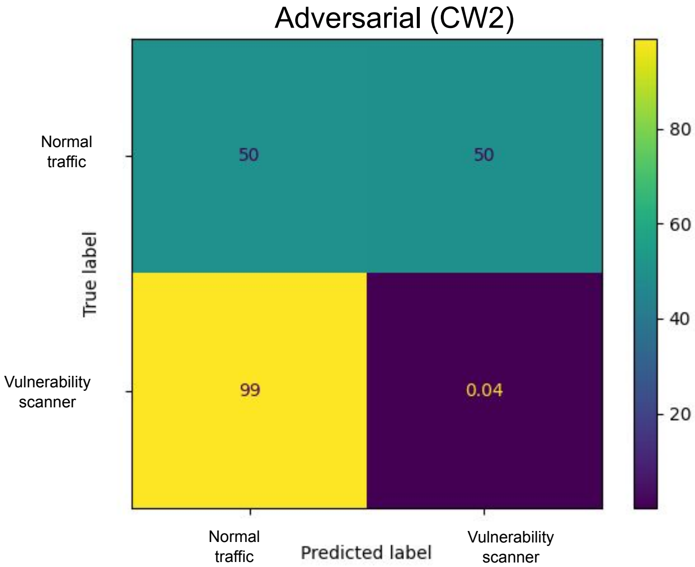
Traffic types

0 - WWW,
1 - MAIL,
2 (UDP)
3 - P2P,
4 - DATABASE
5 - SERVICES

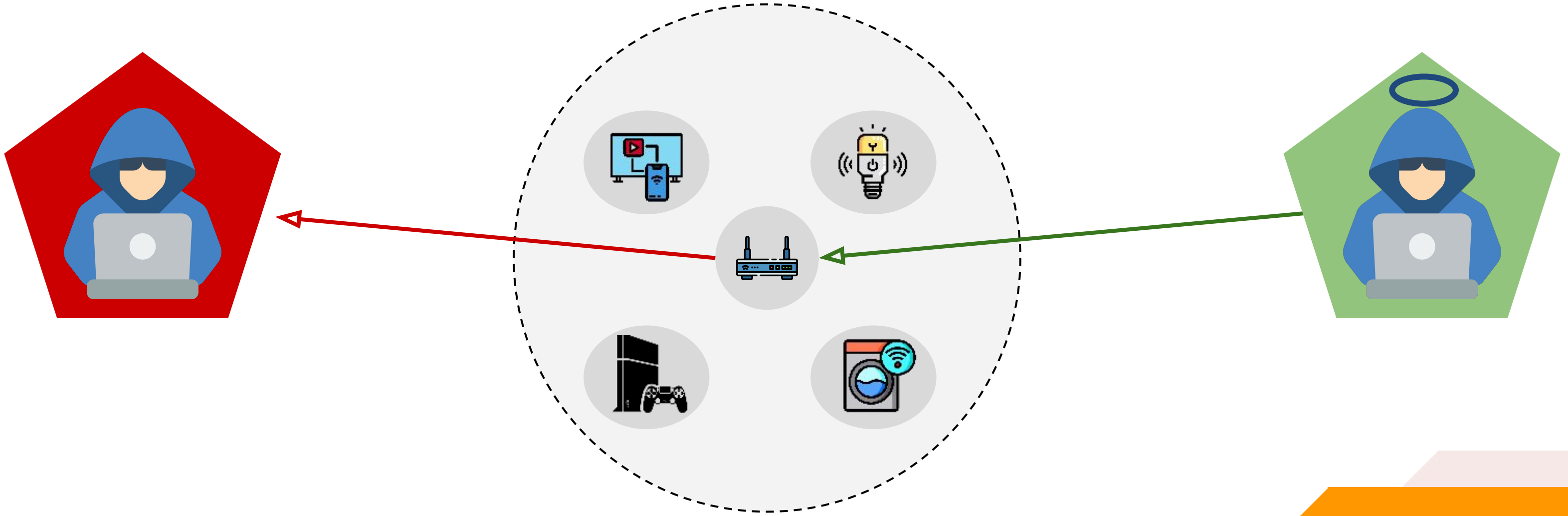
Dataset: Moore, Andrew, Denis Zuev, and Michael Crogan. *Discriminators for use in flow-based classification*. 2013.

DUST Project
Obfuscation against information leakage - results

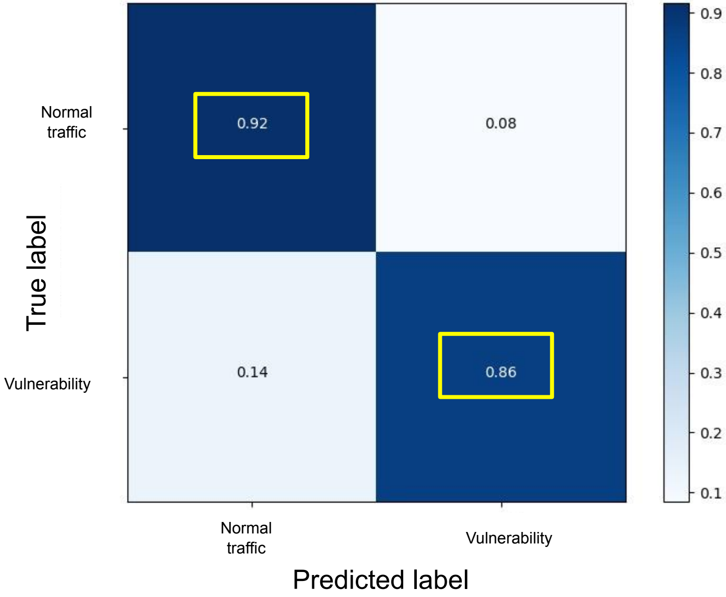




DUST Project
Obfuscation against information leakage - models integration



DUST Project
Obfuscation against information leakage - models integration



True label \ Predicted label	Normal traffic	Vulnerability
Normal traffic	0.92	0.08
Vulnerability	0.14	0.86

DUST Project

Identification and Obfuscation of Security and Behavioral Vulnerabilities in IoT

- 2 research lines
  - Identification of threats
  - Obfuscation of threats using AML
- Behaviour profile - network traffic
- Adversarial samples - network traffic

fernandonakayama@ufpr.br
fernandonakayama@gmail.com

DUST