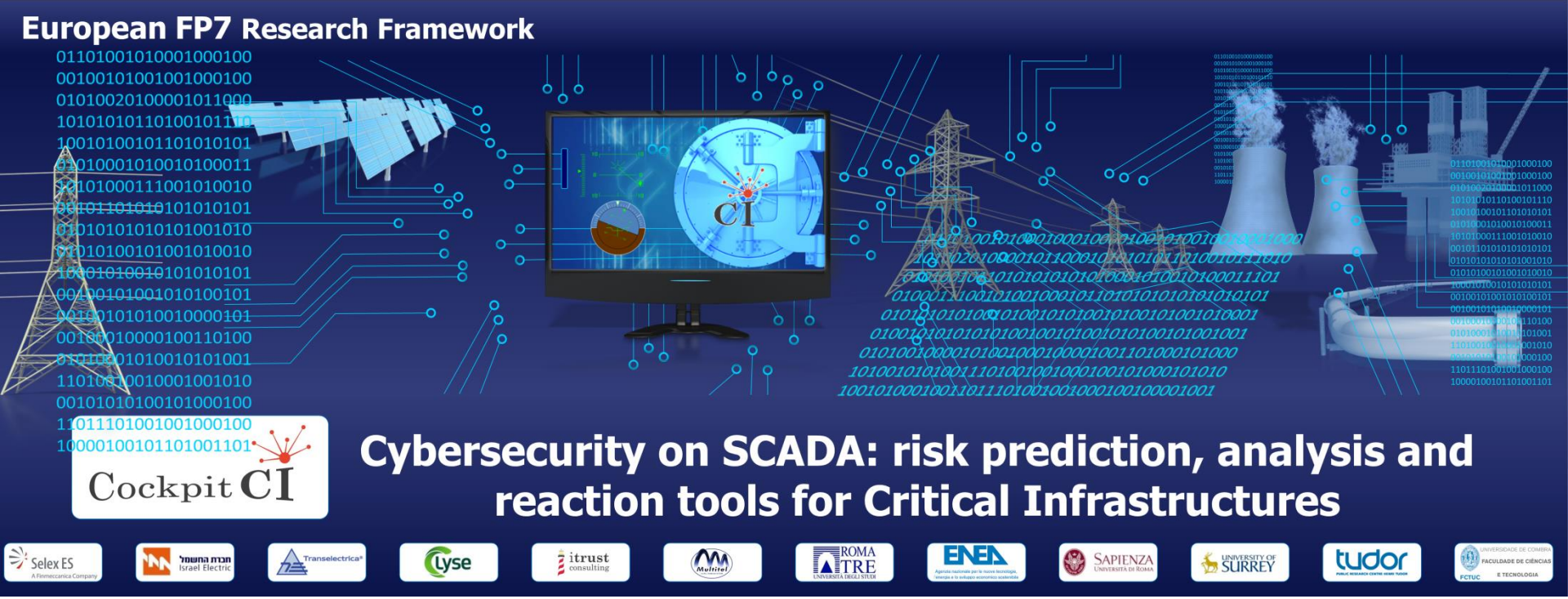


OCSVM detection tool

Leandros Maglaras & Jianmin Jiang

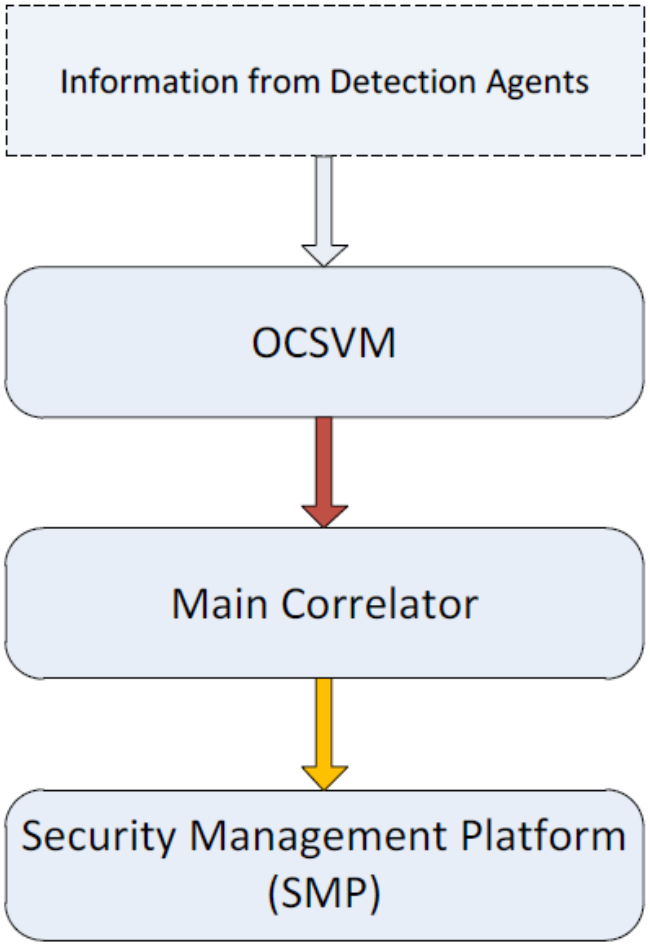
University of Surrey

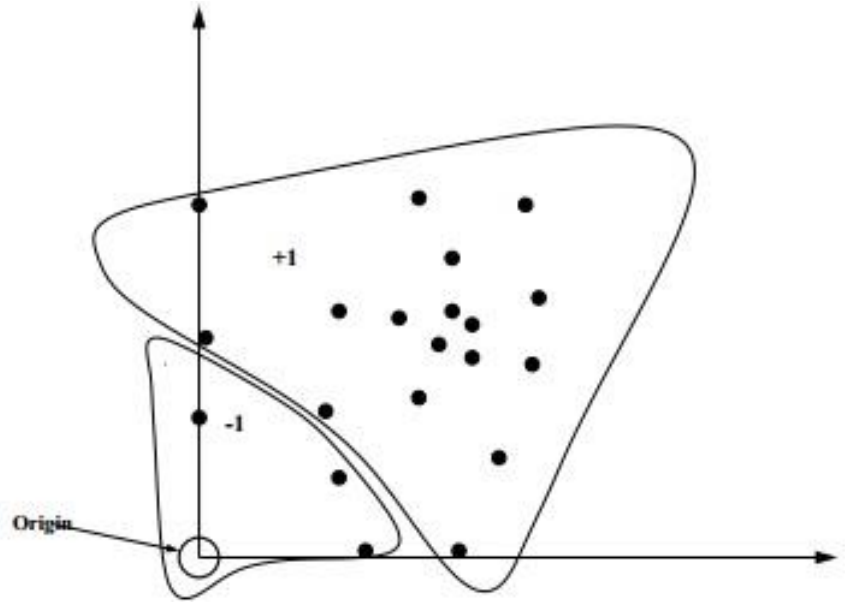
Information Society

Introduction

1. OCSVM
2. Network Dataset
3. Packet ditribution / users
4. Packet distribution over time
5. Rate of packets
6. First testing
7. Format of training/testing file
8. Integration of OCSVM module
9. Conclusions – Discussion

OCSVM





Picture from : Manevitz, Larry M., and Malik Yousef. "One-class SVMs for document classification." *the Journal of machine Learning research* 2 (2002): 139-154.

Dataset analysis

Network Dataset

No.	Time	Source	Destination	Protoco	Length	Info
1	0	AsustekC_b2:	Broadcast	ARP	60	Who has 192.168.1.4? Tell 192.168.1.2
2	0.000017	AsustekC_b2:	Broadcast	ARP	60	Who has 192.168.1.4? Tell 192.168.1.2
3	0.497982	Cisco_70:37:1	Spanning-tree-(	STP	64	RST. Root = 32768/0/08:d0:9f:70:37:12 Cost = 0 Port = 0x8002
4	0.498211	Cisco_70:37:1	Spanning-tree-(	STP	64	RST. Root = 32768/0/08:d0:9f:70:37:12 Cost = 0 Port = 0x8003
5	2.059351	192.168.1.2	192.168.1.3	FTP	60	Request: FREE
6	2.059358	192.168.1.2	192.168.1.3	FTP	60	[TCP Retransmission] Request: FREE
7	2.07154	192.168.1.3	192.168.1.2	FTP	98	Response: 200 free space on SD card: size = 14464000
8	2.071547	192.168.1.3	192.168.1.2	FTP	98	[TCP Retransmission] Response: 200 free space on SD card: size = 14464000
9	2.075051	192.168.1.7	192.168.1.254	DNS	90	Standard query A geoip.ubuntu.com.192.168.1.254
10	2.221634	fe80::d5c8:42	ff02::1:3	LLMNR	84	Standard query A wpad

A. **Protocols used by the attackers is an important issue -> filter this information before OCSVM module.**

B. **Source characteristics (gateway, node) and reputation**

Dataset analysis

Packet distribution / users

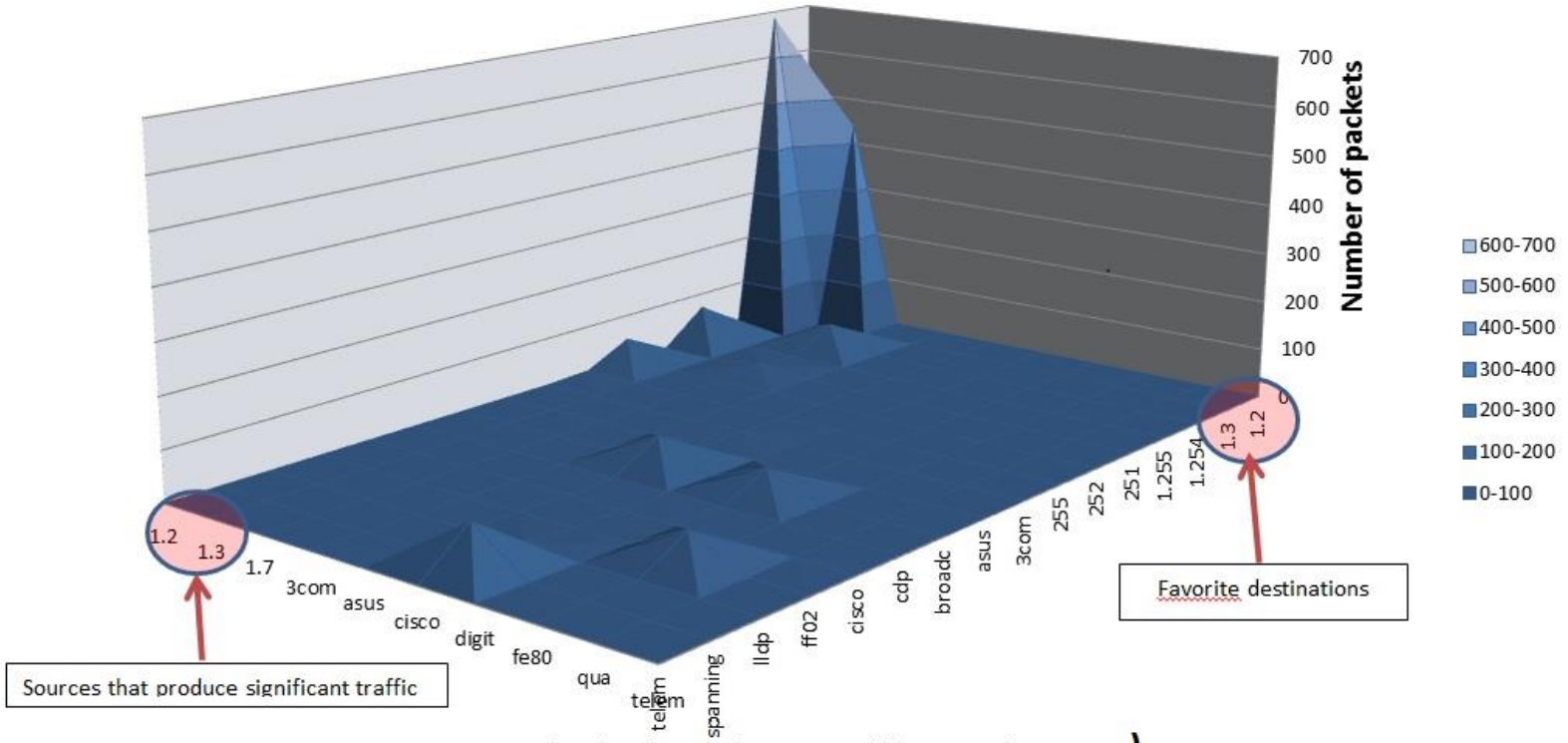


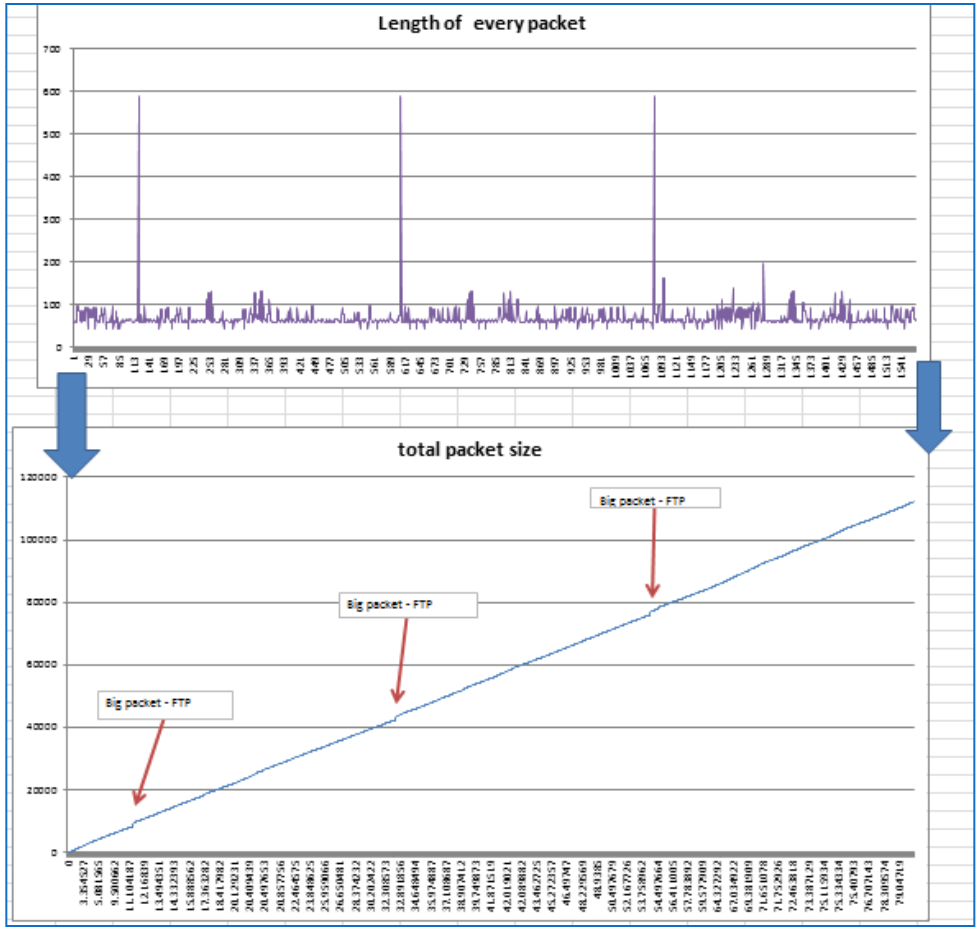
Figure 2. Packet distribution from Romes' file (source/destination)

Dataset analysis

Packet distribution over time

It is evident from figure that the accumulative packet size over time is almost a straight line with some sudden rises only is the instances when big files are circulated in the system. These files are FTP files of size 590 bytes which is ten times larger than the usual send packets in the system.

Packet size is an important factor for our OCSVM program.

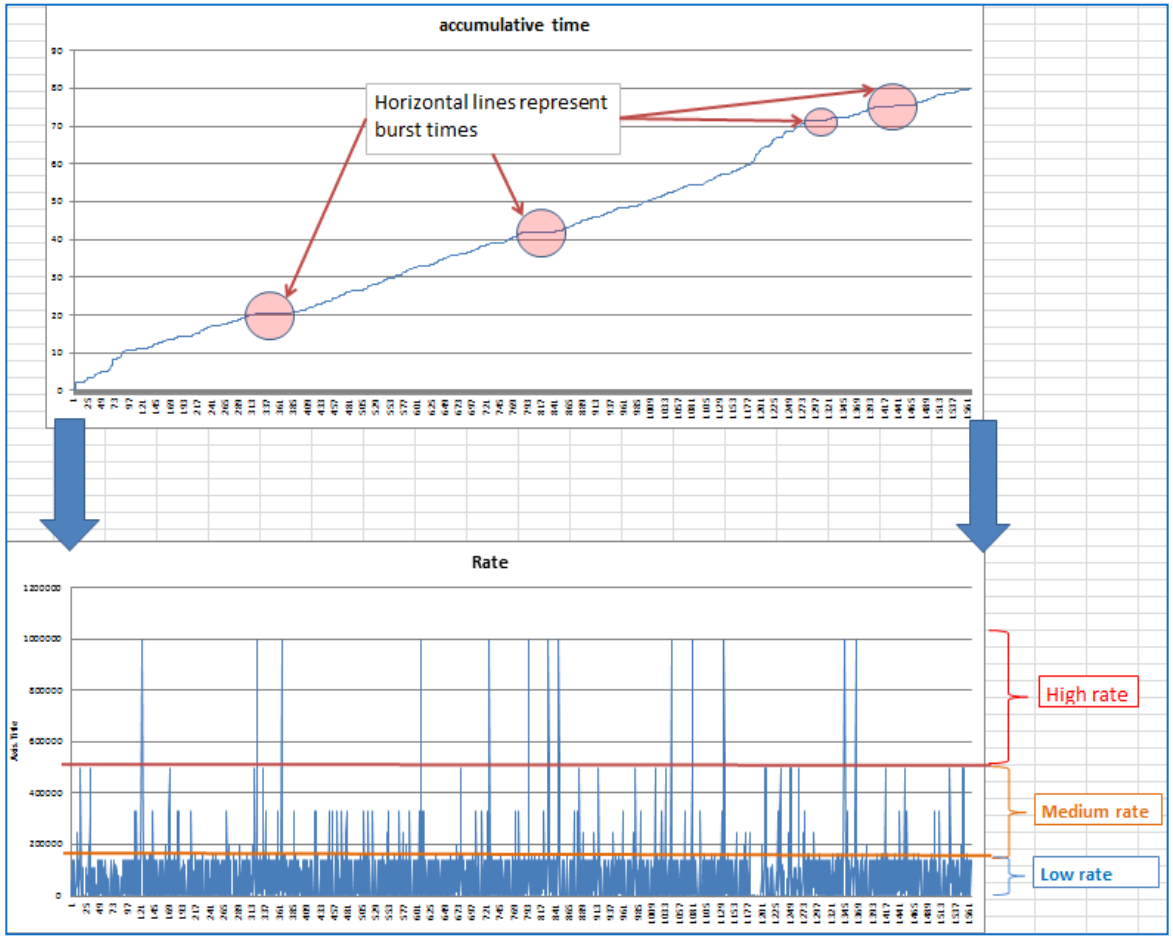


Dataset analysis

Accumulative traffic over time – Rate of traffic

We observe that when in the upper graph there exist horizontal lines then we have a burst of traffic in the system. In the lower graph this burst is more easily observed.

An important attribute for the OCSVM is the rate of the packets injected in the system.



OCSVM training - testing

Based on the above observations we tested the network file with our OCSVM module. The attributes used were RATE & PACKET SIZE.

The rate (1st attribute) was calculated using this equation:

$$Rate = \frac{Time\ difference\ (time\ of\ current\ packet - time\ of\ previous\ packet)}{Max\ time\ \text{difference}}$$

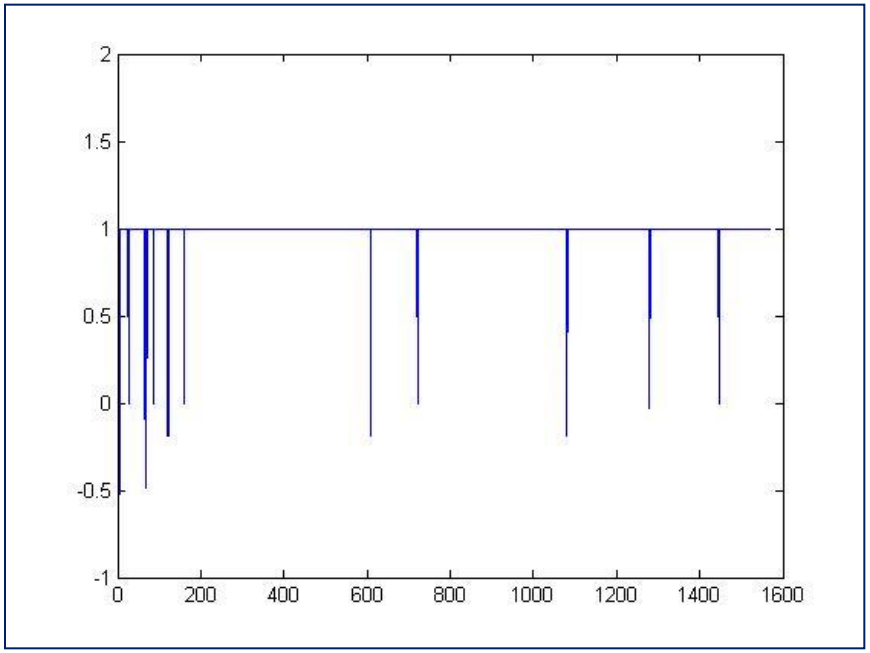
The packet size (2nd attribute) was scaled using this equation:

$$Packetscaled = \frac{packet\ size}{Max\ packet\ size}$$

OCSVM training - testing

The initial training by our OCSVM was conducted using Gaussian kernel v=0.01, sigma = 0.07

1 1: 0 2: 0.101694915254237
1 1: 1.08894782018269E-05 2: 0.101694915254237
1 1: 0.318975236045454 2: 0.108474576271186
1 1: 0.000146687676954043 2: 0.108474576271186
1 1: 1 2: 0.101694915254237



OCSVM Java: Accuracy = 99.04336734693877% (1553/1568) (classification)

Integration

Integration of OCSVM module

In order to cooperate with the other modules we have to create self-executable programs that perform these tasks:

- **reading** of network files: √ **(in cooperation with Coimbra)**
- **transforming** of data to correct format √
- **training** of the model, √ **(based on current attributes)**
- **testing** of data and √ **(based on current attributes)**
- **producing** of output events IDMEF. √ **(in cooperation with Coimbra)**
- **sending of IDMEF messages** √ **(in cooperation with Coimbra)**

Next steps:

- **Test with malicious data**
- **Filter the data**
- **Insert more attributes to the model**

IDMEF message produced by OCSVM module

```
<?xml version="1.0" encoding="UTF-8"?>
<idmef:IDMEF-Message version="1.0">
  - <idmef:Alert>
      - <idmef:Source>
          - <idmef:Node>
              - <idmef:Address>
                    <idmef:address>192.168.1.3</idmef:address>
                </idmef:Address>
                <idmef:location>NET</idmef:location>
                <idmef:name>OCSVM</idmef:name>
            </idmef:Node>
        </idmef:Source>
        <idmef:DetectTime ntpstamp="0x1123111e.0x40000000">2014-02-03T16:03:49Z</idmef:DetectTime>
        <idmef:Classification text="POSSIBLE ALARM"/>
    </idmef:Alert>
</idmef:IDMEF-Message>
```

- **Packet size is an important factor for our OCSVM module.**
- **Rate is important factor for our OCSVM module**
- **Important issue is the protocol used by the attackers – filter**
- **Important factor is the sender of the packet (reputation)**

Conclusions – Discussion

OCSVM Model
Time fragmentation (weekends, weekdays, morning, evening etc.)

Zone fragmentation (high traffic zone, low traffic zone)

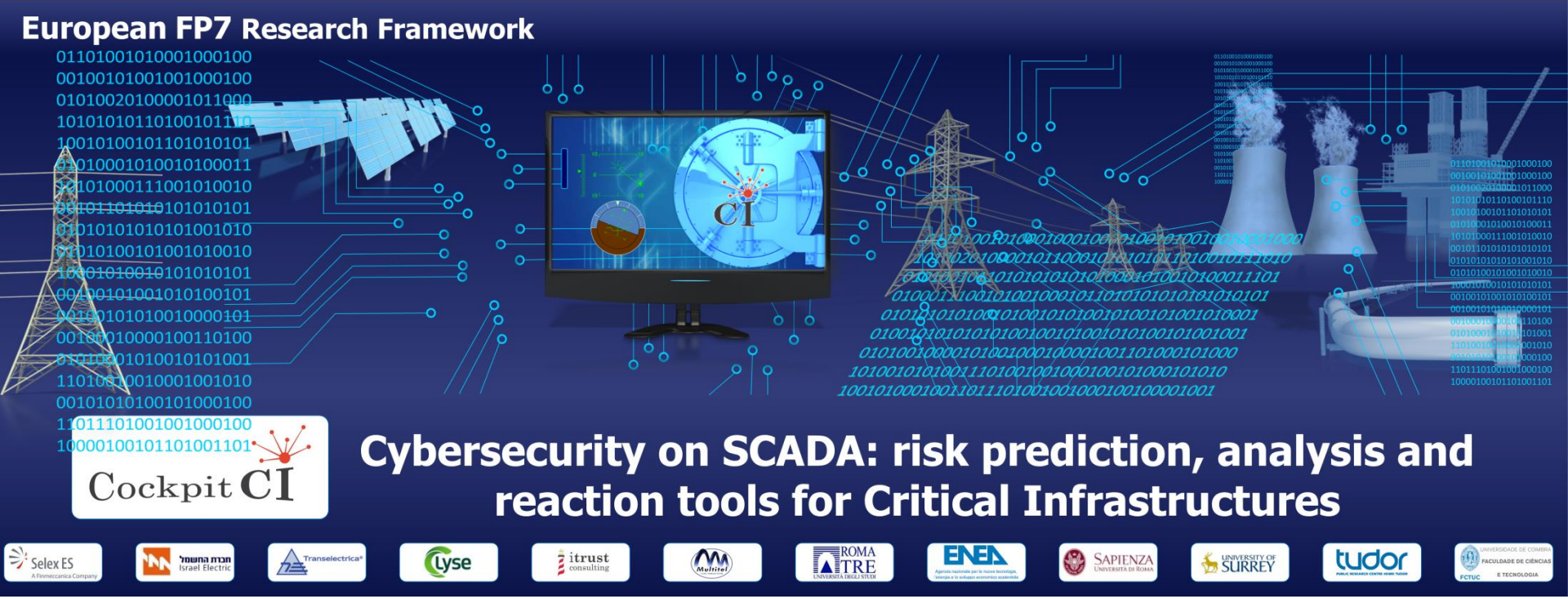
European FP7 Research Framework

Cockpit CI

Cybersecurity on SCADA: risk prediction, analysis and reaction tools for Critical Infrastructures

Selex ES · Israel Electric · Transelectrica · Lyse · itrust consulting · Multitel · Roma Tre · ENEA · Sapienza Università di Roma · University of Surrey · tudor · Faculdade de Ciências e Tecnologia

Any question ?



Thank you for your attention