Using Measurement Data in Network Security Education

Tanja Zseby,
Felix Iglesias, Robert Annessi, Christian Krieg, Davor Frkat, Valentin Bernhard

Institute of Telecommunications, TU Wien

TU Wien Network Security Classes

- Two Security Courses for Master students (EE, some CS)
  - Theory Lectures (6 x 90 min) ➔ written exam
  - Lab Exercises (6 x180 min), Teams of 2 ➔ Report
  - Lab Review (oral exam)
- 1. Network Security
  - Lab: IP Darkspace Analysis
  - Data: CAIDA IP darkspace data
- 2. Network Security Advanced
  - Lab: Network Steganography
  - Data: Modified MAWI Dataset (WIDE)
- Winter School: 1 week compressed content
  - Telecommunications Graduate Initiative (TGI), Ireland, 2016

Educational Objectives

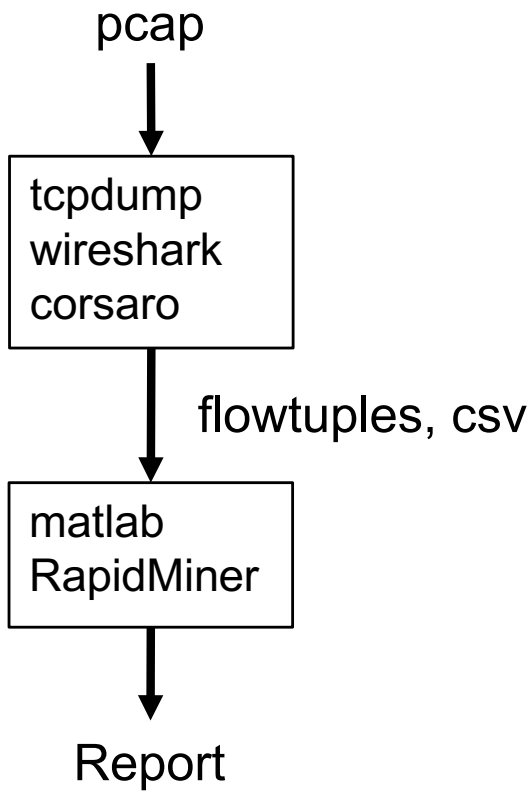
- Research-oriented teaching concept
  - Include current research in the classroom

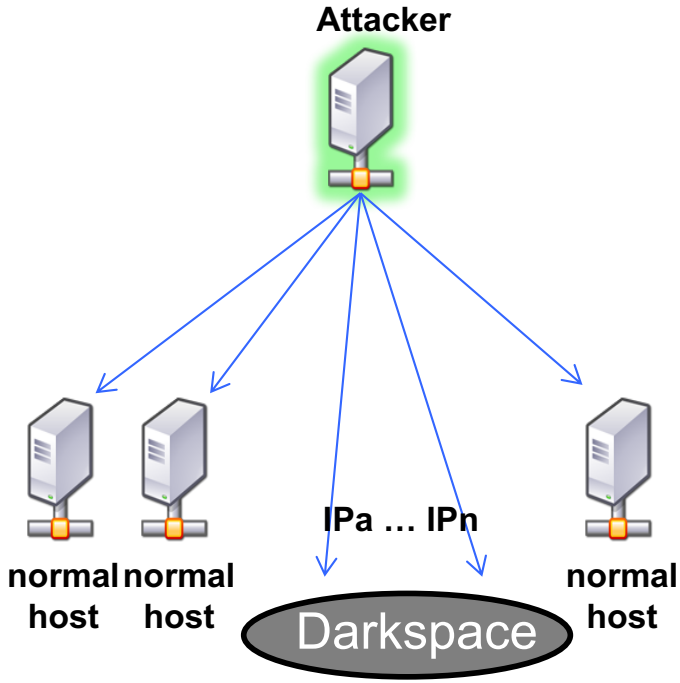
- Class objectives:
  - Familiarize students with network data analysis methods
  - Provide students in-depth understanding of TCP/IP flow behavior
  - Deepen students' network security knowledge
  - Enable students' general scientific work skills
  - Increase exploratory and forensics analysis skill
  - Awaken the scientist in each student

Lab1 (NetSec): IP Darkspace Analysis

- CAIDA IP Darkspace Data (Telescope Data)
  - Each Team gets different set of IP darkspace data
  - Students required to use recommended tools
- Exercises

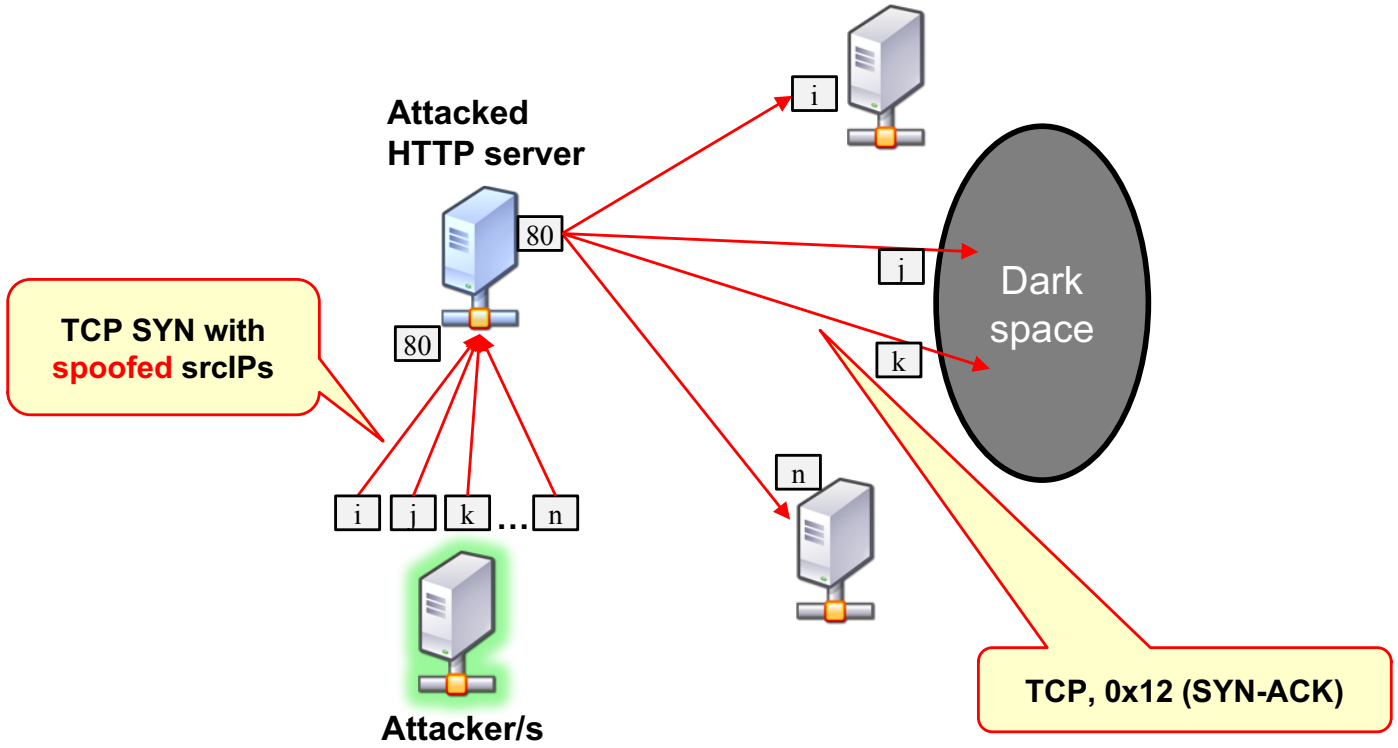
1. preprocessing/ aggregation
2. Uni/bivariate analysis
3. Time Series
4. FFT



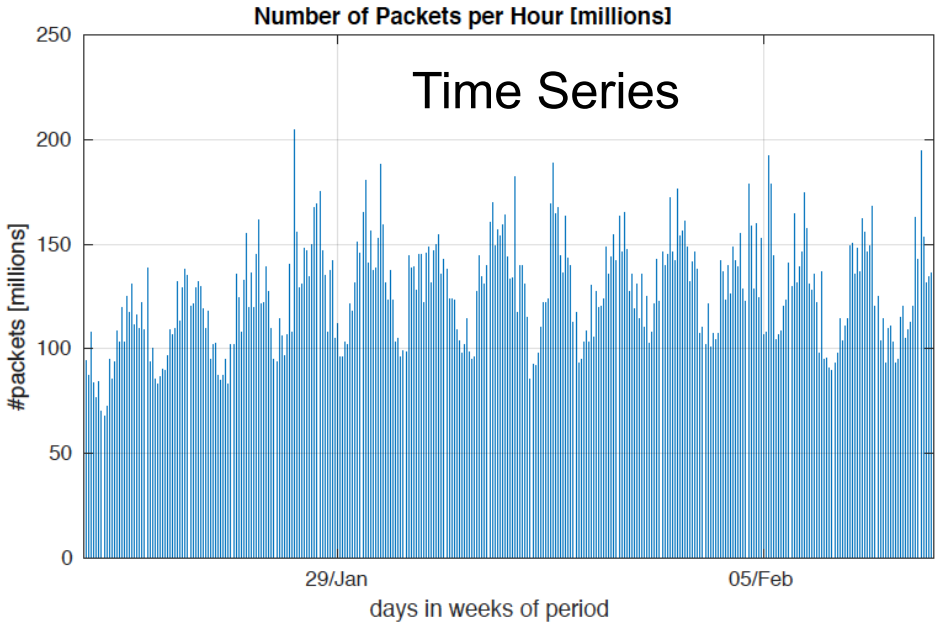


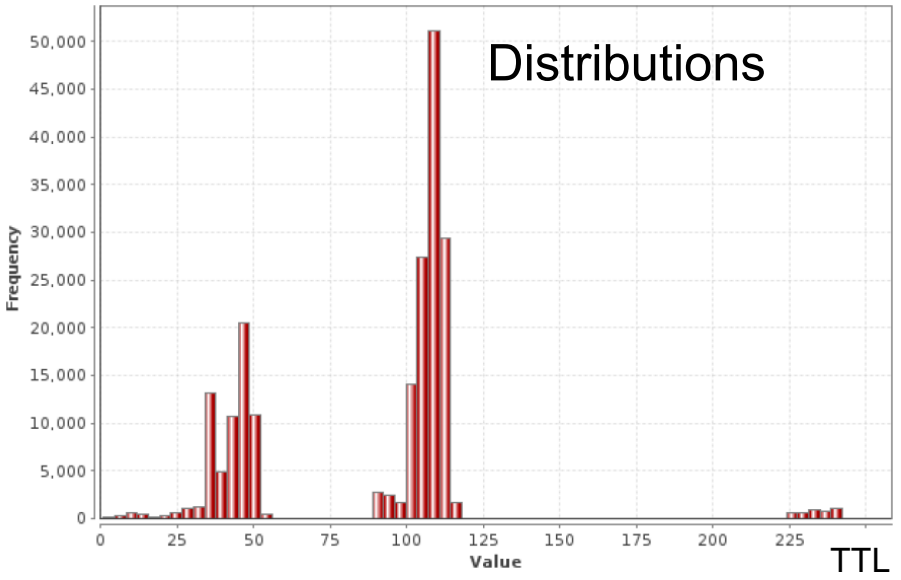
Identifying Backscatter

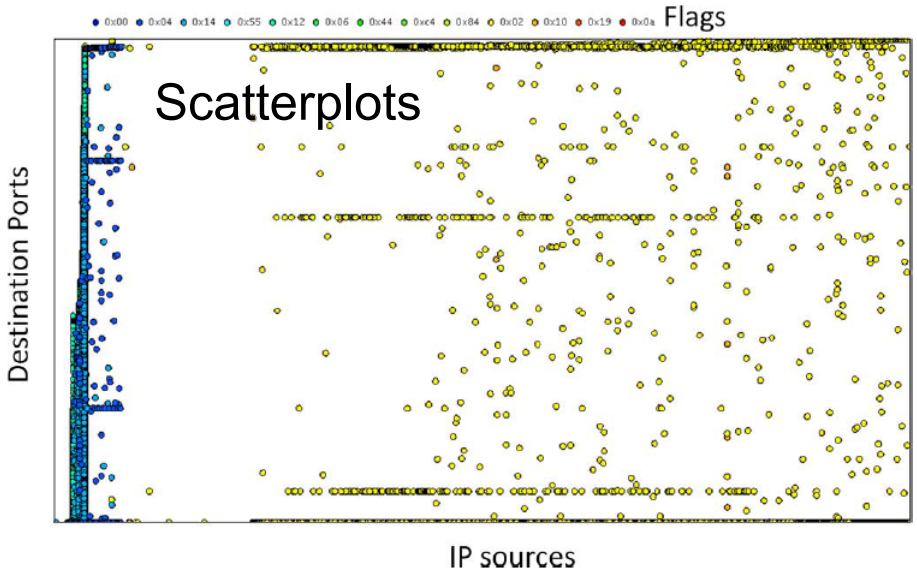
Name	Type	Statistics	Range
srcIP	nominal	mode = 191.191.71.228 (40052), least = 24	191.191.71.228 (40052), 0.11.59.95 (0), 0.119.105.126 (0), (
dstIP	nominal	mode = 0.196.51.3 (2), least = 0.76.243.80	0.128.198.153 (2), 0.129.230.94 (2), 0.130.14.70 (2), 0.130.1
srcPort	nominal	mode = 80 (40052), least = 3 (0)	80 (40052), 0 (0), 10023 (0), 10075 (0), 10081 (0), 10100 (0
dstPort	nominal	mode = 18285 (9), least = 0 (0)	18285 (9), 50218 (8), 17698 (7), 22878 (7), 39225 (7), 5792
Protocol	nominal	mode = 6 (40052), least = 1 (0)	1 (0), 6 (40052)
TTL	numeric	avg = 92.327 +/- 2.845	[38.000 ; 93.000]
flags	nominal	mode = 0x12 (40052), least = 0x00 (0)	0x00 (0), 0x04 (0), 0x14 (0), 0x12 (40052)
len	numeric	avg = 40 +/- 0	[40.000 ; 40.000]

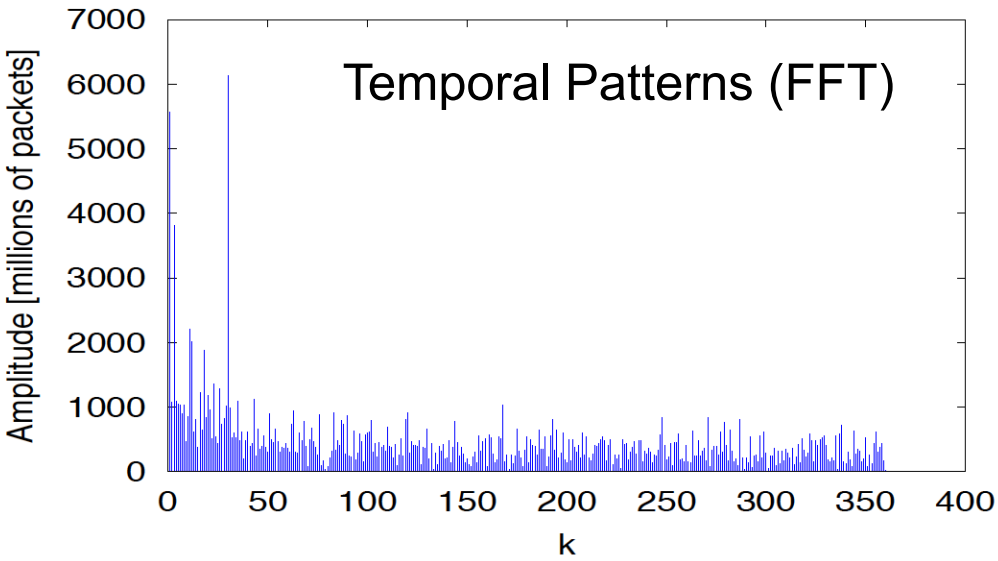


Data Analysis (Examples from Reports)

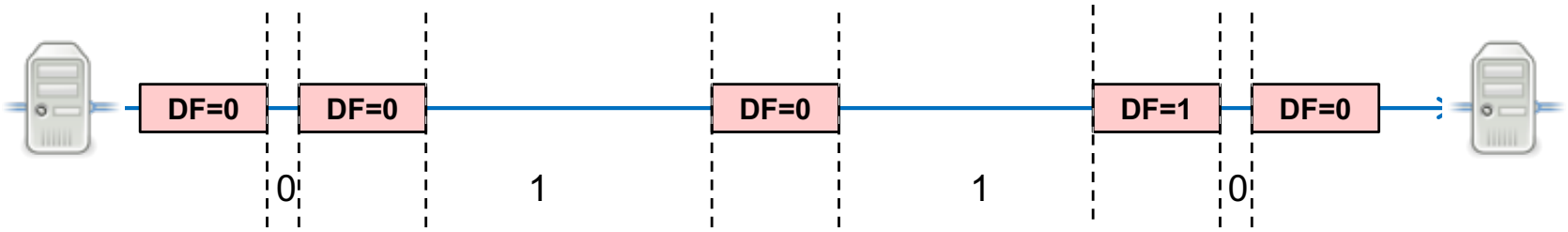








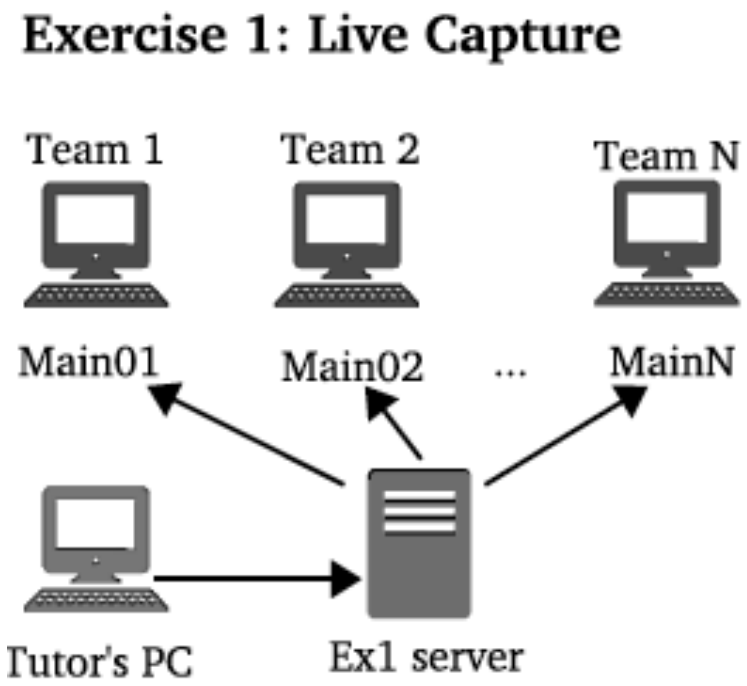
Lab 2 (NetSec Advanced): Network Steganography

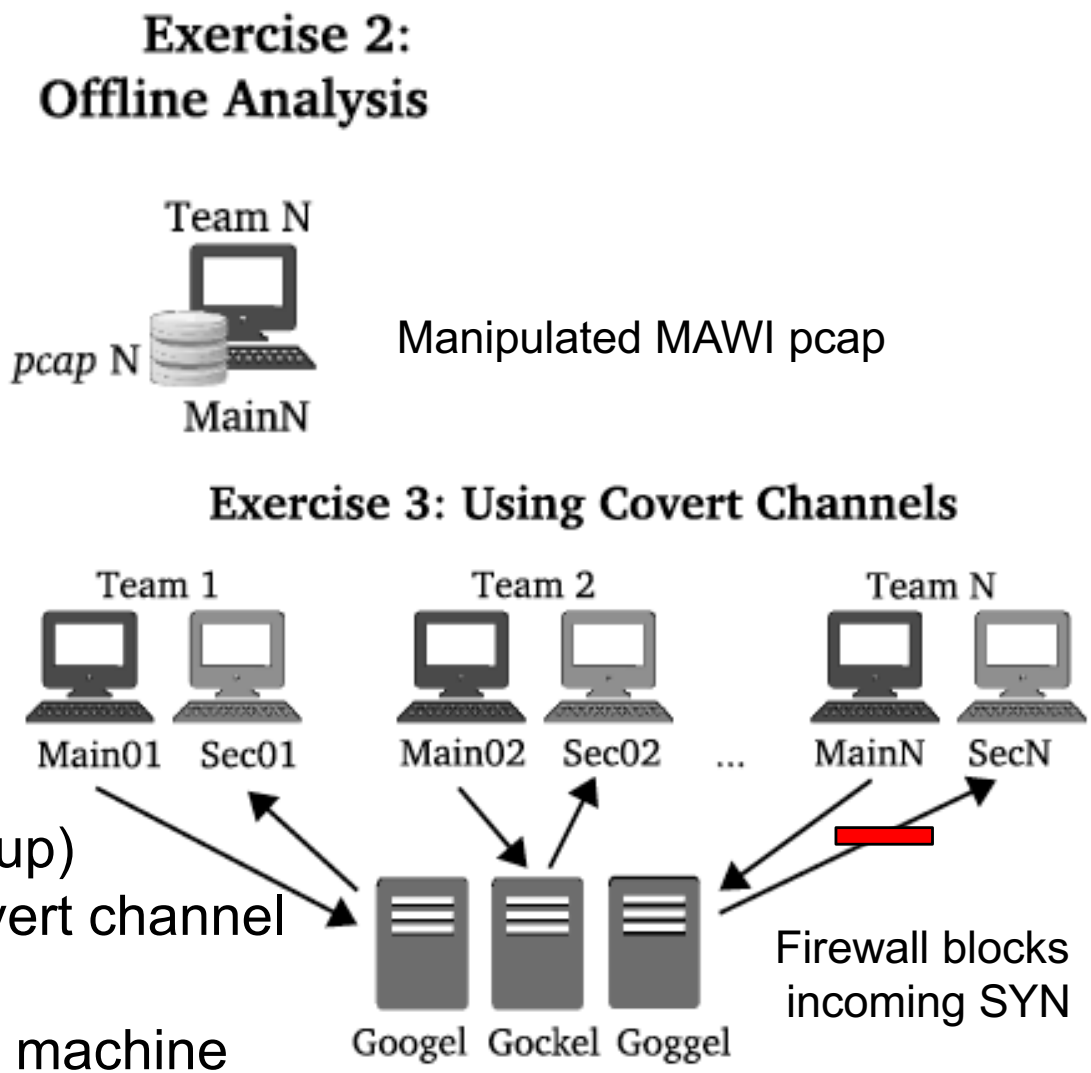


- Find hidden communication (timing, TCP/IP header)
  - Manipulated MAWI pcap file
  - Own tool to insert covert channels
  - Plain and encrypted covert messages
  - Each team different covert channels
- Establish own cover channel
- Exercises embedded in a story.

  "The Ministry of Cyber Affairs suspects that some data leakage has occurred… Find the covert channel in the pcap and identify the sender.
- Free choice of tools
- Hidden "easter eggs" ➔ bonus points

Exercises





1. Capture and analyze traffic (warmup)
2. Analyze pcap with (encrypted) covert channel
3. Establish own covert channel
   - Goal: send commands to remote machine
   - need indirect attack via servers to traverse firewall
   - send SYN with spoofed sIP, hide covert message in TCP ISN

Identifying Potential Sources of Hidden Communication

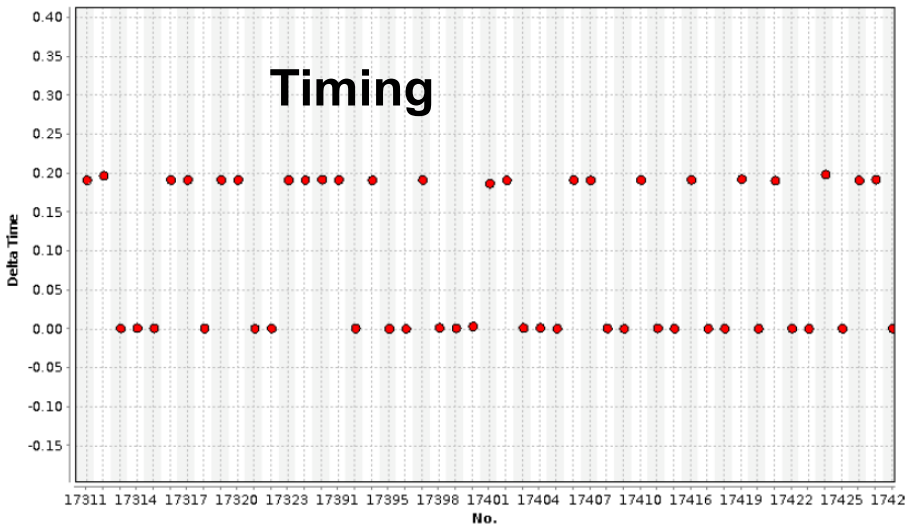
- Number of packets
- Number of states
- Multimodality Estimation

3 suspicious sources
multimodality >1,
enough packets,
24, 2 and 24 symbols

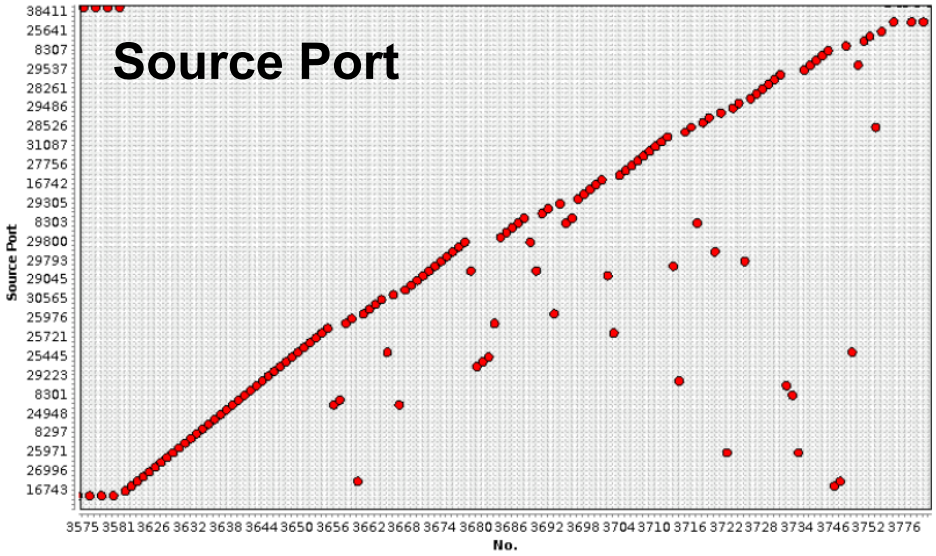
```
ibk@schulung1:~/Desktop/netsec_lab2/team2$ python scripts/srcfeat_power.py --inp
ut team2_clue.csv --feature TTL --sort-by Source
          Source           Packets          States           Symbols
 ------------------------------------------------------------------
 192.168.198.058              1085              24      13.55280957
 192.168.205.001                 3               1       1.00000000
 192.168.202.172                62               1       1.00000000
 192.168.198.049             16909               2       1.00036268
 192.168.198.057              1564              24      15.38429506
 192.168.205.062                 3               1       1.00000000
 192.168.198.001                14               1       1.00000000
 192.168.202.001                19               1       1.00000000
ibk@schulung1:~/Desktop/netsec_lab2/team2$
```

Iglesias, Annessi, Zseby: "*DAT detectors: uncovering TCP/IP covert channels by descriptive analytics*"; Security And Communication Networks, **9** (2016), 15; 3011 - 3029.

Searching for Hidden Communication (Examples)

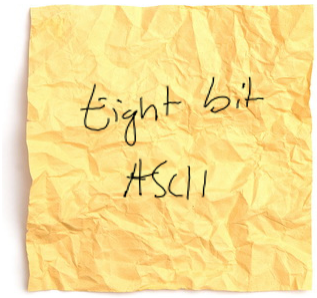


➔ 1100011101100….



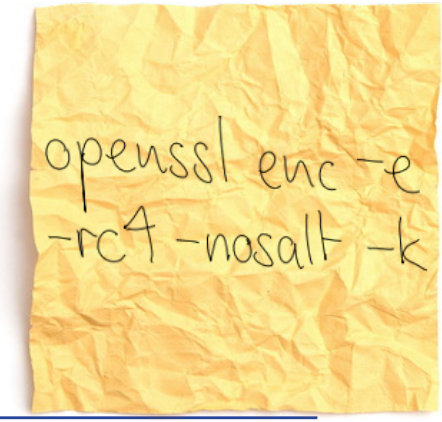
➔ 16743 26996 25971 …

Decoding the Messages (required some hints):



"Tom to Jerry: Starting transmission from Ministry of Cyber Affairs"

"Data acquired! For security reasons, next message will be encrypted. Key: WeOwnTheMinistry"



“Paranoia Mode”

As shown in figure 4, the value of the IPID field is monotonously increasing over time with a varying slope. By calculating the difference of the IPID of two consecutive packets (slope) and adding an offset of 97 to the result (in order to shift the value from the range 7 to 32 to the ASCII characters a-z), we identified the following word in a nonsense looking message:

.....{k[redacted]}.....,

it almost corresponds to the last name of [redacted] K[redacted] .

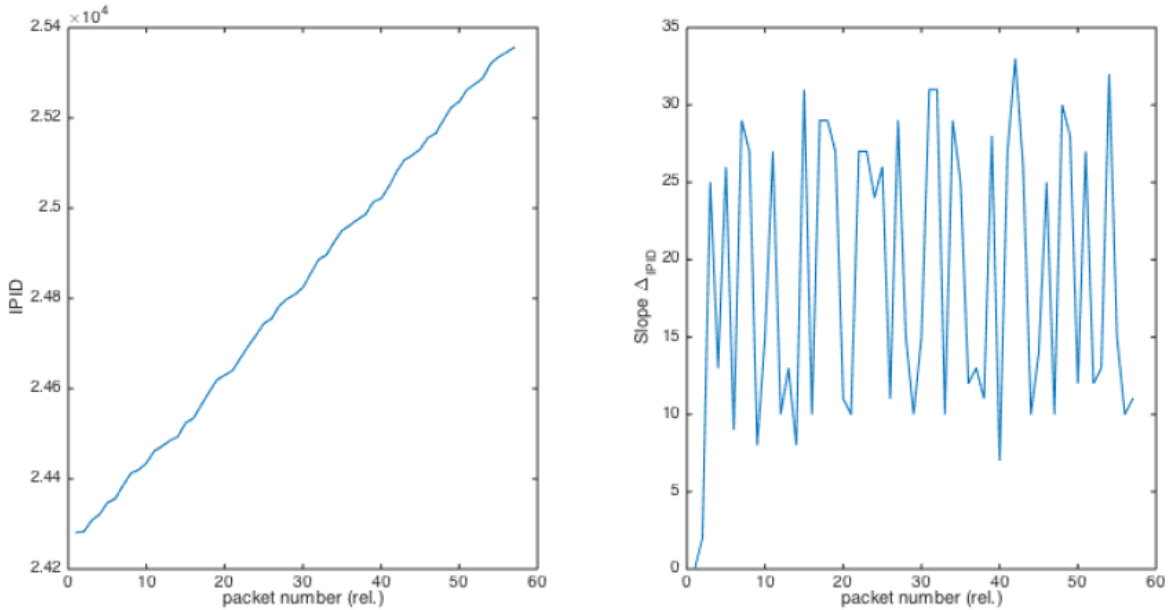


Figure 4: IPID value (left figure: IPID value over time, right figure: difference between IPID value of two consecutive packets)

Student Feedback

STUDENTS SELF-ASSESSMENT OF ACQUIRED SKILLS

1-strongly agree (positive), 5-strongly disagree (negative)

Question	max-min	mean
The course raised my interest in exploring the topic further.	1 - 2	1.29
Information was provided during the course on how I will be able to use the contents in the future.	1 - 3	1.50
The course increased my knowledge.	1 - 3	1.14
I am capable of using the knowledge I gained from the course.	1 - 3	1.21

[feedback provided by 14 students]

what did you enjoy most?

“labs were fun and engaging ”

“the moments: when you successfully finish an exercise”

What could be improved?

“tool-tutorials before the class ”

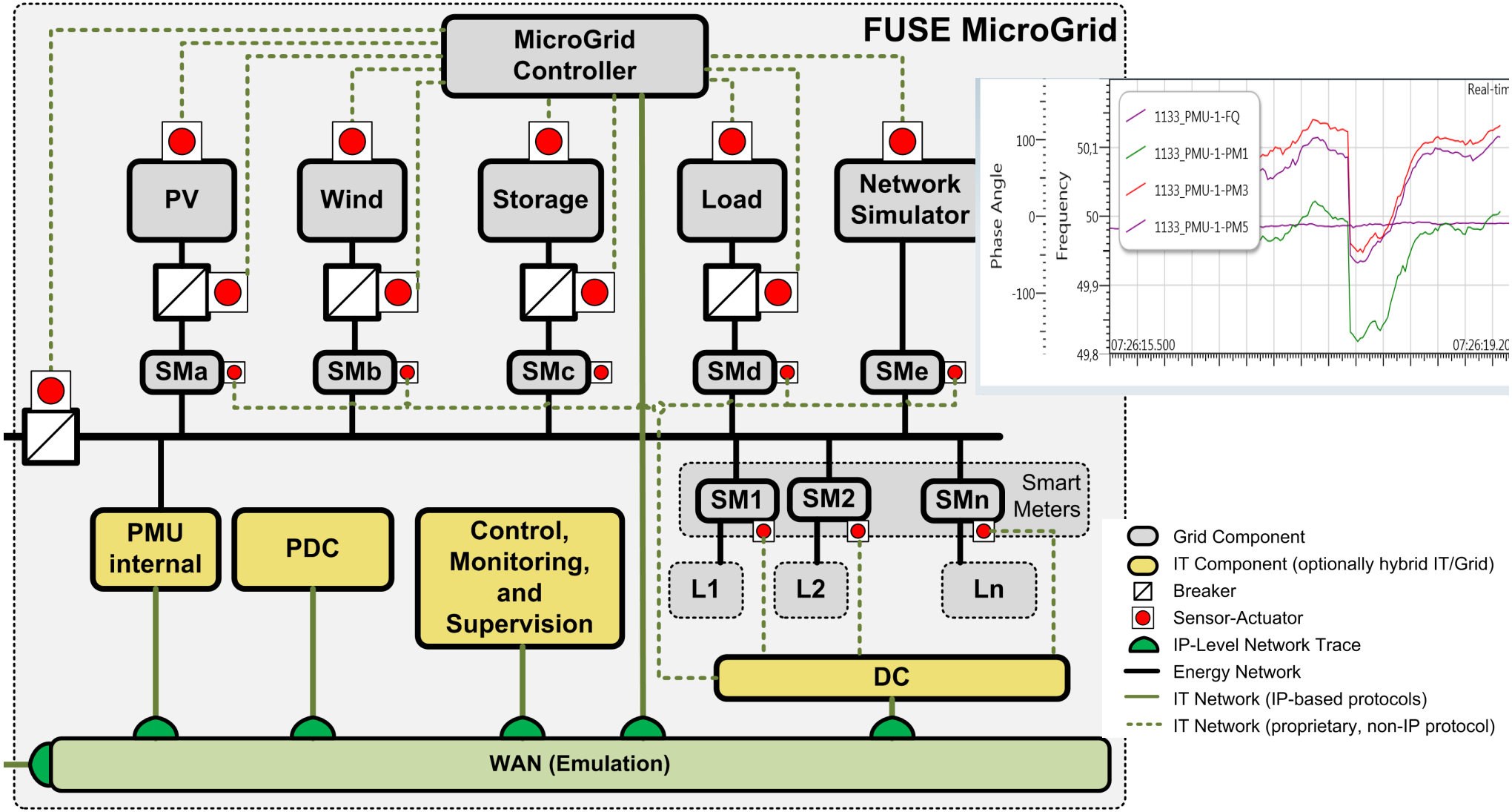
“more free exploration exercises”

“more exercises!; to be honest, I could have done another three exercises, it was fun!”

(Some) Lessons Learned

- Working with real measurement data
  - Boosts motivation
  - Triggers research spirit
  - Encourages to check theory vs. reality
  - Teaches responsible handling of data
  - Unique data set per team ➔ cheating detection
- Challenges
  - Data constraints (use agreement, working in lab)
  - Unexpected effects ➔ need to check data before
- Enforce pre-requisites
- Form heterogeneous teams
- Introduce variety of tools, then allow free choice
- “Keep it Fun!” (story, easter eggs)
- Future: BGP, Smart Grids

Future: Smart Grid Data



FUSE: Future Self-Organizing Energy Networks; Testbed

Xypolytou, Fabini, Gawlik, Zseby: "*The FUSE testbed: establishing a microgrid for smart grid security experiments*"; E&I Elektrotechnik und Informationstechnik, 2017, 1 - 6.

Available Material

- IP Darkspace Data ➔ available at CAIDA

http://www.caida.org/data/passive/telescope-educational_dataset.xml

- MAWI Data: http://mawi.wide.ad.jp/mawi/

- Teaching material ➔ available to other teachers
  - Exercise Sheets
  - Solver scripts
  - Report templates
  - Evaluation and Grading Scheme

http://www.tc.tuwien.ac.at/netsec-lab

Open PostDoc Position at TUWien

Zseby, Iglesias, King, Claffy: "*Teaching Network Security With IP Darkspace Data*"; IEEE Transactions on Education, **59** (2015), 1; 1 - 7.

Zseby, Iglesias, Bernhardt, Frkat, Annessi: "*A Network Steganography Lab on Detecting TCP/IP Covert Channels*"; IEEE Transactions on Education, **59** (2016), 3; 224 - 232.

Thank you!

tanja.zseby@tuwien.ac.at