PeekKernelFlows: Peeking into IP flows

Cynthia Wagner [1] Gérard Wagener [12] Radu State [1]
Alexandre Dulaunoy [2] Thomas Engel [1]

[1]University of Luxembourg
FSTC, Campus Kirchberg
L-1359 Luxembourg, Luxembourg

[2]SES S.A.
Château de Betzdorf
L-6815 Betzdorf, Luxembourg

September 14, 2010

Introduction

- Network flow visualization
  - Popular for human network traffic analysis
  - Evolution of flows per time
  - Identification of protocol patterns
  - Problems
    - Very close view (flow level)
    - Neglects topology information $\rightarrow$ subnet information
- Network flows aggregation
  - Get a broader overview
  - Includes topology information
    - Small thresholds $\rightarrow$ large amount of aggregated network traffic profiles
    - Large thresholds $\rightarrow$ small amount of profiles (information loss)

Aggregated network flows visualization

Spatial and Temporal Flow Aggregation

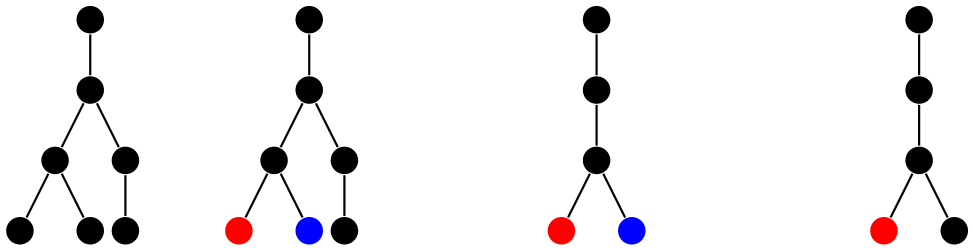
- We use the tool Aguri to create aggregated network profiles
- Overview at subnet level using the CIDR[1] notation

Example

```
%!AGURI-1.0
%%StartTime: Tue Dec 01 13:54:12 (2009/12/01 13:54:12
%%EndTime: Tue Dec 01 13:54:44 (2009/12/01 13:54:44
%AvgRate: 323.40Kbps
[src address] 1293591 (100.00)%
0.0.0.0/5 7351 (0.58%/99.22%)
     10.0.0.0/9 13545 (1.05%/30.79%)
        10.4.0.13 237599 (18.37%)
     10.91.0.0/24 19625 (1.52%/10.09%)
        10.91.0.22 100920 (8.57%)
        10.91.1.4 16664 (1.29%)
    72.0.0.0/5 21618 (1.67%/37.09%)
        74.125.79.91 202791 (15.68%)
        74.125.79.93 214301 (16.57%)
        74.125.79.99 257396 (2.12%)
        74.125.79.104 13649 (1.06%)
    83.231.205.49 324379 (25.08%)
        83.231.205.50 73506 (5.68%)
::/0 10067 (0.78%/0.78%)
```

[1] Classless-InterDomain Routing

Handling sequential Aguri profiles



T_1 T_2 T_3 T_4

Similarity $K(T_i, T_{i-1})$

$K(T_1, T_2)$ $K(T_2, T_3)$ $K(T_3, T_4)$

α_1 α_2 α_3 α_4

Visualize the sequence of similarities $\alpha_{1\ldots n}$

The kernel function $K(T_i, T_{i-1})$

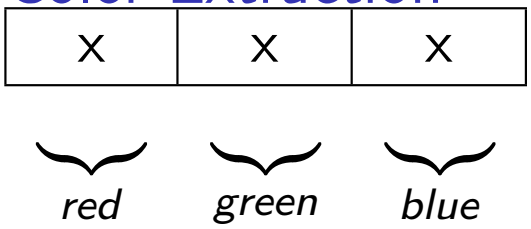
- Purpose
  - Compute similarities between two Aguri profiles
  - Purpose: Get a numerical value that can be mapped into the RGB space
- Similarity needs to take into account
  - Structural aspects
  - Volume information
  - Subnet aggregation
- Use a kernel from Machine Learning
- Take into account these aspects simultaneously

Visualizing Aguri Profile Similarities

- Map numerical values into RGB space
- Put the results into a bitmap
  - Sequentially align colored squares on the x-axis
  - Increment the y-axis with a squares length if the end is reached
- The more rectangles you have the more out-dated the overview is
- Normalize the similarities between 0 and 1
- Extend the space to 0xFFFFFF
- Extract the R,G,B components
- Colors can be manually amplified and shifted

Visualizing Aguri Profile Similarities

Color Extraction



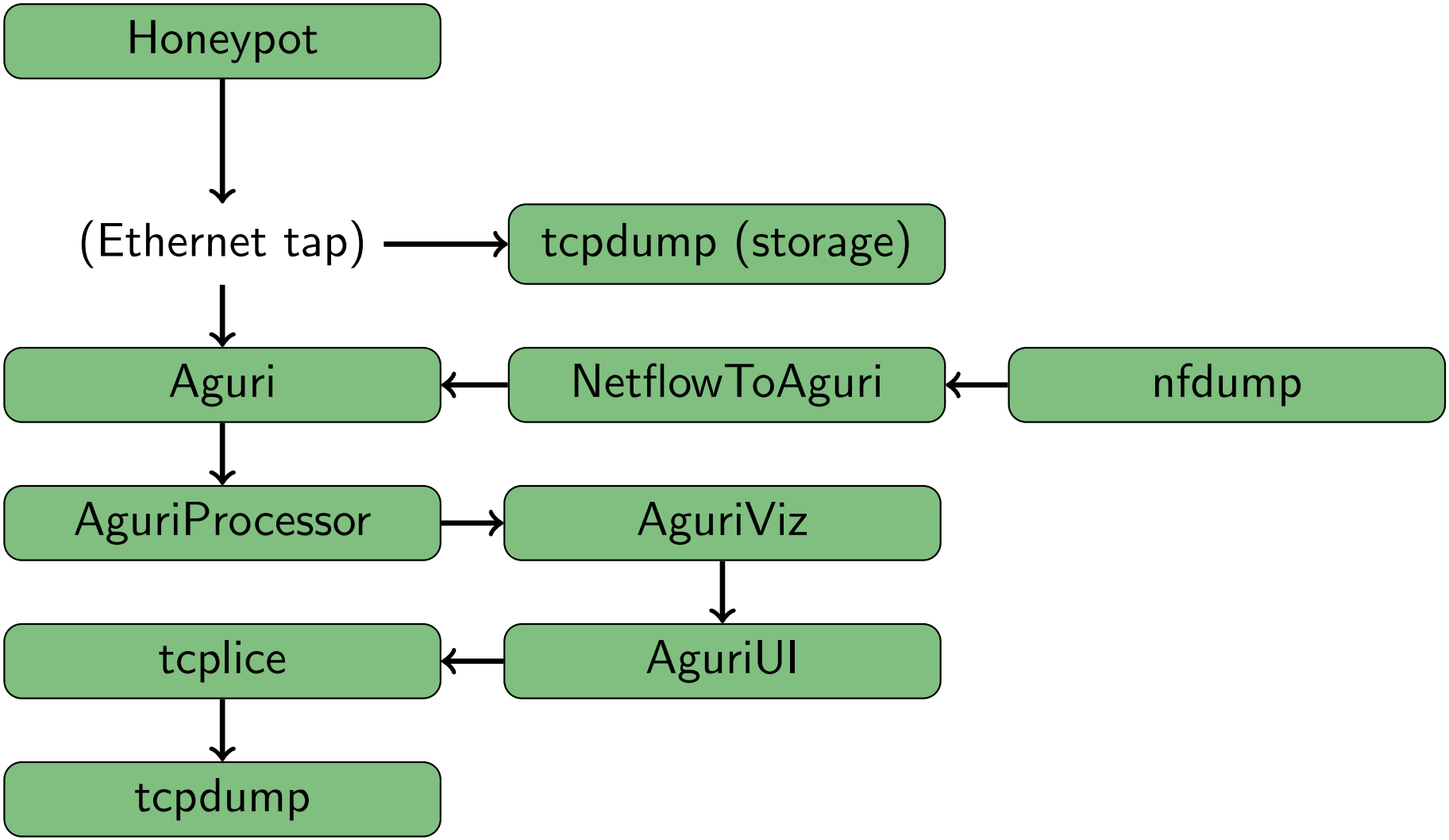
Color Interpretation

- Small similarities $\rightarrow$ bluish colors
- High similarities $\rightarrow$ erythroid colors

Use case - Honeypot operation

- Give the attackers bandwidth to become more attractive
- Avoid them doing real damage and receiving network abuse tickets
- Erythroid colors are not good for the honeypot operator

Implementation



Experiments

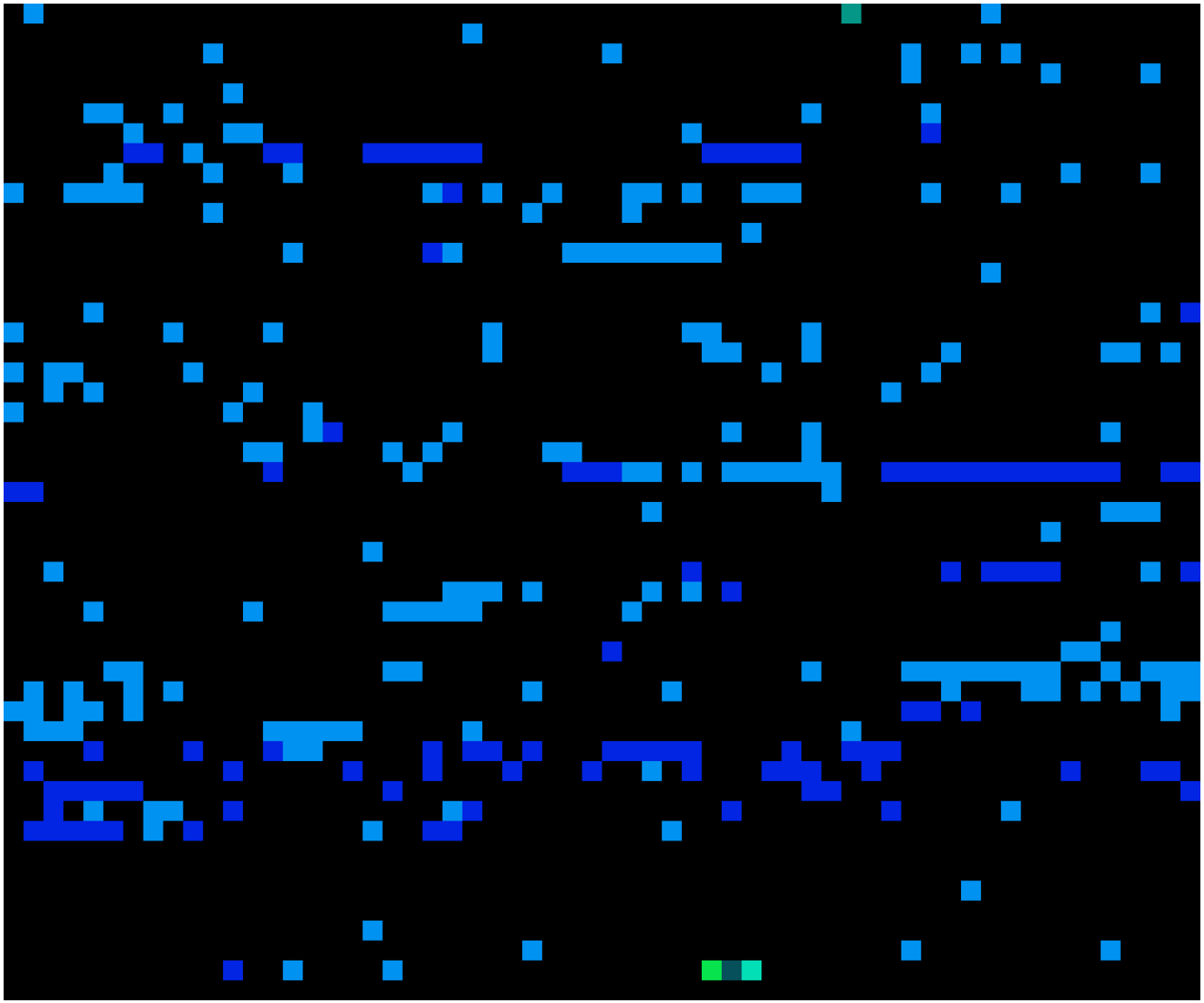
- Setup of a high-interaction honeypot (Linux OS)
- Challenge
  - Permit outbound connections $\rightarrow$ make honeypot attractive
  - Stop attacks when real damage is done

Datasets

Operation time	24 hours
Number of addresses	47 523
Used bandwidth	64Kbit/s
Exchanged TCP packets	1 183 419
α (seconds)	5
Colors (bit)	24

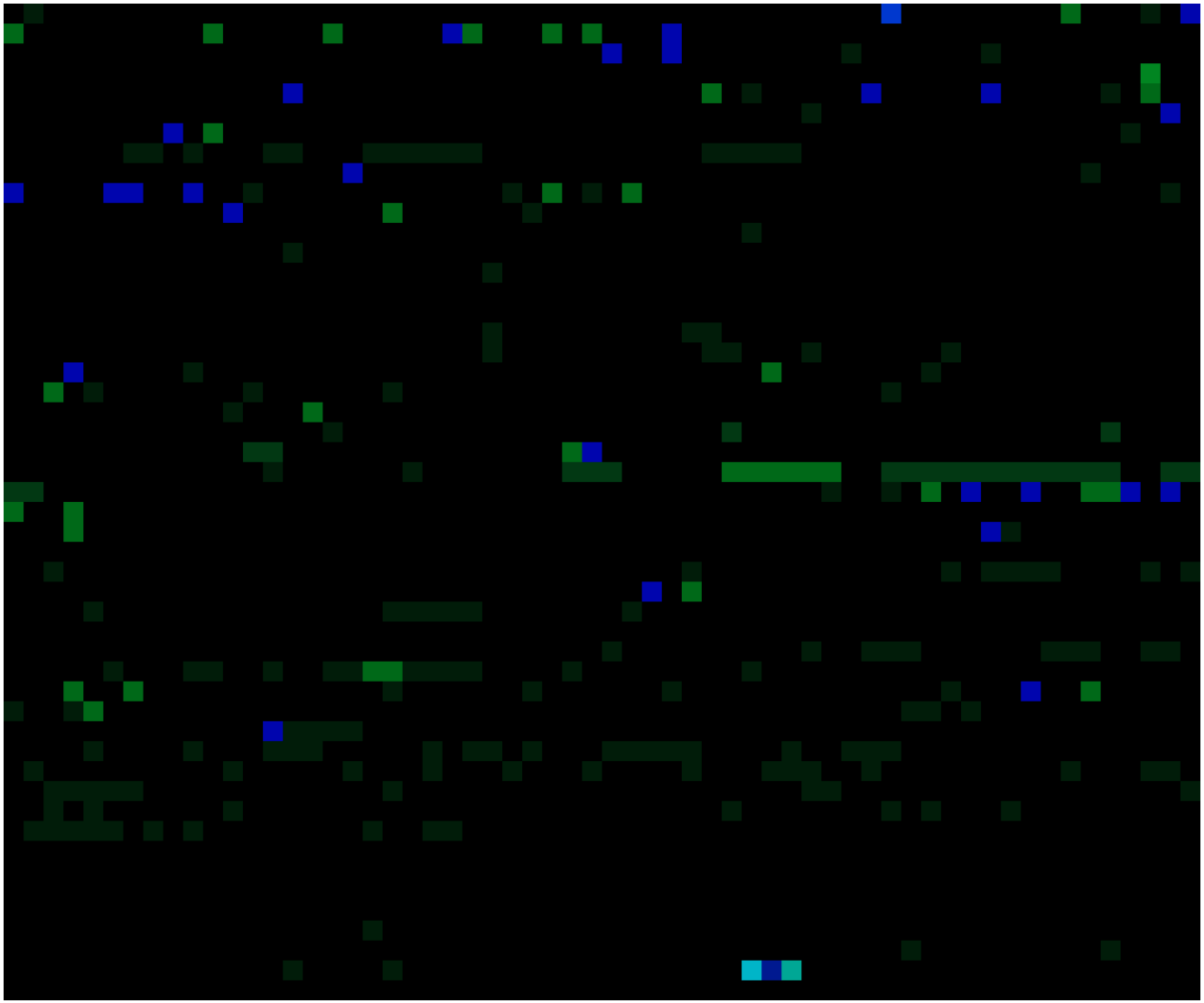
Experiments

Destination profiles



Experiments

Source profiles



Conclusions and Future Work

- Netflow monitoring and visualization is fine grained
- Aggregation parameters must be set in advance
- In this paper we visualize aggregated profiles
- Based on a kernel function to detect similarities
- Which include structural aspects and volume information
- Tested the approach on high-interaction honeypot traffic
- Black colors $\rightarrow$ let attackers play
- Bluish colors $\rightarrow$ attacks become more dangerous for the honeypot operator
- Erythroid colors $\rightarrow$ it is time to stop the attacks
- Need to improve human interaction features

Questions

Acknowledgements

This project is partially supported by the EFIPSANS EU-Project, CSC-SECAN-Lab and SnT. The more we want to acknowledge the National Research Fund Luxembourg, S.E.S.-Astra and RESTENA Luxembourg.