securityandtrust.lu

Automaton Models for Netflow Analysis

Fingerprinting and Classifying Participants

NMRG Workshop, Prague, Czech Republic
Friday, July 24th 2015

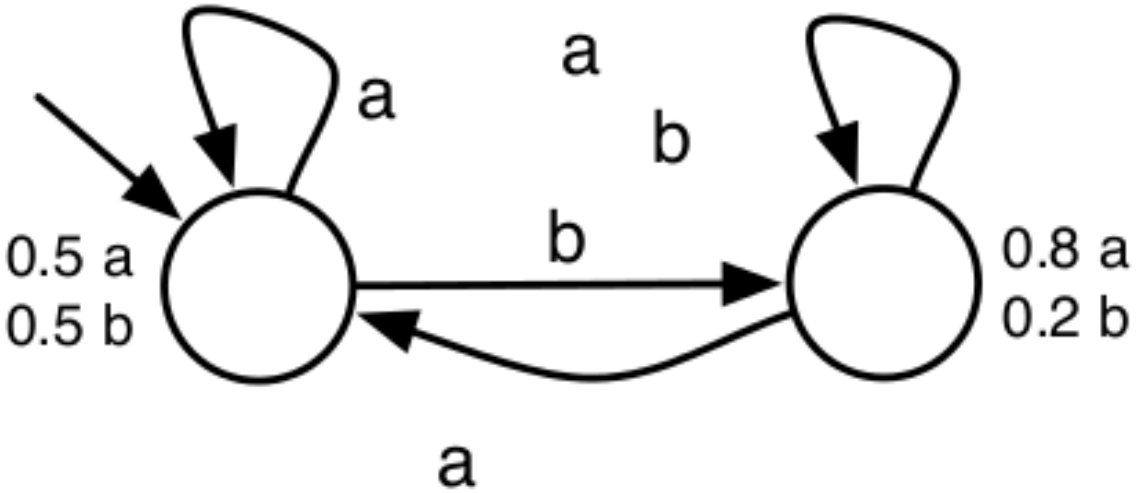
Christian A Hammerschmidt, christian.hammerschmidt@uni.lu

Interdisciplinary Centre for Security, Reliability and Trust
University of Luxembourg

UNIVERSITÉ DU LUXEMBOURG

Automaton Models

Short Overview



Fingerprinting with Automatons

Prediction, Classification, and Visualization (I)

securityandtrust.lu

Prediction

- predicting next states
- detecting outliers and anomalies

unsupervised

Classification

- classifying flows
- identifying type of activity or infection

(semi-) supervised

Fingerprinting with Automatons

Prediction, Classification, and Visualization (I)

securityandtrust.lu

Prediction

- predicting next states
- detecting outliers and anomalies

unsupervised

Classification

- classifying flows
- identifying type of activity or infection

(semi-) supervised

Fingerprinting with Automatons

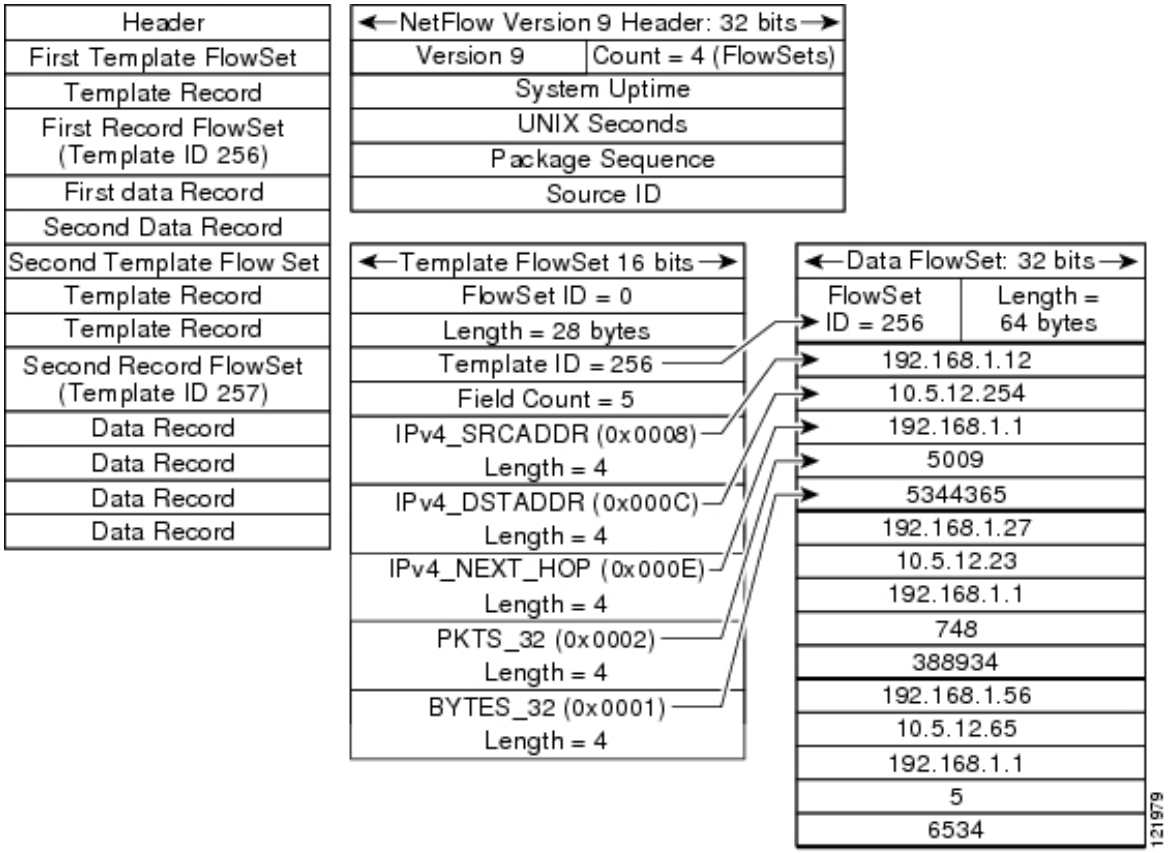
Prediction, Classification, and Visualization (II)

securityandtrust.lu

animation of automaton

Challenges

NetFlow Data as a (Regular) Language



[1]

[1] http://www.cisco.com/c/dam/en/us/td/docs/ios/ipv6/configuration/guide/ip6-netflow_v9.fm/_jcr_content/renditions/ip6-netflow_v9-1.jpg

Challenges

NetFlow Data as a (Regular) Language

From regression of numeric values to classification:

- via clustering to obtain few representatives

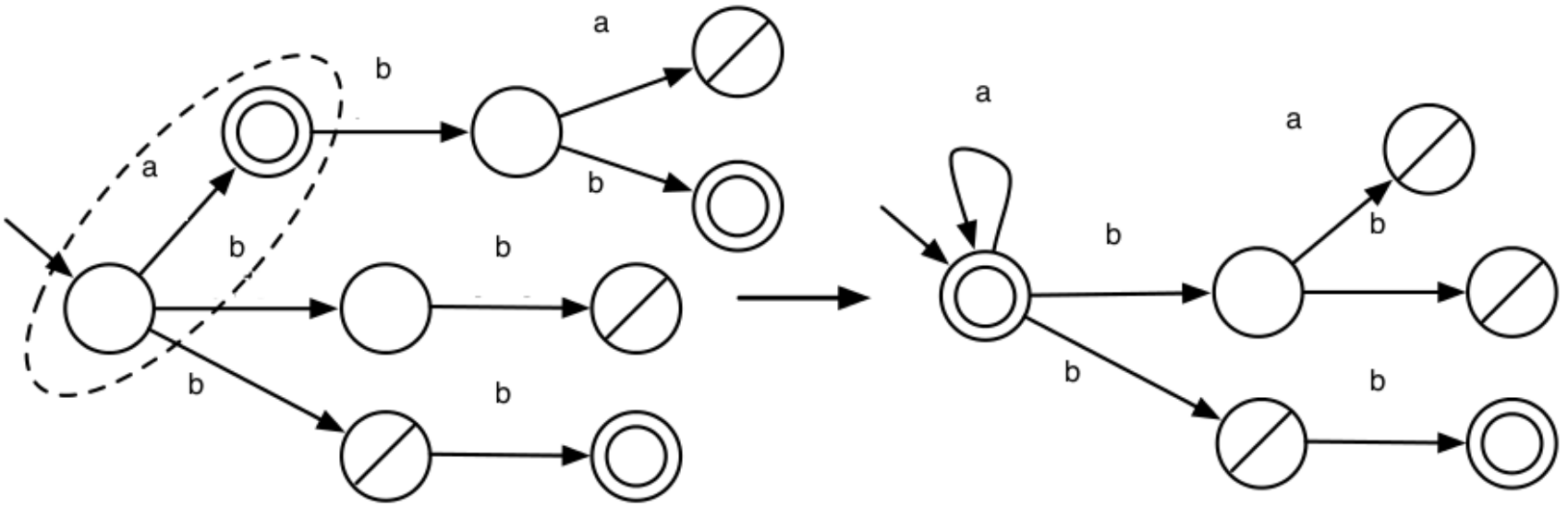
or through discretization

- via binning to obtain a discrete state space

What to choose?

Method

Learning State Structure from Data



2

[2] Taken from [2]

Evaluation

Data Set

securityandtrust.lu

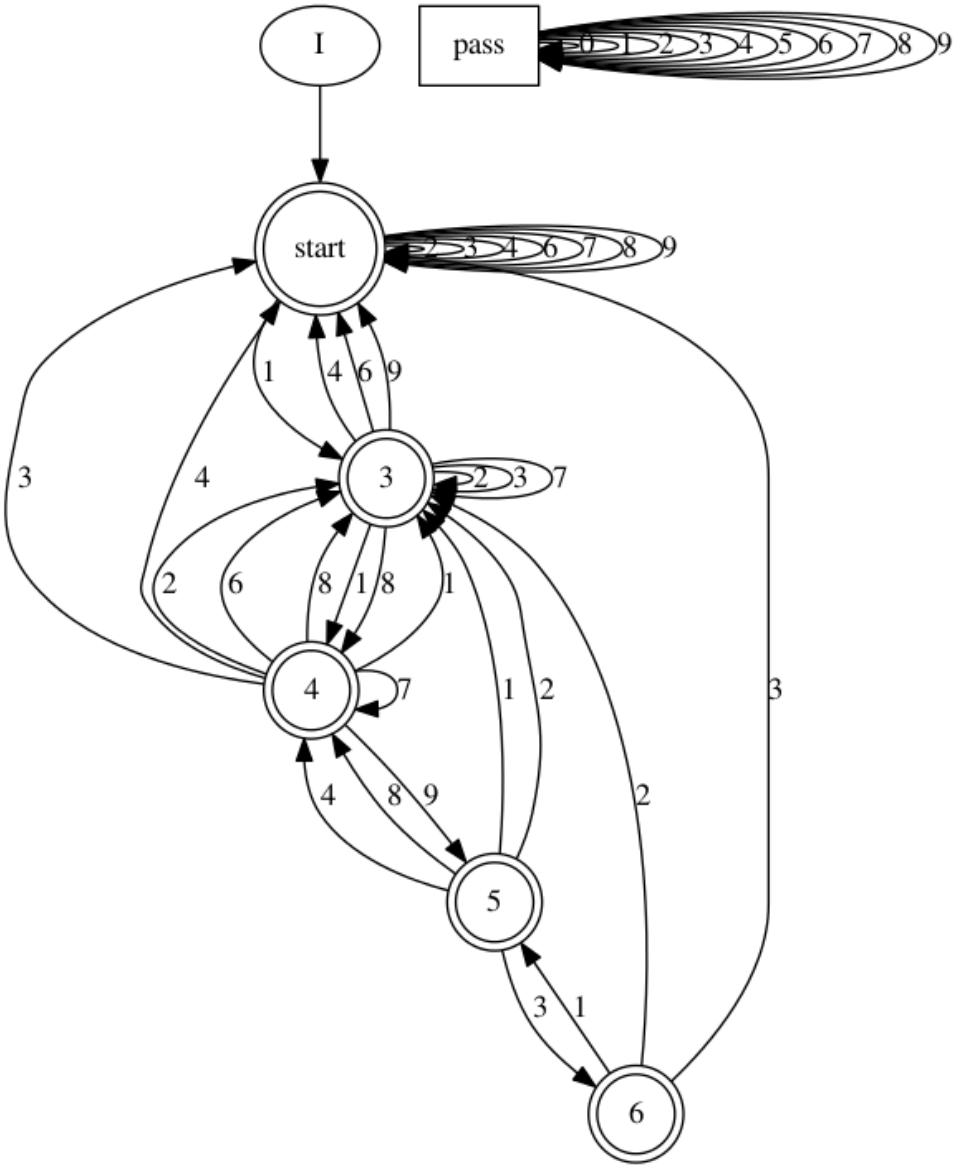
Experiments (on time-aggregated flow data):

1. predicting statistics for next flows
2. classifying flows on unlabeled data
3. classifying flows on labeled data[3]

[3] Using a botnet traffic data set[1]

Evaluation

Generated Automatons



Evaluation

Excerpt

securityandtrust.lu

Data Set	Experiment	Error / F_1 / *FPR*

Conclusion

Conclusion and Future Work

Results

- structure learning on netflow data is feasible
- initial results look very promising
- this is still work-in-progress and offers a number of ways to improve

Further Research

- compare performance to other fingerprinting solutions
- apply a more expressive automaton model

Conclusion

Conclusion and Future Work

securityandtrust.lu

Results

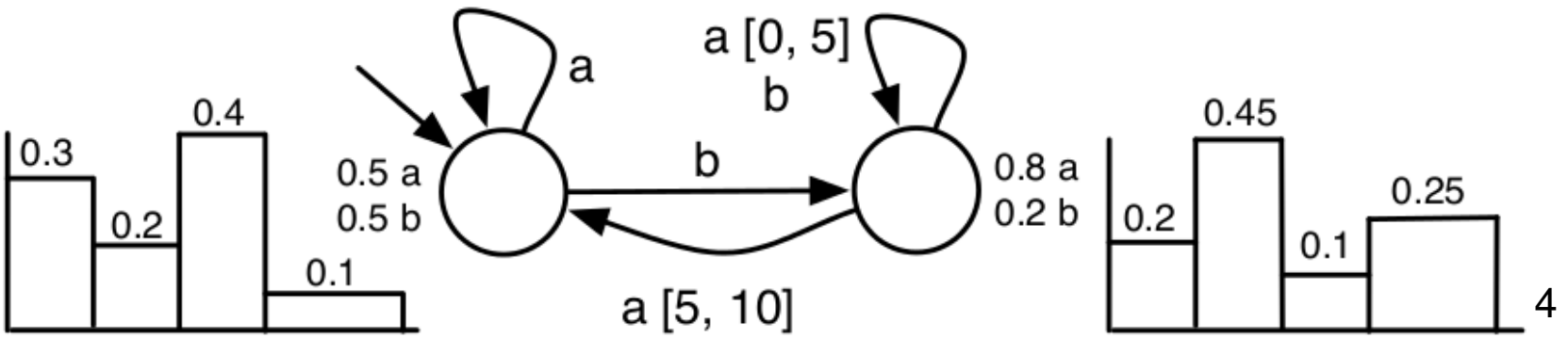
- structure learning on netflow data is feasible
- initial results look very promising
- this is still work-in-progress and offers a number of ways to improve

Further Research

- compare performance to other fingerprinting solutions
- apply a more expressive automaton model

Future Work and Extensions

Currently Ongoing Research



[4] Taken from [2]

Thank You!

Time for questions.

References I

García, S. and Grill, M. and Stiborek, J. and Zunino, A.
An empirical comparison of botnet detection methods
Computers & Security, 2014.

S. E. Verwer, C. Witteveen, M. M. De Weerdt.
Efficient identification of timed automata: Theory and practice, March 2010.

Heule, M.J.H., Verwer, S.,
Software model synthesis using satisfiability solvers.
Empirical Software Engineering 18, 825–856., 2013