TAIPAN:

A Rule Learner for Positive and Negative Rule Generation Applied to Security Alarm Filtering

Pascal Baumgartner
University of Applied Science Biel-Bienne (baump2@bfh.ch)
Swisscom Innovations (pascal.baumgartner@swisscom.com)

TAIPAN

Overview

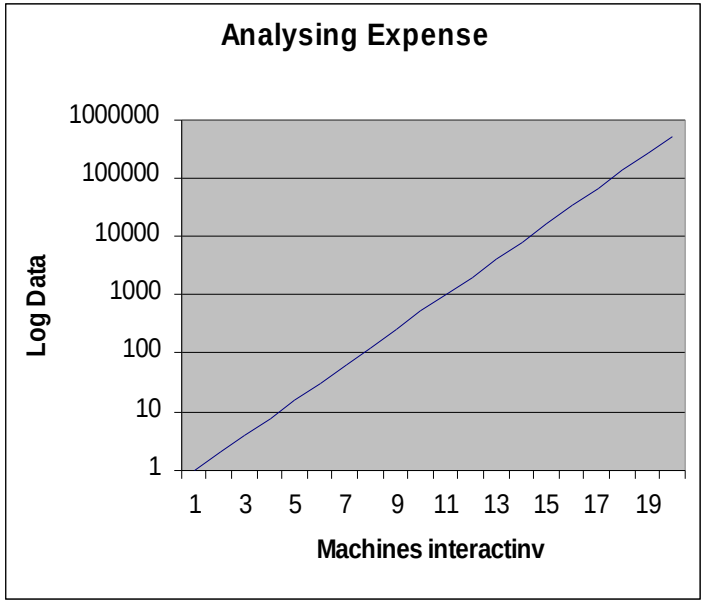
- Motivation
- Filter Problem
- TAIPAN
- Rule Generation
- Classifier
- Results / Performance

TAIPAN

Motivation

Problems in todays Intrusion Management Systems:

- Growing networks, bigger and more logs, human information overload



TAIPAN

Filter Problem

- Admin has to create alarm filters (rules) to avoid False Positives
- Networks are dynamic so filters need to be updated on a regular basis
- Time consuming process
- Subjective Knowhow of Admin reflects quality of alarm filters

Goal:

- Reduce the Admin's workload
- Produce objective and “strong” Rules directly from the data

TAIPAN

TAIPAN

- Rule learning algorithm
- Inherited from FOIL, RIPPER and CPAR
- Can create filter rules directly from training data
- Fast and robust
- Rules set easy to modify for an admin

TAIPAN

Rule Generation

- SELECT all close-to-the-best features and create the initial rules
- TEST coverage of the just created rules
- IF rule covers negative examples -> CONTINUE rule growing
- IF rule covers no more negatives -> CALCULATE quality value and ADD rule to the rule set, ADJUST weight of covered positives
- STOP growing rules when sum of all weigths of the positives examples falls below a specified threshold

TAIPAN

TAIPAN's Classifier

- Applies the rules to new log messages in the same order as the rules were generated
- The first rule that “fires” classifies the new log message
- If no rule “fires” the message to classify is considered as a negative example (default rule)

TAIPAN

Results / Performance

- TAIPAN generated rules were shorter and less complex than human generated rules
- Short rules are favorable for a fast classifier
- On the syslog data the highest percentage of False Negatives was below 0.15‰ of the whole test data set
- Runtime: Training for 1000 logs in 5 seconds / Runtime: Classification for 1000 logs < 1 second

TAIPAN

Conclusion

- TAIPAN is useful in practise for automated filter creation
- TAIPAN outperformed Safeguards Baeysian approach in terms of Rule quality
- TAIPAN is implemented in Perl, a compiled Language e.g. C might accellerate rule generation by a factor of 2-3

TAIPAN

Questions ?

Pascal Baumgartner
University of Applied Science Biel-Bienne (baump2@bfh.ch)
Swisscom Innovations (pascal.baumgartner@swisscom.com)

TAIPAN