Automatic Validation and Evidence Collection of Security Related Network Anomalies

1 Motivation

- Non-negligible false positives in NetFlow based Anomaly Detection tools
- Operators need compact display of all anomalous flows to investigate

2 Apriori algorithm (IMC 2009)

- Anomalies often generate a lot of flows
- Apriori finds large sets of flows (***itemsets***) sharing a combination of features
- Key point: get anomaly indication (set of features, time, duration) and look "around" it (union of features, NOT intersection)

3 Towards Automatic Anomaly Validation

Apriori fine-tuning for GÉANT network

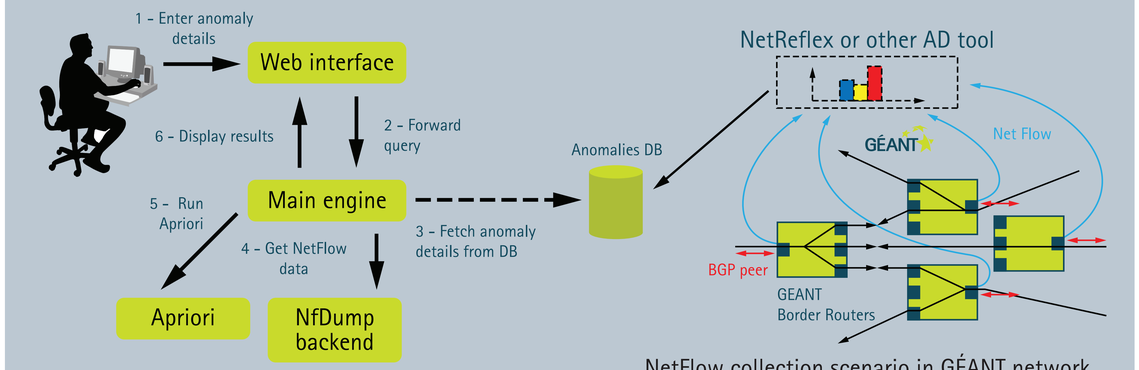
- Apriori misses attacks with a lot of packets but few flows (DoS UDP floods) → Apriori modified version looking for ***itemsets*** with high #packets
- Automatic self-adjusting of Apriori ***minimum_support***
- Filter out Apriori's output (top N)

Using Apriori to help network operators

- Provide all IP flows potentially related to an anomaly in a compact way

	Source IP	Destination IP	Source Port	Destination Port
Portscans	X.191.64.165	Y.13.137.129	55548	*
	X.191.26.33	Y.13.137.129	39573	*
DDoS	*	Y.13.137.129	3072	80
	*	Y.13.137.129	1024	80

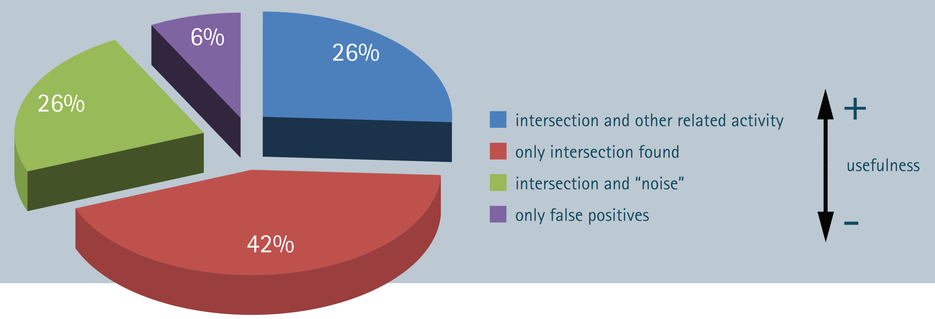
4 Apriori-based GUI



NetFlow collection scenario in GÉANT network

5 Results

- In 94% of the cases Apriori finds the main anomaly under investigation (**low false negatives**)
- In 26% of the cases it also signals other simultaneous attacks to same target as well as attacks to different targets
- Helpful for spotting false positives for DoS



6 Conclusions

- No need to do the drill down anomalies manually
- Additional interesting information
- Quicker and more reliable decisions

www.upc.edu www.dante.net

Ignasi Paredes-Oliva*, Pere Barlet-Ros*, Maurizio Molina‡
Universitat Politècnica de Catalunya (UPC), Barcelona (Spain)*: {iparedes,pbarlet}@ac.upc.edu; DANTE, Cambridge (UK)‡: maurizio.molina@dante.net

This work was supported by a STSM grant from TMA COST Action IC0703, " Data Traffic Monitoring and Analysis"