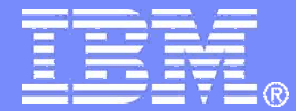


Zurich Research Laboratory

A Two-Layered Anomaly Detection Technique based on Multi-modal Flow Behavior Models

Marc Ph. Stoecklin <mtc@zurich.ibm.com>
Jean-Yves Le Boudec <jean-yves.leboudec@epfl.ch>
Andreas Kind <ank@zurich.ibm.com>

Outline

Part 1

- Background and Motivation
- Observations

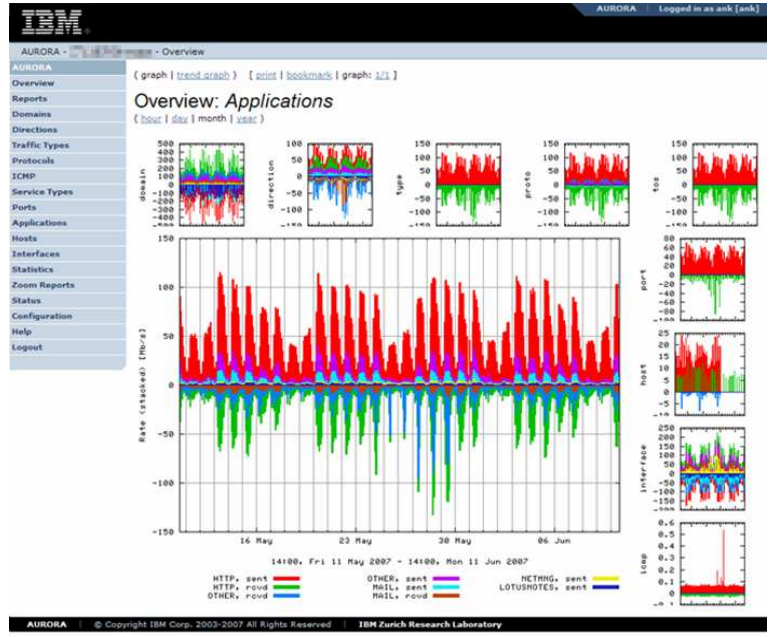
Part 2

- The proposed Technique
  - Methodology
  - Learning phase: behavior mode extraction
  - Detection phase: non-linear dual-layered detection
  - Discussion
- Results

Part 3

- Limitations
- Conclusion and Future Work

Background and Motivation



- **Background**
  - **AURORA** traffic monitoring system
  - **Detect traffic anomalies** and **provide decision support**
  - Anomalies of interest: Attacks/abuse, user behavior changes, failures, configuration errors

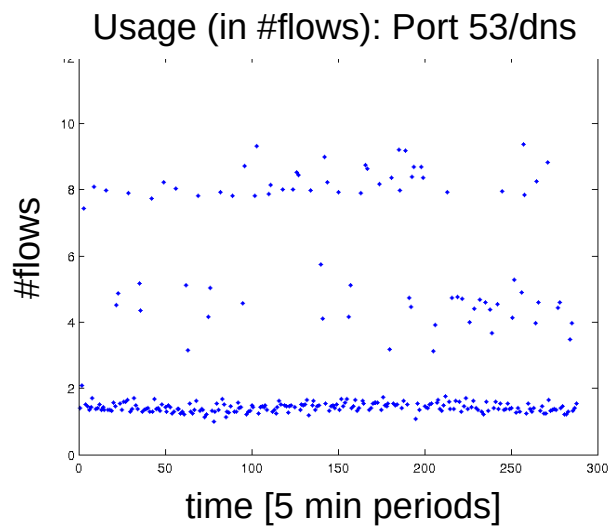
- **Design Goals**
  - **No explicit prior knowledge …**
  - … about what is normal and abnormal traffic. No attack-specific detector!
  - **Dynamic nature of the traffic mix**
  - **On-line detection algorithm**
  - **Interpretable models and results**
    - Understand changes with respect to **baseline models**
  - Incorporation of **administrator feedback**

Characterization and Observations

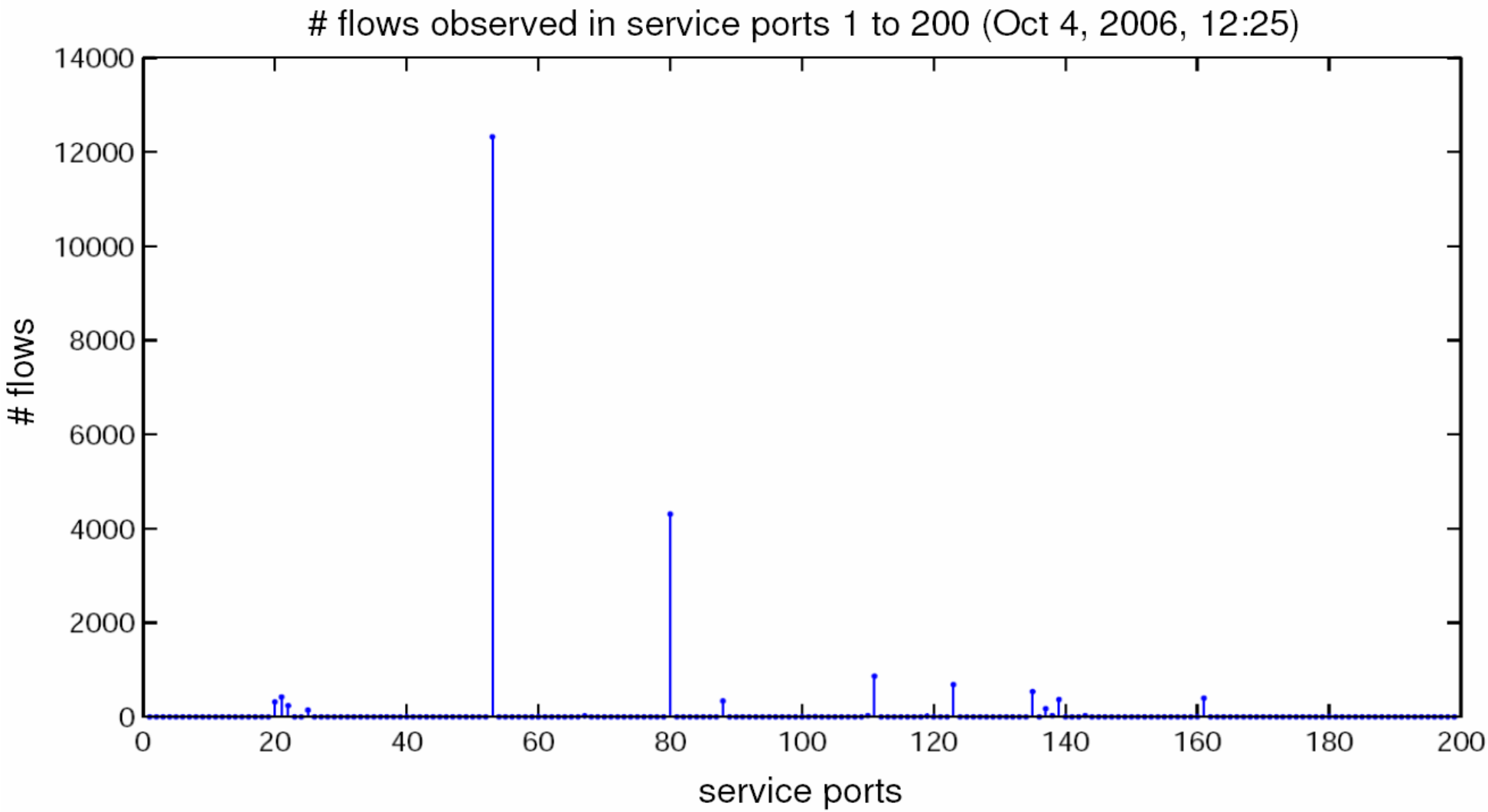
- Traffic anomalies typically leave traces in multiple traffic features
- Anomalies exhibit **different behavior patterns** in multiple traffic features
  - Isolated/single components vs. multiple components

Anomaly	service ports	IP addresses	TCP flags	Other features
Worm outbreak	single dst ↑	multiple dst ↑	SYN, RST/ACK ↑	scan behavior
DoS attack	single dst ↑	single dst ↑	SYN ↑	octs/pkts: single ↑
Host failure	single src ↓	single src ↓	SYN ↑	ICMP

- **Observations**
  - A network has multiple **network behavior modes** (e.g., diurnal, patch)
  - Components (i.e., individual ports, hosts, applications) have their proper **operation modes** (work hours, replication, updates, backups, etc.)
  - **Modes of components do not necessarily coincide with network modes**



Flow-count histogram (5 min period)

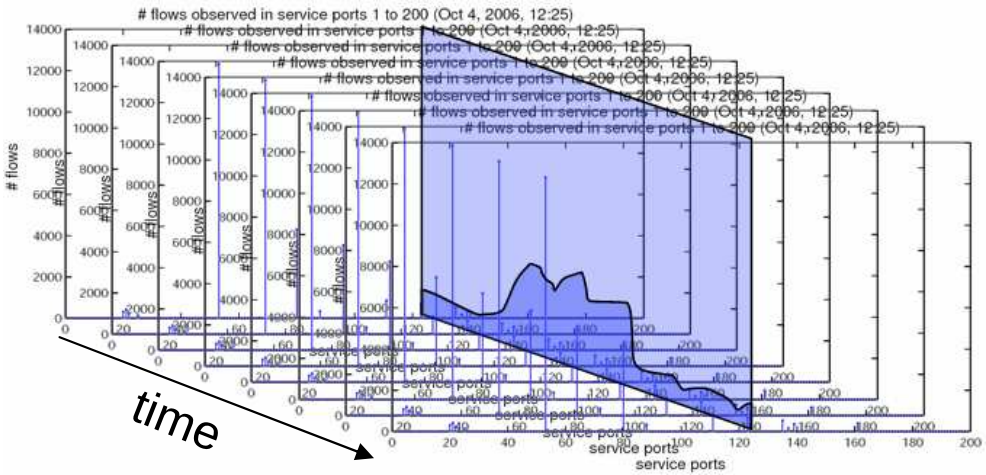


Methodology and Terminology

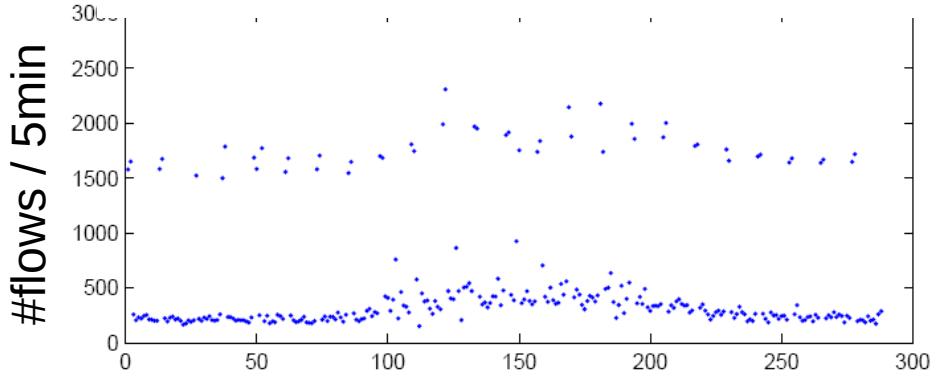
- **Learning phase**
  - Unsupervised learning of baseline behavior from unlabeled data
  - Train the detection logic
- **Detection phase**
  - Compare observed network behavior to the baseline behavior
  - Perform a detection operation

- **Features**
  - IP addresses, service ports, applications, TCP flags, ICMP types, avg. pkt size, etc.
    ... and compounds thereof!
- **Components**
  - Actual values of the traffic features, e.g., 10.1.1.2, port 53, MAIL, SYN/ACK, etc.

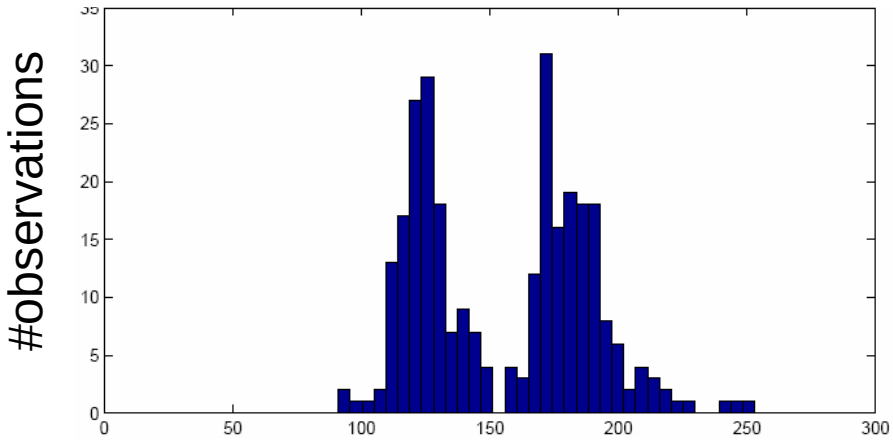
Learning phase: Behavior mode baselining



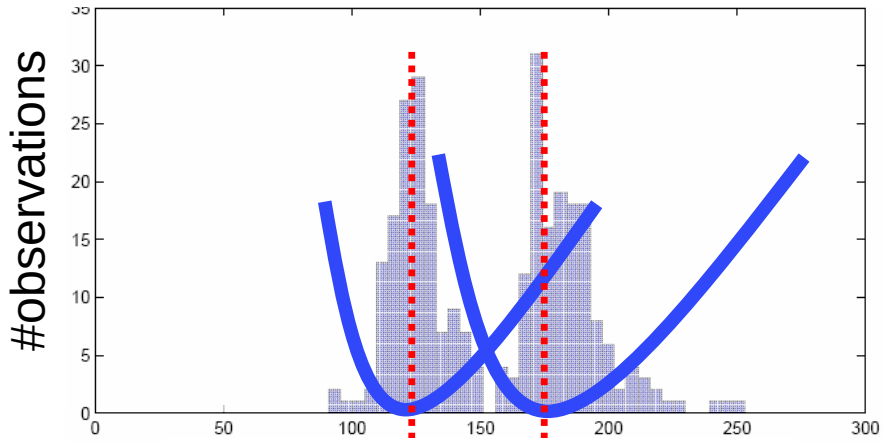
Time series



Frequency distribution



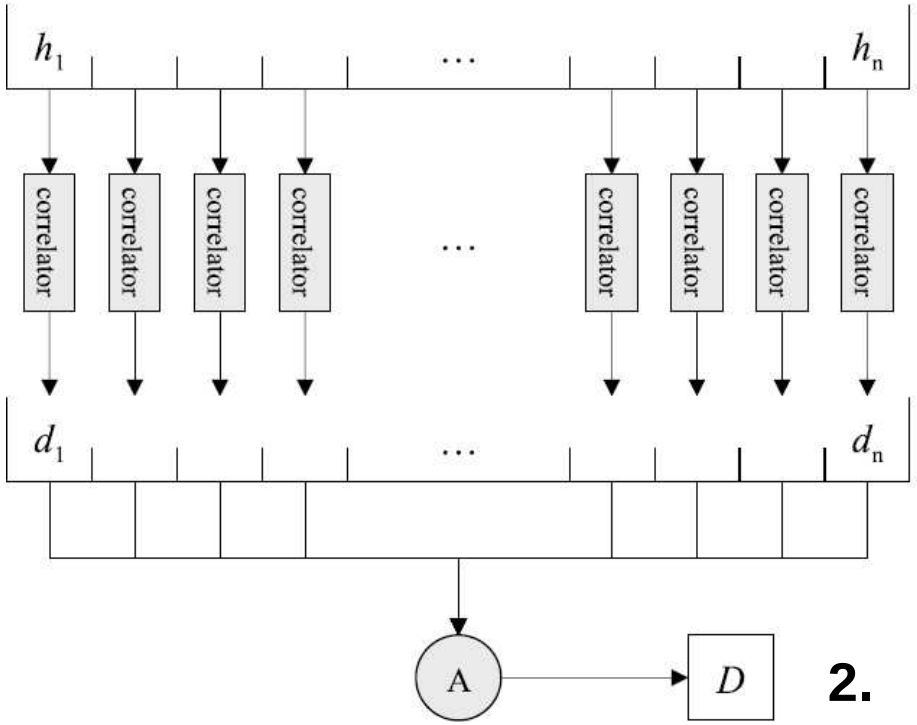
Mode extraction and model derivation



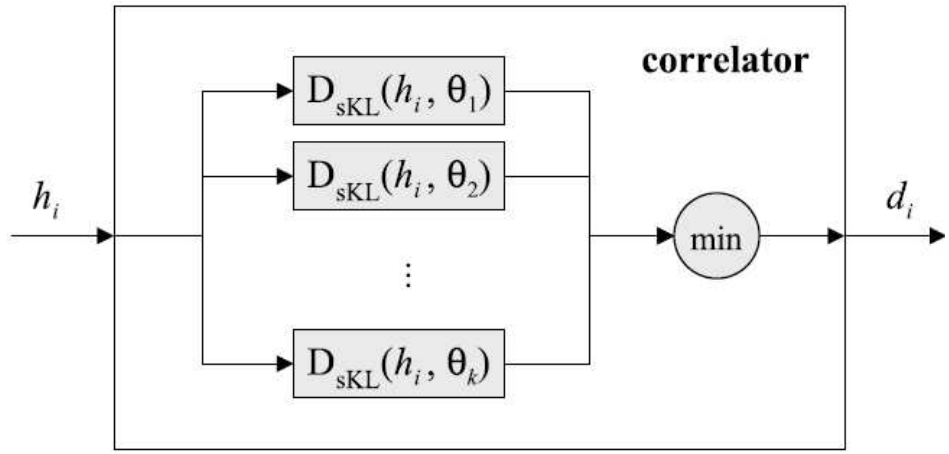
Detection Phase (Adapted Correlation-type Demodulator)

- Construct the **best matching** flow-count **histogram** (non-linear system)
- **Two detection layers**
  1. **Component-wise** baselining and detection (host failures, DoS attacks)
  2. **Feature-wise** detection (scans, worm outbreaks)

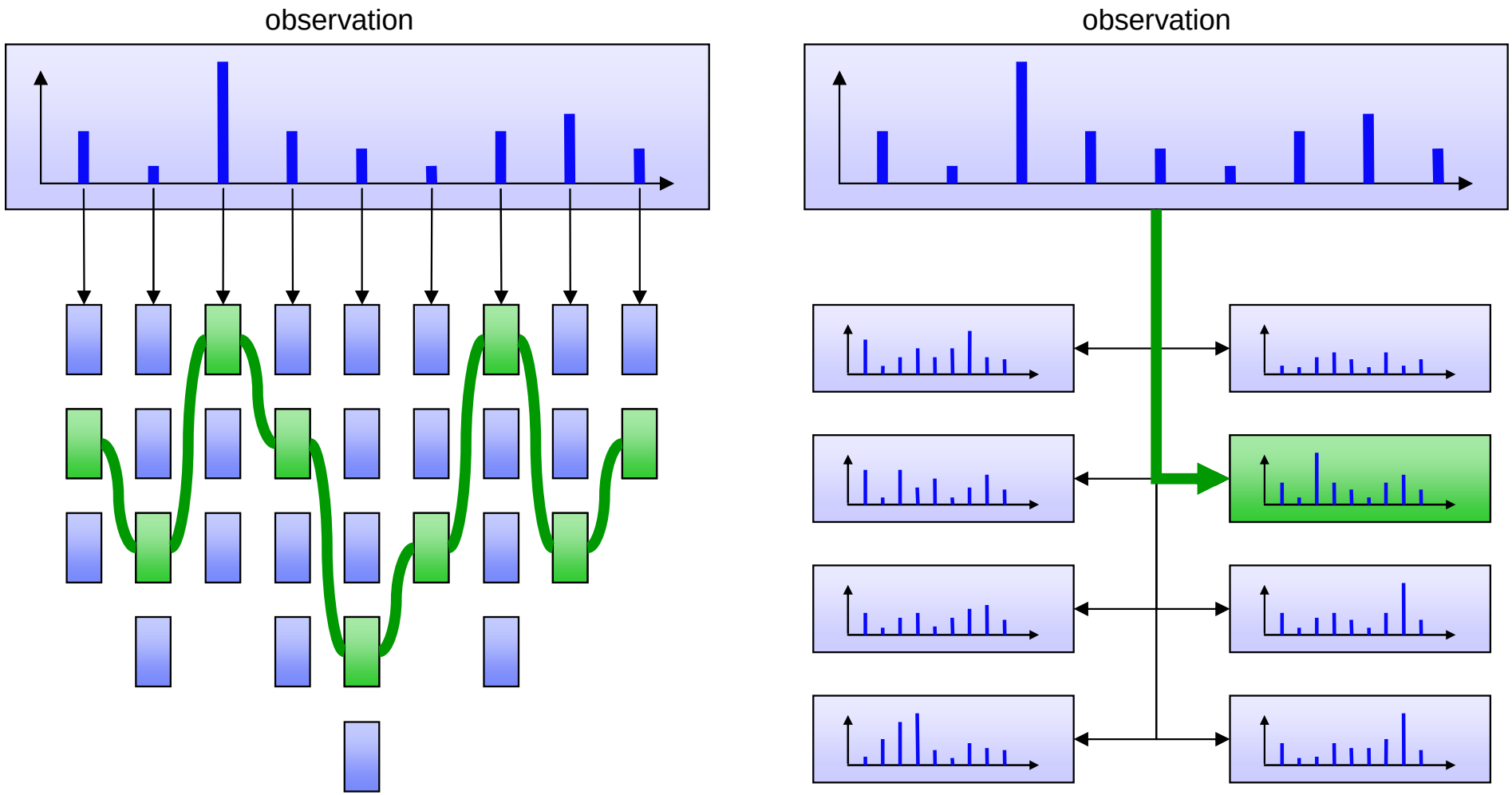
1.



2.



Comparison: Dynamic vs. static histograms



Alphabet of “symbols”

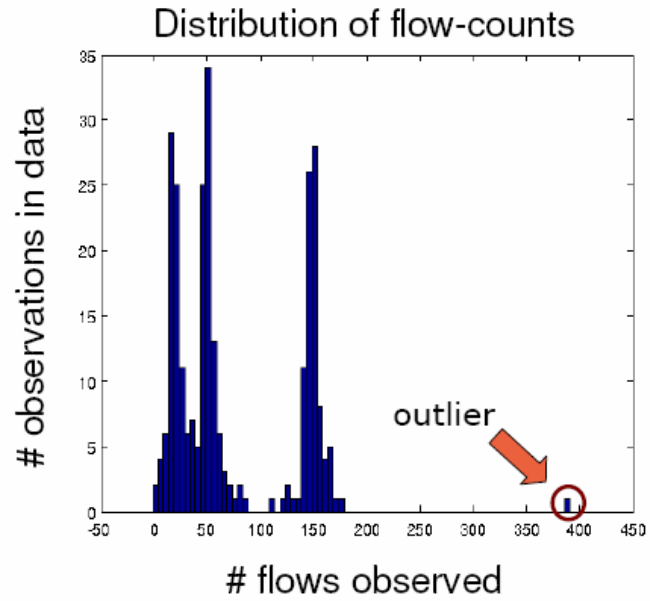
admissible symbols for each block

“Decoding problem”

Dictionary of “words”

signal constellations (known to receiver)

Discussion

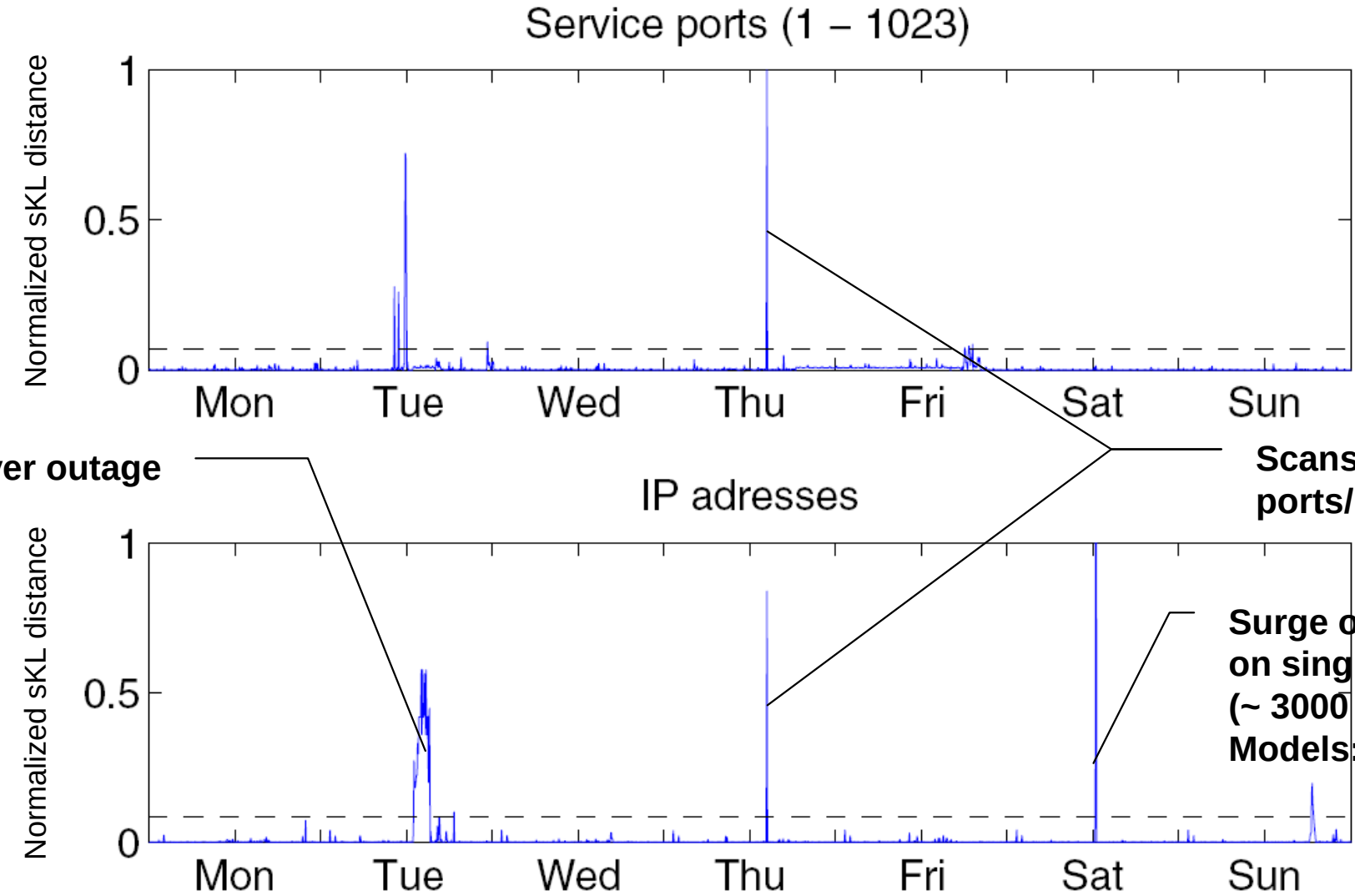


- **Learning phase (unsupervised)**
  - Individual modeling of **component-wise** behavior modes
    - e.g., server and protocol baselining
  - Derivation of meaningful thresholds (per component and feature)
  - **Robustness to anomalies**
- **Detection phase**
  - Two detection levels:
    - ***component-wise*** analysis (host failure, (D)DoS attacks)
    - ***feature-wise*** analysis, e.g., accumulation of small deviations over all service ports (worms)
- **Analysis of suspicious activities ("drill down")**
  - provides a detailed **deviation vector** for each feature and timestamp pair
  - Real life models expressive to administrators
- **Incorporation of administrator feedback in** $O(1)$

Validation

- **Production network**
  - Three weeks: internal and external traffic
  - Heterogeneous traffic patterns: user traffic (web, mail), file transfers, etc.
- **Data center**
  - Router transferring more than 6 TB/d, avg. sending/receiving rates of 550 Mb/s and 100 Mb/s
  - Avg. Flow export rate of about 5K flows/s (peaks export rates > 20K flows/s)
  - Small amount of anomalies: vulnerability scans
- **DARPA 1999 IDeval**
  - Analysis of patterns of labeled anomalies in different traffic features

Production network (internal)

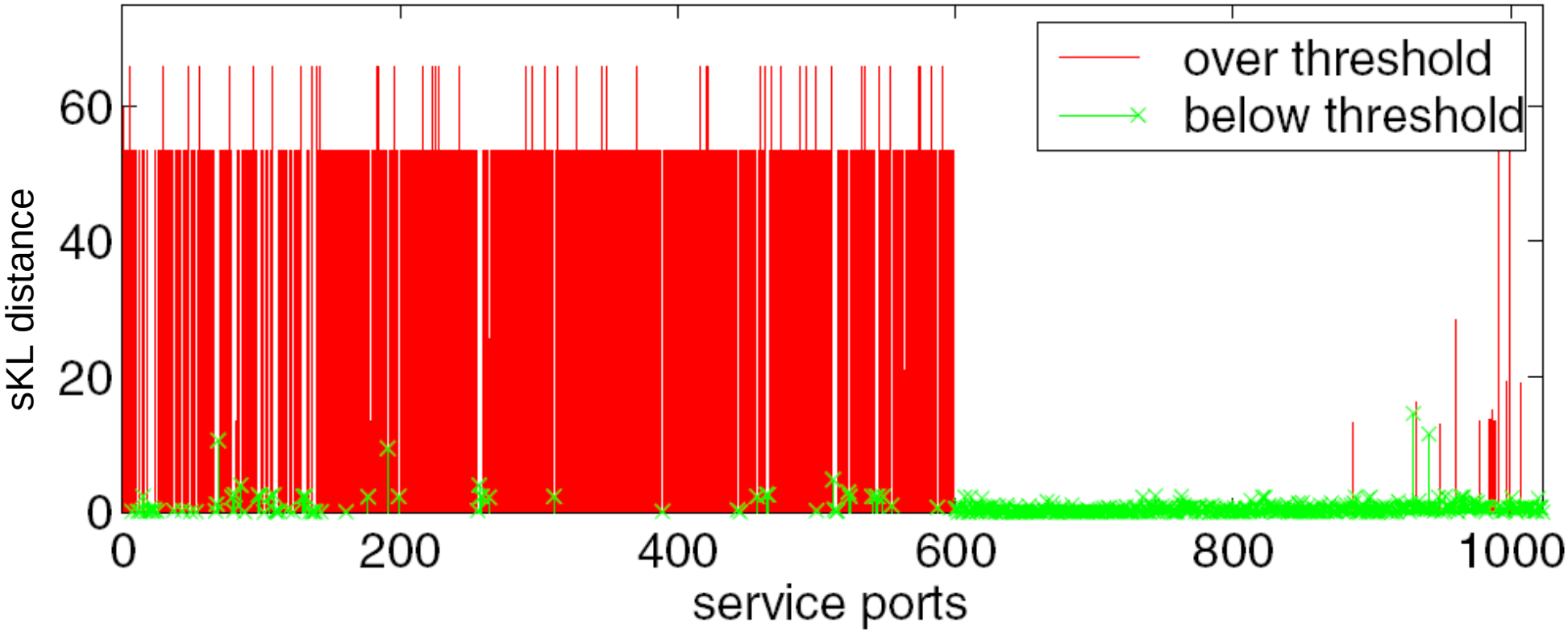


2.5h server outage

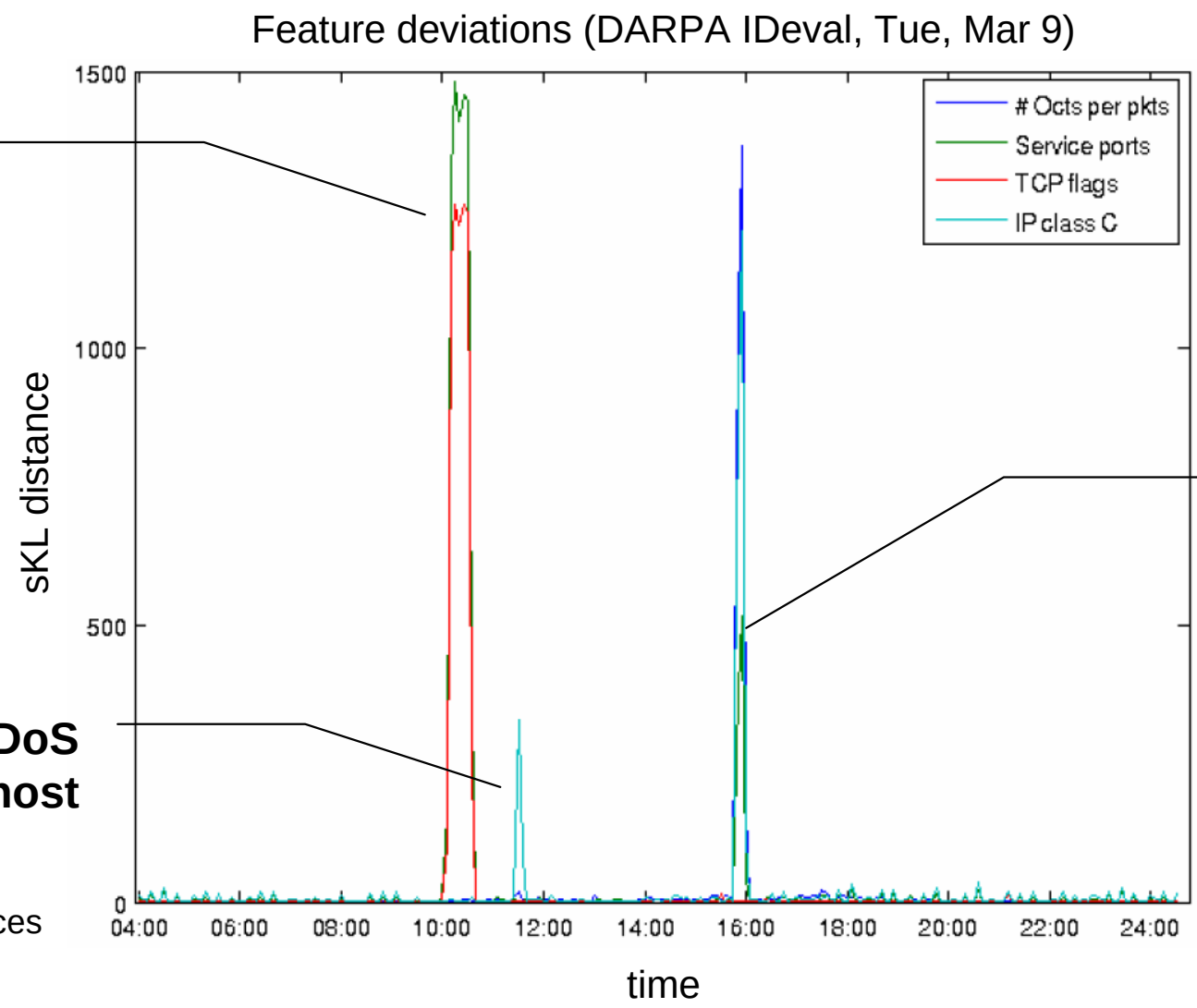
Scans on multiple ports/machines

Surge of traffic on single box (~ 3000 SSH flows, Models: at most 23)

Deviation vector in service ports (Thu, 14:20)

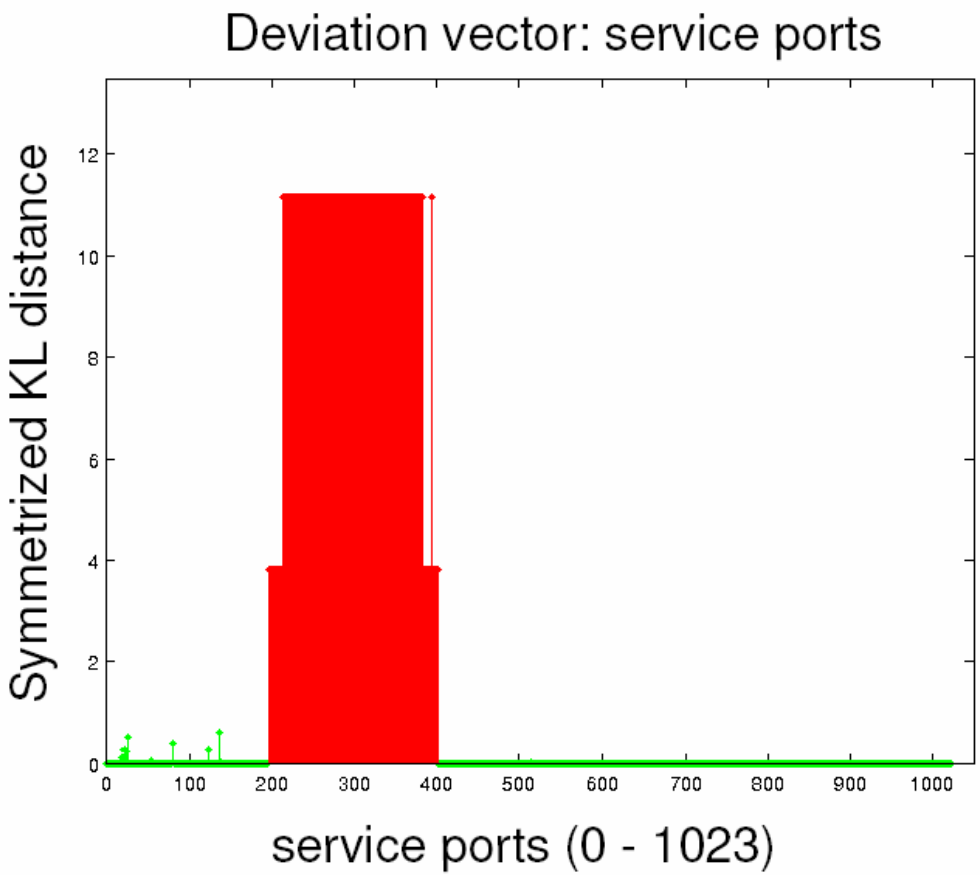


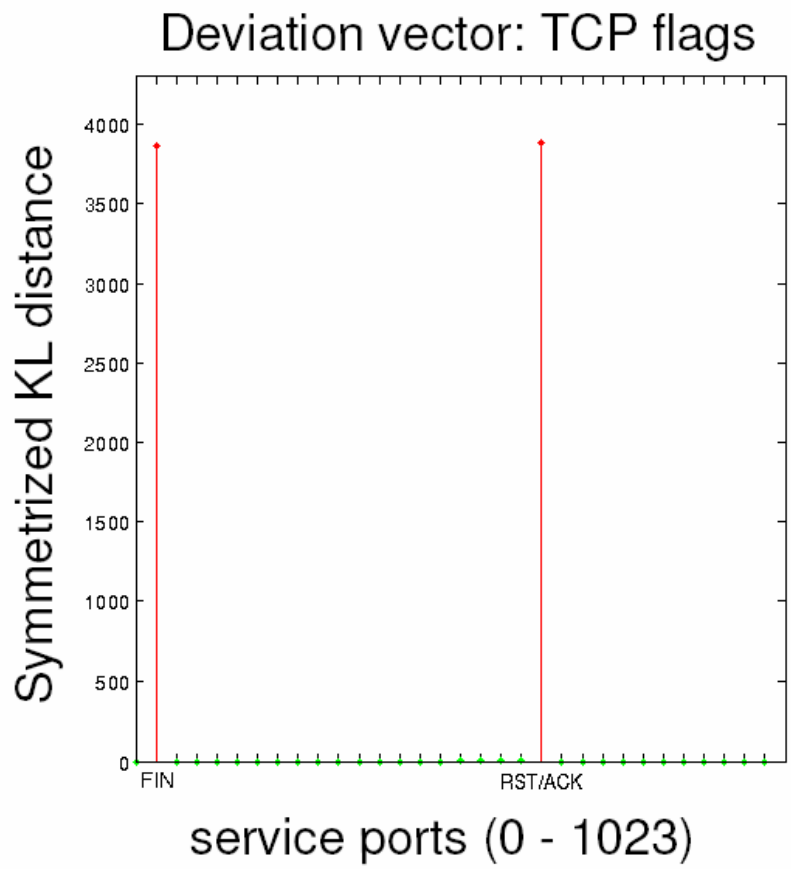
Detection of Anomalies in Multiple Aspects

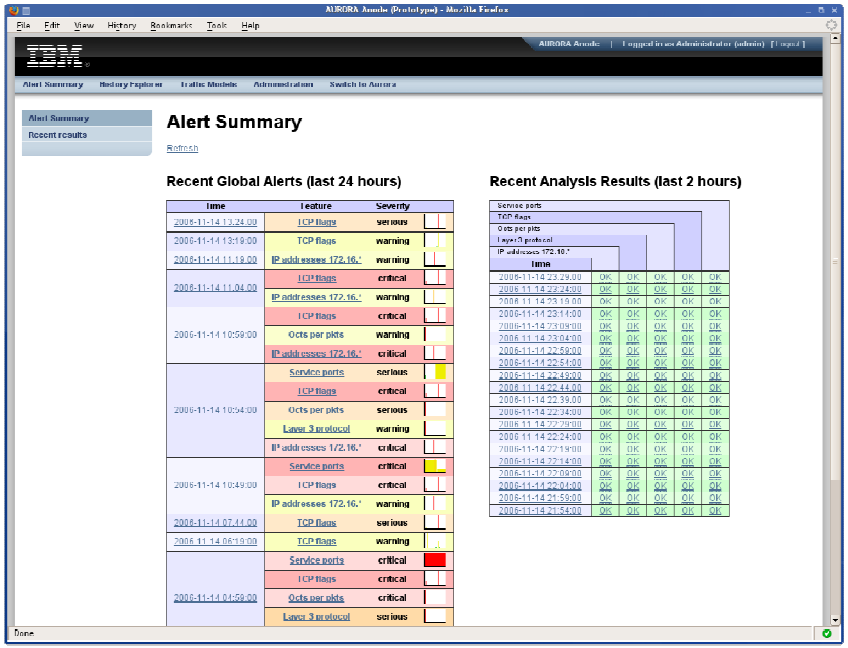


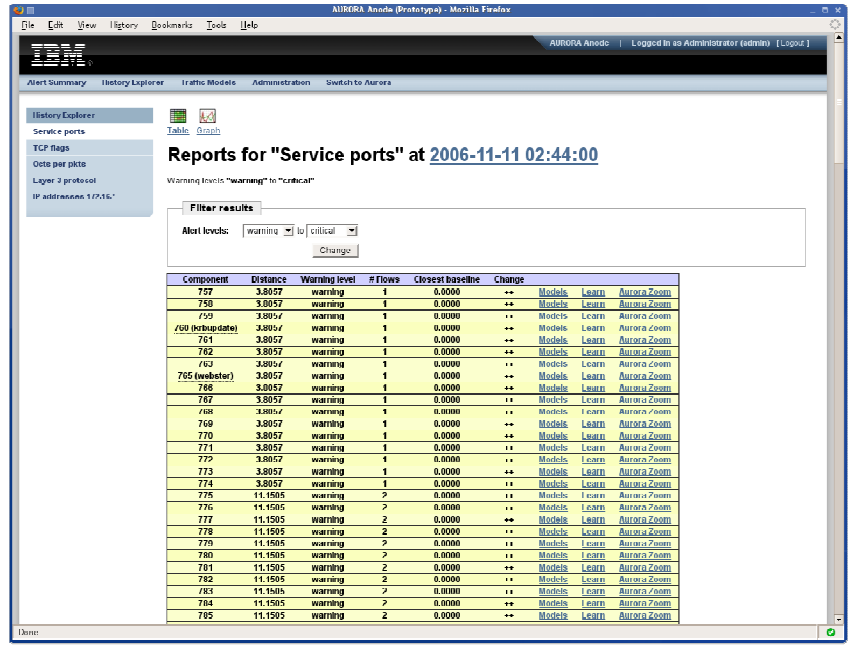
Dataset	inside traces
Learning	first week
Robustness	g = 0.001

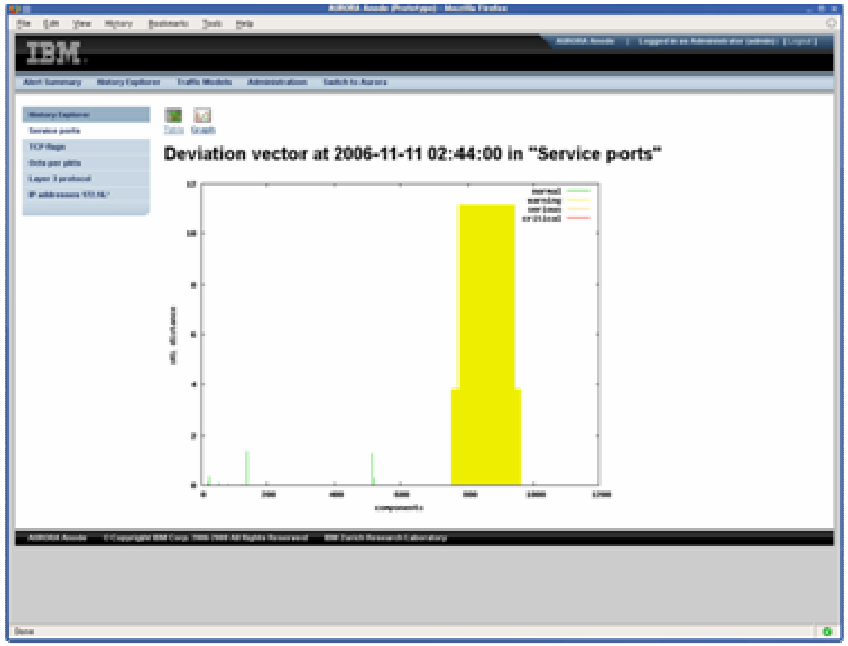
Host Scan (DARPA IDS eval)

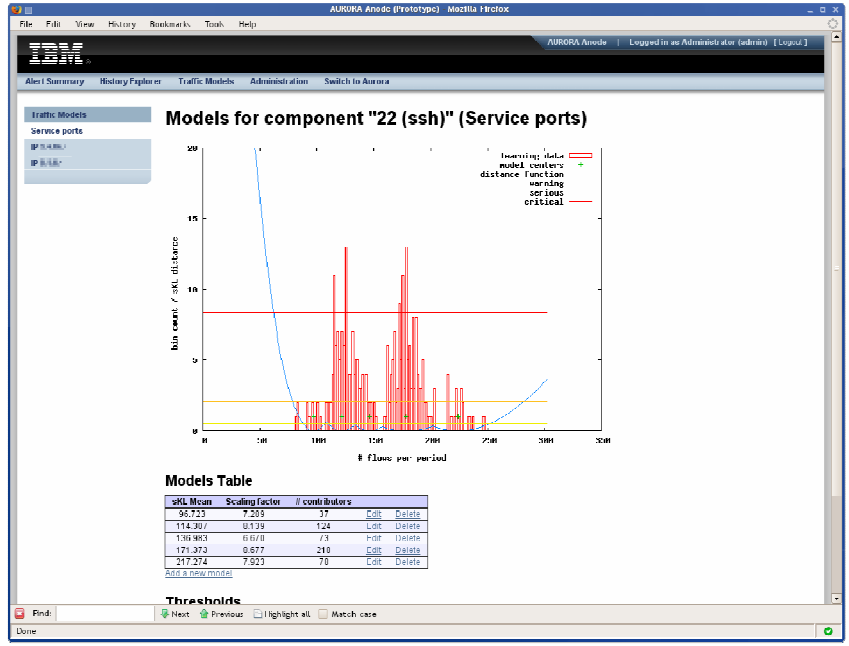










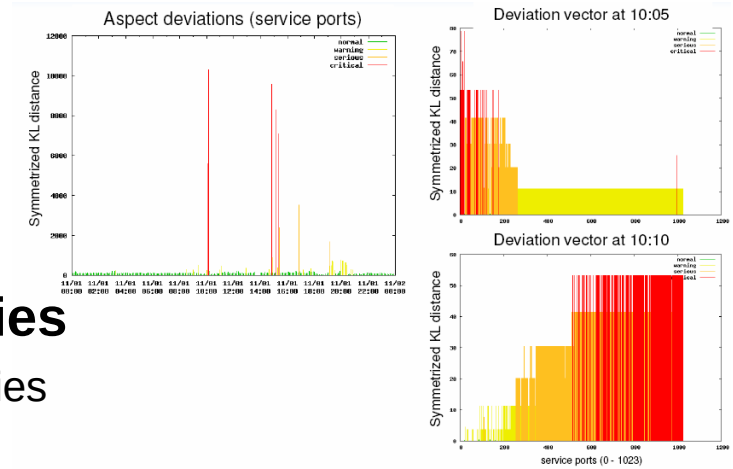


Limitations

- **Detection phase**
  - No explicit signatures of threats
  - Missing semantics of anomalies (compared to signature-based techniques)
  - No access to raw packet data

- **Learning Phase**
  - Learning from anomaly-poor training data
  - No automatic adaptation to long-term changes in network
  - No explicit time models

Conclusion



- **Flow-based approach to detect network anomalies**
  - Detection on **two abstraction levels**: ability to expose anomalies of different natures
  - Interpretable visualization and graphical reports of abnormal events
  - On-line detection, administrator feedback
- **Implementation as a stand-alone module for AURORA**
  - **Scalable to high-speed networks** thanks to flow data analysis (40K flows/s)

- **Future work**
  - **Semantics of incidents (contextualization)**
    - Provide semantic encoding to quantify events
  - **Learning mechanism**
    - Continuous adaptation of models to long-term traffic changes
  - **Model representation**
    - Notion of time in traffic models (e.g., day/night, seasonal effects)
  - **Zoom into anomalous traffic (reactive “Zoom monitors”, FloCon ‘08)**

References

- Stoecklin, M. Ph., Le Boudec, J-Y., Kind, A.: A Two-Layered Anomaly Detection Technique Based on Multi-modal Flow Behavior Models. In: Uhlig, S., Claypool, M. (Eds.): PAM 2008, LNCS 4979, pp. 212–221, 2008.

- Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: ACM Conference on Knowledge Discovery and Data Mining (KDD), pp. 226–231 (1996)

- Stoecklin, M. Ph., Kind, A., Le Boudec, J-Y.: Dynamic Adaptation of Flow Information Granularity for Incident Analysis. In: FloCon '08, Savannah, GA, January 2008.

- IBM Research. Aurora – Network Traffic Analysis and Visualization. http://www.zurich.ibm.com/aurora/