VizSEC 2007
Workshop on Visualization for Computer Security

Graph-based Monitoring of Host Behavior for Network Security

Florian Mansmann, Lorenz Meier, and Daniel A. Keim

Universität Konstanz

Overview

1. Motivation
2. Related work
3. Behavior Graph
4. Automatic accentuation
5. HNMap integration
6. Case Study
7. Evaluation
8. Conclusions

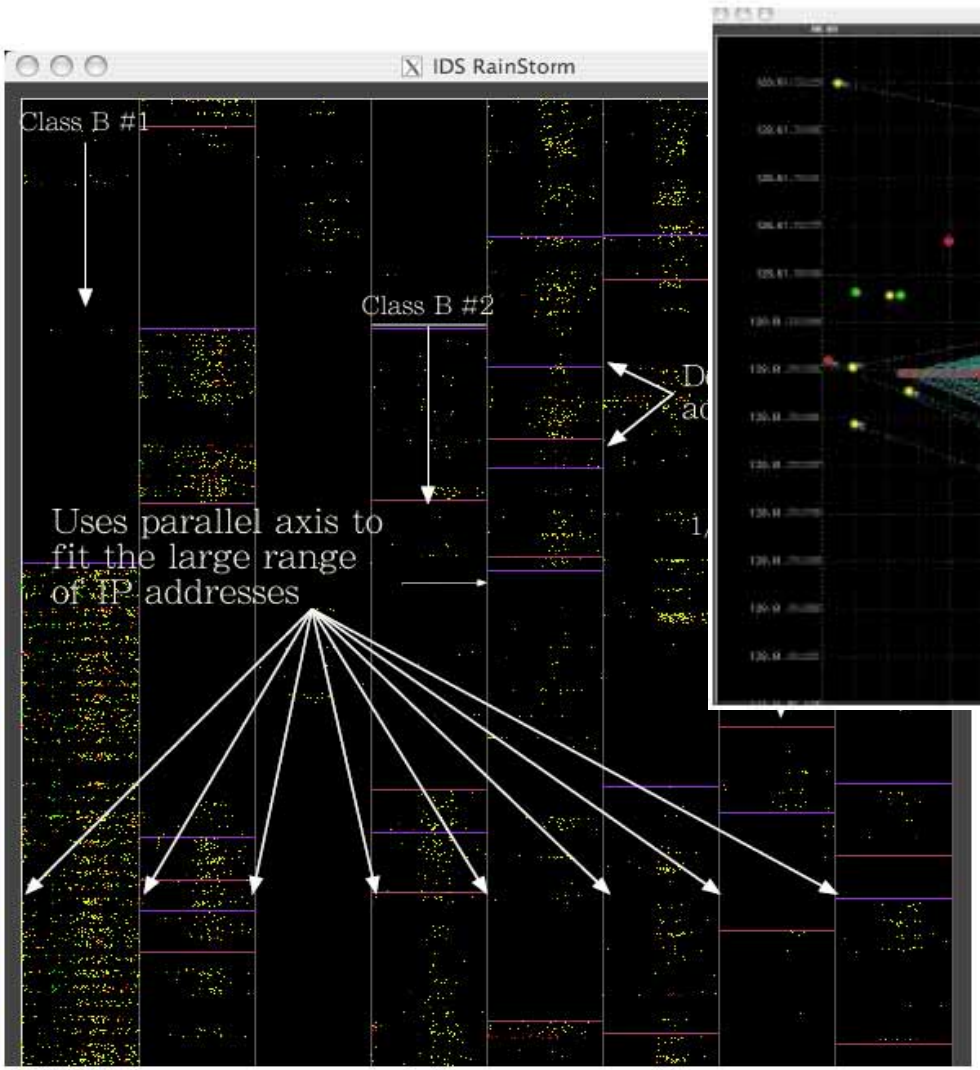
1. Motivation

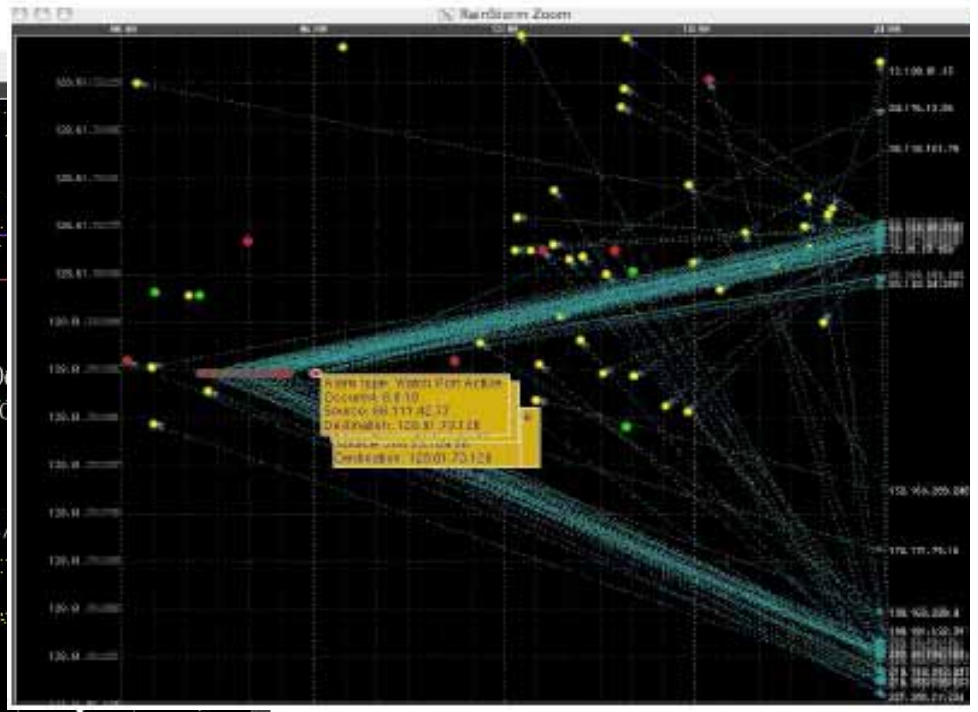
- **Visual analytics research**
  links visualization with automatic analysis methods to combine flexibility, creativity, and background knowledge with the enormous storage capacity and computational power of today's computers.

- **Our goal**:
  to use visual analytics methods to gain more insight into network traffic and intrusion detection data sets.

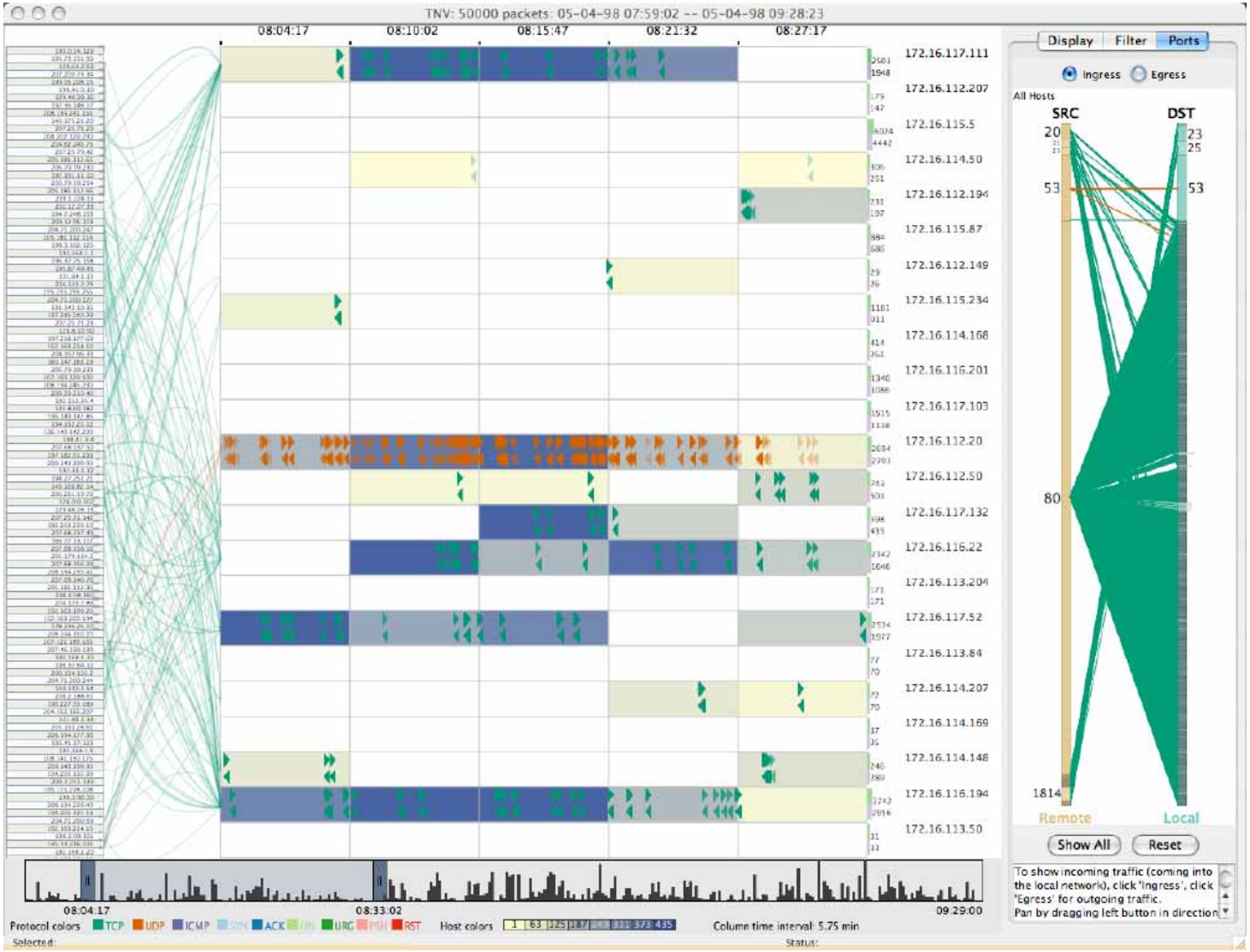
- Focus of this work:
  tracking behavioral changes in network traffic of a collection of **hosts**.

2. Related work: IDS Rainstorm

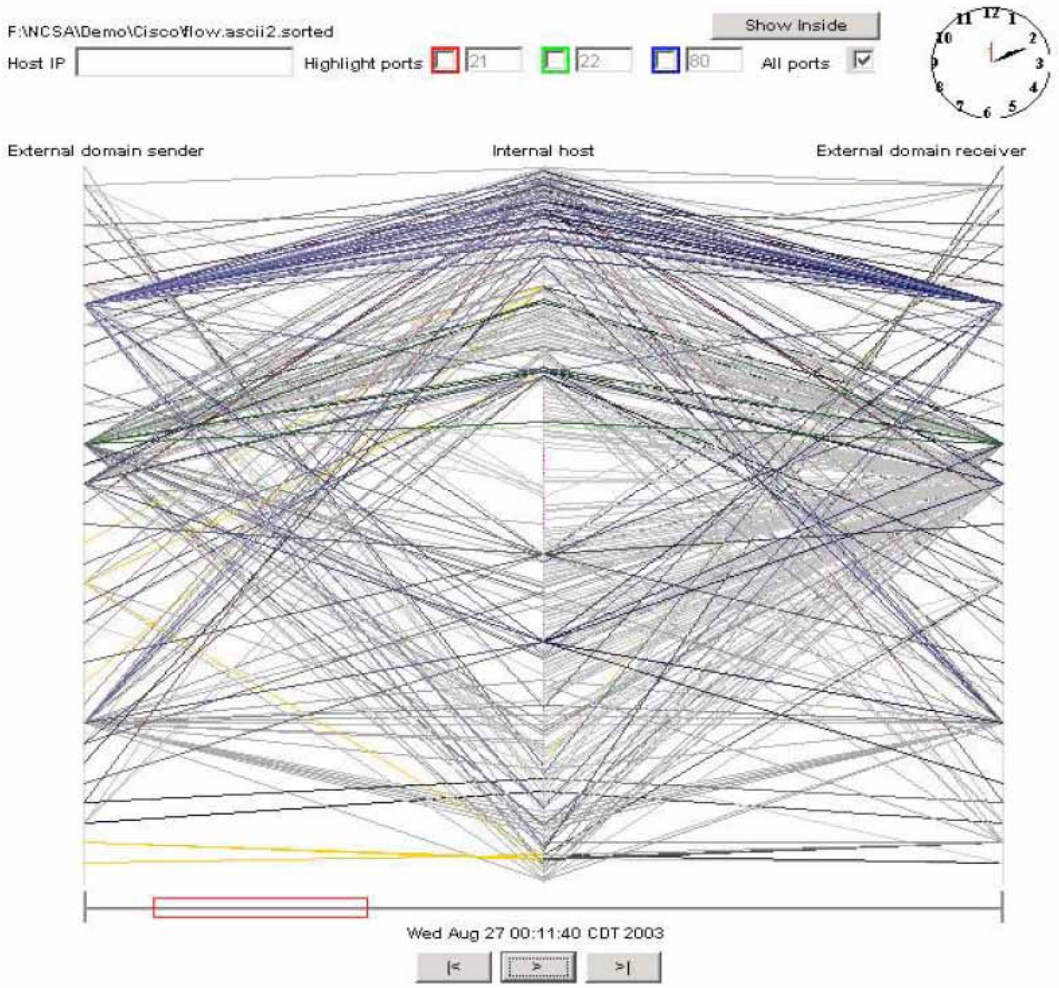




2. Related work: TNV



2. Related work: VisFlowConnect



3. Behavior graph: application design

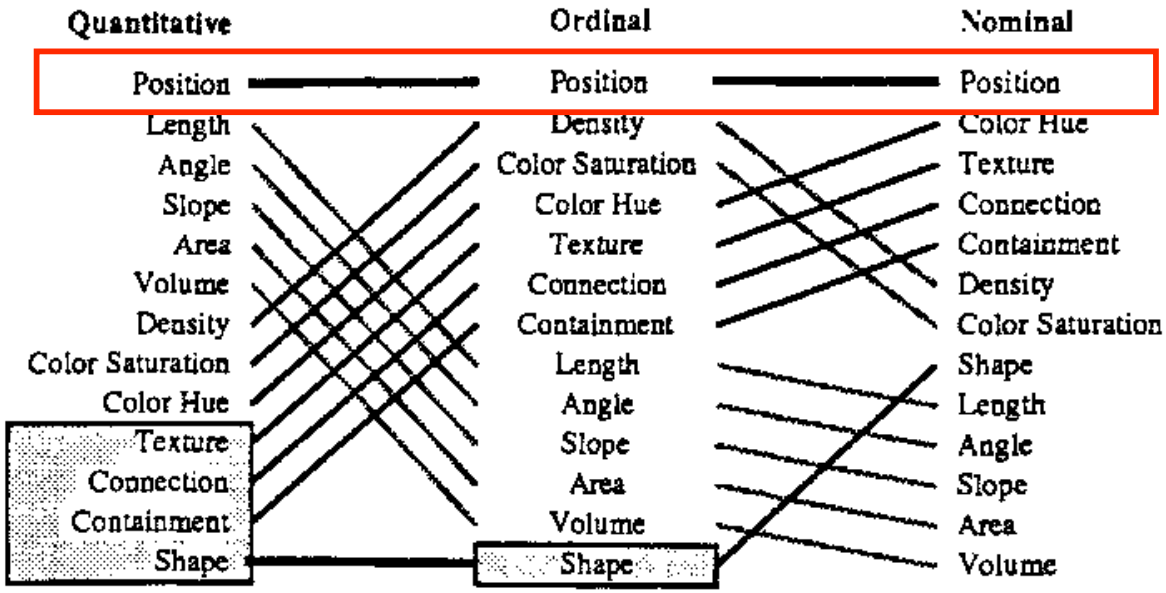
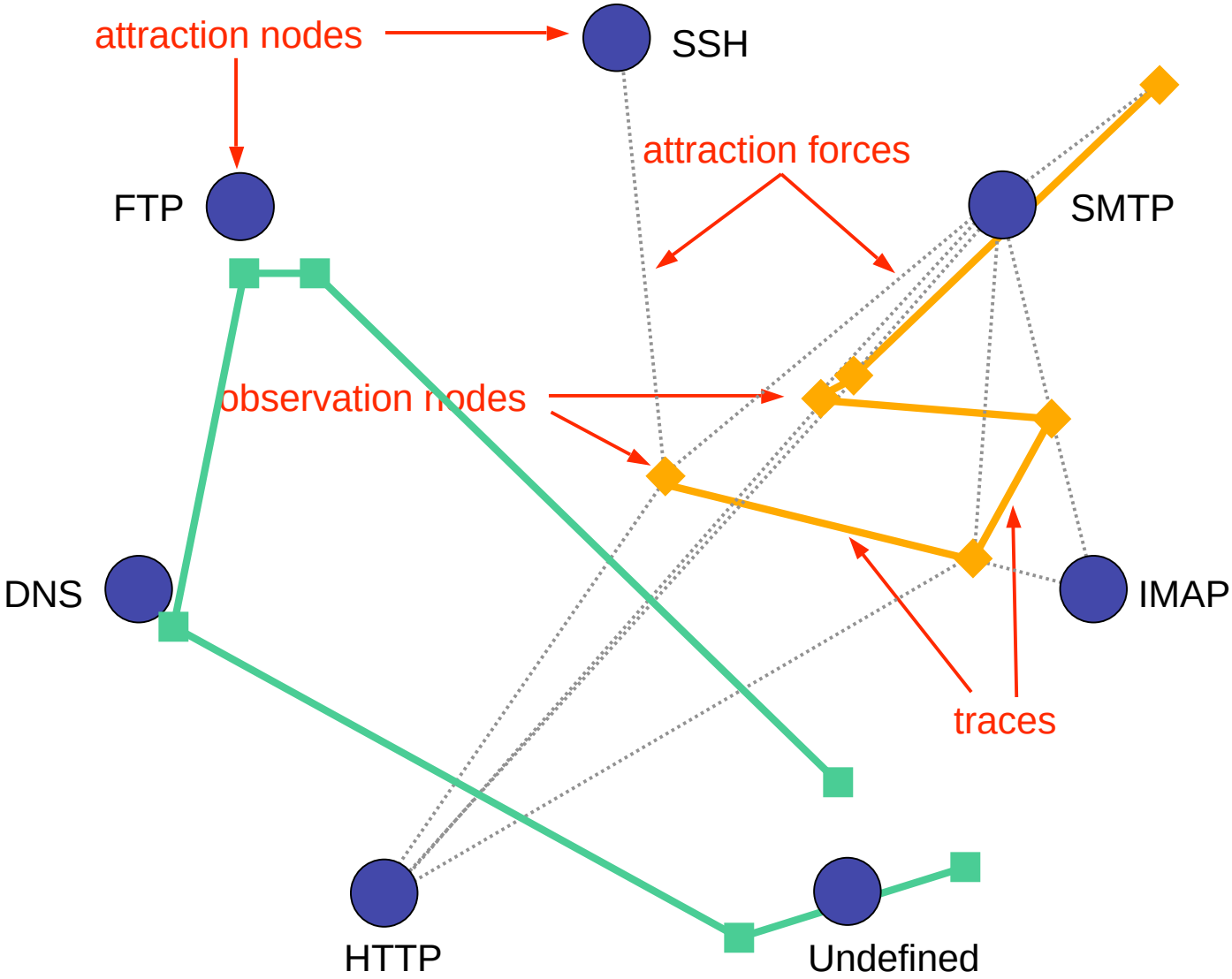


Figure 15: Ranking of Perceptual Tasks. *The tasks shown in the gray boxes are not relevant to that type of data.*

Graph-based host behavior monitoring

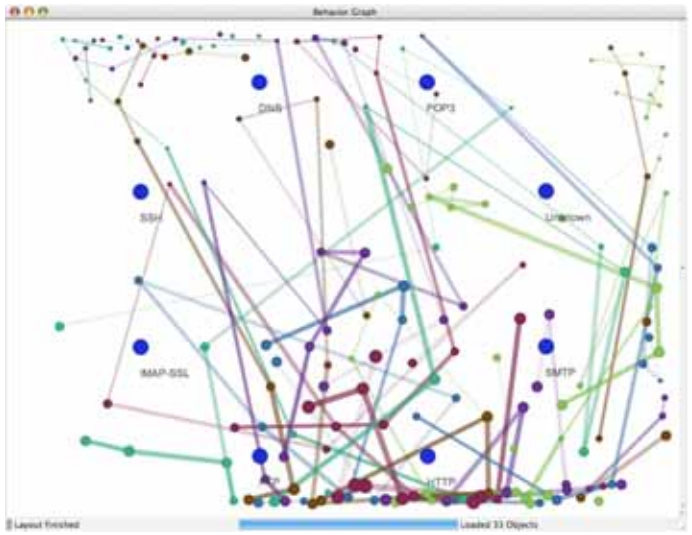


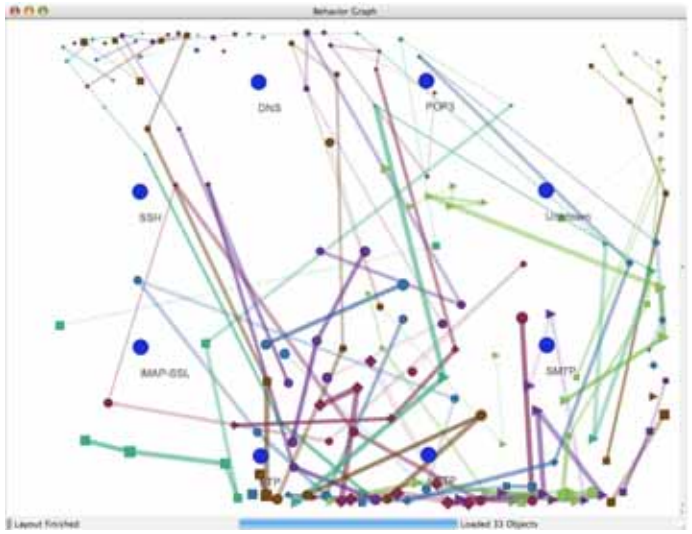
Implementation

- PostgreSQL database to **calculate forces** for each observation node and time span.

- Fruchterman-Reingold **spring embedder** as provided in the JUNG graph drawing library to calculate graph embedding.
  - Efficiency
  - Robustness of force and iteration parameters

Node representations





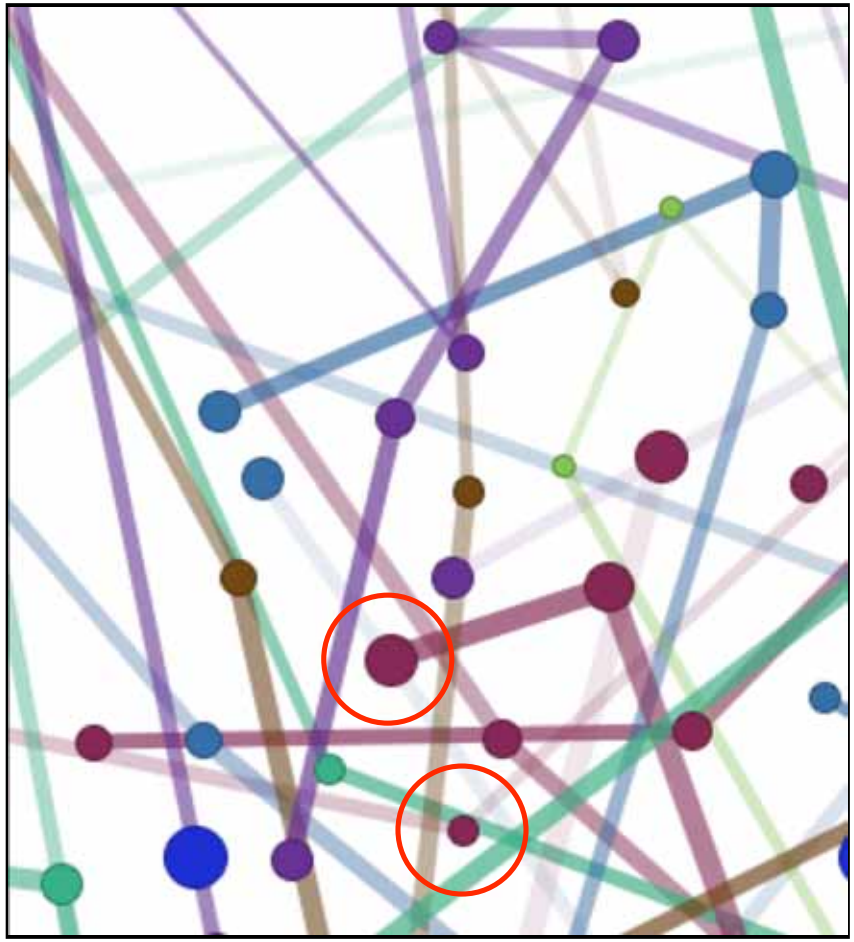
Problem:

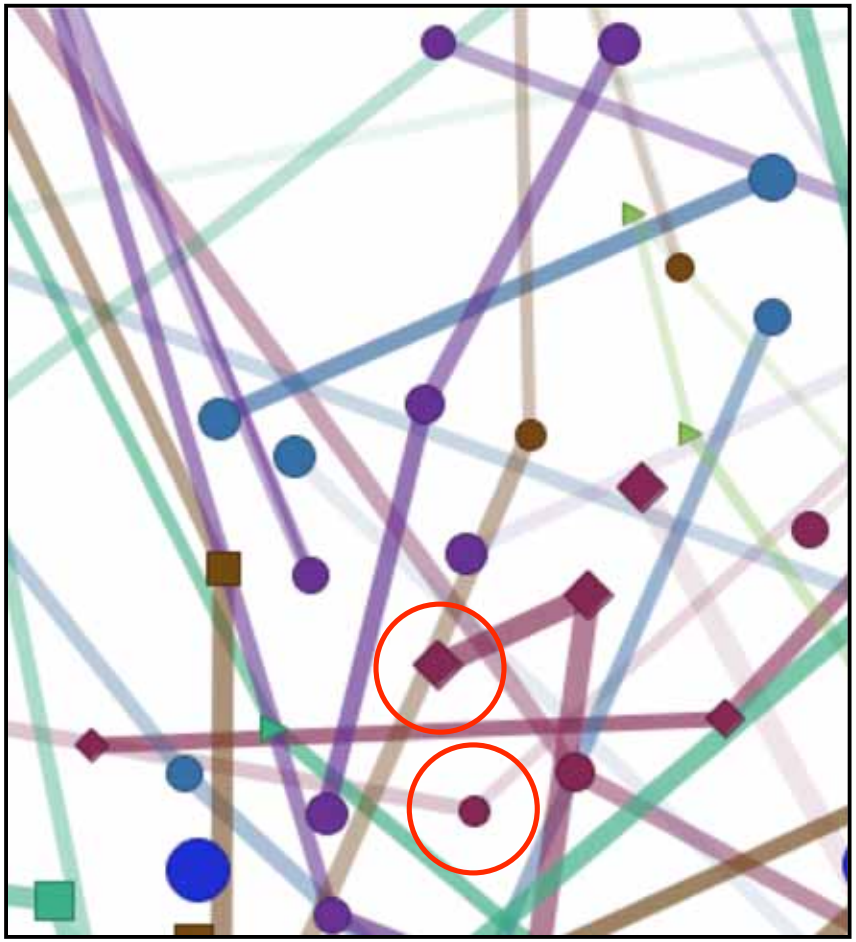
- **Coloring**: only a finit number of colors are distinguishable
- How much traffic?

Solution:

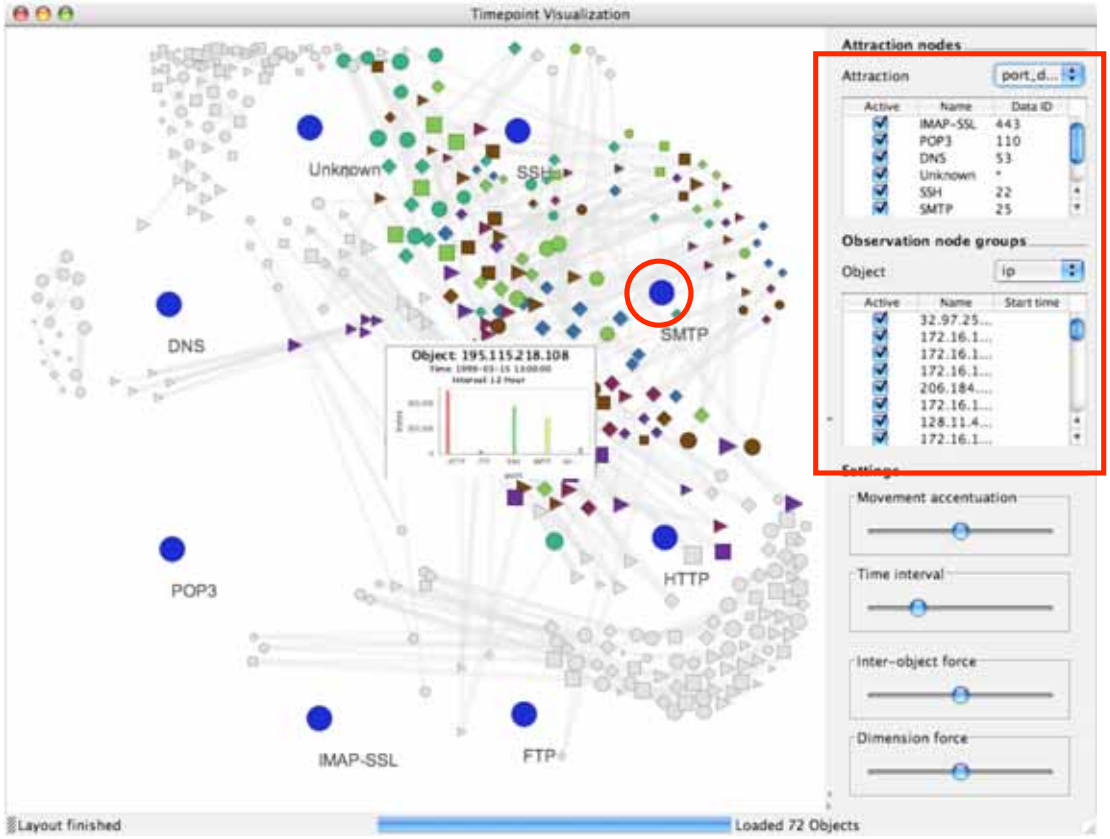
- Combination of **color and shape** (gray-scale usage possible)
- Mapping traffic volume to size (scaling)

Node representations





User interactions

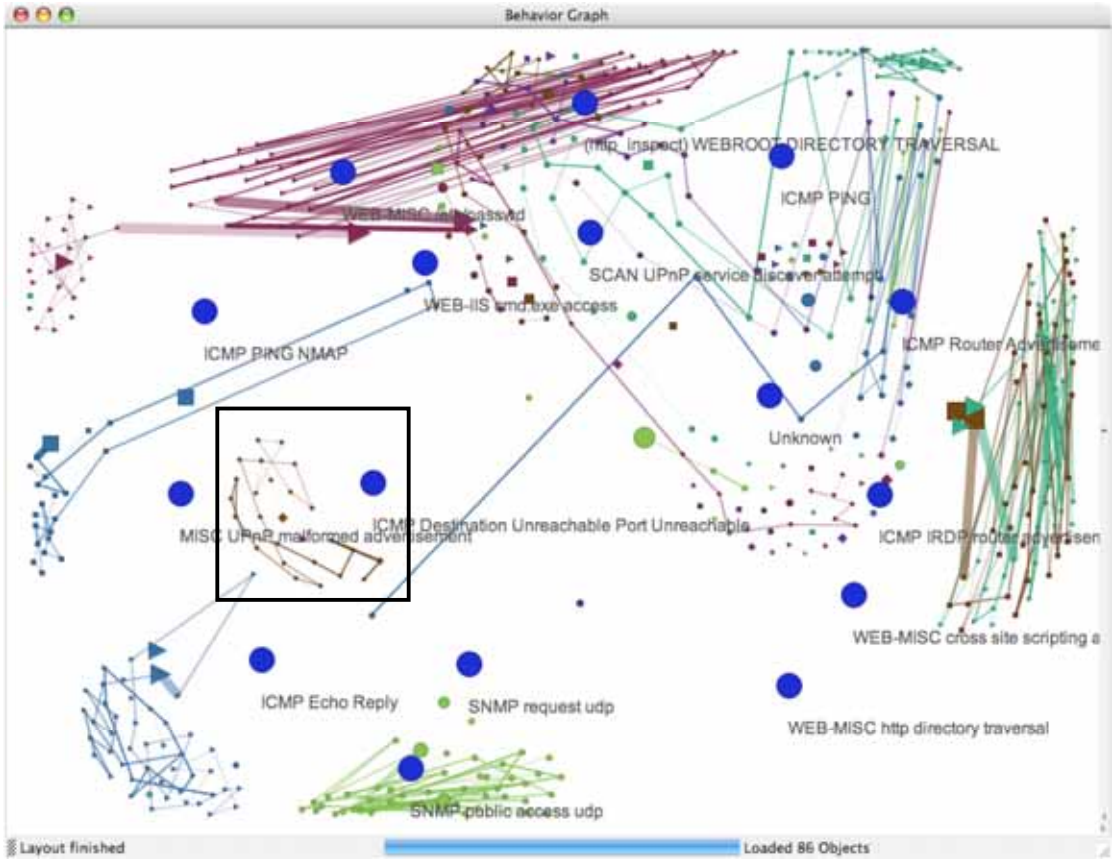


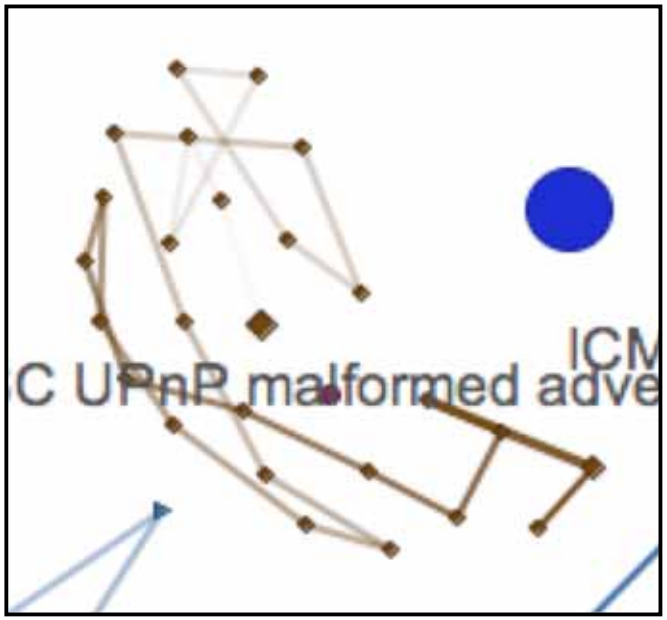
- **Details** (bar chart) on mouse-over
- **Select** node observation group or attraction nodes
- **Add and remove** attraction and observation **nodes**
- **Repositioning** of attraction nodes (drag & drop)

User interactions

- **Movement accentuation** highlights suspicious hosts with highly variant traffic (more later …)
- **Time interval**: increase/decrease the amount of observations per host
- **Inter-object force**: control attraction forces within a node observation group (one group per host)
- **Dimension forces**: fine-tune attraction forces between observation and attraction nodes.

Time visualization





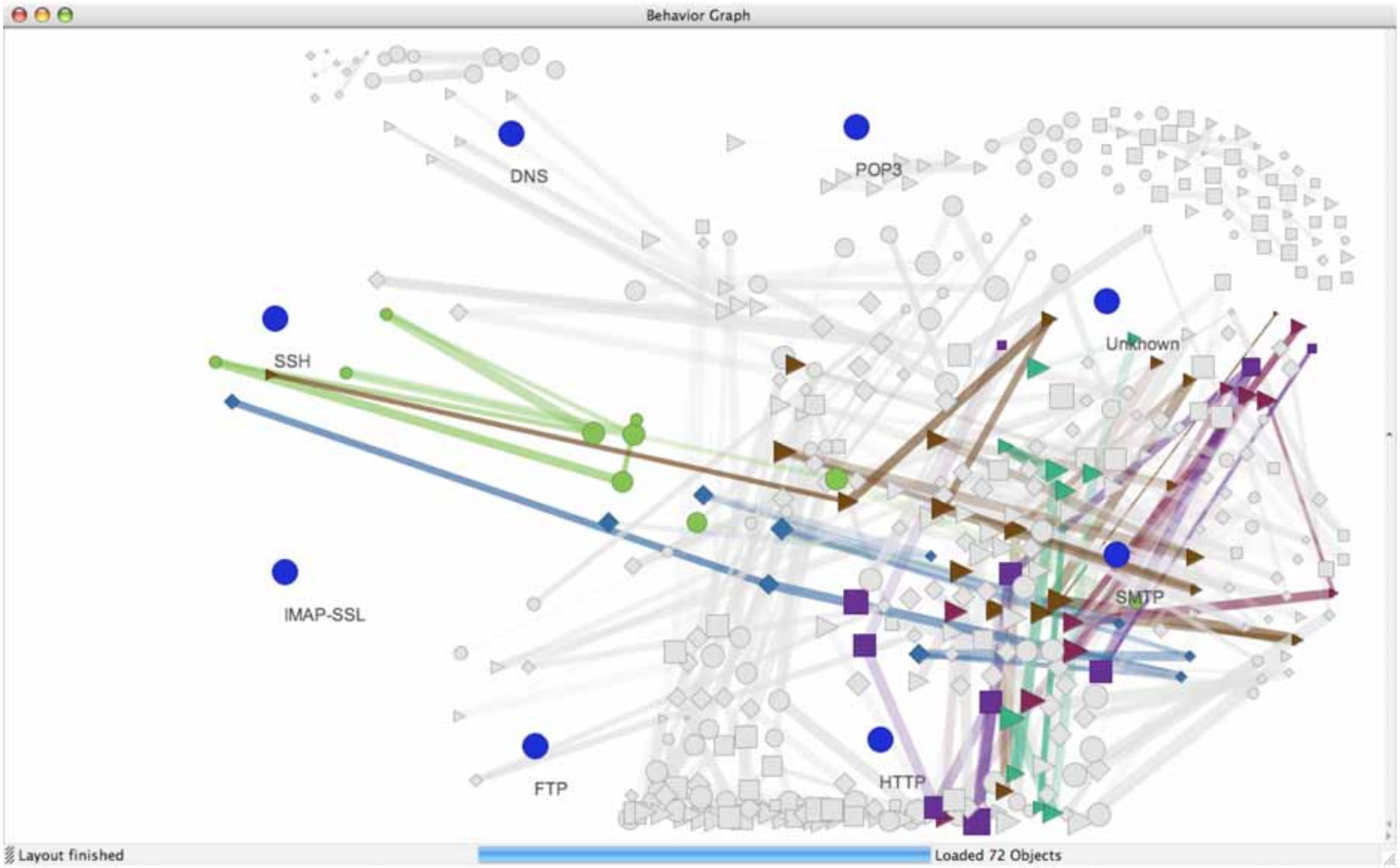
4. Automatic accentuation of node groups with highly variable traffic

Position of a host at a particular timespan in high-dimensional Euclidean space

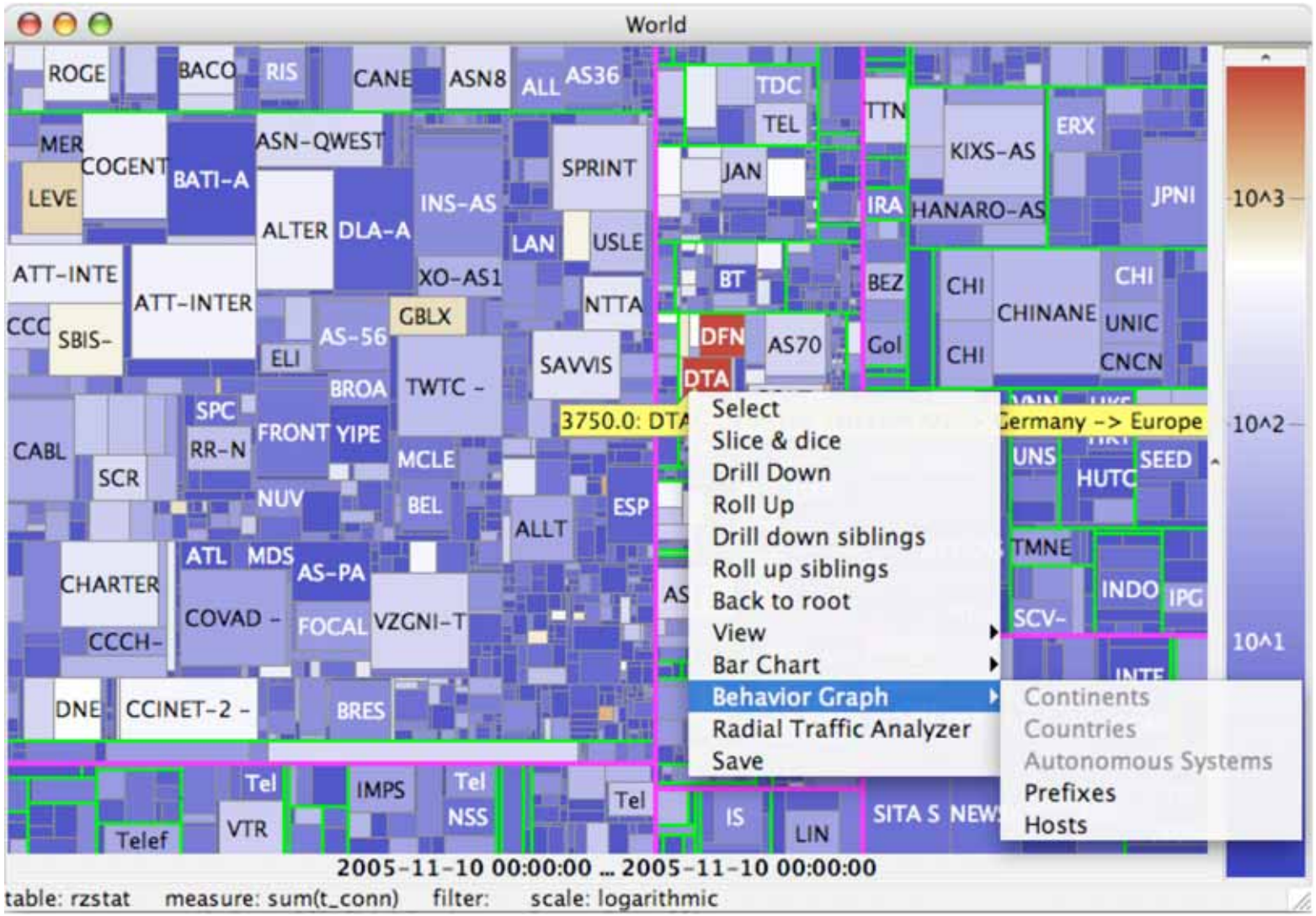
$$\vec{o}^{\,r} = \frac{\vec{o}}{|\vec{o}|}$$

$$pc_{norm} = \frac{\sum_{t=1}^{t_{max}-1} |\vec{o}^{\,r}_t - \vec{o}^{\,r}_{t+1}|}{t_{max}}, \qquad 0 \leq pc_{norm} \leq 2$$

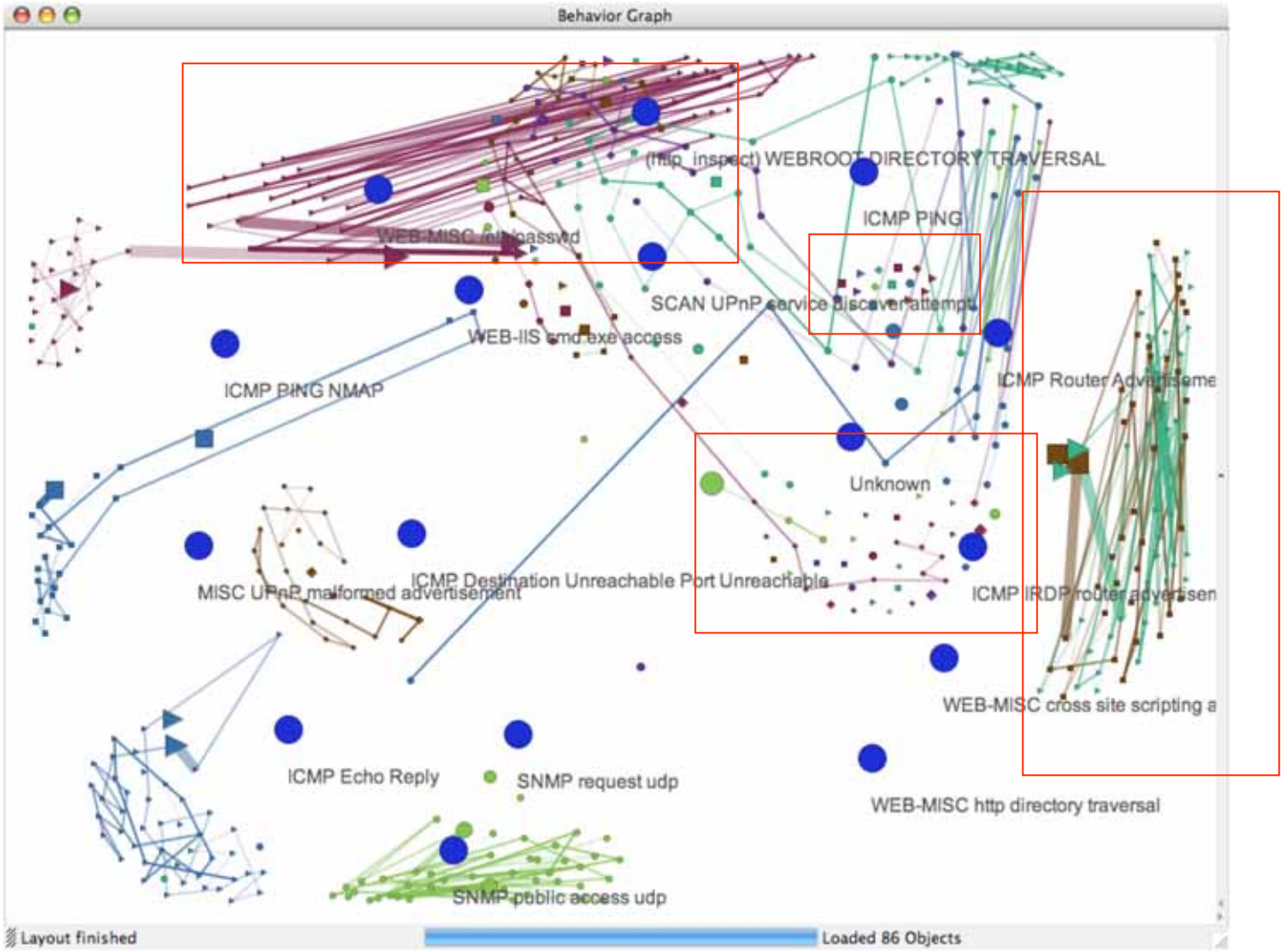
4. Automatic accentuation of node groups with highly variable traffic



5. HNMap integration: Monitor network traffic behavior

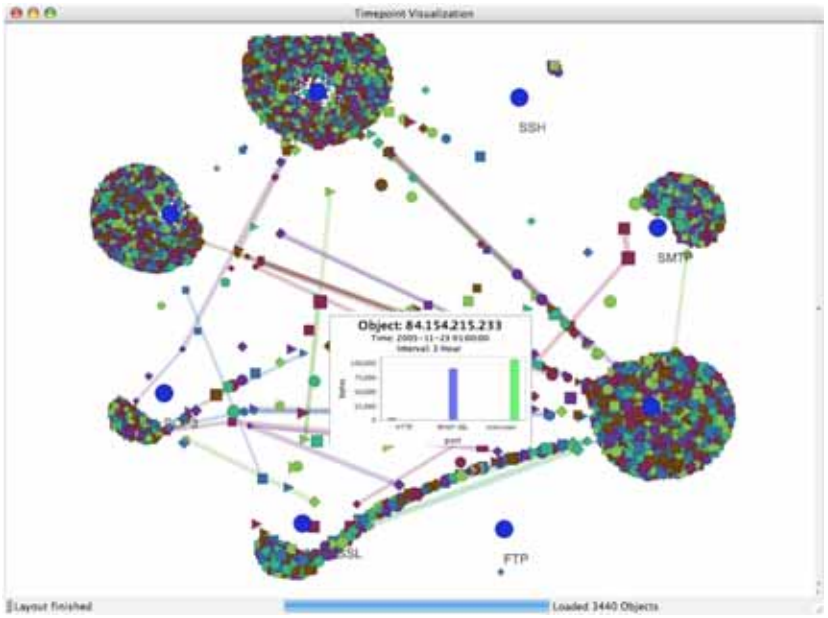


6. Case study: 19 000 SNORT alerts

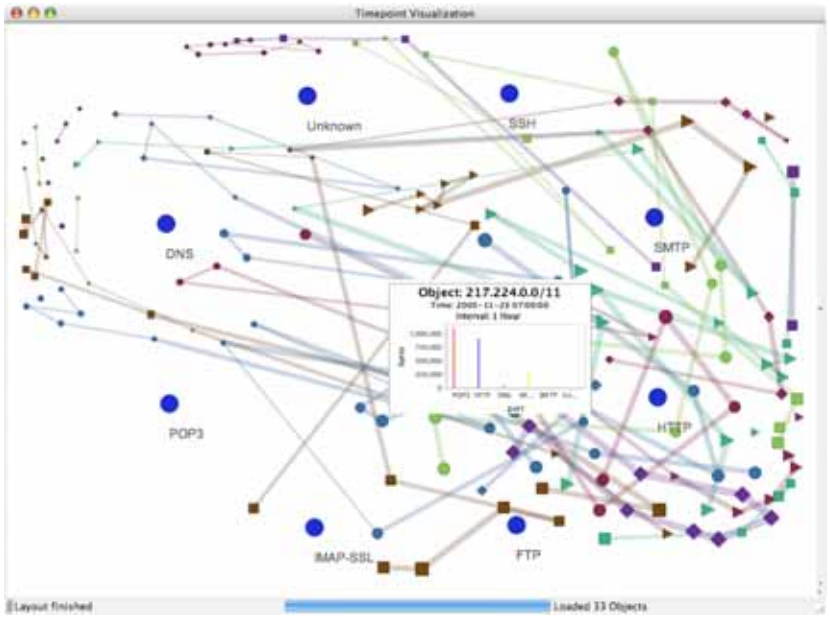


7. Evaluation: Scalability

- Max: 1000 observation nodes
- hosts * intervals



Hosts of AS 3320



Prefixes of AS 3320

7. Evaluation

Drawbacks

- Scalability limitations (1000 observation nodes)
- Only rough approximation of traffic proportions (ambiguity).

Advantages

- Changes in the traffic of many different nodes is made visible.
- Interactive exploration of data sets
- Automatic accentuation of high-variance hosts
- Usefulness of the approach is demonstrated

8. Conclusions

1) We presented a **novel graph-based network traffic visualization** to monitor host behavior using a force-base graph layout

2) HNMap integration: apply graph on **hosts, prefixes, ASes, countries, and continents**

3) Usefulness demonstrated on network traffic (university gateway router, IDS events)

4) Visual analytics feature: **automatic accentuation** of suspicious node groups with highly variable traffic

Future work: spacio temporal data analysis, text visualization

Acknowledgment

- Lorenz Meier for his excellent work
- Data sources:
  - MIT Lincoln Lab (DARPA 99)
  - Fabian Fischer (Snort alerts)
  - Rechenzentrum
- BWFit: information at your fingertips – interactive visualization for Gigapixel displays
- DFG GK 1042: „Explorative Analysis and Visualization of Large Information Spaces“

Questions?

Acknowledgment

- Lorenz Meier for his excellent work
- Data sources:
  - MIT Lincoln Lab (DARPA 99)
  - Fabian Fischer (Snort alerts)
  - Rechenzentrum
- BWFit: information at your fingertips – interactive visualization for Gigapixel displays
- DFG GK 1042: „Explorative Analysis and Visualization of Large Information Spaces“

References

- Kulsoom Abdullah, Chris Lee, Gregory Conti, John A. Copeland, and John Stasko. Ids rainstorm: Visualizing ids alerts. In Proc. IEEE Workshop on Visualization for Computer Security (VizSEC), Minneapolis, U.S.A., October 2005.
- John R. Goodall, Wayne G. Lutters, Penny Rheingans, and Anita Komlodi. Focusing on context in network traffic analysis. IEEE Computer Graphics and Applications, 26(2):72–80, 2006.
- Xiaoxin Yin, William Yurcik, Michael Treaster, Yifan Li, and Kiran Lakkaraju. Visflowconnect: netflow visualizations of link relationships for security situational awareness. In VizSEC/DMSEC, pages 26–34, 2004.
- MacKinlay, J. (1986). Automating the design of graphical presentations of relational information. ACM Transactions on Graphics, 5, 110-141.
- Thomas M. J. Fruchterman and Edward M. Reingold. Graph drawing by force-directed placement. Software - Practice and Experience, 21(11):1129–1164, 1991.