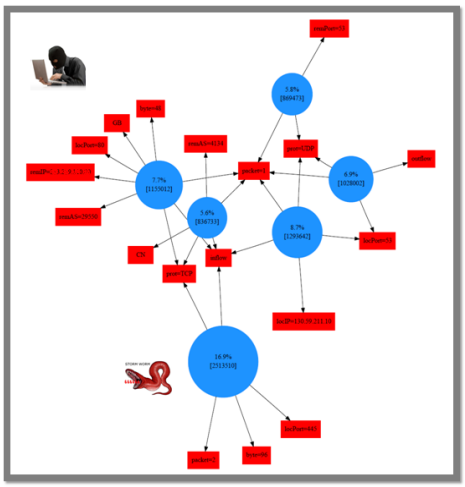
Visual Network Traffic Exploration

Visualization of one hour of traffic
Involving 14,916,092 failed connections



Computer networks are kept working by "masters of complexity", who need to know in detail the infrastructure and configuration of a network, extract useful information from diverse and massive monitoring data, and iteratively test and verify their assumptions about the underlying causes of problems, while they are steered solely by their intuition. Effective visualization techniques for network monitoring data can greatly facilitate slow manual data exploration processes by enabling to absorb large amounts of data quickly as a good picture is worth a thousand words. A key challenge in visualizing network monitoring data, like traffic traces, is their massive volume, which can easily reach the order of petabytes when accumulating traffic over time.

In order to effectively explore large volumes of network traffic data, effective visualization techniques that reduce the cognitive burden of the analyst are essential. This thesis will contribute to a project towards building a visual network traffic exploration tool that can be used to create short movies that show how network traffic changes over time. Recent work by the Communication Systems Group (CSG) of ETH introduced data analytics and visualization techniques for summarizing and visualizing big network traffic data [1]. The introduced techniques exploit frequent pattern mining, which is a well-known data mining approach that the student will have the opportunity to study during this project. Network traffic is visualized as a special type of graph, called hypergraph. The focus of this thesis is to extend the techniques introduced in [1] and in particular to evaluate dynamic graph visualization techniques for showing how network traffic changes over time.

Tasks:
1. Study related work on visualization and literature on frequent item-set mining.
2. **Data analytics**: Evaluate techniques for identifying which patterns are most interesting so that the visualization can focus on these patterns. This may also include automatically assigning labels that characterize different patterns.
3. **Dynamic graph visualization**: Evaluate which dynamic graph visualization techniques and frameworks, e.g., e.g., http://arborjs.org, are most suitable for visualizing how network traffic changes over time using hypergraphs.
4. Contribute to the development of a visual network traffic exploration tool.

We would be happy to provide you more information about the methodology of the project and answer any questions. Please feel free to contact us.

Contacts:
Prof. Eduard Glatz: eglatz@tik.ee.ethz.ch, ETZ H86
Dr. Xenofontas Dimitropoulos: fontas@tik.ee.ethz.ch, ETZ G90
Dr. Bernhard Ager: bager@tik.ee.ethz.ch, ETZ G90

Requirements: Java/Javascript, networking basics. This thesis offers practical and theoretical tasks including the development of analysis software.

Related Work:
[1] E. Glatz, S. Mavromatidis, B. Ager and X. Dimitropoulos "Visualizing Big Network Traffic Data using Frequent Pattern Mining and Hypergraphs" Workshop on Internet Visualization (WIV), Nov. 2012.