Interactive Exploration of the Network Behavior of Personal Machines (Extended Abstract)*

Sascha Simon[1], Mike Sips[1], and John Gerth[2]

[1] Max Planck Intitut Informatik
msips,ssimon@mpi-inf.mpg.de
[2] Stanford University
gerth@graphics.stanford.edu

Abstract. Personal machines are often the weakest points within a large network. Although they run an ever-increasing number of network services, these machines are often controlled by users who are unaware of security threats. Thus, a well-informed attacker can, with modest effort, identify and gain control over personal machines. However, system administrators need to know the tools and techniques used for attacks while simultaneously needing to invest huge analytical efforts to detect malicious behavior in the vast volumes of network traffic. In our research project we investigate the idea that an understanding of the regular behavior of personal machines can improve the chance of detecting the point in time when a machine shows malicious behavior. We propose a visual exploration system based on a data abstraction layer and temporal visual representations of the network traffic. The data abstraction layer enables an interactive change in the level of detail of the network traffic while temporal visualizations help system administrators to detect unexpected network traffic. In the next phase of this project, we will conduct experiments to get a good feel about the limits of our system in detecting malicious behavior in real-world scenarios.

Personal machines are often the weakest links within a large network as they run an increasing number of network services while their owners are often unaware of security threats. Well-informed attackers can exploit these factors to gain control over personal machines.

System administrators currently have two kinds of tools at hand to analyze network traffic: (a) general purpose network visualization systems showing a short snapshot of the raw traffic data (e.g. [1]) and (b) statistical models of malicious behavior (e.g. [2]).

Although general purpose network visualization systems allow one to look at the network traffic from many different perspectives, these systems require a detailed knowledge of the tools and techniques used in attacks in order to identify suspicious hosts. Automated data mining tools based on statistical models

* This work has been supported by the Max-Planck-Center for Visual Computing and Communication

Dagstuhl Seminar Proceedings 09211
Visualization and Monitoring of Network Traffic
http://drops.dagstuhl.de/opus/volltexte/2009/2156

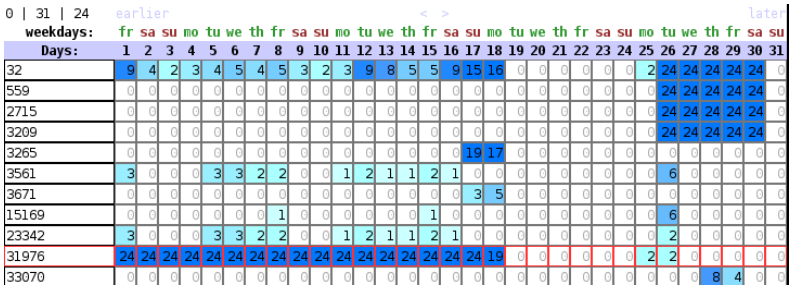

Fig. 1. Temporal and aggregated view of the network traffic – Columns represent the network behavior of a personal machine to an autonomous system (ASN), and rows show the number of hours of a particular day in which a connection to this ASN has been established.

of malicious behavior are indispensable to identity suspicious traffic but often they produce many false network alarms which are cumbersome for a system administrator to manually analyze. Even worse, a well-informed attacker may use knowledge of the internal data structure and algorithms involved in these statistical models to hide suspicious network traffic.

Visualization techniques have been shown to provide insight into large volumes of network traffic (see for example [3]), by combining two powerful information processing systems: the computer to collect network flows and the human visual system to detect unexpected pattern in the network traffic. In our research project we investigate the idea that understanding the patterns of the regular network behavior of personal machines increases the chance of detecting malicious behavior. We propose a visual exploration system based on a data abstraction layer and temporal visualizations of the network traffic. The data abstraction layer enables the administrator to interactively change the level of detail of the network traffic data. It provides an overview of global communication patterns and details on particular patterns. Thus, the system administrator has a powerful tool at hand to check whether the selected pattern might be suspicious behavior. Network traffic is presented in temporal visualizations to support the detection of unexpected network traffic amidst normal diurnal patterns using intuitive visual metaphors.

The abstraction layer has two components: (a) IP-address-space partitioning, and (b) binning of the time dimension. We partition the address space in order to reduce visual complexity. The ip-address-space partitioning is based the Autonomous System Number (ASN) used to route traffic at the wholesale tiers of the internet. An ASN roughly corresponds to an Internet Service Provider (ISP) and so is generally more coherent with respect to organization than the traditional partitioning by IP address octets. Thus, we hope to reduce visual complexity while retaining semantic value. The binning of the time dimension is based on familiar concepts of time such as hour, day, weekday, weekend, etc. The visual interface also provides flexible filtering capabilities to limit results

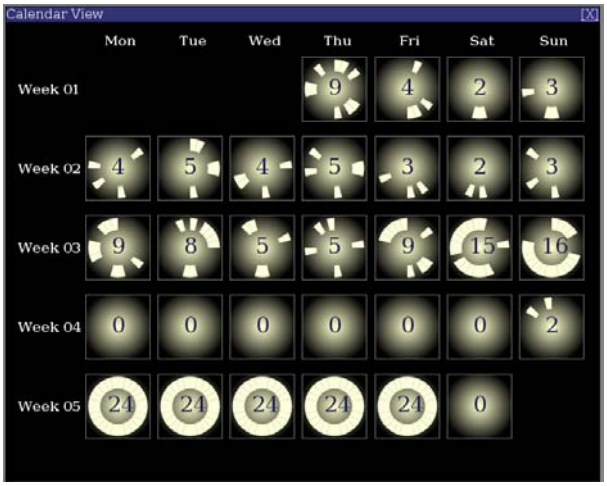

Fig. 2. Calendar View Visualization – the clock metaphor is used to enable a comparison of the network behavior of the local machine with a particular ASN over time

to those ASNs that match (a) a particular time interval (e.g. week day - weekend pattern), and/or (b) to specify a lower bound threshold on the number of network flows in which an ASN is either a source or destination.

Figure 1 shows a view of the aggregated network traffic using a cell per day over a month of traffic for one personal machine. Each row represents the network behavior of this machine and a single autonomous system (ASN), and the cells in each column show the number of different hours in a particular day for which a connection to this ASN was made. In this example, the next to the last row shows the machine starting off the month with regular hourly connections to AS 31976 (Red Hat, Inc). One can easily see that these connections suddenly disappear on Tuesday the 19th amongst other significant changes around this point in time.

The calendar view visualization, as shown in Figure 2, shows finer grain detail by drilling down on the traffic between a single personal machine and a single ASN. The clock metaphor within each day allows a comparison of network behavior of the local machine with a particular ASN during the day. Again, one can easily see the date of the abrupt change in the network traffic patterns of this machine.

In the next phase of this project, we are interested in getting a good feel about the limits of our framework in detecting suspicious behavior. In particular, we are interested in answering the following question: What kind of suspicious activity can be detected? We plan to publish our results in an academic paper.

References

1. NfSen – netflow sensor, last accessed August 2009.
2. Michael H. Cahill, Diane Lambert, José C. Pinheiro, and Don X. Sun. Detecting fraud in the real world. pages 911–929, 2002.
3. Kwan-Liu Ma, Stephen C. North, and William Yurcik, editors. *IEEE Workshop on Visualization for Computer Security (VizSEC 2005), 26 October 2005, Minneapolis, MN, USA*. IEEE Computer Society, 2005.