An Integrated Visualization on Network Events

VAST 2011 Mini Challenge #2 Award:
"Outstanding Integrated Overview Display"

Walter Marcelo Lamagna [1]
Universidad de Buenos Aires

ABSTRACT

To visualize security trends for the data set provided by the VAST 2011 Mini Challenge #2 a custom tool has been developed. Open source tools [1,2], web programming languages [4,7] and an open source database [3] has been used to work with the data and create a visualization for security log files containing network security trends. In this paper, the tools and methods used for the analysis are described. The methods include the log synchronization with different timezone and the development of heat maps and parallel coordinates charts. To develop the visualization, **Processing** and **Canvas** [4,7] was used.

Keywords: visual analysis, security trends, heat map, logs, vast challenge.

1 INTRODUCTION

Log files can be analyzed by using various visual methods to identify trends quickly. This analysis is based on identifying correlated events distributed in time, the importance of an event is sometimes related to the frequency of its occurrence. In this challenge, many network events are logged and some of them are critical, other may not be errors but normal behaviors.

2 DATA SYNCHRONIZATION

Initially the log files were not synchronized because they had different time zones. The interest to synchronize the logs came from the hope to find trends by observing the security events together. To accomplish the synchronization, a unique behavior was searched in the logs that allowed to find the point in time where the same is being logged on different devices. First, the pcap files were exported to text, then a unique combination of network port and address were searched in the firewall log. The next step was to search in the pcap files this unique combination of ports and ip addresses. Once found, the firewall log time was modified to be the same than the pcap files. Secondly, the same procedure was performed between the pcap files and the Intrusion Detection System (ids) log. Third, the firewall and security logs were synchronized (Figure 1.1 and 1.2).

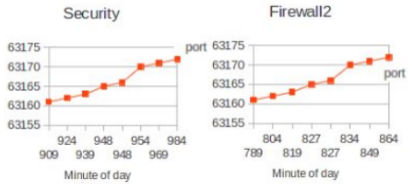

Figure 1.1: Log files before

Firewall2

mod	src	srcport
789	172.20.1.5	63161
804	172.20.1.5	63162
819	172.20.1.5	63163
827	172.20.1.5	63165
827	172.20.1.5	63166
834	172.20.1.5	63170
849	172.20.1.5	63171
864	172.20.1.5	63172

Security2

mod	ip	port
789	172.20.1.5	63161
804	172.20.1.5	63162
819	172.20.1.5	63163
827	172.20.1.5	63165
827	172.20.1.5	63166
834	172.20.1.5	63170
849	172.20.1.5	63171
864	172.20.1.5	63172

Figure 1.2: Log files after the synchronization

3 DATA VISUALIZATION

3.1 Heat map

A heat map, which is commonly used to display and compare variables using a color intensity and position (in time in this case) was used to analyze the logs. It displays the moment and the strength of the event that takes place. The heat map (Figure 2) was developed using canvas [4].

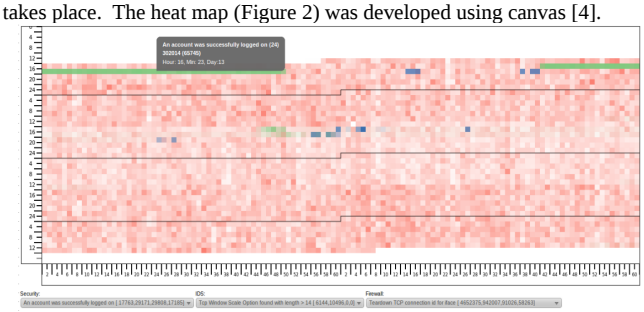

Figure 2: canvas heatmap with 3 selection boxes, each for a log.

Log data was exported to javascript arrays (Figure 3.1), which was processed with perl scripts [5] dumping the data from the database and formatting it (Figure 3.2). This javascript arrays were the input for the heat map.

```
<script type="text/javascript" src="datasec.js"></script>
<script type="text/javascript" src="dataids.js"></script>
<script type="text/javascript" src="datafw.js"></script>
```

Figure 3.1: Importing the javascript arrays.

```
err_4634 = new Array();
err_4624 = new Array();


err_1102[777]=1;
err_1102[3658]=1;
err_4624[777]=18;
```

Figure 3.2: Javascript array formats

The heat map represents in the X axis the minute of the day and in the Y axis the hour of the day. A black thin line separates one day from the other, the heat map has visual space for 4 days. Each square in the heat map is one minute of a day. Each color (red, green, blue) is a log file, being red the security log, blue the IDS and green the firewall log. The intensity of the color is the frequency, the more strong the color, more event instances impacted at that moment.

3.1.1 Informative tools

The heat map has additions to provide on demand information. A tooltip [6] was used to notify which error log is being displayed and the time of day (Figure 4).

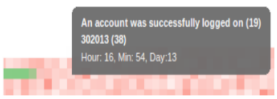

Figure 4: Tooltips on the heat map

Each drop down box provides the visualization of a log event (Figure 5), there are three drop down boxes, one for each log source. If the log files occur in the same time, the colors in the heat map are combined creating a new color.

1 Email: wlamagna@gmail.com

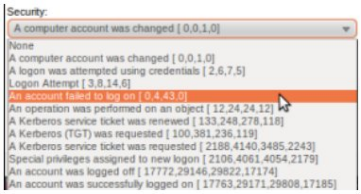

Figure 5: Drop down with log events

The drop down boxes include four numbers between brackets, they inform how many entries of that event appear each day and may help taking some decisions.

When a point in the heat map is clicked, a new browser window is opened and displays a parallel coordinates. The parallel coordinates was developed using **Processing** [7] (Figure 6).

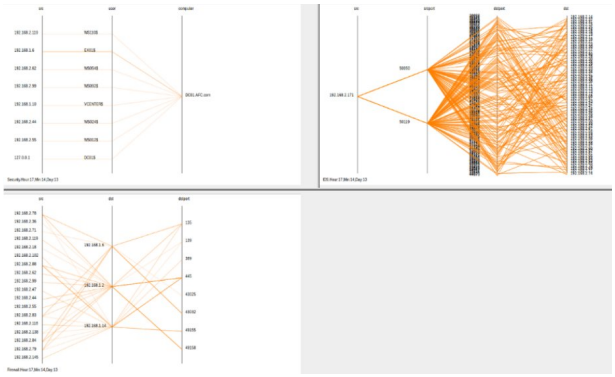

Figure 6: Parallel coordinates

The parallel coordinates are displayed in a web page with 3 frames, each of them for one of the log files (security, firewall and ids). The parallel coordinates include a time stamp for the moment being described. The parallel coordinates read a javascript file, created by dumping the data from the database and then processed with a perl script [5].

4 VISUAL NOTEWORTHY EVENTS

4.1 Denial Of Service on the Web Server

The logs can be browsed and the first identified event is a DoS (denial of service) attack (Figure 7).

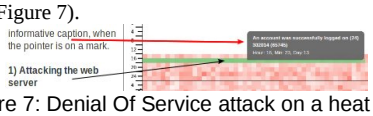

Figure 7: Denial Of Service attack on a heat map

The strong green bar is the firewall log, in this case the event "Built inbound/outbound TCP connection" is selected, it is traffic from the outside to the web server, a stronger color are larger amount of events. The red background are accounts successfully logued in to machines, the stronger the color, more account logins are performed. When the green bar is clicked, at any point but during the DoS, the window with the parallel coordinates graph can be inspected (Figure 8).

Figure 8: Webserver DoS from the outside (parallel coordinates)

4.2 Other threats

The blue points identify the Intrusion Detection System, in this case the event "Tcp Window Scale Option Found" can be selected and inspected through parallel coordinates. (Figure 10). The parallel coordinates do read a javascript file that has been created with a perl script formatting the database output (Figure 9)

```
var firewalltype = [ {src: "category", srcport: "numeric",
dstport: "numeric", dst: "category" }];

var firewall = [
{src: "192.168.1.6", srcport: 43025, dstport: 3394, dst: "192.168.2.174"}
,{src: "192.168.1.6", srcport: 43025, dstport: 3394, dst: "192.168.2.174"}
```

Figure 9: javascript format for the parallel coordinates

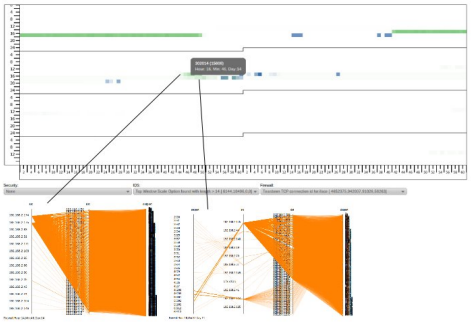

Figure 10: Tcp Window scale option found

5 CONCLUSIONS

For this challenge a custom visualization tool was developed with the hope to understand computer network events. Part of the challenge was to synchronize the logs located in different time zones, then to manage the large data set, and finally to create a proper format for the custom tools (heat map and parallel coordinates). The heat map was useful to identify many of the events; some of them were not identified because extra data processing should have been done, for example, identifying unknown network addresses (for this network in particular) and adding this as another event, or displaying when an unsafe machine was acceded (unpatched machines for example). The benefit of using a custom tool was that it could be adapted for this particular task and may be modified and enhanced.

6 REFERENCES

1. "sed, a stream editor" *www.gnu.org/s/sed*
2. "grep, print lines matching patterns" *www.gnu.org/s/grep*
3. "PostgreSQL: Open source databas" *www.postgresql.org*
4. "Canvas, part of html5" *http://www.w3.org/html/wg/html5/*
5. "The perl programming language" *www.**perl**.org/*
6. "An open source Tooltip" *http://sandbox.scriptiny.com/tooltip/*
7. "Processing, an open source programming language" *http://processing.org*
8. "Parallel coordinates for the VAST Mini Challenge #2" *http://www.magnos.com.ar/vast2011/*