A Firewall Data Log Analysis of Unauthorized and Suspicious Traffic

John Week
University of Nevada, Reno
United States
Email:jweek@weekspace.net
Phone: (775) 741-1555

Polina Ivanova
University of Nevada, Reno
United States
Email:polina_iv@yahoo.com
Phone: (775) 335-4299

Sandy Week
University of Nevada, Reno
United States
Email:smw@weekspace.com
Phone: (775) 784-7054

Alexander McLeod
University of Nevada, Reno
United States
Email:amcleod@unr.edu
Phone: (775) 784-6897

A Firewall Data Log Analysis of Unauthorized and Suspicious Traffic

Abstract

On November 2, 1988, Peter Yee at the NASA Ames Research Center sent a note out to the Internet mailing list reporting, "We are currently under attack from an Internet VIRUS!" As these events were unfolding the firewall was starting its rapid evolution. Management often underestimates the importance of sufficient network security. Remarkably, there is little information available for network administrators to use to analyze the valuable data contained in their firewall logs in order to accurately describe threats to their systems. This paper examines 7,478 attacks logged by a small business Internet Service Provider (ISP) hosting 13 domains. On average, 276 attacks occurred per day. About one half of the attacks are the common Windows RPC and SQL Slammer attacks. Slightly less than one half of those attacks came from ten networks and about 25% of those originated from ten hosts. Results suggest what actions can be taken to strengthen small business network security. Results were compared and contrasted with a similar study called Statistical Analysis of Snort Alarms for a Medium-Sized Network recently undertaken by Chantawut and Ghita (2010.)

Keywords: Network Attacks, Small Business ISP, Origin of Attacks, Time of Attacks, Firewall Data Log

A Firewall Data Log Analysis of Unauthorized and Suspicious Traffic

Many individuals fail to understand the importance of firewalls (Kumar, Mohan, & Holowczak, 2008). Managers are often concerned about company networks connection to the Internet as they are risking the safety of proprietary information. Most businesses have network security policies and practices that dictate how data is to be protected. A firewall provides not only network security - it often plays an important role as a security blanket for management. (Robertson, Curtin, & Ranum, 2004). This report will analyze logs from a small business ISP firewall using Transmission Control Protocols (TCP) and Internet Protocols (IP). Data used in this study was collected from a SonicWall 1260 PRO firewall.

The researchers set out to discern when network attacks occur, what days of the week were showing the highest numbers of attacks and where attacks originated. The outcome should help network administrators prepare systems to withstand most common Internet attacks.

The internet is a Wide Area Network (WAN). For a WAN to work efficiently as a collection of networks, routers pass data packets but do not need to know the exact location of a host for which the information is destined. Routers only know which network the host is a member of and use information stored in their route table to determine how to get the packet to the destination network. Once the packet is forwarded to the destination router, the packet can then be delivered to the appropriate host (Microsoft, 2007). A firewall is a system or group of systems that enforces an access control policy between two or more networks. Although firewall product development has been occurring since the early days of the Internet, they are just barely keeping up with the new applications and services that spring up and immediately become a "requirement" for many Internet users (Avolio, 1999).

This study examined unsolicited inbound TCP/IP traffic on a small business's WAN connections. IP addresses of the originating hosts were analyzed. An IP address is a 32-bit number that uniquely identifies a host on a TCP/IP network. IP addresses are normally expressed in dotted-decimal format, with four numbers separated by periods, such as 192.168.123.132. The number is divided it into four parts of eight binary digits. The first 3 parts identify the network,

the last part identifies the host. If you divide the IP address 192.168.123.132 into two parts the network and the host portions become evident as shown in the table below.

Network	Host
192.168.123	.132 Host

Every IP address will have a number of ports associated with it. Ports either originate or receive connections. In a non-technical sense, one can think of an IP address as the address of an office building and the ports as individual offices within the building. Almost any port can be used to originate a connection. The numbers for originating connections are usually random. Ports which receive connections have to be assigned specific numbers so that network administrators and applications know where to look for them. For example, web server software on a computer would respond on port 80 and mail server software on the same machine would respond on port 25.

Port numbers that are statically assigned are defined as "well known port numbers" and are usually assigned a value below 1024. It is the job of the firewall to ensure TCP/IP traffic enters a network only on authorized specific ports (Nietzsche, 2007).

An Internet firewall has the following properties: it is a single point between two or more networks where all traffic must pass; traffic can be controlled by and may be authenticated through the device, and all traffic is logged.

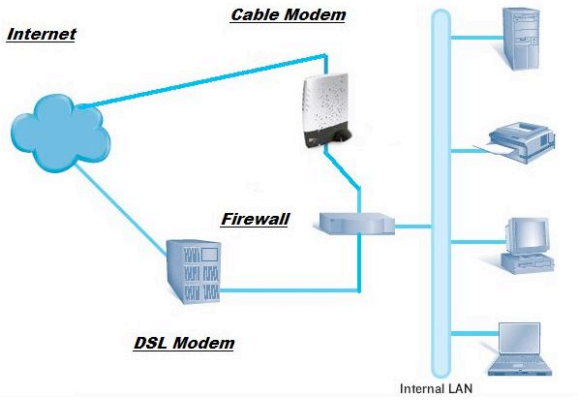


Figure 1 - Firewall placement

As previously stated, there is remarkably little information available to help network administrators analyze the valuable data contained in firewall logs. The problem of processing meaningful information from the data in these logs can be overwhelming. The administrator has to figure out what, where and how logging is accomplished before he/she can even get started on the real work, i.e. making sense out of the data the firewall is producing (Ranum, 2006). In the example above, the small ISP firewall is connected to the Internet by two independent circuits terminating to a SonicWall 1260 PRO firewall. The first connection is a business DSL circuit that provides 64 static IP's and is used for hosting servers. The DSL modem is connected to the WAN port of the firewall. The second circuit is a cable modem, connected to the OPT port and the internal portion of the network connects to the LAN interface. Both Internet connections to the firewall allow deep packet inspection of all incoming traffic and deny unauthorized traffic access to the LAN.

The small business ISP firewall logs unauthorized and suspicious traffic for the network administrator to review. Both circuits are also monitored for availability from San Francisco, CA and Parsippany, NJ. During the data collection period both circuits were available (See Appendix), indicating most attacks were logged and not missed due to circuit unavailability. From March 1, 2009 to March 28, 2009 a total of 7,478 samples were collected.

METHODOLOGY

The small business ISP firewall logs all unauthorized and suspicious traffic attempting to enter the LAN. Each log entry contains the following information:

- Date/Time: The date and time the activity occurred, stored in local (PST) time
- Event type: Notice or Alert, describes level of urgency of event
- Event description: Intrusion Prevention or Network access, describes attacks as either direct or casual
- Action: what action was taken to deny the attack (IP spoof dropped, TCP Syn/Fin packet dropped, UDP packet dropped, Web Access request dropped)
- Source IP: the IP number of the machine making the attack (i.e. 114.121.26.63)

- Source Port: the port number the attack came from (i.e. 2759 Random port)
- Hardware Port: WAN or OPT, which was circuit attacked
- Destination IP: the IP that the attack was destined for ((i.e. 69.239.129.182)
- Destination Port: the destination port that was the target of an attack (i.e. 80 Web Server port)

Table 1 - Firewall data example

03/27/2009 05:24:10.208 - Notice - Network Access - UDP packet dropped - 58.241.69.52, 2759, WAN - 69.239.129.185, 1434, WAN - UDP Port: 1434
03/27/2009 05:25:53.688 - Notice - Network Access - UDP packet dropped - 61.139.54.94, 2150, WAN - 69.239.129.187, 1434, WAN - UDP Port: 1434

Table 1 provides an example of typical firewall data. Firewall logs were collected by the researchers and compiled into a useable format for analysis. The authors then aggregated the firewall data and calculated descriptive statistics. The variables of interest were 1) the number of attacks per week, 2) attack counts for days of the week, 3) target of attacks, 4) attack time of day, 5) ports attacked, and 6) action taken by firewall. ANOVA tests were used to determine whether there was a significant difference between the attacks by hour, day and week.

RESULTS

Results substantiate the need for the additional protection afforded by firewalls. The study found that most attacks on the small business ISP network took place around 02:00 Pacific Standard Time (PST). Table 2 details the average number of attacks by hour of the day.

Table 2 - Attacks by Hour of Day

Hour		
	Freq	Pct
0:00	342	4.6
1:00	350	4.7
2:00	356	4.8
3:00	329	4.4
4:00	341	4.6
5:00	304	4.1
6:00	342	4.6
7:00	334	4.5
8:00	355	4.7
9:00	307	4.1
10:00	293	3.9
11:00	262	3.5

Hour		
	Freq	Pct
12:00	270	3.6
13:00	260	3.5
14:00	285	3.8
15:00	276	3.7
16:00	339	4.5
17:00	311	4.2
18:00	304	4.1
19:00	291	3.9
20:00	277	3.7
21:00	265	3.5
22:00	335	4.5
23:00	350	4.7

The majority of the attacks were between 10pm and 8am. This could mean that the intruders are avoiding normal business working hours or that individual's exercise these attacks during leisure hours. Chantawut and Ghita (2010) stated that "The apparent reduction of the number of attacks detected during office hours might also have been caused by the local Internet traffic congestion as well as packet sniffer (the sensor) missing dropping some of the traffic. Another explanation could be that early in the work week, patches are applied that mitigate the viruses".

On average, there were 312 attacks per hour with a standard deviation of 33. The maximum number of attacks was 356 at 02:00 and the minimum was 260 at 13:00. Figure 2 graphically shows the average number of attacks by hour.

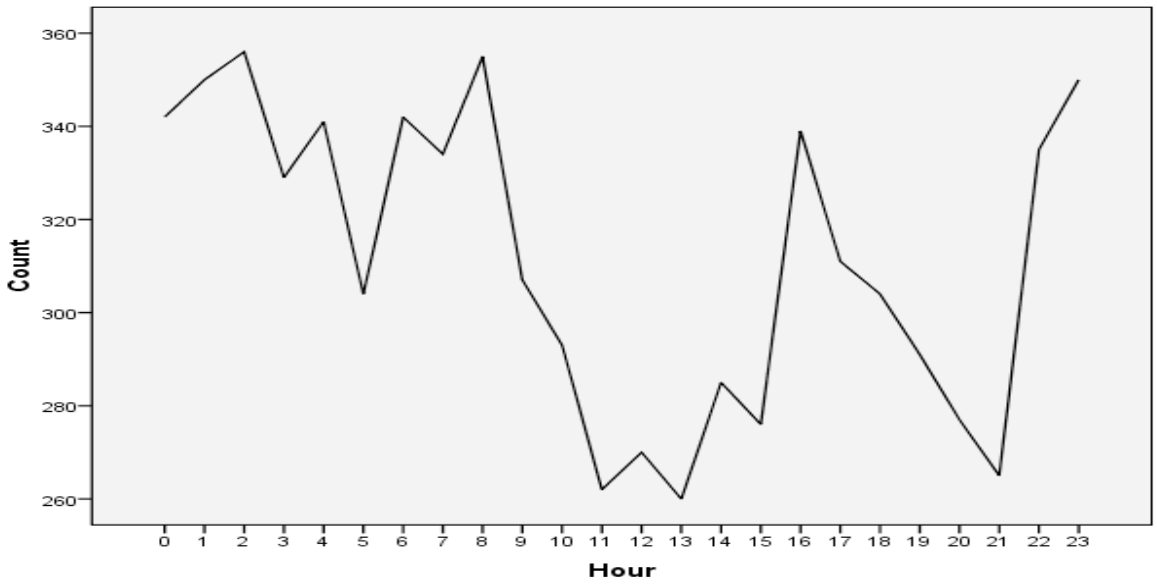


Figure 2 - Number of Attacks by Hour of Day

The researchers found that the busiest day of the week is Monday with 1,549 attacks occurring over the 4 week period on this day of the week. Table 3 shows the number of attacks by day of the week. Consideration should be given to the fact that these were the days and times where the data was collected is located in Pacific Standard Time. Day of the Week assumptions may be impacted by the time zone in the country of origin.

Table 3 - Attacks by Day of the Week

Day	Count	Pct.
Monday	1549	20.7%
Tuesday	1060	14.2%
Wednesday	935	12.5%
Thursday	923	12.3%
Friday	915	12.2%
Saturday	908	12.1%
Sunday	1187	15.9%

An Analysis of Variance was used to compare the means by day of week. Results can be seen in Appendix A. Statistics indicate that there is a significant difference between the mean count by day of week. It appears that Sunday was the busiest day of the week for the small business

network ISP; however, more discussion concerning this and the time zone of attackers is important.

Following the day of week analysis, the researchers examined the data to determine if it was consistent week to week. We ran an ANOVA test to see if there was a significant difference in the number of attacks from week to week, but did not find any significant difference. Figure 3 shows the number of attacks per week.

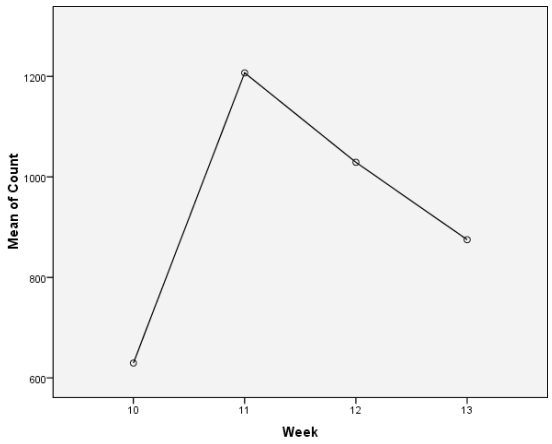


Figure 3 – Number of Attacks per Week

It was discovered there are more attacks from a particular group of IP networks. The count of networks attacking the firewall show that the top ten networks attacking the firewall account for 46.9% of all of the attacks logged. Figure 4 shows the top 10 networks from which attacks originated during our test.

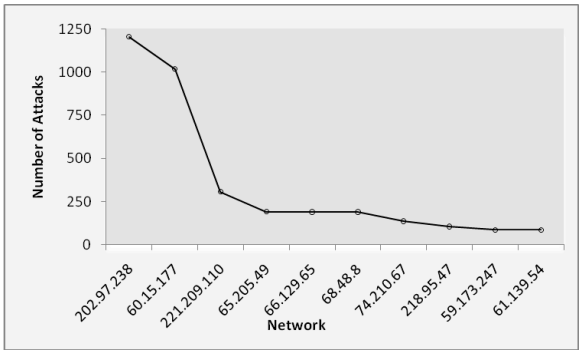


Figure 4 - Top 10 Attack Originating Networks

Thirty-four percent of all attacks originated from just three networks in the same country, China. Network 202.97.238 accounted for 1204, 60.15.177 was second with 1018, and 221.209.110 was

third with 306 attacks. Obviously, a small number of networks are the origin of a large number of internet attacks. Other researchers also found China to be the country where most attacks originated.

The count of IP's attacking the firewall showed that the top ten IP's attacking the firewall accounted for 24.8% of all of the attacks logged. All of the top IP's are in the top ten networks that attack the firewall.

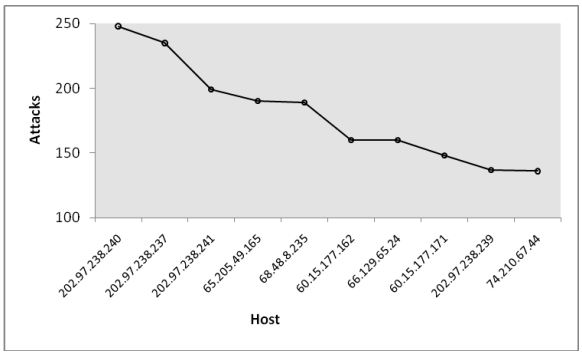


Figure 5 - Top 10 Originating IP

We wanted to know if the small business ISP was experiencing a high number of attacks on a particular interface or to a certain port. We discovered that the WAN interface is experiencing more attacks than the OPT interface. This indicates the DSL circuit that has static IP's is targeted more often than the cable modem.

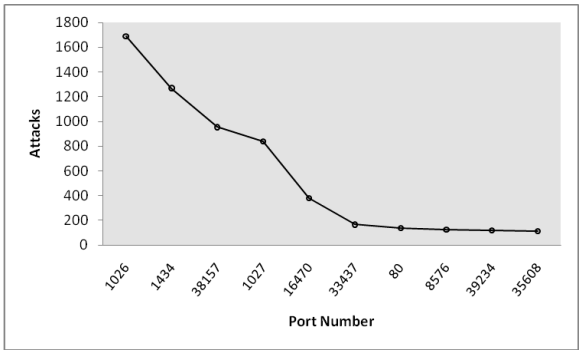


Figure 6 - Attacks by Port

Additionally, the DSL circuit has been established for over five years, while the cable modem has been in operation for three months. It is also possible that the cable modem filtered more attacks prior to them reaching the small business ISP. From the analysis of the ports that are

being attacked it was discovered that the top ten ports account for 77.5% of all of the attacks logged.

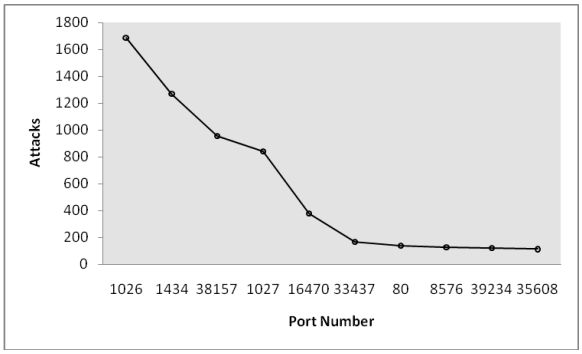


Figure 7 - Top 10 Ports Attacked

Windows RPC and SQL Slammer account for 33.8% and 19.1% of the attacks respectively. This is noteworthy as it appears the source of the Windows RPC traffic is spammers who target the Windows Messenger service that listens for connections on port 1026 and 1027. Windows Messenger has been a target for spammers because it allows anonymous pop-up messages to be displayed on any Windows system running the messenger service. (Stewart, 2003) The SQL Slammer worm looks for vulnerable Microsoft SQL Servers or Microsoft SQL Server Desktop Engine (MSDE) systems to infect. SQL Slammer has the distinction of being the fastest worm ever released on the Internet and had compromised most of its victims worldwide within 15 minutes.

China was the leader in the number of attacks with 53% of all attacks on the small business ISP coming from that country. Second in attacks was the U.S. with 14% and third was Canada. The fourth country with the most attacks was Russia, followed by Brazil at fifth. Figure 8 reports on these results.

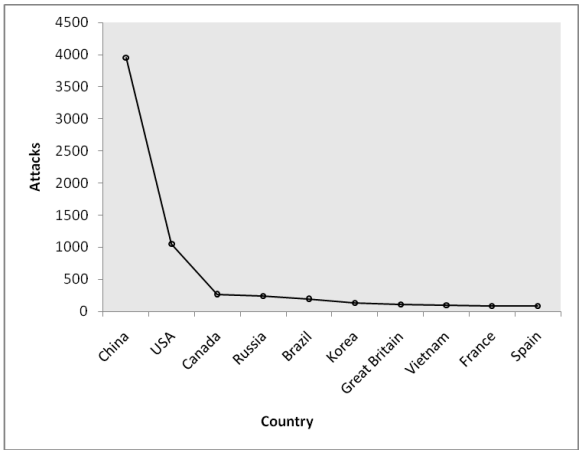


Figure 8 - Attacks by Country of Origin

Differences in time zones may mean that attackers are launching most of their attacks late Saturday evening and early Sunday morning. Based on the country of origin this line may shift and most likely would do so to increase the effect of late night malevolent work on Friday and Saturday.

CONCLUSION

Analysis of the small business ISP firewall shows that the network is attacked about 276 times per day. The firewall was attacked by 2,822 hosts from 2,699 networks. These networks are located in 108 different countries. Approximately half of the attacks are Windows RPC and SQL Slammer. The attacks were fairly evenly distributed over time. There are more attacks during non-business hours and more than 50% of the attacks were on Sunday, Monday and Tuesday Pacific Standard Time. However, due to time zone differences 53% of all attacks originated in China during late night Friday through Saturday. About one half of the attacks are the common Windows RPC and SQL Slammer attacks. Slightly less than one-half of the attacks came from ten networks and about 25% came from ten hosts.

This analysis is valuable information for small business ISPs and researchers. The information provided here offers solid proof of the need for a firewall and network protections and helps network administrators focus on areas that are particular threats. For example, the network administrator may choose to impose extra security measures in preventing Windows RPC and

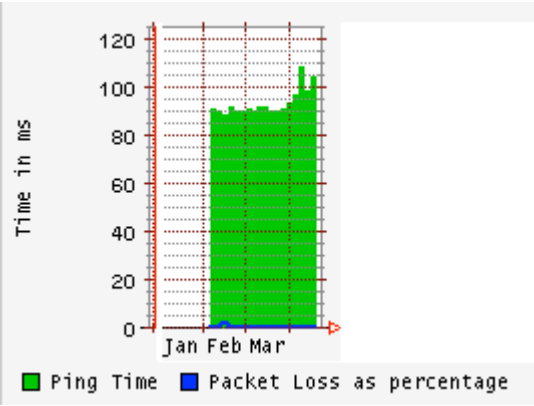
SQL Slammer attacks. In addition, from the analysis the owners may choose to employ more resources during non-regular business hours during Sundays, Mondays, and Tuesdays.

The study could be improved if intruders could be better identified. Although there is an IP for the source, it is not known whether this is the actual IP of the attacker, or the intruder is using a "zombie" to attack the network. Since zombies are remotely controlled, their IP may indicate that poor security measures exist in the country or region of origin. Pirated software may also impact these findings.

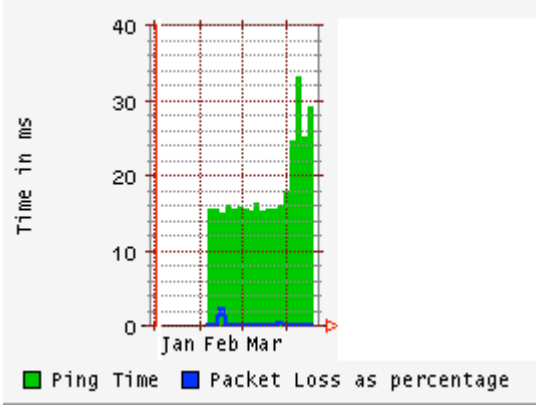
APPENDIX

Circuit Availability from East and West Coast, USA

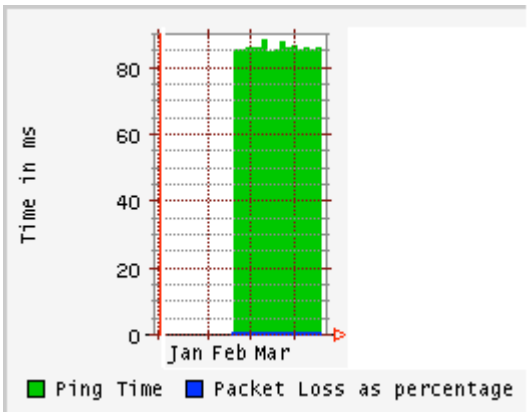
To WAN port from NJ



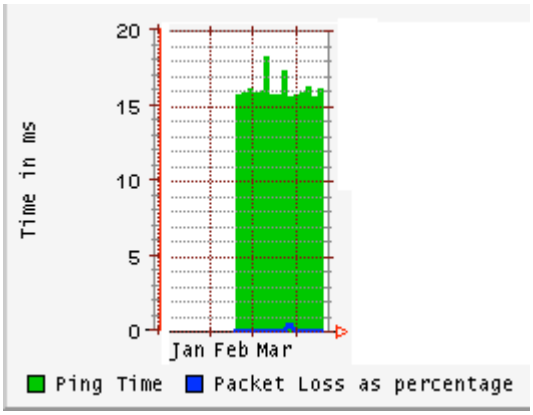
To WAN port from CA



To OPT port from NJ:



To OPT port from CA:



Hour of Day

Descriptives								
HourCount								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
0	342	171.50	98.871	5.346	160.98	182.02	1	342
1	350	175.50	101.181	5.408	164.86	186.14	1	350
2	356	178.50	102.913	5.454	167.77	189.23	1	356
3	329	165.00	95.118	5.244	154.68	175.32	1	329
4	341	171.00	98.582	5.339	160.50	181.50	1	341
5	304	152.50	87.901	5.041	142.58	162.42	1	304
6	342	171.50	98.871	5.346	160.98	182.02	1	342
7	334	167.50	96.562	5.284	157.11	177.89	1	334
8	355	178.00	102.624	5.447	167.29	188.71	1	355
9	307	154.00	88.767	5.066	144.03	163.97	1	307
10	293	147.00	84.726	4.950	137.26	156.74	1	293
11	262	131.50	75.777	4.682	122.28	140.72	1	262
12	270	135.50	78.086	4.752	126.14	144.86	1	270
13	260	130.50	75.200	4.664	121.32	139.68	1	260
14	285	143.00	82.417	4.882	133.39	152.61	1	285
15	276	138.50	79.819	4.805	129.04	147.96	1	276
16	339	170.00	98.005	5.323	159.53	180.47	1	339
17	311	156.00	89.922	5.099	145.97	166.03	1	311
18	304	152.50	87.901	5.041	142.58	162.42	1	304
19	291	146.05	84.236	4.938	136.33	155.77	1	305
20	277	139.00	80.107	4.813	129.52	148.48	1	277
21	265	133.00	76.643	4.708	123.73	142.27	1	265
22	335	168.00	96.850	5.292	157.59	178.41	1	335
23	350	175.50	101.181	5.408	164.86	186.14	1	350
Total	7478	157.95	92.726	1.072	155.84	160.05	1	356

ANOVA					
HourCount					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1870398.884	23	81321.691	9.712	.000
Within Groups	6.242E7	7454	8373.701		
Total	6.429E7	7477			

Average Count by Day of Week

Descriptives								
Count								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	1549	336.99	280.142	7.118	323.03	350.95	1	949
2	1060	664.68	406.493	12.485	640.19	689.18	82	1358
3	935	755.90	406.027	13.279	729.84	781.96	182	1560
4	923	1008.21	399.654	13.155	982.39	1034.03	383	1766
5	915	1249.22	416.266	13.761	1222.22	1276.23	562	1986
6	908	1473.17	416.194	13.812	1446.06	1500.28	750	2193
7	1187	1690.68	428.318	12.432	1666.29	1715.08	936	2413
Total	7477	983.21	613.181	7.091	969.30	997.11	1	2413

ANOVA					
Count					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.680E9	6	2.800E8	1849.850	.000
Within Groups	1.131E9	7470	151375.796		
Total	2.811E9	7476			

Average Count by Week

Descriptives								
Count								
					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
10	1258	629.50	363.298	10.243	609.40	649.60	1	1258
11	2413	1207.00	696.717	14.183	1179.19	1234.81	1	2413
12	2057	1029.00	593.949	13.096	1003.32	1054.68	1	2057
13	1749	875.00	505.037	12.076	851.31	898.69	1	1749
Total	7477	983.21	613.181	7.091	969.30	997.11	1	2413

ANOVA					
Count					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.030E8	3	1.010E8	300.990	.000
Within Groups	2.508E9	7473	335592.455		
Total	2.811E9	7476			

References

Avolio, F. (1999). Firewalls and Internet Security. *The Internet Protocol Journal*, 24-32.

Bouguettaya, A. R. A., & Eltoweissy, M. Y. (2003). Privacy on the Web: facts, challenges, and solutions. *IEEE Security & Privacy, 1*(6), 40-49.

Kumar, N., Mohan, K., & Holowczak, R. (2008). Locking the door but leaving the computer vulnerable: Factors inhibiting home users' adoption of software firewalls. *Decision Support Systems, 46*(1), 254-264.

Microsoft. (2007). Understanding TCP/IP addressing and subnetting basics [Electronic Version]. Retrieved April 23, 2009, from http://support.microsoft.com/kb/164015

Nietzsche, F. (2007). What are TCP/IP ports? [Electronic Version]. Retrieved April 24, 2009, from http://www.tech-faq.com/what-are-tcp-ip-ports.shtml

Ranum, M. (2006). Log Analysis Site Overview [Electronic Version]. Retrieved April 21, 2009, from www.loganalysis.org

Robertson, P., Curtin, M., & Ranum, M. (2004). Internet Firewalls: Frequently Asked Questions [Electronic Version]. Retrieved April 21, 2009,