2010

An Analysis of Malfeasant Activity Directed at a VoIP Honeypot

Craig Valli
Edith Cowan University

DOI: 10.4225/75/57b2b6d840ce5
Originally published in the Proceedings of the 8th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, November 30th 2010

An Analysis of Malfeasant Activity Directed at a VoIP Honeypot

Craig Valli
secau – Security Research Centre
School of Computer and Security Science
Edith Cowan University
Perth, Western Australia
c.valli@ecu.edu.au

Abstract

This paper analyses data collected over a nine month period in a simple VoIP honeypot based on simple design initially put forward by Usken(2009). The honeypot collected 2083 events of malfeasant activity directed towards commonly used VoIP ports. These events resulted in a range of activity being recorded from simple enumeration to advanced probing and attempts to compromise the victim honeypot. The analysis involved traditional statistics from packet analysis, using customised scripts for extraction of data and graphical analysis using i2 Analyst Workstation. The analysis has uncovered an escalation of network activity directed towards the honeypot over a nine month period. Initial geographical IP resolutions also see the majority of traffic emanating from the Chinese IP space. There is strong evidence to suggest that there is a botnet or worm like malcode being directed or developed for VoIP routers.

Keywords

VoIP, honeypot, forensics.

INTRODUCTION

This paper outlines the investigation into data that has been collected by an ongoing project to develop a Voice over Internet Protocol (VoIP) based honeypot system. The research already has produced several outcomes including simple Snort IDS rules, unique NMAP (Fyodor 1998) operating system fingerprints for commodity VoIP routers and some preliminary scripts for topical emulation of a VoIP router administration menu. The research has also produced 2083 candidate events that are potentially malfeasant behaviour directed at the honeypot over a period of nine months.

This paper is not primarily concerned with the construction of a honeypot rather the analysis of its outcomes. The honeypot used in this research itself builds on work by (Usken 2009) in developing simple honeypot systems for VoIP. The design of the particular honeypot used in this research was outlined in (Valli and Al-Lawati 2010).

This paper performs analysis on the 2083 candidate events that were recorded by a actual single honeypot that has operated for the duration of the project. The events that were captured for analysis were triggered by any connection to commonly known and used TCP and UDP ports for VoIP communication. The resulting connection triggered the recording of the event as a network stream in network packet dumps in the commonly used tcpdump (MartinGarcia 2010) file format. The resulting packet dumps have been used as the raw data for the analysis provided in this paper.

RESULTS AND ANALYSIS

Macro analysis was performed on all of the particular recorded packet dumps which were concatenated into one large file for easier analysis and extraction of data. Wireshark (Team 2010) was used for the packet level analysis of the concatenated file of packet traces and also the singular packet traces of each recorded event as necessitated during the research for the paper.

This concatenated data were subjected to traditional statistical measures, geographical IP resolution and also analyzed using the i2 analyst workstation which allows for graphical visualization and query of the data sets. The concatenated capture file was read by tcpdump (MartinGarcia 2010) and then pushed through the Afterglow (Marty 2010) suite of scripts. The information extracted included source and destination IP addresses, source and destination TCP/UDP ports and packet timestamps enabling time lining of the events. The extracted data were then manipulated in a series of Excel spreadsheets, as well as a combination of customised scripts which were used to process and extract information for analysis or representation for example converting the geographical IP data into a Google Keyhole Markup Language (KML) format file for display in Google Earth.

One of the standard measures used by many analysis tools is the top 10 normally by volume or magnitude. The following is a table of the top 10 attacking IP addresses.

IP	Count	Country
121.14.149.145	60	China
60.190.173.30	47	China
117.41.168.235	45	China
117.41.228.243	36	China
113.105.152.220	27	China
121.14.139.34	27	China
121.14.149.175	27	China
121.12.127.168	26	China
119.147.116.249	25	Unknown
117.41.229.74	24	China

Table 1: Top 10 Attacking IPs

The details presented in the table are valid but of little investigative value for most forensic purposes, however they are descriptive and can be used to start a line of analysis or inquiry. The nature of the Internet is such that an attacking endpoint can be singular i.e an IP address or multiple addresses typically from a subnet. The important thing to consider in a forensic analysis of this type is the network endpoint from which the attack emanated or was routed from. The endpoint in the case of compromised networks is the border router for a particular subnet or single real IP. Using this approach this significantly changes the view of the attack epochs. When this logic of border is then applied further to the dataset as geo-located IP addresses a different pattern emerges. The raw data was also subjected to geographical IP query using a customised script and querying a geolocation to IP database. The following table lists the top 10 entries at a level of Country of origin resolution.

Country	IP	
China	1041	63.7%
Unknown/Unassigned	273	16.7%
USA	120	7.3%
India	28	1.7%
Vietnam	27	1.7%
Taiwan	20	1.2%
Korea	16	1.0%
Indonesia	13	0.8%
Russia	13	0.8%
Malaysia	11	0.7%

Table 2: Top 10 Attacking IPs by Country

The second largest item Unknown was composed of either IP numbers that did not result in a resolution in the geographical IP databases as it was unknown or was a spoofed unassigned IP number. The majority of malfeasant traffic at this stage has originated from IP spaces from the Chinese mainland, this is concurrent with observations from other Australian honeypots (Reardon 2010).

Further drilling down with the geographical IP databases for the top ten cities produces the following table of results from the dataset.

Province	City	Country	
Guangzhou	Guangdong	CN	326
Beijing	Beijing	CN	308
Dongguan	Guangdong	CN	152
Zhengzhou	Henan	CN	78
Huzhou	Zhejiang	CN	47
Shanghai	Shanghai	CN	23
Hangzhou	Zhejiang	CN	19
Wuhan	Hubei	CN	13
Changsha	Hunan	CN	12
Nanning	Guangxi	CN	10

Table 3: Top 10 Attacking IPs by City

It is interesting to note that the Guangdong province accounts for almost 50% of China's attributable traffic with 478 events. Also the top three cities within this table with 786 events comprise just under 25% of all events that were logged by the honeypot. The traffic from these regions has been consistent and persistent in probing the honeypot.

Individual micro analysis was performed on candidate events of interest to determine where possible the type of attack or behaviour that was evinced in the execution of that event. The classification is based on the SIP Info and the User-agent details contained within the packets sent to the honeypot. The signatures and handshakes are provided in the Table 4 below.

Scanner (SIP Info)	Events	User-agent
asterisk	700	Asterix PBX
asterisk	419	
sipsscuser	423	sundayddr
sipvicious	331	friendly-scanner
Asterisk	7	friendly-scanner

Table 4: Event Fingerprint/Signature data

The sipvicious(Guac 2010) signatures examined are consistent with use of a default configured sipvicious scanner to enumerate the honeypot. The actual packet traces produced by these particular events are also consistent with default scanning modes provided by the sipvicious tool in duration and magnitude.

The sipsscuser signature based activity detected in the honeypot is interesting and started being recorded on the honeypot on 8 July 2010 and also was the last event extracted from the dataset which occurred on 17 November 2010. The attack profile this particular signature is generating is best described as a persistent threat. The modus operandi displayed by this particular event signature would also indicate similar worm or botnet like activity as detected by Reardon, 2010. The signature provided by this candidate indicates that it is potentially a modified sipvicious scanner or one using the same initial code base.

Sipvicious	Modified sipsscuser
From:"sipvicious"<sip:100@1.1.1.1>; tag=6362613137353765313363340133313536373932363136 Accept: application/sdp User-Agent: friendly-scanner To: "sipvicious"<sip:100@1.1.1.1> Contact: sip:100@127.0.0.1:5064 CSeq: 1 OPTIONS Call-ID: 532778670076825915635622 Max-Forwards: 70	From:"sipsscuser"<sip:100@192.168.1.9>; tag=665879545572395724923248927313319215039211 04449 Accept: application/sdp User-Agent: sundayddr To: "sipssc"<sip:100@192.168.1.9> Contact: sip:100@192.168.1.9:5060 CSeq: 1 OPTIONS Call-ID: 212163834799713241150823794 Max-Forwards: 70

Table 5 - Comparision of sipvicious and sipsscuser

The range of hosts using this tool or piece of code is extensive. The other interesting pattern to note is attributed hosts using this code propagate less than 5 probe events against the honeypot before disappearing. The

geographical origin of this attack profile indicates 82% are coming from host IP addresses originating from the Chinese mainland.

This mode of attack is in stark contrast to some of the concerted efforts by the single attacking hosts or subnets seen elsewhere in the dataset. This event signature also signals an escalation of attack complexity from traditional script kiddies who use a tool in default mode to the same attacker trialling or developing new tools or malware. This *modus operandii* indicates a botnet or some automata for compromise and control of scanning IP. In an online article Reardon (2010) mentions a plausible scenario for propagation of modified VoIP targeted malware via a botnet of victims via SSH compromise and push of payload to the victim computers. The data from this research also supports Reardon's suggested modus operandii.

The analysis undertaken with i2 analyst notebook was exploratory in nature and looked to identify emergent patterns or inconsistency in events through graphical visualisation and analysis of the data. The visualisation of data clearly indicated that the VoIP honeypot was successfully compromised by attackers. The following figure is a snapshot from this graphical analysis

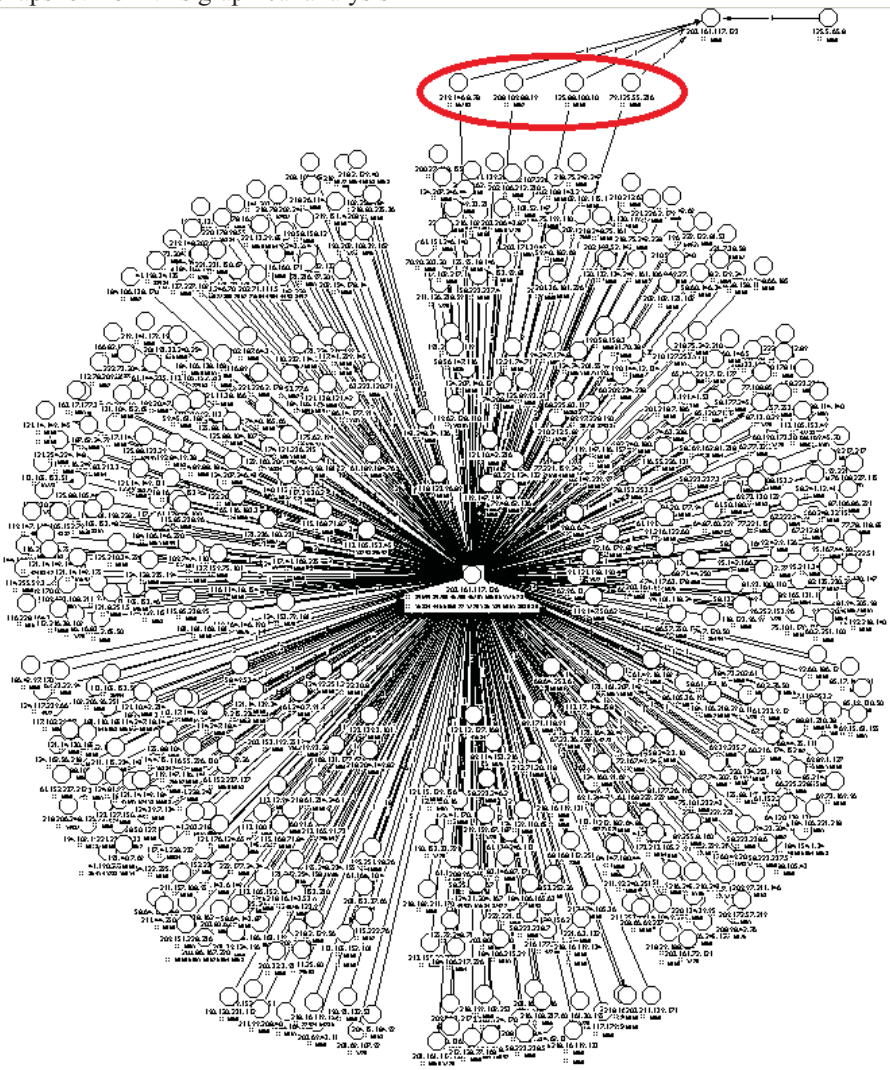

Figure 1: i2 workstation analysis

Upon trace analysis it was determined that there was successful compromise of the emulated VoIP system. The successful compromise of the honeypot saw attackers utilise it to attack another honeypot device within the same physical network subnet. The activity was such that it used four current hosts (circled on image) that the attacker controlled and scanned the next viable target IP in the honeypots subnet. When the scans determined that the newly targeted host was potentially vulnerable to compromise it changed attack vector. The change in attack vector was an attempt to obfuscate their malicious activity as further directed attack against the newly identified victim was from an IP address external to the honeypot subnet. In addition to being external the IP was one not previously used in the widespread attack of the honeypot at that point in time. This particular successful

compromise indicates an increasing maturity of the attacks perpetrated against VoIP systems known to be vulnerable.

KNOWLEDGE, OPPORTUNITY AND MOTIVATION?

The honeypot in this research is basically designed to mimic a domestic VoIP implementation that a standard home user may actually be using. Any crime that is commissioned normally contains three essential hallmarks: knowledge, opportunity and motive. An analysis of these can often aid in the remediation or elimination of the crime.

The analysis of data clearly indicates an escalation in the exploitation attempts targeting VoIP installations. The data, analysis and subsequent evidence clearly indicate that the attackers have knowledge specific to how VoIP is deployed, probed and attacked. There is also clear evidence that their knowledge is growing through development of complex attack methods utilising advanced network exploitation techniques.

As to opportunity, the attackers are exploiting the opportunity afforded to them by the Internet to compromise VoIP implementations that are sitting on relatively insecure default devices. Further opportunity is expanded as detection of malfeasant activity on home systems is negligible due to the naive or novice nature of the home or small business user.

Motivation beyond the criminal intent in this current scenario however is much harder to determine. Large commercial VoIP systems typically have a corresponding large amount of bandwidth on which to send data. It is therefore easier to hide malfeasant activity in a larger volume of data. In essence, a salami attack or slicing of this bandwidth can be used to accommodate extra VoIP traffic which would evade detection as long the footprint was comparatively small as compared to the overall size of the connection. The same corresponding attack on a relatively low bandwidth home-based installation however will in all probability be detected. The detection however may not be as a result of observed degradation of service but via an alternate channel, which is at the next pay period for the Internet or VoIP service. This detection will be possibly as a result of the stolen bandwidth potentially causing an increase in service charge or traffic volume due to the relative size of Internet installation. In addition, there can also be telephony charges associated with the service which again due to the relative low volumes in a household or small business with any significant increase would be easily apparent. Also in the case of the home user they tend to dial a restricted set or known numbers and again matched against the billing may alert the end user to malfeasant activity.

Victims compromised by the use of the identified botnet mechanism may not be in fact targeted for financial gain directly through the theft of telephony as the key motivating factor. Financial gain may instead be garnered as a result of directed distributed denial of service (DDoS) of a third party. The motivation is a sustained DDoS of corporate entities that have critical business exposure on the Internet as a result of service dependencies. The dependencies directly relates to service availability that is one of the core elements for business success and sustainability for example online gambling or online banking. These particular types of network dependent businesses are the targets and consequently enable the primary motivation for compromise of a home or small business users VoIP system.

Furthermore, the control of a large range of victim hosts also has the unfortunate plausible scenario that these telephony devices are used to flood the public switched telephone network with calls in the case of an emergency or a physical attack such as a terrorist incident. This scenario is plausible as the VoIP service on the compromised hosts could be made to make simultaneous and repeated calls to a particular telephone or telephone exchange. Another plausible explanation for motivation is for snooping in on voice communications that are being carried or transmitted through these VoIP routers this could be for a variety of reasons including intelligence gathering or industrial espionage.

FURTHER RESEARCH

This paper is an analysis of one data set taken from this single ADSL based honeypot. The next analysis will be a combination of data taken from various honeypots within the Australian context. This combined analysis will allow correlation of events across a wider network space and also time continuum. Furthermore, the extended dataset will allow for a more in-depth analysis relating to traffic patterns, distribution methods and modus operandi of any botnet or targeted network borne malware that is being used.

There is a definite need to conduct research into this area because unlike other areas of Internet malfeasance this can have significant and direct financial gain for the perpetrator of the attack through theft of service. Unlike many other network borne attacks these may in fact be also readily directed at national critical infrastructure in the form of the telecommunications networks to cause denial of service.

There needs to be research conducted into the methods of targeting in use. This will allow researchers to determine if the compromises that are appearing are totally automated or in fact have human actors. It is indicated from the data in this study and also from Reardon (2010), that there is development of automated attacks being undertaken.

CONCLUSION

This research has uncovered significant patterns of attack that are being directed against VoIP systems. The magnitude of attacks perpetrated is across the full range from simple opportunistic scanning through to full penetration and compromise of a VoIP honeypot. This research also strongly supports other indications of an escalation of attacks from systems located within the Chinese IP spaces with almost 70% of attacks being directly attributable.

The analysis indicates an escalation of attack complexity and sophistication being used. At this stage however, targeting of hosts seems very opportunistic and not discriminatory. Indications are also that the level of attack is reaching a persistent level with several significant attacks per day being experienced by the honeypot. Motivations for escalation of attack on this type of device are also questionable and warrant significant investigation.

Further research is desperately needed to ameliorate the significant threat to this current activity poses to VoIP systems and overall network stability for both the Internet and standard telephony-based services.

REFERENCES

Fyodor. (1998). "Remote OS detection via TCP/IP stack fingerprinting." Retrieved 10 May, 2002, from http://www.insecure.org/nmap/nmap-fingerprinting-article.txt.

Guac, S. (2010). SIPVicious tool suite.

MartinGarcia, L. (2010). tcpdump.

Marty, R. (2010). Afterglow, Raffael Marty.

Reardon, B. (2010). "Sunday (sundayddr) SIP scanning worm. When printers turn bad.." Retrieved 17th November, from http://www.honeynet.org.au/?q=sunday_scanner.

Team, T. W. (2010). Wireshark, Riverbed Technology.

Usken, S. E. (2009). "VoIP - Voice over IP or haVock over IP?", from http://www.honeynor.no/data/honeynet-voip-presentation-anonym.pdf.

Valli, C. and M. Al-Lawati (2010). Developing VoIP Router honeypots. 2010 International Conference on Security & Management,, Las Vegas Nevada, USA, CSREA Press.