Comparative Survey of Local Honeypot Sensors to Assist Network Forensics

P.T. Chen, C.S. Laih
National Cheng Kung University
Tainan, Taiwan
peder@crypto.ee.ncku.edu.tw
laihcs@eembox.ncku.edu.tw

F. Pouget, M. Dacier
Institut Eurécom
Sophia-Antipolis, France
{pouget,dacier}@eurecom.fr

Abstract

This paper intends to illustrate the usefulness of deploying multiple simple honeypot sensors in a large variety of locations. Indeed, a permanent identification of anomalies that occur on a single sensor allows pinpointing abnormal local activities. These can be the manifest of misconfiguration issues or highlight attacks particular to some given environments. Both cases are important for administrators in charge of the networks hosting the sensors. We propose in this paper a comparison of simple parameters that reveal to be an easy way to determine these abnormal and particular activities. On the basis of two identical honeypot sensors that we have deployed for more than 6 months in France and in Taiwan, we detail the analysis of some anomalies that have been found against one unique sensor only. This is a preliminary but useful stage for network forensics and we intend in a near future to deploy the method over a large number of sensors. This is an on-going work and we hope that the illustrations we provide all along the paper will be a good incentive for partners to join this open project.

1. Introduction

Many solutions exist for monitoring suspicious traffic on the Internet. However, they often consist in monitoring a very large range of IP addresses like a whole class A or a large interval of unused IPs. Several names have been used to describe this technique, such as network telescopes [1][2], blackholes [3][4], darknets [5] and Internet Motion Sensor (IMS) [6]. Some other solutions consist in the passive measurement of live networks by centralizing and analyzing firewall logs or IDS alerts ([7][8]). A few websites like DShield, SANS/ISC or MyNetwatchman ([7][9][10]) report such trends. Coarse-grained interface counters and more fine-grained flow analysis tools such as NetFlow [11] offer another readily available source of information.

So far, nobody had investigated the possibility of using a large number of local and similar sensors deployed all over the Internet. However, we strongly believe that local observations can complement the more global ones listed above. A direct analogy can be made here with weather forecast or volcanic eruption prediction, where both global and local approaches are applied. As a consequence, we have deployed many small honeypot sensors in various locations thanks to motivated partners, as part of an academic project. The main objective is to gather statistics and precise information on the attacks that occur in the wild over long periods of time. We have initially used high interaction honeypots. Then, because of the incoming and increasing number of participants in addition to the hard constraints imposed by their implementation, we have considered the idea of deploying low interaction honeypots. We invite the interested reader to have a look at [16] for an in depth presentation of the environment built. [18] offers a study of the limitations induced by the choice of using low interaction honeypots instead of high interaction ones. [12][13] offer some preliminary results based on the initial high interaction platform.

Having a large number of sensors enable us to try to model the various attack processes found in the Internet. This is an ongoing effort which appears promising [14]. It also offers a new way to quickly and easily identify local phenomena that are worth being investigated by the people in charge of the network where the sensor is located. It is the purpose of this paper to highlight, on the basis of a couple of simple examples, the merits of this approach. For the sake of simplicity, we mention in the following only two sensors, one being located in an academic network in France, and the other in an academic network in Taiwan. The Taiwanese sensor is the one where discrepancies have been observed. The French sensor is used as a representative example of what has been seen on the other sensors. It would have been tedious and useless to systematically give all results for all sensors. Thus, when we talk about the French sensor, we actually refer to what has been seen on all sensors, and to make things concrete we focus on that one. It is important to note though, that such an approach can, of course, not be carried out with only two sensors.

Both sensors share the same configuration and have been running for the same period of time. By definition, the honeypot is a non-productive machine, and thus, should not get any particular traffic (in theory). As suspected, it is not the case. In the following, we compare the traffic collected on both sensors over a 6-month period by means of simple parameters, like the targeted ports, the geographical location of the attacks, the attacking domains, etc. We then show that this comparison gives the opportunity to identify activities, which are specific to one particular location only. Their root causes are twofold:

- Some machines can be misconfigured in the network. They participate to the *unexpected and unwanted traffic* collected by the sensors.
- Some attacks are very local and are launched against a small number of networks only.

In both cases, the administrator in charge of the network hosting the sensor needs to be aware of such activities in the network. It is even more important that he is in many cases overwhelmed with IDSs alerts and false positives. This simple technique gives a fast and efficient way to pinpoint local and abnormal traffic thanks to the discrepancies observed with respect to other distant yet identical platforms.

The rest of the paper is structured as follows: Section 2 describes and justifies the setup of the distributed honeypot sensors. Section 3 introduces a comparison of global statistics obtained on two of these sensors, one being in France, and the other one in Taiwan. In particular, we show the similarity of the information provided by the two environments. We also highlight the existence of a few phenomena that are observed on one location only. In Section 4 we take a closer look at some of these phenomena and describe a handful of them. Finally, Section 5 concludes this paper.

2. Description of the Distributed Honeypot Sensors

2.1 Honeypot Sensor: a brief overview

We have deployed a honeypot sensor based on several open source utilities, which emulate operating systems and services. The basic building block used is honeyd [14]. The sensor only needs a single host station, which is carefully secured by means of access controls and integrity checks. This host implements a proxy ARP. This way, the host machine answers to

requests sent to several IP addresses. Each IP is bound to a certain profile (or personality in the honeyd jargon). Thus, the emulation capacity of the sensor is limited to a configuration file and a few scripts. The sensor we are using emulates three Operating Systems, Windows 98, Windows NT Server and Linux RedHat 7.3, respectively. Some service scripts that are available in [14] have been linked to open ports, like port 80 (web server) or port 21 (ftp), among others. A simple sensor architecture is presented in Figure 1. Finally, we connect to the host machine to retrieve traffic logs and check the integrity of chosen files every day.

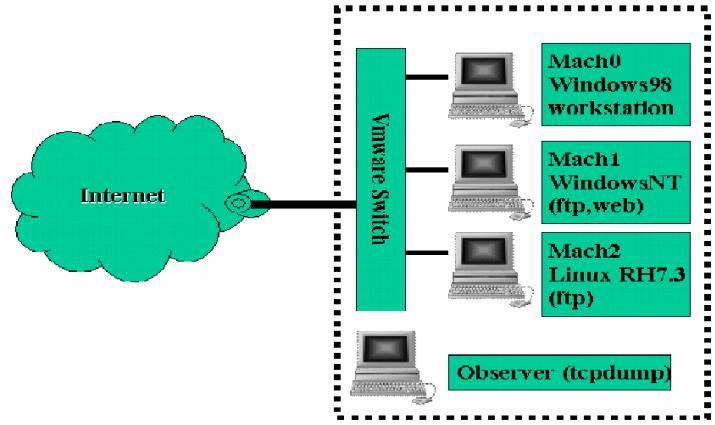


Figure 1. Sensor Architecture

2.2 Deployment

The major objective consists in getting statistical information from the attacks over a long period of time. Therefore, low interaction honeypots like the ones presented above represent a suitable solution. Indeed, we only want to observe the first attack steps in order to get a better understanding of current malicious activities.

The project we have launched aims at disseminating similar sensors everywhere thanks to motivated partners, on a voluntary basis. Partners are invited to join this open project and install a sensor on the premises of their own networks. We take care of the installation by furnishing the sensor image and configuration files. Thus, the installation process is automatic. In exchange, we give the partners access to the centralized database and its enriched information. We are also developing a dedicated web to make research faster and more efficient. The project has started triggering interest from many academic, industrial, and governmental organizations. As of this writing, around 30 platforms are deployed in 20 different countries covering the five continents. We keep installing new ones regularly.

2.3 Data Collection and Analysis

As previously explained, dump files are periodically collected from each sensor and are stored in a centralized database. This one contains, for each attack, a large variety of information, such as:

- Raw packets (entire frames including the payloads are captured with tcpdump);
- TCP level statistics using TCPstat;
- Passive Operating System fingerprinting obtained with Disco, p0f and ettercap;
- IP geographical localization obtained with NetGeo, MaxMind and IP2location;
- DNS reverse lookups, whois queries, etc…

This data needs to be properly organized, as it will be used for further analysis and experiments. In theory, no traffic should be observed from the machines we have set up. As a matter of fact, many packets hit the different virtual machines, coming from different IP addresses. Typically, if an attacker decides to choose one of our honeypots as his next victim, he tries to establish direct TCP connections or to send UDP, or ICMP, packets against it. He can behave differently when targeting each of the three virtual machines. As a consequence, we distinguish in the database three major classes of information:

1. Information that characterizes the attacking source. It includes its IP address, the date it has been observed, the domain and geographical location associated to this address, etc.
2. Information that characterizes the behaviour of the attacking source against the global sensor. It includes the number of virtual machines it has targeted, the global time it has been observed on it, the way it has targeted the virtual machines (sequence vs. parallel), etc.
3. Information that characterizes the behaviour of the attacking source towards one virtual machine. It includes the sequence of ports that have been targeted, the data sent, the number of exchanged packets, etc.

For the sake of conciseness, we do not want to describe the full database architecture here. All details are precisely described in [16]. We just want to point out that most of the comparisons that are presented in the following rely on this efficient way to organize the information.

3. Global Statistics from Two Sensors

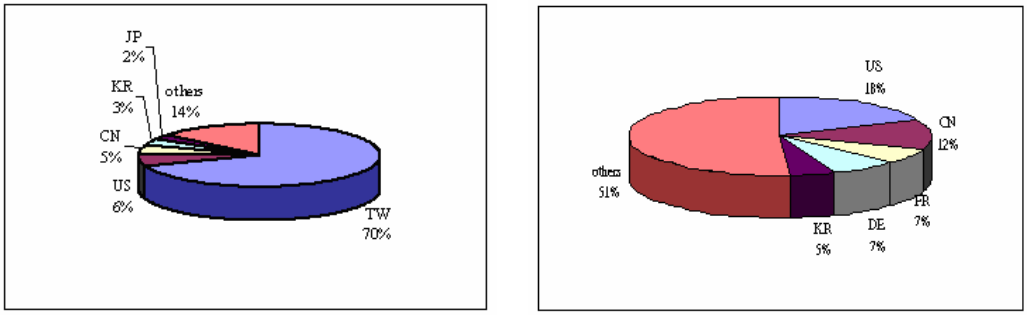
In this section, we present a comparison of two sensors, one being in Taiwan and one in France. In Taiwan, the honeypot sensor was located in the Taiwan Academic Network (TANet) backbone, while the sensor in France was deployed in the national French academic network[1]. In the following, we will call these sensors *Sensor T* (T for Taiwan) and *Sensor F* (F for France) respectively. We provide a comparison based on a small number of parameters, such as attack origins, attackers' Operating Systems, targeted services and attacking time, from the traffic collected during a 6-month period. The results illustrate that there is quite a number of differences between these two sensors. This simple comparison will help pinpointing some anomalies particular to *Sensor T* that will be characterized in Section 4.

3.1 Origins of the Attacks

First, we compare the countries associated to the sources having targeted *Sensor T* with those having targeted *Sensor F*. We observe in Figure 2 that the countries at the origin of the attacks against *Sensor T* and *Sensor F* are very different. The Figure provides the top 5 countries on each sensor, and all the other countries are grouped into the 'others' category. We notice that 28003 distinct IP addresses (70% of the attacks) observed on *Sensor T* are coming from the very same country, Taiwan. This contrasts with *Sensor F* where 53674 distinct IPs, that is 51% of all observed IPs, are found in the 'others' category. Thus, there is

[1] We avoid providing exact locations of the honeypots in order to minimize the risk of introducing a bias in the data collection process.

no clear prevalence of attacking countries on *Sensor F*. Such a particularity is only encountered in the Taiwanese environment.



(a) *Sensor T* (b) *Sensor F*

Figure 2. Attacking countries observed on Taiwanese and French sensors

3.2 Operating Systems of the Attackers

In order to find out the operating system that was running on the attacking machine, we make use of passive fingerprinting techniques instead of active ones in order to minimize the risk of alerting the attackers. Results are summarized in Table 1 for the open-source fingerprinting tool called p0f [17]. With no surprise, the most frequently identified operating system belongs to the family of Microsoft Windows. More precisely, in our analysis, it indicates around 91% (resp. 93%) of the observed attack sources are Windows machines in *Sensor T* (resp. *Sensor F*). A very small number of the attacking machines are running UNIX-like systems, such as Solaris, BSD, and Linux. There is still 7% (resp. 4%) of attacking operating systems that cannot be recognized by p0f. It results either from the situation where p0f does not complete the OS decisions or from the circumstance where some skilled attackers configure their machines to confuse p0f.

Table 1. % of attackers' OS determined on each sensor data

OS_Name	*Sensor T*	*Sensor F*
Windows	90.76	92.81
Unknown	8.5	5.98
Solaris	0.55	0.14
Cisco	0.06	0
CacheFl	0.05	0
FreeBSD	0.04	0

OS_Name	*Sensor T*	*Sensor F*
SunOS	0.02	0.02
OpenBSD	0.01	0
Linux	0.01	0.95
Novell	0	0.06
Eagle	0	0.04

3.3 Targeted Ports

In order to understand the attack trends, a preliminary comparison between the top10 targeted ports against *Sensors T and F* has been conducted. As we show in the following, this preliminary step is very meaningful to find the first indications of local anomalies. Table 2 details the results and leads to the following comments:

1. The usual suspects, i.e. the ports known to be vulnerable on Windows machines (e.g. 135, 137, 139, 445), are found in both environments.

2. Well-known backdoors left open by famous worms are also targeted on both machines (e.g. 5554, 9898).
3. There are several ports that are specific to each sensor (e.g. 8080, 1026, 3128).
4. It is worth noting that most targeted ports on *Sensor T* are associated with web services, such as port 80 (http), 8080 (http-proxy) and 3128 (squid-http). This does not appear to be true on the other one.

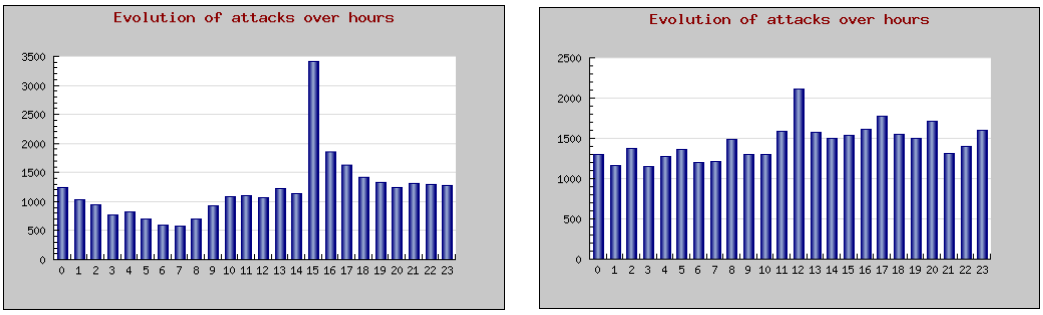
Further investigation reveals that a larger variety of ports have been targeted on *Sensor F*. Indeed, 78% of all attacking IPs against *Sensor T* have targeted at least one of the top10 presented ports. The percentage is only of 59% for *Sensor F*.

Table 2. Target ports on *Sensors T and F*

Sensor T	*Sensor F*
135	445
139	5000
80	135
137	1026
1025	139
8080	1027
1080	1433
3128	9898
5554	5554
9898	1023

3.4 Timing of the Attacks per day

Figure 3 shows the different timing of the attacks per day (local time of the sensors). The x-axis represents the 24-hour intervals. The y-axis represents the sum of the attacks observed during a given hour. It confirms that both sensors are targeted all day long. This seems normal as these attacks are very short and thus, are probably launched by automated robots with an apparent permanent access to the net. From Figure 3, we see that *Sensor F* suffers from more attacks than *Sensor T*. However, although there are globally more attacks on *Sensor F*, values per hour are quite stable (around 1500), except for a small peak between 12 am and 1 pm. If we now compare with the Taiwanese sensor, most values per hour are around 1000, even if the number of attacks seems to decrease during the early morning. There is, however, an abnormal peak of activity between 3 pm and 4 pm. A deeper analysis is required to better understand this surprising temporal pattern.



(a) *Sensor T* (b) *Sensor F*

Figure 3. Different timing of the attacks per hour between sensors

4. Some Features Specific to the Taiwanese Sensor

As we observed above, the honeypot environment in Taiwan presents very specific features. Among others, we have pinpointed some of them, which are:

- An important volume of attacks is coming from Asian countries, and especially from the country hosting the sensor, Taiwan.
- Some frequently targeted ports appear to be very specific to that sensor only.
- Timing of the attacks presents a clear diurnal pattern, and also an abnormal peak of activities around 3 pm.

It would be helpful for administrators to analyze these anomalies. In the next Section, we propose to highlight some of these phenomena according to the information collected on the Taiwanese Sensor.

4.1 Analysis of the Attacking Domains

In order to get more information on the origin of the attacks, we extract the domains of the attacking IPs with *whois* queries. We present in Table 3(a) the top10 domain names with a name level of 2. Around 65% of the attacks are coming from these 10 domains. Table 3 (b) shows the top10 domains with a more granular precision (domain name level of 3). In this case, 51% of the attacks are coming from these 10 smaller domains. It is not our intent to blame any domain for the attacks coming from their machines. Giving the exact name might be misunderstood and would not give any added value to the report. Therefore, we have decided to replace the real names by letters in a consistent way. This means that the same letter in the table will always represent the same name. In the 24352 collected domains, we can say that most of the attacks come from education domains but also from large Taiwanese ISPs. Moreover, we observe that most of the attacks come from the very same domain hosting the sensor (.o.a.b). This indicates that this domain contains many vulnerable computers that try to propagate within the domain by scanning it. Additionally, this confirms that local attack tools, or at least, tools with local propagation strategies are largely targeting *Sensor T*. This is not something we observe on the French sensor, as indicated by the small fraction, 7%, of attacks against it launched from France (see Figure 2(b)).

Table 3. Analysis of the attacks per domain name

Domain Name	Number of Sources	Rate (%)
.a.b	7738	31.77562
.c.d	4009	16.46271
.e.f	887	3.642411
.g.b	819	3.363173
.h.d	552	2.266754
.i.j	546	2.242116
.k.d	331	1.359231
.l.f	289	1.186761
.m.d	284	1.166229
.n.d	236	0.96912
Others	8661	35.56587

(a)

Domain Name	Number of Sources	Rate (%)
.o.a.b	5977	24.5462
.p.c.d	3375	13.86037
.q.c.d	634	2.603696
.r.e.f	525	2.156057
.s.g.b	451	1.852156
.t.k.d	331	1.359343
.u.a.b	322	1.322382
.v.a.b	283	1.162218
.w.l.f	278	1.141684
x.y.d	227	0.932238
Others	11947	49.06366

(b)

In Table 4, we present the attacks on the Taiwanese and French sensors with regard to the number of targeted virtual machines. We remind that each sensor emulates three different machines, and we compare here the attack sources having targeted one, two or three of these virtual machines. The percentage of attacks that target exactly two machines is very low for both *Sensors F and T*. However, in a general manner, ratios are quite different between sensors. On *Sensor T*, we note that most of the attacks target the 3 virtual machines, while ratios are somehow equivalent for *Sensor F* concerning attacks on 1 or 3 targets. A closer look at those attacks having targeted one Taiwanese virtual machine reveals that there is no *"favourite victim"*, despite the fact they are running different OSs and services (mach0: 34.8%, mach1: 31.6%, mach2: 33.6%). As a first conclusion, it indicates that the majority of attacks on *Sensor T* are due to an equal number of random and sequential scans of the local network. The comparison with *Sensor F* together with the indication provided before indicates the virulent activity, on the *Sensor T* network, of machines running localised random scanning tools.

Table 4. Attacks against virtual machines

Sensors	Attacks on 1 virtual machine only	Attacks on 2 virtual machines	Attacks on 3 virtual machines
Sensor T	28.09%	9.90%	62.01%
Sensor F	51.80%	4.25%	43.95%

4.2 Statistics on Ports Sequences

Some attack tools have specific propagation strategies, and each attacking machine probes an ordered list of ports, or *ports sequence* on a virtual machine. Therefore, it is important to gather statistics of ports sequences when we want to identify attack tools for network forensics.

Table 5 represents the top8 ports sequences targeted by the sources observed over months, from December 2004 to May 2005. As an illustration, it has been observed in April 2005, that 50.17% of the attacks have targeted the sole port 135, while 1.63% have targeted the ports sequence {5554,1023,9898}. We want to insist on 3 major remarks resulting from Table 5:

1. It is important to notice that the same set of traditional ports (i.e. 21, 80, 135, 139, and 1025) is observed over many months.
2. There are still very few exceptions, such as ports sequences {3306} or {3127} that are found at very particular dates.
3. The total number of attacks fluctuates a lot. It is mainly due to the single ports sequence {135}, which stems for half of the malicious collected traffic, except on two months, where we note a non-negligible decrease.

The three previous remarks are of interest for the administrator in charge of the network where *Sensor T* is located. Their intrinsic particularity requires specific attention, and it will help understanding many triggered alerts from other network security boxes (Intrusion Detection Systems, firewalls, etc).

We have shown so far that *Sensor T* presents very interesting characteristics, which are unique to this sensor. They are most likely due to some specific malware scanning randomly the local network and its vicinity. In the following, we propose to investigate two of them related to the following ports sequences: {135} and {8080,3128,1080,1813,80}.

Table 5. Ports sequences: % over months

December 2004	**January 2005**	**February 2005**
Total events 5304	**Total events 6873**	**Total events 1856**
135 (58.92%)	135 (55.84%)	135 (24.84%)
139 (12.76%)	139 (12.47%)	80 (12.61%)
80 (3.13%)	445 (3.23%)	139 (8.14%)
5554, 1023, 9898 (2.21%)	80 (2.75%)	1025 (6.14%)
3127 (1.70%)	139, 445 (2.07%)	5554, 1023, 9898 (5.01%)
1025 (1.60%)	5554, 1023, 9898 (1.89%)	42 (3.93%)
22 (1.20%)	1025 (1.50%)	22 (3.29%)
137 (0.92%)	4899 (1.12%)	23 (2.86%)
Others (17.56%)	Others (19.13%)	Others (33.18%)
March 2005	**April 2005**	**May 2005**
Total events 1734	**Total events 2697**	**Total events 3787**
135 (22.71%)	135 (50.17%)	135 (62.34%)
1025 (16.01%)	80 (7.97%)	1025 (7.84%)
80 (7.93%)	1025 (4.15%)	80 (5.04%)
139 (5.39%)	22 (3.34%)	139 (4.33%)
135, 1025 (5.24%)	4899 (3.34%)	22 (1.87%)
135, 1025, 139 (4.88%)	139 (2.15%)	135, 1025, 139 (1.32%)

4899 (3.64%)	5554, 1023, 9898 (1.63%)	3306 (1.08%)
22 (2.69%)	8080 (1.19%)	4899 (1.08%)
5554, 1023, 9898 (1.60%)	Others (26.06%)	Others (15.10%)
Others (29.91%)		

4.3 Analysis of the Port Sequence {135}

As shown in Table 5, Windows ports 135, 139 and 1025 are the most targeted ports. It is worthy to note that the most targeted port is 135, Windows RPC DCOM port, since this port is attacked with a rate over 50% every month. As known to all, this port suffers from several security problems. [19] offers a thorough review of all the problems, exploits and patches related to that specific port.

We observe a rapidly increasing number of attacks on this port since August 2004, and the number reaches its maximum value in October 2004. This is consistent with the various releases of malware exploiting several vulnerabilities on that port, as exposed in [20]. This trend is something that we do observe on all other sensors as well. However, the well known worms are unlikely to be the only cause of all the attacks observed against port 135 on the Taiwanese platform, even if, at first glance, explanations offered elsewhere, e.g. in [19] and [20], are somehow convincing. There are, at least, two reasons that should catch our attention and advocate in favor of the presence of other attack processes, namely:

1. As seen before, most of the attacks on the Taiwanese sensor are coming from a limited number of domains, mostly from a single country. This is not something that we do observe on the other platforms for similar attacks against port 135. As a result, the root causes of these hits are probably different. We probably are not facing the 'classical' worms here but, instead, something, which, from the network viewpoint, shares some characteristics with them, perhaps for the purpose of hiding its activities in the noise created by these other famous worms.
2. The weird peak of activity happening around 3 pm does not correspond to any known behavioural pattern of these exploits. Furthermore, similar attacks observed on other platforms do not share this strange characteristic. Here to, this provides more rationales in favour of the existence of another tool, somehow disguised as a known worm.

The running of the last version of Snort over the associated traffic does not trigger particular alerts. Snort is a rule-based Intrusion Detection System, and its capacity for detecting malware activities consists in matching expert signatures to dump logs [21]. In other words, it means that the observed traffic does not match existing malware signatures. A lookup into web search engines on the collected data payloads is not more successful. Therefore, it is quite likely that we are observing the activities provoked by some malware, which at first glance behave like a known worm. The only differences between this new instance and the known ones become visible only when one can compare the attack profiles, as observed in a large number of viewpoints. This shows the interest for having a large number of similarly configured sensors.

A closer inspection at the peak described in Figure 3 (a) reveals that there is only one ports sequence responsible for this important activity between 3 pm and 4 pm. Figure 4 gives the cumulated activity on port sequence {135} for each hour observed during October 2004. Thus, it appears that another process, targeting the sole port 135, was propagating through

Taiwanese networks at the very same period than the other famous 'classical worms', such as Blaster or Nachi, to name only those two. This tool remains very local and shows very strong temporal patterns. The fact that it is coming from dozens of IPs cannot justify a misconfiguration problem in the network, even if the network implements very dynamic IP allocations. The problem has been detailed to the administrators who investigate the root causes of such a tool within their network.

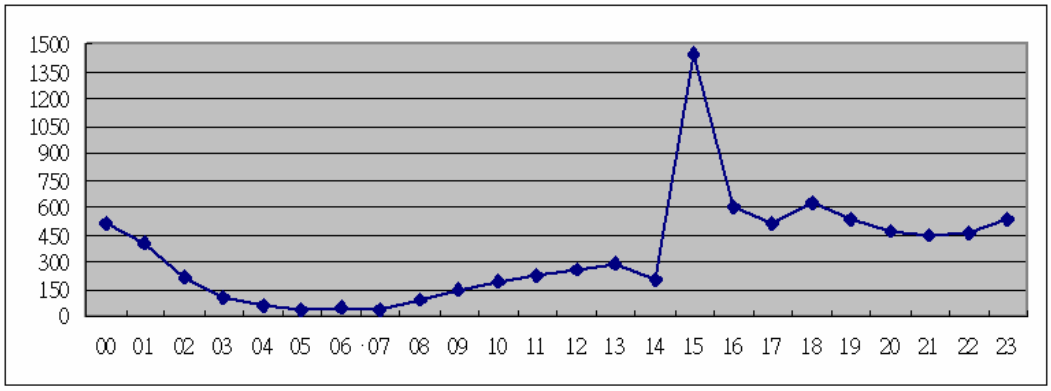


Figure 4. Distribution of Attacks on port sequence {135} over hours

4.4 Analysis of the Ports Sequence {8080, 3128, 1080, 1813, 80}

As another example of information that can be deduced from the comparison between attacks observed on several sensors, we can note the existence of a specific ports sequence made of the following ports: {8080, 3128, 1080, 1813, 80}. Surprisingly enough, that sequence has only be seen on the Taiwanese platform (see Figure 5). These ports are well-known ports; port 80, 3128, 8080 are related to http protocol, and ports 1080, 1813 to SOCKS 5. The very interesting fact here is that all machines that have issued that sequence of ports are located in Taiwan and have all been identified as Windows machines by p0f. Furthermore, all traces are similar in terms of amount of packet exchanged, average inter arrival time, amount of virtual machines they have talked to, payload, etc. There are more than six hundreds hosts that have tried this attack against that platform. They were coming from various networks. It is thus not the case that we are observing, for instance, one specific machine routinely searching for all the open relays on behalf of the network administrators, for good reasons. As a result, it is almost certain that we are seeing the effects of a single tool, which is used on several hundreds of machines but which specifically targets the Taiwanese environment. Due to the type of probed ports, one can speculate that we are seeing a tool, which is looking for open proxy servers in order to issue requests to external web servers in an anonymous way. It is still unclear why we do not observe such activities elsewhere, and what motivates this interest for Taiwan servers only. In any case, this is a topic under current investigation and shows another concrete case where the comparison between two similarly configured sensors led to some practical information that could be used to better understand the threats a given network was facing and to help administrators identifying and removing some of them.

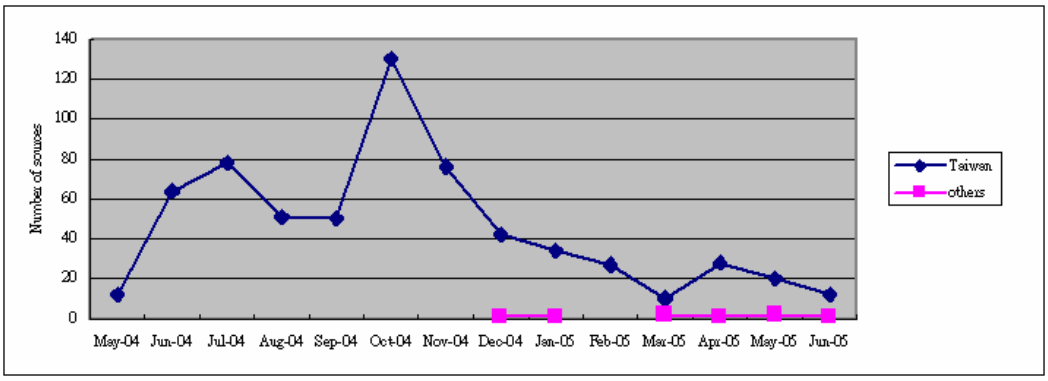


Figure 5. Number of sources, per month and per sensor, having probed the ports sequence {8080, 3128, 1080, 1813, 80}

5. Conclusions

In this paper, we have shown, based on a couple of simple, yet representative examples, the usefulness of deploying similarly configured sensors in various environments. The information gained by comparing the various attack profiles may lead to the discovery of new tools or new methods used against a given site only. As a consequence, it is clear that forensics investigation could greatly benefit from having access to such data sets.

The results presented have been obtained thanks to the Leurré.com project, which aims at deploying as many similar platforms all over the world. Participation to the project is free and we hope that the presentation of such findings may lead new partners to join this ongoing collaborative effort.

Concretely speaking, in the context of the Taiwanese platform, the analysis has shown that Taiwanese Sensor suffers from many specific attacks. As we know, in Taiwan, most of the academic network moderators are not professional technicians. Besides, the attributes of academic networks are different from the security networks for general administration institutes, and hence, no relevant regulations are set for academic networks. The Ministry of Education (MOE) of Taiwan has noticed this problem, and introduced Information Security Management System (ISMS) to deal with it. We believe that there would be some improvement with regard to these phenomena in the very near future. Hopefully, they will be visible on the sensor as well by removing all discrepancies specific to that platform.

6. References

[1] CAIDA, the Cooperative Association for Internet Data Analysis. http://www.caida.org/, 2005.

[2] D. Moore, G. Voelker, and S. Savage. Infering internet denial-of-service activity. In The USENIX Security Symposium, August 2001.

[3] D. Song, R. Malan, and R. Stone. A global snapshot of internet worm activity. Technical report. http://research.arbor.net/downloads/snapshot worm activity.pdf.

[4] B. Gemberling, C. Morrow. How to allow your customers to blackhole their own traffic. http://www.secsup.org/CustomerBlackhole/.

[5] Team Cymru: The Darknet Project. Internet: http://www.cymru.com/Darknet/, 2004.

[6] E. Cooke, M. Bailey, Z.M. Mao, D. Watson, F. Jahanian, and D. McPherson. Toward understanding distributed blackhole placement. In Proceedings of the Recent Advances of Intrusion Detection RAID'04, September 2004.

[7] The SANS Institute Internet Storm Center. The trusted source for computer security training, certification and research, http://isc.sans.org.

[8] V. Yegneswaran, P. Barford, and S. Jha. Global intrusion detection in the domino overlay system. In Proceedings of NDSS, San Diego, CA, 2004.

[9] DShield Distributed Intrusion Detection System, http://www.dshield.org.

[10] myNetWatchman. Network intrusion detection and reporting, http://www.mynetwatchman.com.

[11] Cisco Systems. Netflow Services and Applications, 1999.

[12] M. Dacier, F. Pouget, H. Debar, "Honeypots, a Practical Mean to Validate Malicious Fault Assumptions". In Proceedings of the 10th Pacific Ream Dependable Computing Conference (PRDC04), Tahiti, February 2004.

[13] F. Pouget, M. Dacier, V.H. Pham, Leurre.Com: On the Advantages of Deploying a Large Scale Distributed Honeypot Platform. In Proceedings of the E-Crime and Computer Conference 2005. (ECCE'05), Monaco, March 2005.

[14] E. Alata, M. Dacier, Y. Deswarte, M. Kaaniche, K. Kortchinsky, V. Nicomette, V.H. Pham, F. Pouget, "CADHo: Collection and Analysis of Data from Honeypots", To Appear in Proc. Of the Fifth European Dependable Computing Conference. (EDCC-5), Budapest, Hungary, April 2005.

[15] honeyd Homepage, http://honeyd.org/, 2004.

[16] F. Pouget, M. Dacier, H. Debar, V.H. Pham, "Honeynets: Foundations For the Development of Early Warning Information Systems", NATO Advanced Research Workshop, Gdansk 2004. In the Cyberspace Security and Defense: Research Issues. Publisher Springler-Verlag, LNCS, NATO ARW Series, 2005.

[17] p0f: Passive OS Fingerprinting Tool. http://lcamtuf.coredump.cx/p0f.shtml, 2004.

[18] F. Pouget, T. Holz, A Pointillist Approach for Comparing Honeypots.
To Appear in Proc. Of the Conference on Detection of Intrusions and Malware & Vulnerability Assessment. (DIMVA 2005), Vienna, Austria, July 2005.

[19] Stanford University, "Vulnerabilities, Patches and Exploits for Windows RPC/DCOM", 2003, available at: http://www.stanford.edu/services/securecomputing/rpc-vulns.html

[20] Stanford University, "Windows RPC Vulnerabilities & Exploits", 2003, available at: http://www.stanford.edu/services/securecomputing/win-rpc.html

[21] Snort, the de facto standard for intrusion detection/prevention. http://www.sort.org