2015 48th Hawaii International Conference on System Sciences

The Honey Community: Use of Combined Organizational Data for Community Protection

Keith Harrison, PhD
University of Texas at San Antonio
Keith.Harrison@utsa.edu

James R. Rutherford
Southwest Research Institute
James.Rutherford@swri.org

Gregory B. White, PhD
University of Texas at San Antonio
Greg.White@utsa.edu

Abstract

The United States has US CYBERCOM to protect the US Military Infrastructure and DHS to protect the nation's critical cyber infrastructure. These organizations deal with wide ranging issues at a national level. This leaves local and state governments to largely fend for themselves in the cyber frontier. This paper will focus on how to determine the threat to a community and what indications and warnings can lead us to suspect an attack is underway. To try and help answer these questions we utilized the concepts of Honeypots and Honeynets and extended them to a multi-organization concept within a geographic boundary to form a Honey Community. The initial phase of the research done in support of this paper was to create a fictitious community with various components to entice would-be attackers and determine if the use of multiple sectors in a community would aid in the determination of an attack.

1. Introduction

It is quite obvious that anything connected to the Internet faces a threat. The real question is what is the extent of the threat and what mechanisms are in place to prevent, detect, respond to and recover from the threat. The household PC faces the threat of indiscriminate attacks. The attackers of household PCs are not directly attacking the PC owner but rather it is by happenstance the particular attacker comes upon the victim. The household PC can be protected in many cases by commercial products which can address the indiscriminate threat. The organizational cyber asset faces not only the indiscriminate attacker, but also direct attackers looking to compromise the security of the organization. In many cases the organization has an IT department assessing the penetration attempts and providing a coordinated strategy for protecting the organization. We can use Honeypots to represent the individual and Honeynets to represent the organization. The traditional honey devices have been good at capturing the generic attacker looking to exploit known weaknesses within individual PCs or networks. You can have unexpected attackers attack the Honeypot and you may get a botnet to take aim at your system but how do you draw in the intelligent adversary with a very specific target or target set in mind? Intelligent adversaries do not just use Brute Force methods to penetrate the defenses; they can be very subtle and leave no obvious trace when examined with tools that use known attack signatures or methods.

A sector that has gone largely unnoticed in the efforts to protect the nation are the individual states and communities. In this case, the term community is used to refer to a geographic area as opposed to the sometimes used term "community of interest" which in this paper is referred to as a sector. At the same time the state and community sector has gone largely unnoticed, the impact of cyber events has increased. The best example of this is the attack in the fall of 2012 impacting the state of South Carolina. This attack targeted the systems of the South Carolina Department of Revenue and ultimately involved the loss of financial data of 6.4 million consumers and businesses. The financial loss to the state has exceeded $20 million. [1] This cost reflects not just the expense of securing breached systems and network forensic services, but $10 million to pay for the cost of credit monitoring for individuals affected by the breach desiring the service and $1.2 million to notify those whose information had been included in the data that was lost. States are not the only entity being targeted. Local governments are a target of attackers since they generally have financial accounts to handle city finances and often their security is lacking. An example of what can happen to a city occurred in October 2012 in the city of Burlington, Washington. In the incident, attackers stole over $450,000 from the city's general fund after gaining access to city payroll systems and grabbing passwords and other information they used to make three withdrawals from the city's account maintained by Bank of America. [2]

A frequent concern mentioned by government officials is that the nation's critical infrastructures are

1530-1605/15 $31.00 © 2015 IEEE
DOI 10.1109/HICSS.2015.274

IEEE computer society

targets of enemies (and possible future enemies) of the nation. Often cited is the fear that the nation's energy grid might be targeted by adversaries intent on harming the nation. While an attack on the grid might well be viewed as an attack on the nation, often the way into the grid might be through one of the many local utility companies found in communities across the nation. Interestingly enough, a survey of 219 security practitioners in energy utility companies conducted in 2011 found that 76% had experienced at least one data breach in the previous 12 month period. [3]

At the Blackhat Conference in 2013 Kyle Wilhot [4] presented a paper dealing with the compromise of water plants. He deployed 12 honeypots across 8 countries which attracted 74 intentional attacks with 10 labeled significant. He considered the 10 significant because the attacker had enough control of the facility to take control away from plant operators. The Jan- April 2014 ICS-CERT newsletter reported a public utility had its control system compromised. Nothing was done to the system overtly, but the fact the system was accessed and only after the fact was the intrusion noticed is concerning.

The city of Seattle runs a program called PRISEM[5] which collects IT audit information from all local government entities. The information is collected at a FUSION center to provide a central repository for detecting threats and attacks. The FBI provided the FUSION center with a pattern and addresses that belonged to the Chinese APT1 organization (A part of the Chinese Army tasked with cyber espionage.) The FUSION center detected intrusions at several universities and corporations but more than half of the intrusions occurred at the various port facilities in the Seattle area. It is currently unclear why there was such a significant effort to penetrate the port facilities. However, the implications of the penetrations could be far reaching from disruption of commerce, allowing the smuggling of materials into or out of the United States, or other economic espionage.

Believing that state and city governments are an important element in securing the nation's infrastructures, the Center for Infrastructure Assurance and Security (CIAS) has focused on methods to help states and cities better prepare to prevent, detect, respond to, and recover from cyber security incidents. As part of the center's efforts to help states and communities develop their own viable and sustainable cyber security programs, the center has promoted research in the area of community cyber security incident response. It is important to note that the goal of protecting the cyber assets of a community involves all sectors, both public and private, as an attack on any of the infrastructures will have a detrimental impact on the community overall. In addition, should one sector in a community discover an ongoing attack, there may be a good chance that another sector might also have experienced a similar attack, or may soon be the target of an attack. Due to the interconnected nature of many community services it may also be possible to access internal interfaces which may not be readily available by other means. An effort within the community to treat such incidents as a community incident will allow for a coordinated detection and response that could better protect the community as an entity. With this in mind, the question "how does a community know what an attack on the community looks like?" quickly arose.

2. Honey Devices

The desire to be able to identify and watch the activities of an intruder is not something that is new to the security community. Various methods have been used for more than twenty years to isolate, contain, and view intrusive activity. The first well-known example of this was conducted by Clifford Stoll in 1987 and is described in [6] and [7]. In this incident, which was first detected in 1986 and lasted for many months, Stoll discovered that a computer system at Lawrence Berkeley Laboratory, which he was the administrator for, had been penetrated. In an attempt to determine who the individual was, Stoll spent many months monitoring his system and observing the activities of the individual. He watched the individual not only access his system but others as well. He observed the individual searching information on specific topics by checking for key terms such as "nuclear" and "SDI". (SDI stood for Strategic Defense Initiative which was a US Military program aimed at developing a method to counter a potential nuclear attack from the Soviet Union). Stoll worked with local and federal law enforcement agencies in his efforts to track the attacker to his point of origin. Eventually this led to an international investigation when the trail led to Europe.

The problem that Stoll ran into was the length of time it took to conduct a trace of the connection. This became especially important when it was traced back to systems in Germany operated by the German Bundespost. Technicians in Germany could trace it to the city, but because of the electro-mechanical nature of the equipment then being utilized, a longer connection was needed to be able to trace it to a specific location. The individual, however, never stayed on long enough to complete the trace. Stoll decided if he were to create something on his system of enough interest, then the attacker might maintain the connection long enough to complete the trace. With this in mind, Stoll created the fictitious "SDINET" containing documents he created related to the SDI program. Since Stoll had seen the attacker searching for information related to SDI, he felt there was a good chance he might stay on long enough to view the documents which would allow officials to complete the trace. The plan worked and ultimately the individual was caught, tried, and convicted for his activities. The story is actually much more involved than this simple description and well worth reading more about.

The purpose of the fake network created by Stoll was to help with the identification and capture of the intruder. Another reason to create a fictitious system or network is to observe the techniques used by an intruder in order to gain access to a system or to elevate the level of permissions obtained. An early example of this is described by Bill Cheswick in [8]. In 1991, Cheswick spent several months watching an intruder in his system. In his own words, Cheswick said he did this in order to "trace his location and to learn his techniques." Cheswick went on to describe his motivation as follows:

> I knew there were barbarians out there. Who were they? Where did they attack from and how often? What security holes did they try? They weren't doing any damage to AT&T, merely fiddling with the door. The ultimate fun would be to lure a cracker into a situation where we log his sessions, learn a thing or two, and warn his subsequent targets. [8]

A difference between Cheswick's and Stoll's experience was that Stoll had an intruder in his system when he created his SDINET, where Cheswick created the fake environment knowing that it would eventually be discovered by individuals wanting to gain unauthorized access to systems and when it was, he was ready to watch and learn. Initially, Cheswick simulated a system in real time, reacting to the requests and actions of the intruder but in a way to not release any important files or information. Eventually Cheswick decided it was better to create a "sacrificial machine" from which the intruders activities could be watched. He considered actually setting up a separate machine but didn't have any spare system to use so had to take a software approach to create what he referred to as "the jail". His efforts proved very successful and he was able to watch one intruder in particular for several months. The report of his efforts and lessons learned, as related in [8] are also well worth the time to read.

Neither Stoll nor Cheswick used the term "honeypot" to describe their fictitious setups, but that is the term now generally used to refer to similar systems. A honeypot is a "fake" system deployed to attract potential intruders so that they can be monitored and their techniques observed. Since there is no legitimate reason for the honeypot system to be accessed, any attempt to connect to it should be viewed as an attempt by a potential intruder or malicious insider. The theory is that a honeypot can act as a decoy to divert attention away from production systems and at the same time potentially provide advance warning of an attack. Honeypots have become fairly popular since the days of Stoll and Cheswick and are used both on production networks within organizations as well as in academia and research. Generally, those in production environments are a bit less capable in terms of the amount of information they gather as their role in industry is more to deter and alert rather than to gather information on new methods or techniques. These systems may be virtual machines running on another system rather than separate systems themselves. They will often have limited services running on them and will require less interaction to accomplish their mission. Honeypots deployed in a research environment on the other hand are generally designed to collect a greater amount of data and may require considerably more interaction in accomplishing their purpose of gathering information on the tactics and techniques used by attackers.

The next step in the evolution of the honeypot concept was to place multiple honeypots on the same network thus forming what is referred to as a honeynet. In 1999, the Honeynet Project was created to bring researchers together to cooperatively analyze unusual and intrusive activity. This is accomplished through the use of a specific network designed to be compromised. It isn't a honeypot connected to a production network, but rather a network of systems carefully monitored so that all traffic is captured and logged in a manner that prevents it from being deleted when an intrusion occurs. [9]

The honeypot concept has been extended to other environments with the result that other devices now use the prefix "honey" in order to convey the intent to gather information on attack techniques. For example, Microsoft Research in 2005 coined the term "honeymonkey" to refer to a system that mimics the action of an individual surfing the Internet. The goal is to discover websites that use browser exploits to compromise systems and install malware on them. [10] A separate research effort created the "HoneyBOT" which is designed to open a large number of sockets on a computer system and mimics services that might utilize them. Should an attack attempt to upload a piece of malware on the system, it will be captured for analysis. [11]

In an effort to help develop tools and techniques that can be used by states and communities in their own security programs, the authors decided to extend the concept of honeypots and honeynets to a community. The goal was to gather information on attacks aimed at a fictitious community that would be designed and deployed to allow individuals to attack it. The concept is an extension of the honeynet idea where the simulated networks would all be networks that might reasonably be expected to be found within a community.

3. Implementation

Conceptually, the Honey Community is an extension of the honeynet concept. The goal is to develop an environment that will attract attackers to it and can then record their activities in order to analyze community-based attacks. Specifically, we want to attract attackers that are interested in penetrating and accessing systems within a community. This would not only be the normal hackers we might find attacking a home PC, but rather those specifically interested in attacking the systems and information used in providing services and information within a community. In order to accomplish this, there will be four main issues that will have to be addressed:

- The Honey Community has to look like what the attacker will expect to see. It must look like a real community. This would include services and capabilities as well as information flow.
- Not only will the Honey Community need to look like the websites/systems of a real community, once any of the pages/systems are accessed it must also act like a real community's web site. This is also true should any page/system be compromised, the system needs to act like a compromised host would act.
- Mechanisms need to be provided that will allow for the collection of data and monitoring of the attacker's activities.
- When the Honey Community is penetrated, mechanisms should be in place to prevent the systems from being used to attack another organization.

This last issue was addressed using another honey device, the honeywall. Just like a firewall, the honeywall is designed to monitor (and potentially filter) traffic that enters or exits the network. By doing this, potentially malicious traffic exiting the network can be blocked so that a compromised system that is part of the Honey Community can't be used to exploit other systems on the Internet. One important aspect of the honeywall is it needs to be invisible to systems on the Internet so potential attackers don't notice it to tip them off that their actions are being monitored.

The Honey Community created was called Roadrunner Park. A web site for Roadrunner Park discusses how it grew up around the university campus and many of the citizens are students or employees of the university. The name of the community in fact comes from the mascot of the university. Since the IP addresses used could be traced back to the university, it was decided that a strong link was needed to offer an explanation for the IP addresses – otherwise it could be easily assumed that this was not a real community.

The web sites for Roadrunner Park were patterned after a number of small actual communities within the state. A variety of sites were developed to see what interest there might be in the various functions found in a typical community. This led to the development of sites for the city government with links for the municipal court, city council, election information, information about public meetings, information regarding city ordinances, and the city's finance department, police department, fire and rescue, independent school district (complete with pictures of historical state figures the fictitious schools were named after), and the local chamber of commerce. Traffic citations could be paid online as well as an individual's water bill which allowed for the creation of an ecommerce-like site that would gather personal financial information. A screen capture for City Government page is shown in Figure 1.

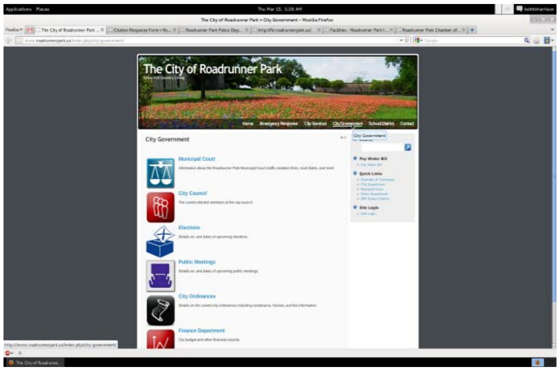

Figure 1. The City of Roadrunner Park

The deployed Honey Community hardware can easily support over 100 virtual machine images. Since the network was designed to be vulnerable, the systems were created to facilitate snapshots to provide an easy rollback mechanism along with remote management. A diagram depicting the organization of the Honey Community is shown in Figure 2.

The authors are happy to state that the Honey Community was very successful and that all hosts were compromised multiple times during the project. The amount of data gathered was considerable, amounting to 1.2GB of raw pcap data. An analysis of the data follows in the next section.

It should be noted that while considerable interesting information was gathered, the authors realize that some of the attacks may not be considered an attack on a community specifically, but rather an attack on a system that was simply available on the

Internet. While this may be true for any number of the attacks observed, the fact remains that had this been a real community, local leaders would have had to contend with it. This effort is considered a success as now local leaders can be shown that their systems are under constant attack and failure to act is not an option.

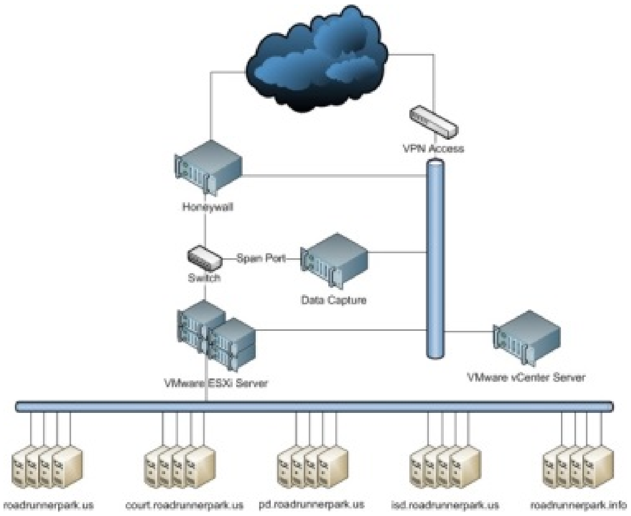

Figure 2. Logical diagram of the Honey Community

4. Objectives

The Honey Community must be useful to researchers or community governments. In order to be useful, the Honey Community should provide the ability to collect and analyze data in both real time and offline. The analysis of collected data must provide insights into threats faced by a community leading to improved threat awareness. Additionally, analysis of data from a Honey Community may help guide the development of improved policies and procedures, information sharing, and detection techniques.

In order to assess the usefulness of this Honey Community, three different analyses were performed. First, a commonly seen pattern of compromises is analyzed over a two day period. Secondly, an analysis of Brute Force attacks, occurring over the same two day period is discussed. Lastly, an aggregate data analysis, comprising three weeks of data is analyzed. It should be noted that the analyses discussed here are not a comprehensive analysis of all data gathered but rather only a subset to evaluate the Honey Community.

5. Compromise Case Study

For our initial observations and analysis, over 1.2 gigabytes of network traffic was gathered over a continuous one month time period. During this time period, the overwhelming majority of attacks consisted of automated Brute Force attacks. Of these attacks, we found that most password guessing attacks were run against all targets, either simultaneously or sequentially. However, vulnerability exploits usually only target one or two of the available targets.

Examining the Compromise timeline we can see a port scan of all five sectors from the same IP address. Then an attack of the School District and Criminal Justice sectors from the same address. The attack on the community began with the port scan, exploit on port 137, and DDoS. Section 8 of this paper will address the use of this information in developing a taxonomy of the attackers that can be used to help define means and motive of the attack.

Compromise Timeline

The following is a timeline of compromises occurring on March 15th and 16th:

March 15th, 1:00am – The Honey Community VMs were reset to a clean state.

March 15th, 9:39am – The Emergency Response Sector was compromised by a worm propagating via port 137. As soon as the machine was compromised, a rootkit was installed, and an IRC connection was established to irc.priv8net.com.

March 15th, 11:02am – A computer with the IP address 64.255.101.197 scans all 5 sectors.

March 15th, 11:43am – The same IP address 64.255.101.197 returns to compromise the School District and Criminal Justice sector. These machines were compromised using a different exploit (and rootkit) on port 137.

March 15th, 11:48am – The School District and Criminal Justice sectors begin a DDoS attack against www.yahoo.com. The packets for this attack were intercepted by the Honeywall and prevented from reaching the intended target.

March 15th, 11:48am – The School District and Criminal Justice sectors begin a DDoS attack against www.icq.com. Again, this attack was thwarted by the interception of the DDoS packets by the honeywall.

March 16th, 9:18am – The Emergency Response sector was compromised by a worm propagating via port 137. As soon as the machine was compromised, a rootkit was installed, and an IRC connection was established to irc.priv8net.com. This is the same exploit and rootkit seen on March 15th at 9:39am. However, the attackers IP address is different.

Compromise Analysis

Analysis of the compromises led to several interesting observations concerning networks in a community and attacks against them.

First, reimaging machines daily has very little effectiveness against most attacks. Our Honey Community was compromised on a daily basis, despite being reset to a clean state every morning. This is a common security measure that is thought to improve security when applied as part of a layered defense strategy. For example, schools, hotels, and libraries are just a few examples of organizations that automatically reimage machines on a daily basis. However, in practice this may provide a false sense of security leading to a less secure environment.

Three of the sectors were not compromised even though they were previously scanned by the same source that compromised two of the sectors. While the attackers reasons are unknown, we can speculate that the attacker may be attempting to keep a low profile to avoid detection. However, regardless of the attacker's motivations, it is clear a single Honeynet only gives defenders part of the picture, exemplifying the need for a Honey Community.

6. Brute Force Case Study

The same 1.2 gigabyte dataset was used for observations and analysis of automated Brute Force password guessing attacks. In this section we look at March 16th to illustrate the Honey Community under attack from multiple types and sources simultaneously.

Brute Force Timeline

Beginning at 4:36am on March 16th, a single source IP address began a Brute Force dictionary attack against MySQL directed at all 5 community sectors. The attacks began with a source port of 6000 for the first round of attacks and continued with similar attacks using gradually increasing port numbers.

Another, more interesting attack, began earlier that morning. At 12:32am, a single port scan was observed originating from a single source IP address with a source port of 6000 targeting 4 of the 5 community sectors. This was determined to be the reconnaissance phase of a larger attack, since only a single connection was made, and the same source IP address was never seen again. At 1:38am the Brute Force phase of the attack begins. The Brute Force attack was a coordinated, distributed attack originating from multiple IP addresses. The first 9 source IP addresses, their country of origin, their targets, and duration of attacks are shown in Table 1.

Table 1. Brute Force attack pattern

Start Time	End Time	Source IP (Country)	Target Sector
1:38am	2:12am	116.236.150.98 (China)	Emergency Response
2:16am	2:30am	117.239.124.194 (India)	Emergency Response
3.00am	3:30am	186.212.14.97 (Brazil)	Criminal Justice
4:45am	5:08am	188.194.170.39 (Germany)	Criminal Justice
6:07am	6:23am	46.225.126.106 (Iran)	Emergency Response
6:10am	6:23am	189.25.200.18 (Brazil)	Water/Sewer Utilities
7:00am	7:30am	114.97.69.235 (China)	Water/Sewer Utilities
8:02am	8:32am	94.179.199.94 (Ukraine)	Emergency Response
8:50am	9:22am	123.127.157.105 (China)	Commerce/ Industry

At 11:13am a similar pattern is seen again. This time the attacks are coming from different sources, but once again, it begins with a connection scan against all targets with a source port of 6000.

Attacks from multiple sources were easily identifiable as being part of a single Brute Force attack by looking for common elements and patterns in the attack type, source and destination port, source and destination IP addresses, and timing. The above Brute Force attacks were correlated as being the same attack by seeing a port scan followed by numerous IP addresses attacking various systems throughout the community using the same method and ports.

Brute Force Analysis

Here we see two different types of attacks against different services, both beginning with a source port of 6000. This is due to the sharing and reuse of common components among multiple pieces of malware. Defenders can exploit these similarities and use them to their advantage. In this case, a source port of 6000 is analogous to a burglar rattling the door knob to a home. If communities know what to listen for, they will be able to more easily recognize an attack before, or as soon as, it begins.

7. Aggregate Data

As part of a related research track, we are developing distributed and real time information sharing and analysis algorithms. These algorithms allow us to better identify attacks on a community by looking for significant increases in information found within intrusion detection alerts. We believe the specifics of the distributed algorithms used are outside the scope of this paper. However, an examination of the end result is a statistically significant increase on all combinations of predefined attribute-value pairs within intrusion detection alerts. For the rest of this section we refer to these significant increases as attacks, because that is what they best represent. However, more than one of these "attacks" may actually correspond to a single real world attack.

For this analysis, new data was gathered over a three week time period. The pcap data was run through snort, generating 229,529 IDS alerts. This equals about 7.95 IDS alerts per minute for the entire Honey Community on average. The highest number of IDS alerts received during a one minute time period was 654.

From these IDS alerts, we identified 3,060 attacks on the community during the three week time period. Each attack may be against one or more sectors, and/or the community as a whole:

Table 2. Sectors per attack

Number of Sectors	Identified Attacks
*	1,402
1	1,430
2	151
3	52
4	16
5	9

Table 2 depicts the number of community sectors under attack for each of our 3,060 identified attacks. " * " sectors indicate that although no single sector could be identified as being under attack, we could identify the entire community as being under attack. This differs from the case where we could individually detect all 5 sectors as being under attack.

To further illustrate this fact, imagine that a particular sector sees a small increase in a particular type of scan. On its own, this small increase might seem insignificant; however, when this data is gathered and analyzed across sectors, it may become evident that something significant is occurring across the community.

Table 3 depicts the number of attacks detected against systems of each sector and the systems spread across the community. A large number of attacks were identified as being spread against the systems in the community as a whole. This spreading of attacks across a community is of particular interest. First, if the community is a target of opportunity, then examining the vulnerability being exploited and passing the information to other sectors can help detect the attacks as well as protect vulnerable systems. The second area of interest is, if the community is the target of attacks that are directed with multiple attempts to penetrate the community. This may be a concerted effort to damage the community and could be a precursor to either more direct cyber or physical events.

Table 3. Attacks per sector

Sector	Identified Attacks
Community	2,319
Water and Sewer	369
Criminal Justice	345
Emergency Response	398
Education	381
Commerce	504

Aggregate Data Analysis

Table 2, one can see that about half (55 %) of the 3,060 attacks would be detected as being significant by at least one individual sector. However, 45 percent of the attacks (1,402/3,060) are only seen as significant when considering the data from all 5 sectors. This illustrates the need for information sharing, not only within sectors, but across sectors as well.

What is interesting to note is, these figures assume that all sectors of a community are conducting roughly the same level of intrusion detection activity. It is the experience of the authors that this is seldom the case in actual communities. What this might mean is for those attacks that targeted more than one sector, not all sectors might actually detect the attack thus increasing the benefit of information sharing across all sectors within a community. Therefore, as shown in Figure 3, in 55 percent of the attacks, the community would benefit from an active information sharing program.

This view of the data only considers simultaneous attacks on multiple sectors. We have also found that attacks which affect only a single sector at a time commonly move from sector to sector. This means that communities also benefit from sharing all information on attacks, even if only a single sector has been affected. This finding is confirmed by the Seattle incident where information from an external attack was used and multiple organizations were able to recognize previously unknown attacks and benefit.

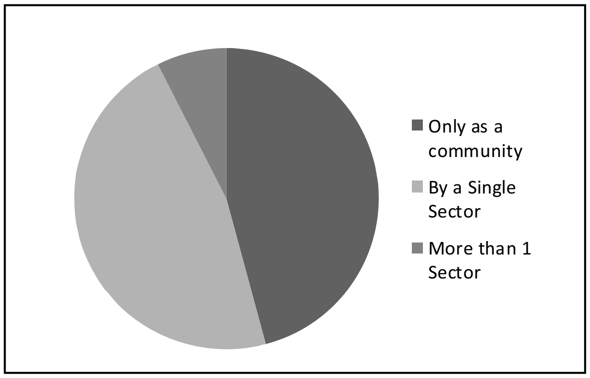

Figure 3. Detection locality

Additionally, from Figure 3, it is clear that the majority of attacks target either the community as a whole or a single sector, at any specific time. Based on this data, and our previous analysis, we have identified two main types of Brute Force attacks leading to this phenomenon:

Shorter, high intensity attacks are generally directed against one sector at a time, and finish attacking one sector before moving on to the next. In this case, information sharing across sectors is not necessary to detect these attacks. However, since other sectors are likely to be attacked and due to differing IT strategies amongst organizations, information sharing could significantly benefit other sectors within the community.

Lower and slower attacks are usually directed against multiple sectors at a single time. In this case, information sharing across sectors can aid greatly in determining a significant attack is occurring.

In Table 3 we can see that the community identified a much greater number of attacks than any individual sector. The total number of identified attacks in Table 3 is actually greater than the 3,060 attacks identified in Table 2. This is because each identified attack in Table 2 may correspond to multiple sectors and possibly the community as a whole.

It is interesting to note, the commerce sector was attacked almost half again as much as the others. We are unable to determine the exact reason for this difference, but it may be due to the fact that the commerce sector was the only sector with a top level domain of .info.

Regardless, the Honey Community was able to detect an abnormal amount of interest in not only a particular sector, but also in the community as a whole. If the Honey Community concept was expanded, matured, and deployed into real communities it could provide defenders a much broader insight into the threats faced by their community.

8. Examination of the Taxonomy of an Attack on a Community

Looking at the Taxonomy of the attacks on the community with respect to the paper by Harrison and White [12] it can be seen that combining information can help define not only the attackers' methods and techniques, but the objectives of the attacks. The taxonomy first looks at the Event Vector which consists of the agent, motivation, objective, method, and technique. If we just examine a single instance of an attack we will get a basic idea of the technique and method of attack, but by correlating the data from multiple organizations within a community we can more precisely determine the methods and technique being employed as well as getting a clearer idea of the objective by looking at the systems within the community that are being brought under attack. Looking at the case in Seattle, we can see the attack vector was isolated by the FBI and provided to the Regional IT office which then was able to determine that several universities were hit and 6 different port facilities were penetrated. While the end objective is not currently clear it can be seen that the attackers had a specific interest in the Seattle ports. It would be interesting to take the known attack vector for the Seattle Ports (Chinese APT-1, Unknown Motivation, Port Facilities, FBI defined method and technique) and apply that to other US/European Ports to see if they have experienced the same Event Vector. By using the information, we can endeavor to determine the motivation as well as the second half of the Taxonomy which is the Effect Vector.

The Effect Vector is composed of the Cause, Services affected, Disruption impact, and Impact metrics. Again, looking at the Seattle Ports, we can see a direct Cyber Disruption within the Port Facility. The services directly affected would be shipping, but indirectly it could affect all of the merchandise and cargo that flows, not only through those facilities, but every facility that Port works with. The cargo could be placed on the wrong ships, the ships could be misdirected, on-load and off-load schedules could be tampered with to cause confusion. This would result in impacts that could run through all three impact areas of economic, population, and government. Paralyzing of the ports could cause financial losses and thus economic impact. The population could be affected by the loss of goods and materials, some of which may be vital such as coal and oil. Finally, the government would suffer both a loss in economic terms and reputation in trying to recover from this cyber event.

Bringing this back to the Honey Community, by bringing together a wider variety of information from the respective sectors we can see an Event Vector with the attacker using similar methods and techniques to hit multiple parts of the community. From these multiple attacks and an understanding of their methods, we can start to see their objectives. Are they only hitting one organization or are they hitting multiple organizations such as attacking the Police Department and the Courts of Law? Depending on the parts of the objectives they are attacking we can work to draw a motivation and possibly an agent.

9. Conclusion

In this paper, we first introduced the issues facing Communities and the cyber threats they are facing. A new Honey Community framework, as a logical extension of honeynets, was introduced as a way to help communities collect and disseminate information among the various community organizations. Secondly, we analyzed a typical compromise timeline from our Honey Community collection. Next, we analyzed a commonly seen Brute Force password guessing attack. Lastly, we presented and analyzed aggregate results by looking at the number of community sectors attacked simultaneously and the number of attacks per sector.

It can be seen that by aggregating the information between organizations with similar IT policies and strategies, we can improve our detection rates. It can be surmised that by sharing attack and probe information amongst organizations with dissimilar IT strategies, we can significantly improve the overall preparedness of the community. Also, due to the interconnected and related nature of community functions, by sharing information between the various organizations, the weakest links can be strengthened and improve the overall security posture.

Investigating successful attacks on the Honey Community led to several interesting conclusions. First, compromises occurred very frequently. In fact, solutions that rely on resetting computers to a clean, uncompromised state provide very little security by themselves and should only be used as part of a layered defense strategy. Secondly, a single honeynet is not sufficient to detect many attacks. In one example, all 5 community sectors were scanned, but only 3 were attacked, even though all 5 were vulnerable.

We identified two completely different types of Brute Force attacks. Interestingly enough, both began with a source port of 6000. One attacked all 4 community sectors simultaneously, while the other attacked all 5 sectors at different times. By investigating patterns, such as the source port of 6000, or the timing of attacks, better detection techniques may be developed. In some cases it is possible to detect the impending Brute Force attack during the reconnaissance phase of the attack.

When looking at the aggregate data collected over a time period of three weeks, almost half of the attacks are only identifiable as significant when looking at the community as a whole, underscoring the need for community wide information sharing across sectors. Additionally, the commerce sector was found to be attracting more attacks than the other sectors, providing useful insight into the security posture of the community our Honey Community is modeled against.

Just as honeynets are comprised of multiple honeypots, a Honey Community is conceptually made of multiple honeynets. Just as with honeynets and honeypots, any network activity is suspicious and care must be taken to ensure that compromised machines do not engage in outgoing malicious activity. We monitored and protected the Honey Community by using a honeywall. By using virtualization we are able to support over 100 VMs on a single server. Furthermore, virtualization features, such as snapshots and automatic rollback, made the time cost of administration very low.

10. Future Work

The Honey Community deployed in this project simulated websites that could be found in a typical small community almost anywhere in the nation. A problem with this approach is it is easy to verify the validity/existence of a community. Since the community only exists in the Internet domain, the attacker can determine it is not real and they are being monitored and researched.

In this paper the authors have shown that combining the information from different organizations can help defend and respond to attacks. The future work of the Honey Community should be aimed at providing civic and business leaders with the information they need to understand the need for cooperation and data sharing. This can be done through the expansion of the Honey Community in several areas:

1. Work with leaders in select communities to embed Honeypots within local organizations
2. Expand the capabilities of the Honey Community to include a more realistic environment as well as making honeypots more active within a network.
3. Building of an intelligent process to correlate and act on data from the Honey Community in order to make the defense of the community be a more proactive rather than reactive process.

The next phase will expand the Honey Community to deploy Honeypots into a distributed honeynet within a real community. It provides three benefits:

1. Honeypots configured within the IT policies of real organizations will provide a realistic understanding of the vulnerability of the various organizations.
2. It will be more difficult for attackers to determine they are attacking a network defense device. Since the Honeypots would be within the community we would collect information on actual attacks on the real community.
3. Advanced attackers, such as the Chinese APT1, are being called out for attacking targets and they will become more selective in their target selection. Placing the Honey Community within real community organizations will place it in the path of these advanced cyber threats making it more likely we will detect their activities.

This stage of deploying sensors within a real community is a large step in terms of the complexity of the Honey Community. This is true from the technical standpoint, as well as a political standpoint. Deploying honeypots and honeynets throughout a community, coordinating their activities, and having them communicate their observations is much more difficult task than deploying multiple systems within the same network. This, however, is doable and is much less of a challenge then the political element. Convincing organizations throughout a community that they should allow a system to be attached to their network, even if it resides outside of their firewall, is not a trivial matter. A certain amount of trust must exist between the security professional and the organization for this to take place. In addition, individuals need to understand the benefit to their organization to take this step. This also involves the issue of information sharing as organizations need to be comfortable with the fact that the information that they share at some level may be transmitted to other organizations. There are a number of approaches to accomplishing this in a manner that ensures anonymity and they will need to be explored and explained to organizations that would be included in the next stage of the Honey Community.

11. REFERENCES

[1] Shain, A., SC working on security, notifying victims of data breach., *The State,* January 6, 2013, www.thestate.com/2013/01/06/2578924/the-latest-on-sc-hacking-costs.html.

[2] Langeler, J., Hackers steal more than $450,000 from City of Burlington., *King 5 news*, October 13, 2012, http://www.king5.com/news/Hackers-steal-more-than-450000-from-City-of-Burlington-174045801.html.

[3] Schwartz, M. 76% of Energy Utilities Breached in Past Year, *Informationweek.com*, April 6, 2011, www.informationweek.com/security/attacks/76-of-energy-utilities-breached-in-past/229401071

[4] Wilhoit, Kyle, The SCADA That Didn't Cry Wolf- Who's Really Attacking Your ICS Devices- Part Deux! *Blackhat Conference 2013*, July 27-August 1, Las Vegas, NV.

[5] Higgins, Kelly, Hacking The Threat Intelligence-Sharing Model, DarkReading.com, September 25,,2013, http://www.darkreading.com/vulnerabilities---threats/hacking-the-threat-intelligence-sharing-model/d/d-id/1140553?

[6] Stoll, C. The Cuckoo's Egg: Tracking a Spy through the Maze of Computer Espionage, Pocket Books, New York, 1990.

[7] Stoll, C. Stalking the Wily Hacker. Communication of the ACM, vol 31, No.5, May 1988, 484-500.

[8] Cheswick, W., An Evening with Berferd in Which a Cracker is Lured, Endured, and Studied. *Proceedings of the Usenix Winter 92 Conference*, January 1992, http://web.cheswick.com/ches/papers/berferd.pdf

[9] The Honeynet Project, *Know Your Enemy: Revealing the Security Tools, Tactics, and Motives of the Blackhat Community*, Addison-Wesley, Boston, 2002.

[10] Microsoft Research, Strider HoneyMonkey Exploit Detection, January 20, 2010,

[11] Atomic Software Solutions, HoneyBOT, http://www.atomicsoftwaresolutions.com/honeybot.php

[12]Harrison, Keith, and Gregory White. "A Taxonomy of Cyber Events Affecting Communities." *System Sciences (HICSS), 2011 44th Hawaii International Conference on*. IEEE, 2011.