

KNOW YOUR NETWORK

Join the Discussion
Connect

How Understanding the Traffic Transiting Your Network Can Improve Your Information Security Posture

By Josh Goldfarb – ISSA member, Baltimore, USA Chapter

This article describes a methodology for network traffic analysis and provides some practical insights and techniques that can be used to monitor a network.

Abstract

This article describes practical techniques for the cyber security professional to efficiently sift through the voluminous amounts of network data. These techniques leverage different views of the data to discern between patterns of normal and abnormal behavior and provide tangible jumping off points for deeper investigation.

Operational status quo

Donald Rumsfeld once remarked, “there are also unknown unknowns – the ones we don’t know we don’t know.”¹ Although Mr. Rumsfeld was not talking about securing a network, the point is still a valid one. Nowadays, most organizations have perimeter defenses and some form of network monitoring. Unfortunately, many organizations look for traffic on their network using only static, signature-based alerts. For example, many organizations run Intrusion Detection Systems (IDS) or Intrusion Prevention Systems (IPS). Unfortunately, IDS/IPS only looks for traffic it knows, or is told (by us) to look for. In other words, organizations are looking for known threats that might be on their network. This is extremely important and must be done, but how can an organization find the threats to their network that they don’t know they don’t know about? Network traffic analysis is one way of getting a handle on this question. In

this article, I will describe a methodology for network traffic analysis, as well as provide some practical insights and techniques that can be used to monitor a network.

What is network traffic analysis and why is it important?

I consider network traffic analysis to be a toolbox of methods that together can be used to understand the traffic transiting a network. The approach I use is similar to the Unix philosophy: a number of simple tools that together can be used to accomplish some very powerful things. In the Unix world, one can pipe together two or more simple commands to produce a more sophisticated output. For example, by piping `ls -l` to `wc -l [ls -l | wc -l]`, one can very easily count the number of files in a directory. This is a simple example of how combining two simple tools can create a more powerful tool that can be used in scripts and elsewhere. Similarly, in the network traffic analysis world, one can pipe together two or more simple queries to produce a more sophisticated query that is more likely to yield analytically actionable results. For example, instead of searching network flow data for just an IP address, one can search that same network flow data for an IP address communicating on unusual/unexpected protocols, which will yield more interesting and actionable results.

Since most networks are extremely complicated, I essentially treat the entire network as a black box. I seek to interrogate the data through queries specially crafted to exploit certain

¹ <http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636>.

nuances of the data. I have aimed to illustrate this point through specific examples provided later in this article. The point of this approach is to look at the black box that is the network from many angles (which I call jumping off points) to ascertain the best picture of what is truly going on inside the box. Jumping off points provide tangible starting points for network monitoring professionals to latch onto suspicious activity and analyze it until ground truth is reached. These tangible starting points make all the difference in my experience. With this approach, an interesting phenomenon occurs. As time moves on, one begins to know the network through an iterative process of learning and interacting with the data. I've seen this occur first hand in multiple different Security Operations Center (SOC) settings. Why is this important? Because once you understand what belongs on your network, you can begin to look for the opposite of that behavior.

Some analytical concepts

Naturally, knowing what is normal for your network is easier said than done. Getting to a point where you are comfortable identifying what is normal and what is anomalous or abnormal can take some time and is an iterative learning process. I have been fortunate enough to have spent over a decade analyzing traffic transiting real live networks. My experience has taught me to identify useful methods for analyzing vast quantities of network traffic data, and I am excited to share some insights with the community. I will provide a few tangible, illustrative examples of jumping off points here to get us started down the path. In my experience, an organized, well-structured approach to network traffic analysis goes a long way towards tackling what is a very difficult challenge. Finding a way to make sense of the vast quantities of network traffic data collected on operational networks is a key component of a successful network monitoring program.

Example analytical techniques discussed in this article include:

- Monitoring unused network real estate – the darknet
- Computing ratios of outbound to inbound traffic – data exfiltration
- Watching for packets not obeying defined protocol standards
- Checking for periodic traffic – beaconing
- Using aggregation to your analytical advantage
- Trending over the long term
- Looking for the uncommon

Power of the darknet

One of the best ways to gain intelligence about reconnaissance being conducted against your network is by examining the traffic headed to your *darknet*. Team Cymru defines a darknet as “a portion of routed, allocated IP space in which

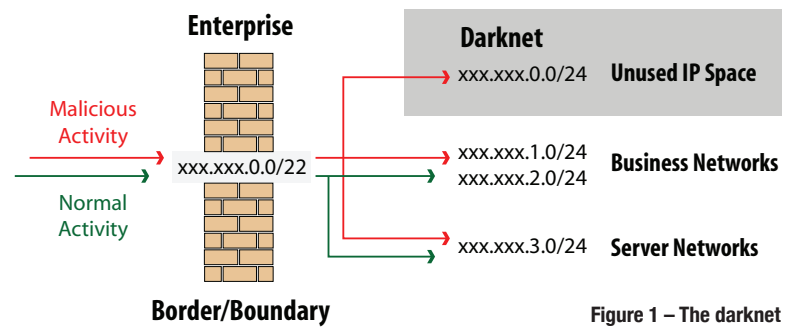


Figure 1 – The darknet

no active services or servers reside.”² If you have IP space that you own but do not use, you have what you need to get started with darknet analysis. Since all traffic destined for the darknet is suspect, it makes trending much simpler – no legitimate traffic to cloud the analysis of the suspect traffic (see Figure 1). For example, running statistics against the top source addresses, source and destination ports, protocols, etc., can help you pick up on an emerging threat before it is widespread.

When I was at US-CERT, we noticed a huge uptick in port 2967 traffic inbound to the darknet we were monitoring. At the time, this puzzled us, as it was extremely anomalous. What we later learned was that there was an unpatched vulnerability in Symantec’s anti-virus software. Virus writers were apparently probing networks, conducting reconnaissance. The virus writers were doing this as research/information gathering for the virus they were about to release named *Big Yellow*. Two weeks later, *Big Yellow* was big news.³

What types of new threats against your network might you be able to identify by studying your darknet? Here are some simple analyses you might find useful when monitoring your darknet data:

- Top source IP addresses of traffic to the darknet
- Top targeted (destination) ports of traffic to the darknet
- Top protocols of traffic to the darknet
- Top byte counts of traffic to the darknet (in other words – is there a threat that uses multiple ports but perhaps maintains a consistent byte size)

Outbound/inbound ratios

The concept of using ratios of outbound traffic to inbound traffic to detect malicious traffic is not a new one. Peiter

2 <http://www.team-cymru.org/Services/darknets.html>.

3 <http://www.scmagazineus.com/big-yellow-worm-avoids-microsoft-applications-targets-symantec-products-says-eyeye/article/34269>; <http://www.networkworld.com/news/2006/121506-big-yellow.html>.

Once you understand what belongs on your network, you can begin to look for the opposite of that behavior.

“Mudge” Zatko first described it in a Usenix *login* article in 2003.⁴ It’s a good analytical technique and one that fits well within the jumping off points methodology for network traffic analysis. The basic concept is fairly straightforward: examine outbound traffic from desktops (not servers) on your network heading out to the Internet (overlook traffic from internal hosts to other internal hosts for the purposes of this method).

The expected behavior for a desktop communicating with the Internet would be some data going out (e.g., an HTTP GET request), but much more data coming back in (e.g., a web page download following the HTTP GET request). If one or more desktops is observed sending far more data out than it is receiving back, then it would be indicative of data leaving the network. This is one way to discover exfiltration occurring on your network. For example, say Sally works in marketing and one Thursday morning at 3:00 AM her desktop sends 2GB of data to a host on the Internet. Furthermore, say that data was transferred as encrypted data over a protocol that is clear text (e.g. HTTP). This would probably be considered to be quite anomalous and would be something that would warrant further investigation. Monitoring the ratio of outbound to inbound traffic would allow the analyst to spot the anomaly and jump off to investigate the suspicious traffic.

Packets not conforming to standards

Fortunately for network traffic analysis professionals, TCP/IP conforms to well-defined standards. That gives us another great jumping off point – looking for packets that do not conform to standards. For example, the minimum number of bytes for a TCP packet is 48 (20 for the IP header and 28 for the TCP header). Similarly, the minimum number of bytes for a UDP packet is 40 (20 for the IP header and 20 for the UDP header). What if one was to monitor for TCP or UDP packets less than 48 bytes or 40 bytes, respectively? Those packets would not be expected on a network and, if they were observed, would be indicative of something suspicious. All legitimate traffic should conform to standards. Traffic that does not would cause the inquisitive analyst to investigate further and seek to determine the true nature of the non-conforming traffic. For example, I have seen non-standard UDP packets on the order of 24 bytes (sometimes a little more, sometimes a little less) on a few different networks. I have never been able to fully identify the cause of that traffic, and it has always bothered me. I am very curious to know what other organizations have in the way of traffic that does not conform to defined standards. I think we could all learn from each other here.

Beaconing

Human beings tend to be very irregular in their activity. In other words, when I’m reading articles on a news website, I probably will not click on a new article exactly every 120 seconds. If I did, I would be behaving more like a machine, e.g.,

a botnet. By looking for highly periodic traffic on a network (called a beacon), one can identify network traffic that is indicative of a machine source rather than a human source. For example, if an IP address inside an enterprise network sent an HTTP GET request or a DNS request to a location outside the network exactly every five minutes, that would be indicative of traffic initiated by a machine (and more specifically, could be caused by malicious code). Traffic coming from a machine source would be suspicious and would be something an analyst would want to investigate further.

Aggregation

Aggregation can be a powerful analytical tool for creating jumping off points. For example, consider examining network flow data by this aggregate:

```
source port | destination port | number of bytes
| count (sorted in descending order)
```

Results near the top of this list will show network transactions in which the same source port and destination port were used together with the same number of bytes. Theoretically, source (client) port selection should be somewhat random and uniformly distributed when examined over a large number of transactions. A large number of identical source port selections, particularly when coupled with the same byte counts, would indicate an anomaly that should be investigated further. For example, suppose an IP address inside an enterprise network was using the same source port (e.g. port 4444) to communicate repeatedly with a web server outside the network so much so that it bubbled up to the top row of our aggregation. Furthermore, suppose that IP address inside the network was also repeatedly sending the same number of bytes (e.g., 237 bytes) to that web server. That would be highly anomalous, as a typical human user would send and receive packets of different byte counts, as well as use a somewhat randomly chosen source port. In this example, the top row of our aggregation would look like this:

```
4444 | 80 | 237 | 10,000,000
```

Having 10 million network transactions with those exact parameters (source port = 4444, destination port = 80, and number of bytes = 237) would be highly anomalous. In this example, a jumping off point is created by using aggregation as another way to view the black box that is the network.

Long-term trending

When done wisely, long-term trending can help identify anomalies occurring on a network. For example, suppose one trends the volume of ICMP traffic hourly. Suppose one Wednesday morning at 2:00 AM, the volume of ICMP traffic jumps 100x over the expected volume for the 2:00 AM hour and then returns to normal levels. This would be indicative of something unusual that should be examined more closely and serves as another jumping off point. The key to trending, as mentioned above, is that it must be done wisely. What does this mean? In part, it means that you need to make sure you are comparing apples to apples. In other words, in the

⁴ <http://www.usenix.org/publications/login/2003-12/pdfs/mudge.pdf>.

accelerated learning • total immersion • immediate results • locations nationwide



All industries need a trained and effective IT workforce to combat hackers, attackers and security threats.

Maintain the integrity of your organization's communications, infrastructure and operations with certifications in Information Security.

As an authorized training partner of (ISC)², Training Camp is the only approved accelerated training provider for CISSP certification on a global scale.

Through our Certified Ethical Hacker and Forensics programs, Training Camp equips information security professionals with the knowledge to identify and correct exploits that make information systems vulnerable to attack.



IT & Management Training
Corporate & Government
Solutions Available

www.trainingcamp.com
CALL 800.698.5501



CompTIA A+
CompTIA Network+
CompTIA Security+
CompTIA Linux+
CompTIA Project+



We have successfully certified over 25,000 industry professionals in the last 10 years and look forward to adding you and your organization to our list of success stories.

above example, the volume of ICMP traffic in the 2:00 AM hour is only interesting when compared with the volume of ICMP traffic normally seen in the 2:00 AM hour. Here are a few examples of trending that compare data relative to other like data, which helps yield more interesting and actionable jumping off points:

- Volume of traffic by protocol for a time period (e.g., hourly, daily) compared to the expected volume for that time period
- Volume of traffic by port for a time period (e.g., hourly, daily) compared to the expected volume for that time period
- Volume of traffic by sensor for a time period (e.g., hourly, daily) compared to the expected volume for that time period
- Daily volume of traffic by port compared to the daily average volume (30-day moving average) for that port

Uncommon protocols

Most organizations need to route a handful of protocols for business purposes (e.g., TCP, UDP, ICMP, GRE, VPN, etc.). IANA defines 256 protocols (numbered 0-255), most of which are not typically used by an organization in the course of its daily business operations. By looking for protocols that are uncommon, one can gain insight (and a jumping off point) into an unusual protocol that may appear on the network. The analyst can then investigate that traffic to understand whether the protocol is legitimate and necessary for business purposes, or whether the protocol is being used for a nefarious purpose. For example, IANA numbers TCP as protocol 6, which is a protocol organizations would obviously route large volumes of for business purposes. Protocol numbers 143-252 are unassigned by IANA. Perhaps you have a proprietary protocol inside your organization that you route on

one of those protocol numbers for business purposes. If not, and you observe traffic on one of those protocols, it would cue you to some traffic that should be investigated further. As an aside, it is an interesting exercise to try and determine (based on the network traffic data) how many protocols you actually need to route for business purposes. Protocols that are not necessary for business purposes probably should not be routed. Unfortunately, in many organizations, that is not the reality. I am often surprised at how many protocols are actually being routed inside organizations, most of the time unbeknownst to the organization.

Conclusion

Why guess at what your network is actually doing? Analyzing the traffic transiting your network can give you a firm and scientific understanding of what your network is up to. The number of jumping off points one can employ to move towards knowing one's network is seemingly limitless. My hope is that this article has provided some useful starting points and will pique interest and raise awareness in the community. Through professional dialogue, we can move the state of analytics forward and increase our collective information security posture. I often post thoughts on analytical methodologies and jumping off points on my blog (see bio below). Please have a look and add to the dialogue!

About the Author

Josh Goldfarb is Principal Security Analyst at 21st Century Technologies and focuses on helping organizations build and enhance their network monitoring programs to improve their information security posture. Read his blog at <http://ananalyticalapproach.blogspot.com> or reach him directly at jgoldfarb@21technologies.com.



Connect Learn Advance...Join Today!



ISSA
Information Systems Security Association

For less than \$10 a month become an ISSA Member and take your career to the next level through:

Local Chapter Meetings • Face-to-Face Networking

The ISSA Journal • ISSA Web Conferences

Trusted Online Member Community

Discounts to Industry Conferences • Certification Study Courses

Continuing Professional Education (CPE) Credits

The Preeminent Trusted Global Information Security Community

www.ISSA.org