Network-based Threat Detection

Version 1.5
Released: June 19, 2015

Securosis, L.L.C. 515 E. Carefree Highway Suite #766 Phoenix, AZ 85085 **T** 602-412-3051
info@securosis.com www.securosis.com

Author's Note

The content in this report was developed independently of any sponsors. It is based on material originally posted on the Securosis blog, but has been enhanced, reviewed, and professionally edited.

Special thanks to Chris Pepper for editing and content support.

This report is licensed by Vectra Networks, whose support allows us to release it for free.

All content was developed independently.

www.vectranetworks.com

Vectra Networks™ is the leader in real-time detection of in-progress cyber attacks. The company's automated threat-management solution continuously monitors internal network traffic to pinpoint cyber attacks as they happen. It then automatically correlates threats against hosts that are under attack and provides unique context about what attackers are doing so organizations can quickly prevent or mitigate loss. Vectra prioritizes attacks that pose the greatest business risk, enabling organizations to make rapid decisions on where to focus time and resources. The company's headquarters are in San Jose, Calif., and it has European operations in Zurich. More information can be found at www.vectranetworks.com.

Copyright

This report is licensed under Creative Commons Attribution-Noncommercial-No Derivative Works 3.0.

http://creativecommons.org/licenses/by-nc-nd/3.0/us/

Network-based Threat Detection

Table of Contents

Overcoming the Limitations of Prevention 4

Looking for Indicators 8

Prioritizing with Context 12

Operationalizing Detection 17

Summary 21

About the Analyst 22

About Securosis 23

Overcoming the Limitations of Prevention

Organizations invest heavily to block advanced attacks, on both endpoints and networks. Despite all this investment, devices continue to be compromised in increasing numbers and high-profile breaches continue unabated. Something doesn't add up. It comes down to psychology: security practitioners want to believe the latest shiny widget for preventing compromise will finally work and stop the pain.

Of course we are still waiting to see effective prevention. So Securosis has been advocating a shift in security spending — away from ineffective prevention, and towards detection and investigation of active adversaries within your networks and systems. We know many organizations have already spent a bunch of money on detection — particularly intrusion detection, its big brother intrusion prevention, and SIEM.

Many organizations have already spent a bunch of money on detection — particularly intrusion detection, its big brother intrusion prevention, and SIEM. But these techniques haven't worked effectively either, so now is time to approach the issue with fresh eyes.

But these techniques haven't worked effectively either, so now is time to approach the issue with fresh eyes. By taking a new forward look at detection, not from the standpoint of what we have already done and implemented (IDS and SIEM), but instead in terms of what we need to do to isolate and identify adversary activity, we will be able to look at the kinds of technologies needed right now to deal with modern attacks. Times have changed and attackers have advanced, so our detection techniques need to evolve as well.

Threat Management, Reimagined

Let's revisit how we think about threat management. As we documented in Advanced Endpoint and Server Protection, threats have changed, so you need to change the way you handle them. We believe threat management needs to evolve as follows:

- **Assessment:** You cannot protect what you don't know about — that hasn't changed and isn't about to. So the first step is to gain visibility into all devices, data sources, and applications that pose a risk to your environment. Additionally you need to understand the security posture of anything you have to protect.

- **Prevention:** Next try to stop attacks from succeeding. This is where most of the effort in security has been over the past decade, with mixed (okay, lousy) results. A number of new tactics and techniques are modestly increasing effectiveness, but the simple fact is that *you cannot prevent every attack*. It is now a question of reducing attack surface as much as practical. If you can stop the simplistic attacks reliably, you can focus on advanced ones.
- **Detection:** You cannot prevent every attack, so you need a way to detect attacks after they get through your defenses. There are a number of different detection options — most based on watching for patterns that indicate a compromised device. The key is to shorten the time between when a device is compromised and when you *discover* it has been compromised.
- **Investigation:** Once you detect an attack, you need to verify the compromise and understand what it actually did. This typically involves a formal investigation — including a structured process to gather forensic data from devices, triage to determine the root cause of the attack, and a search to determine how widely the attack spread within your environment.
- **Remediation:** Once you understand what happened, you can put a plan in place to recover the compromised device. This might involve cleaning the machine, or more likely reimaging it and starting over again. This step can leverage ongoing hygiene activities (such as patch and configuration management) because you can and should use tools you already have to reimage compromised devices.

This reimagined threat management process incorporates people, processes, and technology — integrated across endpoints, servers, networks, and mobile devices. That makes a huge matrix of combinations, to manage threats across the entire lifecycle for all device types. Whew! That would be a lot of work (and a really long paper). The good news is that this paper will focus specifically on network-based detection.

Why Not Prevention?

From reading this far you might think we have surrendered and given up on preventing attacks. Not true! We still believe in the value of restrictive application-centric firewall policies and looking for malware on ingress pipes. But you cannot count on your prevention tactics — they are insufficient.

Adversaries have made tremendous progress in being able to evade intrusion prevention and malware detonation devices (sandboxes). Remember that your devices aren't always protected by your network perimeter or other defenses. Employees take devices outside your network and click things. So devices come back onto the corporate network infected.

That doesn't mean these perimeter security gateways (firewalls, IPS, and UTM devices) don't catch stuff — they do. But they cannot catch *everything*. If you are questioning the importance of detection, think of it as Plan B. Every good strategist has Plan B (and Plans C, D, and E). Detection effort (Plan B) gives you a fallback position when prevention (Plan A) doesn't pan out.

*In a nutshell, it is not prevention **or** detection. It is both.*

Why Not Existing Monitoring?

You have probably already spent a bunch of time and money implementing intrusion detection/prevention and SIEM to monitor network segments. So why isn't IDS and SIEM good enough? It comes down to a fundamental aspect of these technologies: they need to know what you are looking for. Basically, you define a set of conditions (rules/policies) to match typical attack patterns in your network traffic or event logs. If an attacker uses a common attack that has already been profiled, and you have added that rule to your detection system, and your device can handle the volume (because you probably have 10,000 other rules defined on that device), you will detect that attack.

So why isn't IDS and SIEM good enough? It comes down to a fundamental aspect of these technologies: they need to know what you are looking for.

But what if the attacker is evading your devices by hiding traffic in a standard protocol, and communicating by proxying through a legitimate network? What if they are using a pattern you haven't seen before? Yep — then you'll miss the attack.

Again, it's not that you no longer need to monitor systems and networks. Compliance mandates that you still need IDS and SIEM. It's still critical to collect data and analyze it to find attacks you know about. And to be fair, many IDS/IPS and SIEM platforms are adding more sophisticated analysis to their standard correlation capabilities to improve detection. But these approaches still require a lot of tuning and experimentation to get right, and nobody has time to waste on a noisy security monitor.

The Answer Is…

Unfortunately we haven't found sustainable cold fusion, or a magic bullet that identifies every attack from every adversary every time (cold fusion actually seems a bit more likely). That would be nice though, right? But a couple capabilities have come together to enable better and more accurate network detection:

1. **Math:** Actually math has been around for a while. But advanced analytics provide improved ability to find patterns among a variety of data sources, which has made a big difference in the effectiveness of detection. Vendors call this "Big Data Analytics" and "Machine Learning". Shiny buzzwords aside, these capabilities improve your ability to find anomalous traffic, earlier in the attack chain.

Unfortunately we haven't found sustainable cold fusion, or a magic bullet that identifies every attack from every adversary every time (cold fusion actually seems a bit more likely).

2. **Context:** Anomaly detection has been around as long as detection algorithms, but it offered limited value because it threw off a lot of false positives. An anomaly could be legitimate or malicious, and you had no way to tell the difference without a fairly deep investigation. So being able to evaluate other types of data such as identity and content/payload, and to prioritize anomalies based on which are more likely to be attacks, helps you eliminate now-obvious false positives.

Network-based detection has evolved to the point where you can identify devices that look like they have been compromised. To be clear, this is still suboptimal because the device is under the adversary's control and our inner security purist still wants to block every attack. But a breach doesn't happen until exfiltration occurs, and if you are able to respond faster and better you can contain the damage. That's what better detection is for.

Looking for Indicators

We need to collect and analyze network telemetry to determine whether communications between devices, and communication contents, are legitimate or warrant additional investigation. Modern malware relies almost exclusively on the network to initiate the connection between the device and the command and control nodes, download attacks, perform automated beaconing, etc. Fortunately these activities typically exhibit a deterministic pattern, which enables you to pinpoint malicious activity and identify compromised systems.

Attackers bet they will be able to obscure their communications within the tens of billions of legitimate packets traversing enterprise networks on any given day, and on defenders' general lack of sophistication

Regardless of whether the attack happens as a result of malware, stolen credentials, social engineering or any other means of compromising the device, the attackers need to actively communicate with the device. Attackers bet they will be able to obscure their communications within the tens of billions of legitimate packets traversing enterprise networks on any given day, and on defenders' general lack of sophistication preventing them from identifying giveaway patterns. But if you can identify the patterns you have an opportunity to detect the attacks.

Command and Control

Command and Control (C&C) traffic is communication between compromised devices and C&C nodes/controllers. Once the device executes malware (by whatever means) and the dropper is installed, the device searches for its controller to receive further instructions. There are two main ways to identify C&C activity: traffic destination, and communication patterns between device and controller.

The industry has been using IP reputation for years to identify malicious destinations on the Internet. Security researchers evaluate each IP address and determine whether it is 'good' or 'bad' via automated means, based on activity observed across a massive network of sensors. IP reputation turns out to be a pretty good indicator that an address has already been used for malicious activity *at some point*. Traffic to known-bad destinations is definitely worth checking out, and perhaps even blocking. But malicious IP addresses (and even domains) are not active for long, as attackers cycle through addresses and domains frequently to avoid detection.

Attackers also use legitimate web sites as C&C nodes, which can leave innocent (but compromised) sites with bad reputations. The downside to blocking traffic to sites with bad reputations is the risk of irritating users who want to use 'safe' sites. Our research shows enterprises are becoming increasingly comfortable with blocking sites, because the great majority of addresses with bad reputations have legitimately earned them.

Keep in mind that IP reputation is not sufficient to identify all the C&C traffic on your network — many malicious sites used in targeted attacks don't show up on IP reputation lists. Thus you also need to look for other indications of malicious activity on the network, which is usually a result of compromised devices attempting to find their controllers.

With the increasing use of domain generating algorithms (DGA), malware doesn't need to be hard-coded with specific domains or IP addresses — instead it cycles through a set of domains according to its DGA, searching for a dynamically addressed C&C controller; addresses cycle daily. This provides tremendous flexibility for attackers to ensure newly compromised devices can establish contact, despite frequent domain takedowns and C&C interruptions. But these algorithms look for controllers predictably, making frequent DNS calls in specific patterns. So DNS traffic analysis has become critical for identification of C&C traffic, as has monitoring of packet streams.

Outliers

Network-based anomaly detection was reasonably effective, but as adversaries got more sophisticated, detection needed to dig more deeply into traffic.

Identifying C&C traffic before a compromised device becomes a full-fledged member of a botnet is optimal. But if you miss, once the device is part of the botnet you can look for indications that it is being used as part of an attack chain. You do this by looking for outliers: devices acting atypically or suspiciously.

Does this sound familiar? It should — anomaly detection has been used to find attackers for over a decade, typically using Netflow. You profile normal traffic patterns (source/destination/protocol) for users on your network, and then look for traffic variations from your baseline which exceed tolerances.

Network-based anomaly detection was reasonably effective, but as adversaries got more sophisticated, detection needed to dig more deeply into traffic. Deep packet inspection combined with better analytics now allows network-based detection offerings to assess network traffic context. Attack traffic tends to occur in a few stages:

1. **Command and Control:** As described above, devices communicate with nodes/servers under attacker's control.

2. **Reconnaissance:** After compromising a device and gaining control, attackers communicate with internal devices to map the network, identify the location of their target, and determine the most efficient path to their target.

3. **Lateral Movement:** Once the best path is identified, attackers systematically move through the network to both solidify their presence within the target organization, as well as reach the intended target by compromising additional devices.

4. **Exfiltration:** Once the target device is compromised, the attacker needs to accumulate the data, and then move it outside the network. This can be done using tunnels, staging servers, and other techniques to obfuscate activity.

Each of these stages involve patterns you can watch for to detect attacks, and can indicate the presence of attackers. But this still isn't a smoking gun — at some point you need to apply additional context and correlate the activity happening during multiple stages to both create more evidence of malfeasance and understand intent. Analyzing content in the communication stream is the next step in identifying attacks.

Content-based Indicators

One way to glean more context in network traffic analysis is to understand what kind of data is being moved. With deep packet inspection and session reassembly you can perform file-based analysis of content as well to improve detection accuracy. Then you can compare current traffic with baselines to look for anomalies in data movement as well.

1. **File size:** For example if a user moves 2GB of traffic in 24 hours, but they normally move no more than 100MB per day, that should trigger an alert. Perhaps it's nothing, but it should be investigated.

2. **Application Patterns:** Many web-based applications have known and predictable transaction patterns (http get/post and other database transactions) that can be profiled. With an established baseline you can look for anomalies.

3. **Time of day:** Similarly, if a user doesn't normally work in the middle of the night, but does so two days in a row, that could indicate malicious activity. Of course it might be just a big project, but it merits investigation.

4. **Simple DLP:** You can fingerprint files to look for sensitive content, or regular expressions which match account numbers or other protected data. Of course that isn't full DLP classification and analysis. But it could flag something as malicious without the overhead of full DLP.

> Malware crossing the perimeter does not necessarily mean it executed on any devices. That is a weakness of network-based sandboxes, which just look at and alert on files coming into the network.

Content analysis won't provide a smoking gun either. But combined with network traffic detection it provides more detail to discern intent. This can help explain behavior that would otherwise be flagged as anomalous, to reduce false positives.

Endpoint Confirmation

Malware crossing the perimeter does not necessarily mean it executed on any devices. That is a weakness of network-based sandboxes, which just look at and alert on files coming into the network. They fire an alert whenever they see malware, even if the target device was totally protected from the attack. One way to further identify real attacks is to integrate endpoint telemetry into security analysis, to verify and validate what actually happened. We increasingly see a drive for coordination of network-based detection with endpoint detection.

This is useful both from the standpoint of confirming the attack on the targeted device, but also looking for indicators of the attack on *other* devices in the network. For example, malware could be detected on the perimeter targeting the CFO. The first step is to see if the CFO's device is compromised, but it's also important to look for other devices potentially exposed to the malware, since maybe the CFO's assistant opens all the email and could therefore be hit by the attack first.

Of course you still want to know malware entered the network, but you need some way to prioritize whether or not it needs to be dealt with *right now*. Which brings up the much larger issue of prioritization: deciding which potential attack to handle first. It comes down to understanding what presents the most clear and present danger (risk) to your environment, and being able to confirm if the endpoint has in fact been compromised is a key aspect of that determination.

Prioritizing with Context

During speaking gigs we ask how many in the audience actually get through the to-do list every day. Usually we get one or two jokers in the crowd between jobs, or maybe just trying to troll us a bit. But nobody in a security operational role gets everything done every day. The critical success factor is to make sure you are getting the right things done, and not burning time on activities that don't reduce risk or contain attack damage.

Underpinning this paper is the fact that prevention inevitably fails at some point. Along with a renewed focus on network-based detection, that means your monitoring systems will detect a bunch of things — including possible exploitation. But which alerts are important? Which represent active adversary activity? Which are just noise and need to be ignored? Figuring out which is which is where you need the most help.

The critical success factor is to make sure you are getting the right things done, and not burning time on activities that don't reduce risk or contain attack damage.

To use a physical security analogy, if you are monitoring a physical security fence you will get alerts regularly. But you need to figure out an alert is caused by a confused squirrel, a wayward bird, a kid on a dare, or an attack. Just looking at an alert won't tell you much. But if your analysis provides other details and therefore additional context, you can figure out which is which. The stakes are high for getting this right, as the postmortems of many recent high-profile breaches indicate alerts did fire — in some cases multiple times from multiple systems — but organizations failed to take action… and suffered the consequences.

Earlier we listed network telemetry as one indicator you could evaluate to indicate potential malicious activity. Let's say you like that approach, and decide to implement it in your own monitoring systems. So you flip the switch and alerts come streaming in. Now comes the art: separating signal from noise and narrowing your focus to the alerts that matter and demand immediate attention. You do this by adding context to general network telemetry, then using an analytics engine to crunch the numbers.

For context you can leverage both internal and external information. Here we will focus on internal data, because you already have that and can implement it right away. Later in this paper we will tackle external data, typically accessible via a threat intelligence feed.

Device Behavior

You start by figuring out what's important — not all devices are created equal. Some store very important data. Some employees have access to important data, typically executives, regardless of who owns the device (factoring in BYOD). But not all devices present a direct risk to your organization, so categorizing them provides your first cut at prioritization. You can use this hierarchy to kickstart your efforts:

1. **Critical devices:** Devices with access to protected information and/or particularly valuable intellectual property should bubble to the top. Fast. If a device on a protected and segmented network shows indications of compromise, that's bad and needs to be dealt with immediately. Even if the device is dormant, traffic on a protected network that looks like command and control constitutes smoke, and you need to act quickly to ensure any fire doesn't spread. Or enjoy your disclosure activities…

2. **Active malicious devices:** If you see device behavior which indicates an active attack (perhaps reconnaissance, moving laterally within the environment, blasting bits at internal resources, or exfiltrating data), that's your next order of business. Even if the device isn't considered critical, if you don't deal with it promptly the attack might find an exploitable hole to a higher-value device and move laterally within the organization. So investigate and remediate these devices next.

3. **Dormant devices:** These devices at some point showed behavior consistent with command and control traffic (typically communication with a C&C network), but aren't doing anything malicious at the moment. Given the number of other fires raging in your environment, you may not have time to remediate these dormant devices immediately.

These priorities are coarse but should be sufficient. You don't need a complicated multi-tier rating system which is too involved to use daily. Priorities should be clear. Of course this last bucket might show malicious activity at any time, so you still need to watch it. The question is when you remediate.

This categorization helps, but within each bucket you likely have multiple devices. So you still need additional information and context to make decisions.

Who and Where

Not all employees are created equal either. An important source of context is user identity, and there are a bunch of groups you need to pay attention to. The first is people with elevated privileges, such as administrators and others with entitlements to manage devices that hold critical information. They can add, delete, and change accounts and access rules on servers, and manipulate data. They have access to tamper with logs, and basically can wreck an environment from the inside. Moreover, with the access most administrators have there is little need to use additional malware (and potentially trigger malware detection alerts) making administrators very high value targets. There are plenty of examples of rogue or disgruntled administrators making a real mess, so when you detect anomalous behavior on an administrator's device, that should rise to the top of your list.

The next group of folks to watch closely is executives with access to financials, company strategy, or other key intellectual property. These users are attacked most frequently via phishing and other social engineering, so they need to be watched closely — even trained, they aren't perfect. This may trigger organizational backlash — some executives get cranky when they are monitored. But that's not your problem, and you need this kind of context to do your job. So dig in and make your case to the executives for why it's important. As you look for indicators that devices are connecting to a C&C server or performing reconnaissance, you are protecting the organization, and executives should know better than to fight that.

The next group of folks to watch closely is executives with access to financials, company strategy, or other key intellectual property. These users are attacked most frequently via phishing and other social engineering, so they need to be watched closely — even trained, they aren't perfect.

Keep in mind that a sophisticated attack typically involves a variety of targets during the mission. So the adversary may gain a foothold on your network via a low value target, but then move laterally to areas of the network more of interest. Which is why constantly assessing all of the traffic on your networks is an ongoing priority to detect the behavior before the mission is completed.

The location of your critical data also provides input for prioritization. Critical data lives on particular network segments, typically in the data center, so you should be making sure those networks are monitored. But it's not just PII you need to worry about. Your organization should isolate segments for labs doing cutting-edge R&D, finance networks with preliminary numbers from last quarter, and anything else demanding special caution. Isolation is your friend — use different segments, at least logically, to minimize data intermingling.

You can get contextual information from a variety of sources, which you likely already use. For instance identity information (such as Active Directory users and groups) enables you to map a device to a user and/or group. Then you can profile typical finance department activity to know how it differs from the way marketing and engineering groups communicate with each other and the broader Internet. You could go deeper and profile specific people.

Additionally, network topology is important in attack path analysis, to understand the blast radius of any specific attack. That's a fancy term for damage assessment in case a device or network is compromised: what else would be directly exposed? Once you figure out which other devices on the network can be reached from the compromised device (during lateral movement), and what potential attacks would succeed, you can further prioritize activities.

Content

The next area to mine is the content flowing through the network — of course not all data is equally sensitive. You need to be able to analyze the content stream within network traffic to look for protected data, or data identified as critical intellectual property. This rough data classification can be very resource-intensive and hard to keep current (just ask anyone trying to implement DLP), so make it as simple as possible. For instance private personal information (PPI) may be the most important data to protect in your environment. But intellectual property is the lifeblood of most high-tech organizations, and typically their top priority. It doesn't really matter what you put at the top — just reflect your organization's priorities.

Compliance remains a factor for many organizations, so potential compliance violations bubble up when figuring out priorities, especially right around assessment time.

The importance of various specific types of content depends on the organization, and you need to do the work to understand how they need to be protected and monitored. That will entail building consensus with executives because you need clear marching orders for what alerts need to be validated and investigated first.

> We recommend that you include a feedback loop in your security alerting process. Assess the value of your alerts, identify gaps, and then tune based on what is really happening in the field.

Math

Armed with network data identifying attack indicators and information that provides additional context such as identity, location, and content, now you need to figure out what is at greatest risk and react accordingly. This involves crunching numbers, analyzing the results, and identifying the highest priority alert. You are looking to:

1. Get a verdict on a device and/or a network: whether it has been compromised and how badly.
2. Dig deeper into the attack to figure out the extent of the damage and how far it has spread.

This requires math. We aren't being flippant (okay, maybe a bit), but this type of analysis requires fairly sophisticated algorithms to establish a general risk assessment. You will hear a lot of noise about "risk scoring" as you dig into the current state of network-based detection. A quantified risk score can be rather arbitrary, so it is useful to understand how the score is calculated and where the

numbers come from. Make sure your numbers pass the sniff test and you can defend where they come from, because they will drive decisions.

As discussed above, your organization will have its own ideas about what's important, and different risk tolerances than other organizations. So you should be able to tune algorithms and weight factors differently to get more meaningful alerts. Your environment is not static — it will change constantly, so you need to tune the alerting systems on an ongoing basis. Sorry, but there is not much set-it-and-forget-it any more. We recommend that you include a feedback loop in your security alerting process. Assess the value of your alerts, identify gaps, and then tune based on what is really happening in the field.

Operationalizing Detection

To finish off we will discuss additional context and making alerts operationally useful.

Leveraging Threat Intelligence for Detection

Our analysis so far has been restricted to your organization. You are gathering data from your networks and adding context from your enterprise systems. That is great but not enough. Factoring data from other organizations into your analysis can help you refine it and prioritize your activities more effectively.

For prevention, threat intel can help decide which external sites should be blocked on your egress filters, based on reputation and possibly adversary analysis. This approach helps ensure devices on your network don't communicate with known malware sites, bot networks, phishing sites, watering hole servers, or other places on the Internet you want nothing to do with. Recent conversations with practitioners have indicated much greater willingness to block traffic based on threat intel — so long as they have confidence in the alerts.

But this series isn't called Network-based Threat Prevention, so how does threat intelligence help with detection? TI provides a view of network traffic patterns used in attacks on other organizations. Knowledge of these patterns enables you to look for them (Domain Generating Algorithms, for example) within your own environment. You might also see indicators of internal reconnaissance or lateral movement typically used by certain adversaries, and use them to identify attacks in process. Watching for bulk file transfers, for example, or types of file encryption known to be used by particular crime networks, could yield insight into exfiltration activities.

As the burden of proof is far lower in civil litigation than in criminal litigation, the bar for useful detection accuracy is much lower than for prevention.

As the burden of proof is far lower in civil litigation than in criminal litigation, the bar for useful detection accuracy is much lower than for prevention. When you are blocking network traffic for prevention, you had better be right. Users get cranky when you block legitimate network sessions, so you should be conservative about what you block. That means you will inevitably miss something — the dreaded false negative, a legitimate attack. But an alert offers more leeway, so you can be a bit more expansive.

That said, you still want to be close — false positives are still expensive. This is where the approach we mapped out earlier comes into play. If you see something that looks like an attack based on external threat intel, you apply the same contextual filters to validate and prioritize.

Retrospection

What happens when you don't recognize an attack when the traffic enters your network? This happens every time a truly new attack emerges. Obviously you don't know about it, so the active controls on your network miss it and your security monitors don't know what to look for. No one has seen it yet, so it doesn't show up in threat intel feeds. You miss, but that's life. Everyone misses new attacks. The question is: *how long* do you miss it?

> One of the most powerful concepts in threat intelligence is using newly discovered indicators to look retrospectively through security data, to see if an attack has already hit you.

One of the most powerful concepts in threat intelligence is using newly discovered indicators to look retrospectively through security data, to see if an attack has already hit you. When you get a new threat intel indicator you can search your network telemetry (using your fancy analytics engine) to see if you saw it before. This isn't optimal — you already missed. But it's much better than waiting for an attacker to take the next step in the attack chain. It shortens the window between compromise and detection, and that's what you are trying to achieve. In the security game nothing is perfect. But the hard-won experience of other organizations can make your own detection faster and more accurate.

A Picture Is Worth a Thousand Words

At this point you have alerts, and have done some analysis to prioritize the alerts presenting the biggest risk to the organization. But any organization of scale is still going to have a bunch of these alerts to work through. But one of the most difficult tasks is deciding how to navigate through the hundreds or thousands of alerts from a network at scale. That's where visualization comes into play. A key criterion for choosing a detection offering is presentation of information in a way that makes sense to you and will work in your organization's culture.

Some like the traditional user experience: a Top 10 list of potentially compromised devices, with a grid showing details of each alert. Another way to visualize detection data is a heat map showing devices and potential risks visually. Either way you will need to drill down into indicators and alerts to find the root cause of the attack. There is no right or wrong when it comes to user experience — just a question of what will be most effective for your security operations team.

Operationalizing Detection

As compelling as network-based threat detection is conceptually, integration with other enterprise systems is required before you can provide value and increase your security program's effectiveness. There are two sides to integration: data you need for detection, and information about alerts that is sent to other operational systems. For the former (security data), integrating with identity management systems and external threat intelligence provide the raw security data for detection analytics. The latter includes the ability to pump the alert and contextual data into your SIEM or other alerting system to kick off your investigation process.

If you get comfortable enough with your detection results you can even configure active controls such as IPS blocking rules based on these alerts. You might also quarantine compromised devices (via integration with NAC), block C&C traffic (egress firewall), or stop exfiltration (firewall or DLP). As described above, you always have to worry about false positives blocking legitimate traffic, but disrupting attackers is extremely valuable.

For network forensics you might integrate with a full packet capture/network forensics platform. In this use case, when a device shows potential compromise, traffic to and from it could be captured for forensic analysis. Such captured network traffic may provide a proverbial smoking gun. Providing the actual attack packets could also make you popular with the forensics folks. Prioritized alerts enable you to be more precise and efficient about what traffic to capture, and ultimately what to investigate.

We will increasingly see security controls reconfigured based on alerts, network traffic redirected, and infrastructure quarantined and pulled offline for investigation. Attacks hit too fast to do it any other way.

Automation of these functions is still in its infancy. But we expect all sorts of security automation to emerge within the short-term planning horizon (18-24 months). We will increasingly see security controls reconfigured based on alerts, network traffic redirected, and infrastructure quarantined and pulled offline for investigation. Attacks hit too fast to do it any other way, but automation scares many security professionals. We expect to see this play out over 5-7 years, but have no doubt it will happen.

When to Remediate, and When Not to

It may be hard to believe, but there are scenarios where you might not want to immediately remediate a compromised device. The first — and easiest to justify — is when the attack is part of an ongoing investigation; HR, legal, senior management, or law enforcement may mandate the device be observed but otherwise left alone. There isn't much wiggle room in this scenario. With the remediation decision no longer in your hands, and the risk of an actively compromised device on your network declared acceptable, you take reasonable steps to monitor the device closely, prevent it from accessing critical assets, and ensure it is unable to exfiltrate data.

Another scenario where remediation may not be appropriate is when you need to study and profile your adversary to learn about the malware and its command and control apparatus through direct observation. You need a sophisticated security program to undertake a detailed malware analysis (as described in Malware Analysis Quant), but understanding and identifying indicators of compromise can help identify other compromised devices, and enable you to deploy workarounds and other infrastructure protections such as IPS rules and HIPS signatures.

That said, in most cases you will just want to pull the device off the network as quickly as possible, pull a forensic image, and then reimage it. That is usually the only way to ensure the device is truly clean before letting it back into the general population. If you are going to leave the device active and monitor it, ensure the criteria for that decision are documented and agreed on as part of your incident response plan.

Summary

Prevention of attacks has proven insufficient to keep pace with adversaries. Regardless of the billions spent on existing technologies, including intrusion prevention/detection and SIEM, endpoint and server devices continue to be compromised and critical corporate data exfiltrated. It's clear that threat detection needs to evolve to more effectively catch modern adversaries. The path of least resistance to implementing new detection techniques is the network, which doesn't require deployment to thousands of devices and can monitor key ingress and egress points.

But detecting attacks from network traffic can be challenging. Attackers work diligently to "hide in plain sight" by obscuring their attack traffic within the tens of billions of legitimate packets on the network. So a deeper analysis of network traffic to identify behavioral aspects of sessions is required, as well as an integration with endpoints to confirm whether malware actually executed on a device. This helps to reduce false positives and improve the efficiency of your response.

The success criterion of any detection process is to ensure that effective action results from alerts. That involves ensuring alerts are legitimate, providing some measure of context to allow prioritization of the remediation given the risk to your organization, and visualization of threats to drive prioritization decisions. This context comes from both internal data (similar behaviors on multiple devices can indicate an outbreak) and external data (new indicators can retrospectively identify ongoing adversary campaigns within your environment).

Over time, given the significant (and likely insurmountable) security staffing constraints, organizations need to embrace automated actions based on alerts from detection."Trustable automation" will require detection to continue to evolve in both accuracy and scale. With new technologies described in this paper, detection can make the requisite improvements to provide the basis for this critical automation.

If you have any questions on this topic, or want to discuss your situation specifically, feel free to send us a note at info@securosis.com.

About the Analyst

Mike Rothman, Analyst and President

Mike's bold perspectives and irreverent style are invaluable as companies determine effective strategies to grapple with the dynamic security threatscape. Mike specializes in the sexy aspects of security — such as protecting networks and endpoints, security management, and compliance. Mike is one of the most sought-after speakers and commentators in the security business, and brings a deep background in information security. After 20 years in and around security, he's one of the guys who "knows where the bodies are buried" in the space.

Starting his career as a programmer and networking consultant, Mike joined META Group in 1993 and spearheaded META's initial foray into information security research. Mike left META in 1998 to found SHYM Technology, a pioneer in the PKI software market, and then held executive roles at CipherTrust and TruSecure. After getting fed up with vendor life, Mike started Security Incite in 2006 to provide a voice of reason in an over-hyped yet underwhelming security industry. After taking a short detour as Senior VP, Strategy at eIQnetworks to chase shiny objects in security and compliance management, Mike joined Securosis with a rejuvenated cynicism about the state of security and what it takes to survive as a security professional.

Mike published The Pragmatic CSO <http://www.pragmaticcso.com/> in 2007 to introduce technically oriented security professionals to the nuances of what is required to be a senior security professional. He also possesses a very expensive engineering degree in Operations Research and Industrial Engineering from Cornell University. His folks are overjoyed that he uses literally zero percent of his education on a daily basis. He can be reached at mrothman (at) securosis (dot) com.

About Securosis

Securosis, LLC is an independent research and analysis firm dedicated to thought leadership, objectivity, and transparency. Our analysts have all held executive level positions and are dedicated to providing high-value, pragmatic advisory services. Our services include:

- **Primary research publishing**: We currently release the vast majority of our research for free through our blog, and archive it in our Research Library. Most of these research documents can be sponsored for distribution on an annual basis. All published materials and presentations meet our strict objectivity requirements and conform to our Totally Transparent Research policy.
- **Research products and strategic advisory services for end users**: Securosis will be introducing a line of research products and inquiry-based subscription services designed to assist end user organizations in accelerating project and program success. Additional advisory projects are also available, including product selection assistance, technology and architecture strategy, education, security management evaluations, and risk assessment.
- **Retainer services for vendors**: Although we will accept briefings from anyone, some vendors opt for a tighter, ongoing relationship. We offer a number of flexible retainer packages. Services available as part of a retainer package include market and product analysis and strategy, technology guidance, product evaluation, and merger and acquisition assessment. Even with paid clients, we maintain our strict objectivity and confidentiality requirements. More information on our retainer services (PDF) is available.
- **External speaking and editorial**: Securosis analysts frequently speak at industry events, give online presentations, and write and/or speak for a variety of publications and media.
- **Other expert services**: Securosis analysts are available for other services as well, including Strategic Advisory Days, Strategy Consulting engagements, and Investor Services. These tend to be customized to meet a client's particular requirements.

Our clients range from stealth startups to some of the best known technology vendors and end users. Clients include large financial institutions, institutional investors, mid-sized enterprises, and major security vendors.

Additionally, Securosis partners with security testing labs to provide unique product evaluations that combine in-depth technical analysis with high-level product, architecture, and market analysis. For more information about Securosis, visit our website: <http://securosis.com/>.