CHAPTER 6

Emerging Threats and Vulnerabilities: Reality and Rhetoric

Curiosity is lying in wait for every secret.

—Ralph Waldo Emerson

These days it's hard to read an online news source, pick up a newspaper, or watch TV without seeing reports of new threats: cybercrimes, data breaches, industrial espionage, and potential destruction of national infrastructure. These reports inevitably leave the impression that we are drowning in an inexorable tide of new and terrifying threats.

One has to question how much of this is rhetoric, and how much is reality. There are political and profit-driven motives for making threats seem bigger and more imminent than they really are. US government officials have warned that cyber attacks potentially can be "devastating, approaching weapons of mass destruction in their effects" (Levin 2010). Such warnings have been used to justify requests for increased national cybersecurity funding, as well as proposed restrictions on private networks. It's not surprising, therefore, that some experts have expressed skepticism about the real extent of the threat. In fact, academics at the George Mason University Mercatus Center have warned, "the United States may be witnessing a bout of threat inflation similar to that seen in the run-up to the Iraq War" (Brito and Watkins 2012).

On the other hand, common sense tells us new cyber threats really are emerging and growing. More data is online and vulnerable to attack, and millions of new Internet-connected devices are inevitably introducing new risks. Malware production has matured into a sizable industry. Government agencies and businesses have suffered real attacks attributed to nation-state actors: in 2014, for example, the US Government charged five members of the Chinese military with stealing information from SolarWorld and other companies, during a trade dispute over solar-energy products.

M.W. Harkins, *Managing Risk and Information Security*,
DOI 10.1007/978-1-4842-1455-8_6

Given the flood of often-conflicting information, how can we get an accurate picture of the threat landscape so that we can develop an appropriate security strategy? How do we determine which threats directly affect our organizations, and distinguish them from those that are irrelevant? How do we decide which threats require immediate defensive measures, as opposed to those that attract attention but don't yet present significant risks?

In this chapter, I'll describe methods for identifying the real threat and vulnerability trends among the rhetoric. I'll also discuss some key areas of threat activity that have been analyzed using these methods. My goal is to help information security groups stay ahead of the attackers and focus their limited resources on mitigating the most important threats.

Structured Methods for Identifying Threat Trends

To identify the real trends in emerging threats among the mass of news and speculation, we need to carefully examine the available information using a structured, analytical approach. Unfortunately, many security groups absorb information about emerging threats using methods that are unstructured and sometimes almost haphazard.

A typical process looks something like this. The security team relies on external sources, such as news feeds and alerts, as well as informal anecdotes, to gather information about emerging threats. Based on this information, the team holds brainstorming sessions to review the threat landscape. The output from these sessions is a list of "top risks." Security resources are then focused on mitigating the items on the list.

There are several problems with this approach. Information comes from a narrow, limited range of sources, resulting in a blinkered security perspective that tends to stifle creative thinking. Also, the information is usually fragmented, making it difficult for the team to identify trends and gaps in the data. These deficiencies continue through security planning and implementation. Because the team lacks a full view of the threat landscape, it's hard to determine which threats require immediate attention and how much of the limited security budget they deserve. As a result, risks are incorporated into plans on an ad hoc basis, and not all risks are adequately mitigated. Finally, security teams often don't have a structured process for communicating threat information to other people within their organizations. Because of this, people outside the security group remain unaware of emerging risks and don't know how to respond when they experience an attack.

I realized the limitations of this approach several years ago, and began trying to inject more rigor into the risk-sensing strategy. Over time, those efforts progressively developed into a more structured risk-sensing process that helps identify threats, prioritize them, plan responses, and deliver actionable information to those who may need it. Through continued use, risk sensing can become a systemic process within any organization.

The process for analyzing emerging threats includes several valuable techniques that may be unfamiliar to some security groups. I have used a product life cycle analogy to track threats as they mature from theoretical risks into full-blown exploits. I have also used nontraditional analysis techniques, such as war games and threat agent profiles, to encourage creative thinking and identify threats that might otherwise be missed. I'll discuss these methods in more detail later in this chapter.

The process can be managed by a small core team, supplemented by a broad set of experts (including people outside the security group) across an organization. This arrangement ensures continuity while enabling the team to mine a diverse variety of sources to get a more complete picture of immediate and future threats.

Security team members should research a wide range of security topics in depth. This diversity of perspective and discussion essentially creates a crowd-sourcing of intelligence and reduces the influence of any single person's bias. Team members use typical sources, such as external feeds and analysis; they also mine academic research and hacker discussion forums, and connect with security professionals at other organizations. Other team members may scan the regulatory horizon to identify upcoming laws and regulations with potential impact, or analyze internal investigations and other near-miss incident data.

The team should hold regular meetings to analyze the threat landscape. At these meetings, each security domain expert explains his or her findings to other members of the security team. For each security topic, the discussion should include a review of recent events and a look ahead to the future. This helps identify the key trends and the factors driving those trends, provides context that can be used to analyze the current state, and predicts the likely evolution of each threat. The structured evaluation uncovers emerging risks that the team might otherwise miss. It's also useful to look back at previous predictions to see which ones were accurate, and to analyze the reasons why threats may not have materialized in the way that was expected.

It's important to communicate the findings to stakeholders across your organization in regular reports and briefings, including a wide-ranging annual assessment of the threat landscape. This communication provides further opportunities to get feedback from across the organization and its business units, which can then be used to refine your risk-sensing analysis.

The Product Life Cycle Model

I have found that a product life cycle model is a useful way to track and prioritize emerging threats as they evolve and begin to present real risks to the enterprise. Almost all security groups have a limited budget, so they need to focus their resources on effectively mitigating the highest-priority threats.

This model, shown in Figure 6-1, recognizes that many threats initially emerge as theoretical risks, but are on a path to exploitation, and we need to evaluate and monitor them.

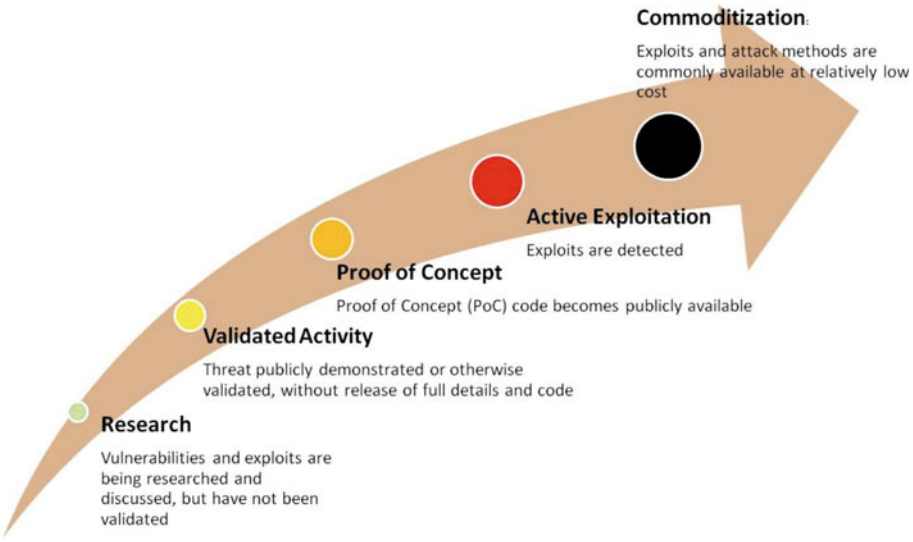

Figure 6-1. *The product life cycle model for tracking the evolution of threats. Source: Intel Corporation, 2012*

Often, researchers or hackers first reveal a possible attack or vulnerability at a security conference or publish information about it online. Next, attackers begin testing the use of this technique and making their results publicly available. Once the method has been proven, the threat enters the production phase as attackers start exploiting it in earnest. Ultimately, the threat becomes a mature commodity—source code is often freely available, many variants exist, and organizations treat the threat as part of the everyday landscape and build defenses accordingly.

This life cycle model enables security teams to systematically track the evolution of threats. It helps us determine when to allocate resources to fighting each threat. As each threat approaches maturity, we can examine how it is likely to affect our organizations and plan appropriate mitigation.

In addition, this model provides a great way to communicate actionable information to business groups using terminology they already understand (the product life cycle). When we provide regular threat landscape assessments to stakeholders, each security topic should include a description of the activity at each life cycle phase, thus providing a context that helps the security team inform business groups about how they should act on each of these emerging risks.

Let's examine some examples showing how this model can be used in real life. Figure 6-2 illustrates the evolution of threats targeting smartphones and other handheld devices. Researchers and hackers began to take notice of handheld devices almost a decade ago, demonstrating weaknesses and theoretical avenues of exploitation. Initially, they focused on what were then known as personal digital assistants. As smartphones

took off, attackers shifted their attention to this bigger market, which rapidly became a major area of threat activity. Monitoring trends at these earlier stages enables organizations to prepare. As threats mature and employees begin using smartphones more widely at work, well-prepared organizations are in a better position to develop risk mitigation measures including technical controls and incident response plans.

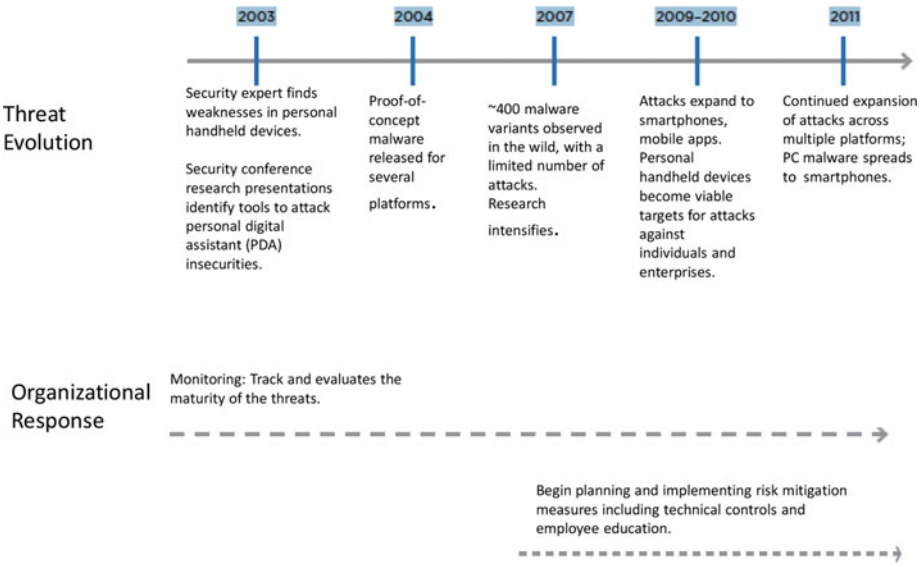

Figure 6-2. *How an organization could use the product life cycle model to track and respond to smartphone security threats*

By visually comparing activity across multiple threat areas, as shown in Figure 6-3, we can quickly identify major areas of activity and see the likely timing and extent of their impact. This chart also shows us areas in which there are numerous proof-of-concept tests and other activities that suggest major problems in the near future. And it indicates areas of focused research that may ripen into active exploitation over the long term.

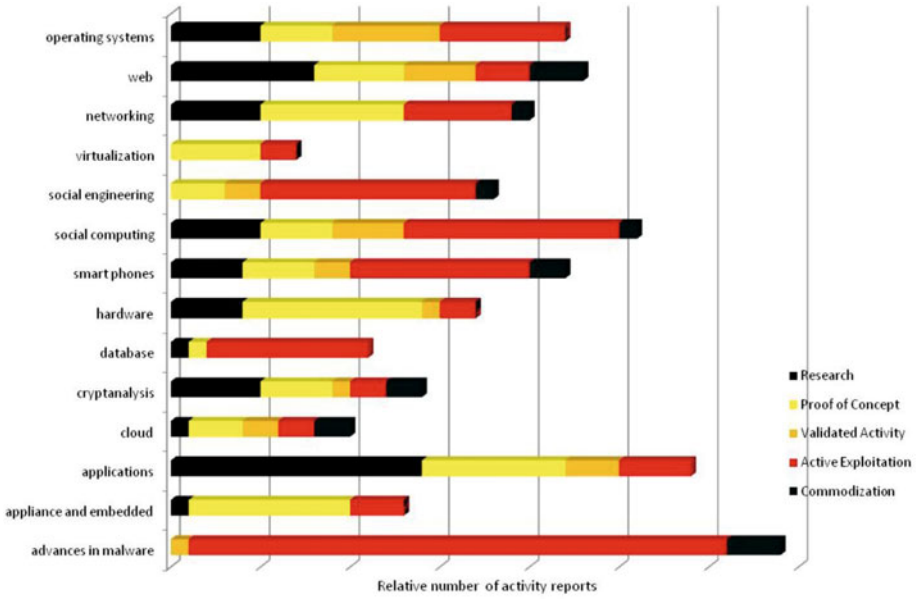

Figure 6-3. *A visual comparison of security-related activity across different technology areas. Data are for illustration purposes only. Source: Intel Corporation, 2012*

Although the depth of detail in Figure 6-3 is valuable to the security team, I have found a simpler, consolidated view such as the chart in Figure 6-4 can help communicate the essential trends to a broader audience, supplementing other threat analysis materials. These simpler charts are based on the activity identified using the product life cycle model, but add further trend analysis and group the activity areas into four main clusters, depending on their level of activity and maturity potential and on their potential impact to the company. These clusters are

- *Sustained drivers*: These are areas that already have a high impact or otherwise cause considerable concern. Typically, they are characterized by commoditized distribution and active exploitation by multiple threat agents. Today, examples include malware and web attacks.
- *Critical trends*: These areas have begun undergoing active exploitation, with growing adoption beginning to shift toward commoditization. Current examples include social computing and smartphones.
- *Emerging trends*: These areas have a low current level of exploitation, but considerable research and proof-of-concept activity. Examples include embedded and cloud computing.

- *Disruptive trends*: These are areas with little or no active exploitation, but significant research activity and the disruptive potential to cause a major security problem. Frequently, they are discussed as theoretical risks, and because of this, many people in the industry would be caught off guard by a significant event. Examples include virtualization, an area in which potential threats and vulnerabilities have been exposed and a successful exploit could cause far-reaching damage.

I have found that clustering threat analysis information in this way enhances communication with stakeholders. Representing the information in easy-to-understand charts helps to convey the key trends and their potential impact to a broad cross-section of people, helping them quickly assess whether they need to make adjustments to security strategy.

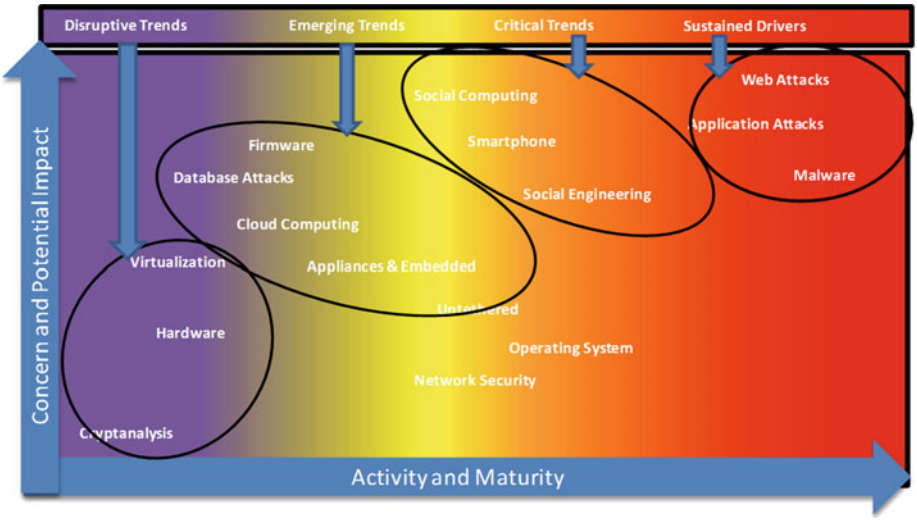

Figure 6-4. *Clustering areas of threat activity to highlight trends.*
Source: Intel Corporation, 2012

ASSESSING HOW TO RESPOND TO A NEW THREAT REPORT

A continuous stream of new threat reports emerges from agencies, intelligence services, and vendors. It can be hard to determine what to do with all the new information—especially since most security organizations have limited resources. Here are five questions you can ask yourself the next time you see a published threat report.

1. Are we immediately affected? Are the indicators of compromise shared in the report found in our environment? If so, we have an incident that we must deal with.
2. If we're not already affected, what is the likelihood that we will be a future target? We're more likely to be targeted if we work in the same industry as a previous victim, or if we are connected to them in another way (as a supplier, customer, or partner). If the attackers are hactivists or politically motivated threat actors, we are more likely to be targeted if we align with the victim's philosophy. Note that we may be a target even if there's no obvious linkage to the victim.
3. How were the victims attacked? What compensating controls do I have in my security stack to mitigate the risks across the kill chain of a similar attack?
4. Have we seen the same malware used, or families derived from it, against our assets?
5. Were any interesting tools, techniques, or procedures used that I should capture and share with my security team? This part of the report can be used to educate responders, architects, and risk managers so they can make better decisions.

Based on a blog post by Steve Mancini, Director of Information Security at Cylance (Mancini 2016).

Understanding Threat Agents

Besides the product life cycle analogy, there are other techniques that can help us think creatively about threats and identify risks we might otherwise miss.

Behind every threat is a human agent. To effectively plan defenses, it helps if we can understand why and how these agents operate: their motives, typical methods, and targets. However, I realized several years ago that we lacked agreed-upon definitions of threat agents, as well as a clear understanding of which agents actually pose the biggest risks to us.

Some agents and their activities attract considerable publicity, resulting in the "TV news effect" in which the most-publicized agents appear to be the biggest threat, so they often receive a disproportionately large percentage of limited mitigation resources. In reality, a wide spectrum of threat agents exists, some of which may be less well-known but pose bigger threats. For example, hactivists often want to publicize their activities as much as possible to draw attention to their cause. This publicity makes them appear to be a bigger threat than other groups, such as organized crime syndicates, which try to conceal their exploits.

In addition, terms often are used without clear agreement about what they mean. The phrase *advanced persistent threat* has become a buzzword whose exact meaning depends on who is using the term. It usually implies adaptive, long-term strategies employing a

variety of stealthy techniques and used by attackers with considerable resources. However, it's important to remember that a variety of agents may be capable of generating this type of threat. One thing that all these threat agents have in common is the use of malicious code to achieve their goals. But to understand and predict their likely motives and methods, it is useful to clearly define the agents, whether they represent nations or other powerful groups, such as organized crime. To solve this problem, Tim Casey, a member of my security team at Intel at that time, developed a standard threat agent library that provides a consistent, up-to-date reference describing the human agents that pose threats to our information assets (Casey 2007). The library helps risk management professionals quickly identify relevant threat agents and understand the importance of the threats.

The library acts as a collection point for information about each agent, making it easier to share information across your organization. It includes profiles of agents such as disgruntled employees, opportunistic employees, industrial spies, and politically motivated attackers. The library also catalogs agents' typical targets, objectives, skill levels, current activity, and exploit outcomes. When used as part of regular threat assessments, this model can help determine which agents pose the biggest risks to your organization. The security team can then use the information about their typical methods and exploits to help plan its strategy. The library helps the team understand why specific events and attack trends occur and what might happen next.

NSA'S CHIEF HACKER EXPLAINS HOW TO DEFEND AGAINST THREATS

It's hard to imagine someone who is better placed to provide advice about defending against advanced adversaries than Rob Joyce, who heads the National Security Agency's Tailored Access Operations (TAO) elite hacking unit. So the audience listened closely when he took the stage for an eye-opening talk at the 2016 Usenix Enigma conference. "My talk is to tell you, as a nation-state exploiter, what can you do to defend yourself to make my life hard," he said.

Joyce said that six intrusion phases comprise what is typically referred to as the "kill chain:" reconnaissance; initial exploitation; establish persistence; install tools; move laterally; and collect, exfiltrate, and exploit the data. Organizations can thwart attackers by disrupting the transition between any of these phases. For example, to help prevent reconnaissance turning into initial exploitation, you can reduce the attack surface by locking down or disabling devices that are unused or don't need to be open to access. "Don't assume a crack is too small to be exploited," he said. "We will look for that esoteric edge case."

Contrary to popular belief, advanced adversaries don't rely exclusively on zero-day exploits, Joyce added. Most intrusions occur via easier vectors: e-mail, web sites (using techniques such as waterholing—infecting web sites that are frequently accessed by users at the target organization), and removable media like USB drives. Joyce noted that you can't rely on users not to click, even with the best security policies and education (see my Irrefutable Laws of Information Security in Chapter 1), so you need technical controls that will prevent the execution of malicious code.

Once advanced attackers have established a beachhead, they try to steal credentials that enable them to maintain a presence, install tools, and move laterally to the prized assets they seek. Techniques such as segmenting the network, limiting administrator privileges, and forcing two-factor authentication can help make this more difficult. Joyce also said that he liked some of the new ideas emerging from the industry such dynamic privileges, which is analogous to the granular trust model described in Chapter 7: the level of access provided depends on factors such as the device you're using and your location.

Finally, he stressed the need to continually evaluate and improve your defenses. An organization with static defenses will drift to the back of the herd, where it is easily picked off by a predator (see Irrefutable Law #6). "Don't be that easy mark," he said.

Playing War Games

I like to conduct war games a few times a year. War games are intense role-playing exercises in which employees take on the role of attackers and attempt to compromise key assets using any feasible methods. I have found war games are particularly valuable for analyzing threats that may have major consequences but whose vulnerabilities are not well understood.

This technique provides the most comprehensive method of assessing threats to key assets because the people playing the role of our adversaries are essentially allowed to use any method to achieve their goals. However, because of this, it is also resource-intensive and should be used selectively.

Typical war games that I have overseen take one and a half days and may involve eight to ten staff from a variety of roles, such as factory workers, business process leads, salespeople, and technical experts. Some war games can take much longer; in *Wargaming for Leaders*, written by wargaming experts at management-consulting firm Booz Allen Hamilton, (Herman, Frost, and Kurz 2009), the authors discuss games that may last weeks and involve many more players across an organization.

A typical game focuses on a specific target or scenario, such as disabling a key facility or stealing trade secrets. You can use war games to examine potentially catastrophic events that have a low probability of occurrence, but a high probability of causing damage if they do occur. Team members are instructed about the threat agents involved and draw on archetypes from a threat agent library or descriptions provided by the game architect. Led by a facilitator, the team takes on the attacker's perspective and postulates ways to achieve the attack's objectives.

Because the team can propose any attack method, they often identify risks that might be overlooked using conventional methods. As the authors of *Wargaming for Leaders* put it, "We create the environment, the players engage, and what comes out of team play often surprises and even stuns everyone involved." For example, a malicious group might attempt a devastating attack by purchasing a small but essential technology provider and inserting malware into their products in order to infect their customers. After each game, security analysts examine the results to determine how to address newly identified vulnerabilities.

I also like to examine the cyber consequences of large physical events as part of disaster recovery planning. These could include earthquakes and tsunamis that damage data centers, or even solar flares that disrupt the communications that the business relies on. Exercises can include drills that last a day or more.

A large organization can justify the considerable effort involved in conducting these exercises because of the enormous potential benefit of mitigating the threats. In fact, some organizations hire professionals to create and facilitate these games. Booz Allen Hamilton, for example, has an extensive war gaming practice covering diverse subject areas including market dynamics, cybersecurity, geo-political events, and even real war scenarios.

But smaller organizations can also benefit by considering extreme events and formulating response plans. If you prepare for the extreme, you'll be more prepared to deal with everyday events. Planning doesn't need to be as resource-intensive as a full-blown war game. It can be as basic as bringing team members together to discuss likely scenarios and responses in a shorter tabletop exercise lasting just a few hours. This method enables members to get a feel for what it would be like to work together in the event of a real disaster. Considering these extremes can also provide motivation for introducing simple yet effective measures to reduce the risk that catastrophes will occur. You might realize it is worth increasing investment in user education to reduce the risk of social engineering compromises, or becoming more diligent about analyzing logs and network traffic to identify patterns that indicate suspicious activity.

Trends That Span the Threat Landscape

I've described some of the methods that can be used to analyze emerging threats. Now I'd like to turn to some key themes that have emerged from such threat analysis. These themes paint a broad-brush picture of threat and vulnerability trends spanning multiple technologies across the threat landscape.

Trust Is an Attack Surface

As the technology industry erects new technical defenses, attackers seek to bypass these controls by exploiting user trust, typically using social engineering techniques such as phishing.

If an attacker can win a user's trust with a sufficiently convincing e-mail or fake web site, the user will make it easy for the attacker by clicking a link or downloading a file. These actions usually undermine even the most rigorous system-level controls, initiating a chain of compromises that ultimately can result in major damage.

Whenever users place their trust in a new technology, attackers quickly follow. Studies have shown that users trust social media services more than other information sources. A user is more likely to click a link if it appears to have been sent by a social media "friend." Exploiting this trend, attackers have spread malware via social computing circles of trust such as friend networks.

Attackers have also been quick to take advantage of the trust users place in their smartphones and in other appliances such as game consoles. The exploitation of trust also extends to the relationships between systems. Once configured, communications

between systems often operate autonomously, without manual oversight. Smartphones are set to automatically update applications from trusted app stores; other systems blindly trust firmware updates and dutifully install them. This automation provides convenient opportunities to insert malicious code, abusing trust without the need to directly involve the user.

In the near future, I anticipate trust will become a commodity that is bought and sold. The digital reputation of systems and services will become critically important. In the past, tokens of trust, such as digital certificates and social computing credentials, were stolen for immediate use. In the future, they will be stolen so they can be sold in underground markets. The value of these tokens depends upon the access they grant and the other circles of trust they can be used to penetrate. Already, attackers are using stolen digital certificates to sign their malware in an attempt to avoid detection by operating system defenses.

I expect social engineering attacks will continue to present significant risks because they exploit human weaknesses and will adapt to take advantage of new technologies. So we, as security professionals, need to focus on the role of users as part of the security perimeter, as I discussed in Chapter 5. To reduce the risk to the enterprise, we need to make users more security-aware and influence them to act in more secure ways. But it's also important to note that a successful phishing exploit is also ultimately a technology failure that allowed malicious code to execute.

Barriers to Entry Are Crumbling

Our adversaries gravitate toward the path of least resistance. They tend to select targets that are easy to access and analyze, and they typically use the most readily available and cheapest tools.

They are much less likely to use methods with high barriers to entry such as the need for specialized expertise, expensive hardware or software, or access to extensive compute capacity. However, several of these barriers have begun to crumble as a result of trends such as cloud computing, lower-cost communications components, and commodity malware toolsets. This trend ultimately is likely to result in new types of attack.

A key factor is that security researchers are sharing not only their knowledge but also the tools they design as part of their research. Recently publicized tools, such as rogue base stations and Bluetooth sniffers, provide attackers with more accessible, low-cost ways to intercept network traffic. Researchers have uncovered vulnerabilities in femtocell devices (miniature, low-cost cell towers) that can be used to take control of the devices, lowering the barriers to attacks targeting cell phone data traffic.

Ultimately, lower barriers to entry mean increased risk to enterprises. However, because several of these areas are still at the research stage, it will take time for them to mature into active exploitation.

The Rise of Edge Case Insecurity

Each day, the environment becomes more complex with millions of new devices, each running its own operating system and collection of applications. This complexity generates new edge cases—problems or situations that occur only in unexpected or extreme situations.

Edge cases can include unlikely interactions between two familiar objects. A hacker team recently demonstrated that, with a popular smartphone, a paperclip (used to pop out the phone's SIM card at the critical moment), and a little patience, it's possible to gain access to contact information, phone call logs and voice mail, e-mails, and other information stored on the phone.

Overall, the growing number of third-party plug-ins and widgets introduce edge cases that are hard for developers to anticipate even if they use secure design techniques.

Interoperability between programs has resulted in a new category of hybrid attacks where malicious objects are concealed in innocent-looking ones to thwart detection. One proof of concept in 2011 demonstrated it was possible to conceal a fully functioning Trojan in an e-mail plug-in.

Some of these hybrid attacks have shown they can circumvent new security features. As web browsers and search engines try to protect users from malicious links, attackers are responding by hiding links in image search results, where they cannot be detected using standard tools. Research into network intrusion methods has discovered over a hundred methods of evading detection by manipulating traffic to remain functional but undetectable by typical tools.

There is no silver-bullet solution for eliminating edge-case insecurities. It's unlikely that even the most rigorous testing could ever uncover them all. The best approach may be to exercise caution when adopting new technologies with the potential to generate edge cases.

The Enemy Knows the System

The technology industry has often relied on security through obscurity: the idea that if attackers can't see the insecurities in code or other technology, they won't exploit them.

Over time, it has become clear that security through obscurity is poor security. To quote the maxim coined by Claude Shannon, one of the founders of modern computing, "The enemy knows the system."

It's now relatively easy for attackers to get access to the same tools enterprises use, such as web hosting services and smartphone application development tools. Hackers can now more easily engineer malware and attacks that take advantage of these elements. The fact that static platform controls tend to become less effective over time (one of the Irrefutable Laws of Information Security noted in Chapter 1) is partly due to the ability of malware authors to pretest their malicious code against technical controls. They can do this by obtaining code from malware repositories that have already been tested against existing controls, or by actually purchasing the technical controls.

Even the success of social engineering demonstrates that the attackers' knowledge of the target greatly increases the likelihood of successful deception. Today, competitors and other threat agents learn a great deal about a company and its employees by simply searching information publicly available on web sites or social media accounts.

Because we cannot assume insecure technology is safe just because it is hidden, we need to design with security in mind. The ineffectiveness of security through obscurity is also an argument in favor of standards and open-source solutions. This idea may initially seem counterintuitive, but the fact that open source is exposed to public scrutiny requires it to be secure. At a minimum, we should ensure devices are rigorously tested against industry standards because the attackers will do so.

Key Threat Activity Areas

Threats are evolving in many technology areas, from embedded systems to cloud computing. I'd like to discuss a few areas experiencing significant developments with implications for enterprise IT.

The Industry of Malware

Malware has become a profitable industry that increasingly resembles the legitimate software market, with market leaders, mergers, licensing agreements, real-time support, and open source. The organized business activity in this market reflects the extent to which well-crafted malware has become a viable career pursuit for members of the criminal underground.

Today, malware development and malware use may in some cases be distinct activities carried out by different groups or individuals. Malware authors are producing standardized toolkits, which have made life much easier for would-be attackers. These attackers can now simply buy or acquire a toolkit rather than expending the effort to identify vulnerable web sites and develop their own exploits.

The Zeus malware family provides a useful case study showing how complex this industry has become and how hard it is to accurately track developments. Sold mainly in underground forums, Zeus has been used extensively for theft by creating botnet nodes. During 2011, a code merger was reported between Zeus and another popular crimeware kit, complete with assurances of future support for the customers of both products. Around the same time, Zeus toolkit source code was made publicly available. Since then, multiple new variants have appeared and been used for a variety of attacks. At one point, security researchers attempting to monitor Zeus exploits discovered a server they believed was the hub of a Zeus botnet. However, the server was the equivalent of an espionage honey pot, allowing the botmasters to turn the tables by spying on the researchers who were attempting to analyze the hub.

Ransomware has also become a profitable activity for some organized crime elements. Ransomware was mostly at the validated proof of concept stage when I wrote the first edition of this book in 2012; it has since progressed to active exploitation with some commoditization. Today's ransomware exploits typically exploit system vulnerabilities using Trojans and other methods, then lock or encrypt information so users cannot access it and hold people and organizations hostage until they pay. In February 2016, a Los Angeles hospital paid a ransom in bitcoin after staff were locked out of the hospital's own network for more than a week; during the same month, one ransomware variant was reported to be infecting more than 90,000 PCs per day (Fox-Brewster 2016).

The Web Expands to the Internet of Things

The Web continues to present a huge attack surface. And this attack surface is growing rapidly as it expands to include the Internet of Things, encompassing nontraditional devices such as appliances and control systems, cars, wearable and medical devices, and the "smart" grid. Each of these is a potential source of risks.

Recent headlines have highlighted the growing threat activity focused on IoT. Researchers hacked into a Jeep via its Internet-connected entertainment system and remotely controlled the vehicle's functions, including turning off the transmission and brakes while someone was driving (Greenberg 2015). Other researchers showed that thousands of devices in hospitals are vulnerable to attack, including x-ray machines, MRI scanners, and drug infusion pumps, partly because medical equipment is increasingly connected to the Internet so that data can be fed into electronic patient records systems (Pauli 2015a). Yet another researcher demonstrated the ability to hack into FitBit fitness trackers via Bluetooth (Pauli 2015b). Many IoT devices, including cars, wearables, and home appliances, include wireless capabilities, so exploitation doesn't require a physical network connection.

Clearly, we should expect continued growth in IoT threat activity. However, should be noted that the activity to date has generally has been at the research or early proof-of-concept phase (see Figure 6-1). As the IoT expands and matures, we will see a progression to advanced active exploits over the next few years; given the rapid pace at which IoT is evolving, if companies don't use good privacy and security design principles when building their products, the time from research to active exploitation could be much shorter than has typically been the norm.

Many embedded devices that are already installed in businesses are similarly vulnerable. Companies have a history of deploying specialized devices without adequate security controls, often because of the perception that specialized devices are "dumb" and do not have a full set of capabilities. In reality, the opposite is often true: devices marketed for a specific function are often capable of much more. Printers contain processors, use wireless connections, and may be capable of acting as file servers, for example. As a result, embedded devices can introduce as much risk, or more, to an organization as a traditional computing device since they lack security controls and administrators are generally unaware of the danger. New devices may be vulnerable to new attack methods: recent research showed that the sounds 3D-printer nozzles make as they cross the machine bed can be recorded by smartphones, analysed, and then used to duplicate prototypes (Nelson 2016).

The vulnerabilities in embedded industrial control systems were exposed by the widely publicized Stuxnet malware, which was used to sabotage the systems that supported Iran's uranium enrichment capabilities. The incorporation of computer-based control and automation technology into the existing electrical power infrastructure—resulting in the "smart grid"—is another source of potential vulnerabilities. The US government has warned of increasing threats to the grid, noting that many embedded systems lack adequate security controls and are susceptible to known techniques such as cross-site scripting attacks (US GAO 2012).

We might also see logical attacks as precursors to physical attacks. On a macro scale, a nation state might attack another nation's cyber infrastructure before staging a physical attack. This approach might also be applied at a more personal level. A burglar might remotely disable an Internet-connected alarm system before sneaking into a house, or perhaps even use the system's video cameras to watch the owners and note when they leave the house unattended.

Here are two more potential future IoT scenarios in which innovative technology designed to do good could be exploited for harm, unless designed with strong security and privacy protection. Last year, doctors for the first time inserted an artificial "eye" that enabled a blind person to see. The device is a retinal implant that receives signals from a video camera integrated into eyeglasses. Think ahead a few years, to a time when the implants are more sophisticated and can see in much higher resolution, and also include

software to automatically interpret visual information, such as QR codes. Then imagine that a malicious actor creates a QR code that triggers the vision system to download malware. Like the PC malware that paralyzed Sony's network in 2014, the malware then demands a ransom to re-enable the person's vision. Now consider the example of a cement company that's embedding sensors in the concrete mix used to build a new road, thus enabling local authorities to monitor traffic patterns and adjust signals to optimize the flow of vehicles. If the technology is not securely designed and implemented, all that a malicious person needs is the ability to execute malicious code, in order to falsify the traffic pattern in such a way that vehicles converge on the scene of a planned bomb attack.

Smartphones

Smartphones are attracting almost as much malicious interest as desktop and laptop platforms. However, even though smartphone sales have outstripped PC sales, smartphone malware isn't yet as prevalent as PC malware and doesn't cause the same kind of widespread damage. That's partly because most valuable corporate and personal data is still held on PCs and servers. Another factor is that smartphone vendors have somewhat greater control over applications, since users generally access them via vendor-controlled app stores.

Just as in legitimate software markets, malware authors are likely to maximize the value of their code by using tools that allow their software to run on multiple devices. They are increasingly targeting applications, a trend also seen on other platforms. Attackers have purchased copies of applications, incorporated their malicious content into the otherwise legitimate software, and then redistributed their code under a new name or as a "free" version of the original. On one smartphone platform, autodialing malware was found in more than 20 applications. Variations of a Trojan were found in dozens of applications and are believed to have been downloaded by at least 30,000 users.

A further development is the use of smartphones as bridges to traditional networks, resulting in the potential for enterprise network attacks that originate from within mobile networks.

In the future, we could see greater exploitation of location-based services to deceive users. Because smartphones contain location sensors such as Global Positioning System (GPS) chips, knowledge of the phone's location can be used to present targeted ads and useful information. For example, a user in a supermarket aisle might be presented with online coupons for products on nearby shelves. But this information could also be exploited to present fake coupons that are all the more convincing because they suggest that the sender knows the user's preferences.

Attackers could also exploit other smartphone capabilities to take advantage of the fact that the devices are carried into confidential meetings and other highly sensitive situations. Imagine being able to remotely control a device that has a microphone, a camera, or other recording capabilities. Or think about a vulnerability in any of the popular web-conferencing services that people use for confidential discussions and to exchange information.

Current trends in the mobile platform space indicate that attackers are most interested in stealing personal data. This trend is partly due to the increasing use of smartphones for financial and banking transactions, which provides new opportunities for identity thieves and other criminal groups. As a result, it is now important that smartphone hardware and software developers focus on protecting personal data.

Software developers should adopt the same discipline and commitment to following secure design principles as traditional platform developers. Today, more and more people are becoming app developers, creating software, and posting it online for others to use. One has to question how much security testing and validation has been applied to these applications. As users move more of their everyday activities onto smartphones and other small devices, the consequences of poor or insecure designs will have greater impact on individuals and their employers.

Web Applications

Web applications, primarily comprising client browsers and server-based applications, continue to be heavily attacked. Threat analysis indicates that this area is experiencing full exploitation activity and moving toward commoditization. There is also considerable research in this area, suggesting the number of attacks will continue to grow.

Attackers have adopted new techniques to hide their intentions and deceive users long enough to achieve their aims. As web browsers and search engines try to protect systems from malicious links, attackers are instead obfuscating their links in image search results, where they may not be detected.

Techniques for hiding messages within images have been used within the security realm since long before the invention of information technology. Now, this technique, known as *steganography*, is being used to hide malware and botnets on publicly used image hosting sites.

Search poisoning has also become a common method. Attackers using search poisoning tend to focus on events and topics of popular interest, optimizing their web pages to achieve high search engine rankings. After a search query, the victim clicks a link among the search results. They are redirected multiple times and eventually land on a page that is used as a vector to deliver malware.

Conclusion

In this chapter, I've outlined some of the real threat trends and described methods information security groups can use to analyze the threat landscape as it continues to evolve.

No doubt, new and more sophisticated types of exploitation will continue to emerge, and we need to stay aware of them. As Mustaque Ahamad, director of Georgia Tech Information Security Center, noted in 2011, "We continue to witness cyber attacks of unprecedented sophistication and reach, demonstrating that malicious actors have the ability to compromise and control millions of computers that belong to governments, private enterprises, and ordinary citizens."

Yet, as we try to make sense of the deluge of news about attacks and vulnerabilities, it's essential to retain a sense of perspective. Most threats do not take place using exotic, obscure methods. Instead, they take the path of least resistance, exploiting well-known vulnerabilities. Therefore, business can mitigate many of these threats by implementing basic, established security measures. To put it another way: when you hear hoof beats, think horses—not zebras.

Social engineering will continue to be a key attack method because it takes advantage of user trust and is hard to prevent using technical controls. Therefore, as I discussed in Chapter 5, we need to continue to focus on educating users to become more security-aware. By doing so, we can reduce the risk to the enterprise.

Ultimately, while doing our best to prevent compromises and breaches, we must remember we cannot control the threat actors and their exploit attempts. Because all threat categories use malicious code in some way, advanced preventive tools that effectively stop the execution of malicious code can greatly reduce the potential of compromise. But all organizations face the possibility of some level of compromise, making defense in depth as essential as ever. Losers ignore the trends. Winners survive by being able to predict, prevent, detect, and respond.