New Paradigms in Incident Management

Tom Perrine & Abe Singer

tep@sdsc.edu
+1.858.534.8328

abe@sdsc.edu
+1.858.534.8305

The San Diego Supercomputer Center
University of California San Diego
9500 Gilman Drive
La Jolla CA 92093-0505

Abstract

Our current security paradigms are almost entirely drawn from the technical work in the area in the last four decades. This legacy of protecting data and systems has provided a rich set of tools for preventing, discovering and recovering from security failures. As we have gotten better at detecting security incidents, whether they are successful or not, we have not seen an equivalent increase in our capacity for analyzing and tracing these abuses to their source. In fact, it is not uncommon for a large site to be able to detect many more intrusion attempts than it can analyze and trace in any timely fashion.

In short, we are being out-scaled by the intrusion community. We need incident management tools that inter-operate, scale, and decrease security investigator workload. Also, due to the multi-site (multi-country, multi-cultural!) nature of investigations, we need better ways for sites to communicate about security incidents, past and present. Secure software agents or other technologies are needed to allow the small number of qualified investigators to extend their reach to remote sites, and also to provide a mechanism for educating more investigators.

This paper is an attempt to identify, at a minimum, some of the new technologies that we will need in order to address these issues. In some cases, we need refinements, or wider deployment of existing technologies. In other cases we need completely new tools and methods of working.

This paper created quite a lively discussion at the workshop. It became obvious that, if anything, the original paper was too cautious in some of its recommendations, and was not as aggressive in suggesting more radical new paradigms. Rather than re-write the paper based on the discussion, we decided to follow the excellent example of Greenwald[1] and add an epilog that incorporates the discussion comments and our "second thoughts" based on those comments.

1 Introduction

The Internet community continues to see acceleration in the rate of reported Internet security incidents. Not only is the number of incidents increasing, the rate at which they occur is increasing as well[2]. This is causing a problem of scale in attempting to respond to incidents. Responding to Distributed Denial of Service (DDoS) attacks and some other incidents require near-real-time response and easy but trusted communication between sites. These requirements call for inter-operability and automation of tools, and an architecture for integrating them. The problem of developing these technologies is not currently being addressed comprehensively in the security community.

We suggest that what is required is the development of an architecture and tools for scalable management of security incidents. This architecture should include standards for interoperability, inter-site communication, and agents for remote data gathering, investigation and analysis. Developing such an architecture raises numerous issues which require discussion in the security community. These issues include trust, legal jurisdiction, privacy, liability, and policy. Also, developing distributed, inter-operable (standards-based), trustable technology raises interesting questions. Finally, the over-arching issues of scale—what it means and how to effectively scale technology and human resources- will be of great interest.

2 Background

Operational (as opposed to research) computer security has evolved in several independent directions over the last decade. We have seen an attempt to shift from pure "response" (intrusion detection) to "prevention"(source code audits, string configuration management). Our capabilities have increased: from securing and monitoring single hosts, we have "graduated" to dealing with multiple nearly identical hosts within a site, and are heading towards being able to successfully manage and secure heterogeneous collections of hosts within a site. However, almost all multi-site problems remain beyond our grasp.

Also, after a slow painful period, security tools are beginning to evolve and mature. Unlike five years ago, today one can actually purchase security software instead of having to create their own. In addition, the separate "point solutions" of firewalls, "intrusion detection" and vulnerability scanners are being combined into "product suites" that cooperate and inter-operate, as long as all the tools come from the same vendor. There are emerging draft standards for exchanging "intrusion alerts" between "sensors" and management stations[3].

These trends are cause for optimism but are creating both new problems and new opportunities. More opportunities than we can manage, unfortunately. Despite the growth in our abilities, we are still being "out-scaled" by the attack and intrusion methods, tools and sheer number of security events. Wide-scale, remote controls Distributed Denial of Service (DDoS) tools are only the opening salvo in wide-scale attack tools. Combine this technology with lessons from the distributed crypto-cracking[4] and other "wide-scale cooperative computing[5] projects and you have large-scale password cracking, mobile-agent intrusion tools, "robotic" automated intrusion worms and the like. Primitive versions of some of these tools are showing up in the wild even now.

These tools are "open source" within the black-hat community leading to propagation and mutation like computer virii. This makes detection using passive network monitoring problematic at best and impossible at worst. As our detection methods have improved, the number of incidents requiring investigation has increased. There are actually more probes, and we are capable of detecting more kinds of probes. This has led to capable security monitoring becoming its own "denial of service" attack on staff time. In essence, we are making more work than we can handle.

3 New Paradigms

What is needed is a revolutionary jump, instead of the current slow evolution. We need to leapfrog over the current situation into one where the defensive side actually has more capabilities and capacity than the attack side. This "break" will require, in part, new paradigms in security tools and capabilities.

So, what are the new paradigms that we need to address? Here is a partial list of new technologies, issues and interesting questions that bear on this question.

- Incident management software - smart "assistants" or "wizards" that can be programmed to either investigate the simplest probes and intrusions, or assist a human in performing analysis of more complex incidents. This addresses the sheer number of simple probes, which usually require only a simple response, and also helps develop a true forensics process for incident analysis.

- Mobile agents to assist in collecting incident data and to make near real time tracing of intruders practical. Currently, a major problem is the widely varying monitoring capabilities of various network service providers. During an investigation, as soon as you trace back to a site that has no monitoring capability or knowledge, the trail is broken. Mobile agents that can easily be downloaded and installed by less-capable sites, but then be securely remote-controlled by an expert will allow the trace to continue and additional evidence to be gathered. Research at UCSD[6] is developing some of the secure and trustable agent infrastructure that will be needed to solve this problem.

- Inter-site cooperation - How can multiple sites running heterogeneous (or no) security software exchange incident data? Aside from the technical issues, what are the policy and legal implications? When there are multiple legal systems involved, how can cooperation with law enforcement agencies from multiple countries be accomplished?

4 Why is this a new paradigm? Why is this interesting?

In the past the focus has been on having a human respond to an alarm of unknown severity and urgency. We are proposing a management approach, with prevention, documented procedures, and automation to assist humans in responding (when necessary).

Using existing vulnerability scanners plus local automated configuration management, such as the open source cfengine, or Microsoft's SMS, coupled with better user authentication (such as Kerberos, SSH and SSL-based methods) provides a significant level of prevention. However, once it has been determined that an incident has occurred, we are immediately back to a manual, labor-intensive, poorly defined process. There are at least two new technologies that would ameliorate this situation. The first is a semi-automated "investigator's assistant", the second is a "remote control" agent that allows an expert to assist in real-time.

Using a software tool to guide a non-expert person through a well-defined incident analysis process would be a significant improvement. The current situation, requiring an expert using primitive tools to proceed through some ad-hoc process, does not scale, and is difficult to reproduce, teach or improve. Providing some automated guidance in the form of a software tool would allow process improvement, better education and distribution of the collective knowledge-base of security experts. This begins to move computer forensics towards the more traditional true forensic sciences; a necessary step to begin to better interface with the legal system.

Using mobile agents as part of a semi-automated process allows those who do not already have the appropriate security tools to be led through the investigation process, by secure remote control, over the network. A person at a highly-knowledgeable site A, while tracing an intrusion through a less-knowledgeable site B, could have site B install the security investigation agent software over the Internet, to be controlled by the expert at site A.

All of these new approaches require cooperation between sites. It is a fact in today's Internet that all intrusions will involve multiple sites, usually in multiple countries. Deciding to use any security software requires the user to place some trust in the software's authors, and the method of acquiring the software. Using mobile agent software as we propose requires a very high level of inter-site trust and cooperation. The issues of trust, liability, due diligence, site and organizational policy, and law will all need to be addressed. This level of trust requires technical and social initiative to succeed, perhaps by legislation in most jurisdictions.

5 Conclusion

The security problems of the Internet will continue to get worse before they get better. The increasing number of sites, the decrease in the average security capabilities of these sites, and the increasing sophistication of intrusion and attack tools guarantee this. The problem will worsen at the high end (more sophisticated attacks) and at the low end (ever larger numbers of probes and unsuccessful nuisance attacks).

The current situation does not scale. Proper incident management requires highly trained security staff to investigate and analyze incidents. These people are rare and expensive. There will never be enough to go around. The answer is to leverage their contributions through software assistance: incident analysis tools that embody their expertise (and handle the simplest problems with no human input) and mobile agents that extend their reach.

Since these experts, and their software, will need to affect computers "owned" by others, and because all incident investigation requires data to be gathered at multiple sites, the issues of inter-site (actually inter-organizational) trust must be addressed. These issues are not technical, but social in nature, requiring new paradigms in social, technical and legal interaction.

6 Epilogue

As mentioned in the abstract, we decided to leave this paper practically untouched after NSPW 2000, with the exception of this section. From the comments and discussions, it seemed that some of our arguments were more compelling than others. The presentation we made was not the presentation originally written for the paper. The presentation was changed dramatically, based on the discussions of papers that were presented earlier in the workshop. It seemed that some of our topics were more or less important than we had originally thought. This epilogue is a summation of our revised presentation and the discussion comments. Where possible, we have attributed comments and thoughts as recorded by NSPW's incredibly able scribe, Bob Blakley.

6.1 The Threat Triangle

Our first slide, the threat triangle (Figure 1), was intended to call attention to the varying sophistication of would-be intruders, as well as the number of miscreants at each level of sophistication. It also reinforces the ideas that not only must you protect against the "high-end hacker" but you must also be prepared to endure tens, hundreds or perhaps even thousands of "low-level" attacks. Even though those low-end attacks will hopefully be unsuccessful, they are still incidents worthy of some attention and will require some resources to address. The discussion at this point turned to the issue of attacks against "end users". In the past, attacks were primarily against service providers, such as companies, organizations and Internet Service providers. With the proliferation of broadband to the home (cable modems and DSL), we are seeing many more attacks against unprepared, unsophisticated home users. Tom Daniels and Mike Williams reported that their home systems are scanned regularly (sometimes more than once a day), and that some of these scans are from their own service providers. It is not clear if this is a service, or signs of an intrusion at the service provider. This correlates with similar anecdotal evidence from SDSC and UCSD's home users.

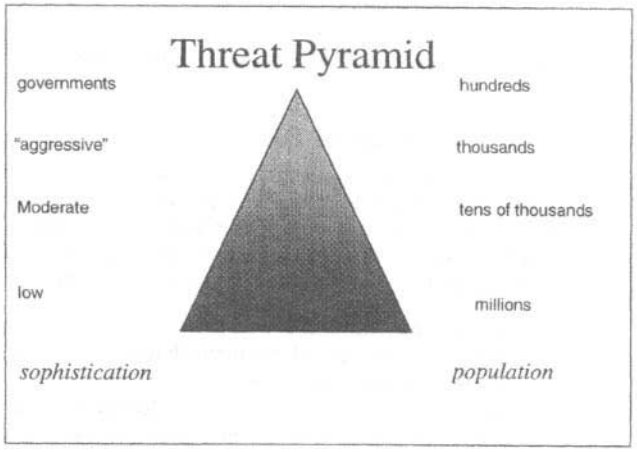

Figure 1. Threat Pyramid

The pyramid can also be used as a representation of the "hacker food chain", where sophisticated hackers at the top of the chain produce tools which are eventually traded down the pyramid to the unsophisticated. Bob Blakley pointed out that this is really the opposite of a food chain, as the predators are at the top. It's really more of a waterfall with tools flowing from top to bottom.

The pyramid can also be used to represent the current shortfall of skilled, seasoned "veterans" in the system administration staff pool. Instead of system administrators with years of experience in systems programming and security lessons learned "the hard way", we have administrators who have followed the career path from "graphics designer" to "web master" to "system administrator", often all within one or two years. Tom Daniels reported that he is working on getting Purdue's associated technical schools to teach courses in security and system administration. We pointed out that UCSD's Extension has finally started teaching security classes, but that it took us several years to get this started, and we had to write and teach the classes ourselves to get them into the catalog.

One interesting side topic led to a discussion of "protocol Darwinism", and the similarities between the current security environment and an ecological system. If the hackers are predators and the victims are herbivores, will we see (are we seeing?) a stable ecology? And since domestication seems to be a very successful strategy for herbivores, what does this suggest about our future strategies? This theme would come up later, as well.

Our second slide included what we see as the underlying problems in the current state of security practice:

- the cost of defense is always greater than the cost of offense, especially in the arenas of mass scanning and Distributed Denial of Service (DDoS);
- the hacker food chain (which should be the hacker "tool chain") promotes more sharing and faster communications than the defensive community at present
- the need to defend all systems against all threats balanced against the attacker's simpler goal of finding any one vulnerability on any one host

All of this leads to our basic premise: Scale or Die.

6.2 Scale or Die

It seems obvious that our current methods are not scaling well enough to prevent the increase in the rate of security incidents. But what does it actually mean "to scale"? In our presentation we argued that there are at least three things that need to scale, in order to turn the tide:

- Knowledge – information on current and past vulnerabilities, current popular exploits, how to perform a proper intrusion analysis, etc.
- Activity – the limited number of skilled security experts need to be able to leverage their knowledge and be able to perform some analysis functions on computers not normally under their control while in "hot pursuit".
- View – currently analysts are limited to seeing only activity at their own site, or activity that has been publicly reported. This limits their situational awareness to only a small part of the overall intrusion activity that might concern them and virtually prohibits correlation of wide-scale hacker activity.

To address these areas, we proposed the following new tools and technologies.

SCALE	TOOL
Knowledge	Incident Management Co-pilot
Activity	Mobile Agent
View	Inter-site Cooperation

6.3 Incident Management Co-pilot

As we described it, the co-pilot would act as a limited expert system to assist a user in the analysis of a security incident. The overall strategy is to use the ideas of "open source" to capture the combined knowledge and experience of many security experts into a single tool. Various incarnations of this technology could be produced, ranging from a simple "smart checklist", to some form of expert system "software guide". Any such tool should have the ability to capture the decision tree of a trained analyst for use by others, as well as the actions taken by an untrained user, which would allow the software to be improved based on actual end-user experiences. Any such tool should also capture the investigative process, as this would be an aid in any prosecution.

Tom Daniels pointed out Pascal Meunier's report management database at CERIAS, which is tied to the Common Vulnerabilities and Exposures (CVE) Database [7]. The CVE and the report database (or their successors) could provide the beginnings of the subject matter knowledge base required for the co-pilot tool.

The issue of scale was raised again, as John McHugh reported that CERT/CC is still seeing a doubling of attack reports every year. He mentioned that although CERT/CC does perform triage on incoming incidents, all the originally reported data is being kept in a "data landfill". The data is not organized, but could be searched, or a curator could organize it. CERT/CC is beginning some "data mining" activities on the "landfill" but there are no results yet. He also pointed out that in most cases the attacks reported were using well-known vulnerabilities and that in many cases the rate of reports of attacks goes up **after** patches are made available. John McDermott pointed out that CERT triage may cause people to stop submitting reports.

A long discussion ensured, much of which revolved around all the various sources of raw data that could be incorporated into a co-pilot tool. We also got back into the ecosystem discussion, which seemed to be a recurring theme. Brad Wood pointed out that current security efforts seemed to be breeding a "super bad guy", we are only catching the slow and stupid, and the smarter specimens are surviving long enough to breed (teach others their superior techniques). There are even elements of protective coloration, as lots of little attacks are background noise, and the superior attacks proceed under cover, with few consequences for getting caught.

Victor Raskin pointed out that perhaps if there are enough predators, they might attack each other. We pointed out that there was already anecdotal evidence of this happening; in 1985 a university was over-run by competing hacker groups, which spent much of their time trying to lock the other groups out of the networks. This is also seen in the Attrition.org defaced web page archive [8], where competing groups will deface and re-deface victim web pages in attempts to claim technical superiority.

We commented that AI and "expert systems" seemed to have fallen out of favor as neither had lived up to the hype of the 80's and 90's, but that this seemed to be more a case of over-hype instead of a true technical failure. Bob Blakley and Victor Raskin pointed out that expert systems were actually the part of AI that "worked". Limiting systems to a reasonable problem domain seemed to be the key. They pointed out that "knowledge engineering" is still hard, but the more limited form of just eliciting knowledge from expert humans is tractable and should be sufficient for this limited case. Tom Daniels pointed out that there had been successful expert systems, especially in the medical field, in well-specified problem domains.

6.4 Mobile Agent

Our presentation of a mobile software agent to allow experts to perform security investigation and analysis functions at a remote site elicited a discussion that avoided the technical hurdles and jumped straight to societal and legal issues.

In our incident investigation experience, it is very common to trace an incident to a service provider or site that has little or no security awareness, knowledge or infrastructure. In fact, at least twice we have called a compromised site to report a problem and gotten the response, "I have no idea what this means, but if I give you all my passwords, can you fix it?"

Our mobile agent is an attempt to address this problem. In general, we could provide a mobile agent via a web page; "click here and I'll help you." This agent would install itself on the remote site's computer(s) and be controlled by a skilled investigator. In practice, this is similar to Back Orifice and other "remote control" software, although hopefully more secure and more trustable. Presumably it could perform security critical functions such as network monitoring, disk forensics and perhaps contain a network or host vulnerability scanner.

All participants seemed to be confident that a secure, trustable mobile agent could be produced, and that most of the interesting questions were not technical in nature.

Mike Williams pointed out that not only were we expecting the remote site to trust us, but we were also trusting them. This could be problematical since it is sometimes the case that the victim organization might itself be engaged in dubious or criminal activity. He also pointed out that security technology providers (e.g. anonymizers) might in a legal sense be seen as co-conspirators in criminal acts.

We spent a fair amount of time discussing the possible legal liability issues. One thing that became apparent was that there is a crying need for basic legal training for the computer security community. Although we were able to formulate lots of legal questions and theories, we were often undecided or disagreeing on legal principles, practice and law. When we got to questions of multiple jurisdictions and international law we were obviously in new territory for almost all of us.

Cynthia Irvine asked why we didn't just use a prepositioned agent for remote control instead of a mobile agent. We had considered this, but it wold require a global infrastructure to be pre-installed and waiting to be activated. In practice, the sites that would be least likely to pre-position such an agent are those that we encounter the most often: those that have no security plans, infrastructure or tools. It seems that in this area, tools that can be deployed when and where needed are more useful. This also lead into an additional discussion of "scale to oppression".

During our work examining the Federal Bureau of Investigation's Carnivore program Steve Bellovin reinforced the danger of security tools that could be "scaled up to oppression". This could also be a concern with a widely deployed investigative infrastructure, especially if that infrastructure could be taken over by a national-scale effort. This concern is an additional argument for using ad-hoc tools that require some non-trivial manual effort to install.

6.5 Inter-site cooperation

All of the discussion of a mobile agent led back to issues of trust and inter-site cooperation. One recurring theme in all our security efforts has been the need for better inter-organizational cooperation. In today's security environment, just detecting and analyzing attacks against your own site is too narrow a view to be truly effective.

For example, a port sweep against our supercomputer center networks is only mildly interesting. But a sweep against multiple .EDU supercomputer centers, followed by a sweep of DoD supercomputer centers, followed by a sweep of DoE supercomputer centers would be very significant. It is possible, perhaps likely that we are missing nationally significant threats due to a lack of communications among organizations. This is an issue that should be addressed through the Critical Infrastructure Protection programs.

We suggest that the current efforts to standardize on "alert" formats[3] are a good start, but are at too low a level to address the larger issues. These proposed standards allow low-level alert data to be communicated, but are too simplistic in their current forms. We suggest that an interesting set of messages to be communicated might be as follows.

- Have you seen X? (X is a host, subnet, site, person or exploit)
- Has Y probed you?
- What vulnerability was used to compromise your system Z?
- I am under attack. Are you? From who?

We suggested that there are many barriers to cooperation, even on this most rudimentary level. Most of these barriers are non-technical. For example, one barrier is the military community's desire to classify most incidents involving their systems. Another barrier is the desire of companies to avoid disclosing any information concerning intrusions into their systems, either to avoid embarrassment, shareholder lawsuits or other legal liability. In many cases, however, both these communities would be more likely to contribute information concerning incidents if the data could be sanitized to remove all references to their organizations. Such sanitization would have to include IP addresses, host and domain names, user names and perhaps even timestamps.

Bob Blakley pointed out that Cliff Stoll[9] would have caught his hacker much earlier if there had been any form of inter-site cooperation.

Once again, as we proceeded through the discussion, it became apparent that the legal issues are not simple. We waded through everything from liability, libel, slander, non-disclosure agreements and the Uniform Commercial Code.

7 The Aftermath

It appears that this discussion topic served its purpose. The discussions were lively and wide-ranging. In a few cases, the discussions pointed out prior work that we were not aware of.

Our work and viewpoint is that of people in the trenches, performing security incident response on a regular basis. The tools and ideas are those that were sparked by discussions during real investigations, usually along the lines of "I sure wish we had a..." By nature, we have been infrastructure builders, and when we see broken or missing infrastructure, it is usually pretty obvious to us. Coming from the production supercomputing community, we also seem to see problems of scale pretty quickly. Having hundreds of incidents each year for the past three years may have contributed to that vision as well.

Do we believe that each and every one of the technologies we present should be immediately implemented in a research or commercial product? Not really. There are practical considerations as well as societal impacts to be considered first.

One thing that has become rather obvious is that there is insufficient legal and societal infrastructure to support some of these technologies. The questions of liability and responsibility for actions or lack of actions in the online community seem to be almost completely uninvestigated. Even if the research were to be finished today, it would be years before any required changes would be reflected in legislation in any jurisdiction. It would take years beyond that to have good case law interpreting that legislation.

After reviewing the papers presented at the workshop, it almost seems as though we have at least asked many of the hard technical questions, even if we don't yet have good answers, let alone running code or deployed solutions. It seems that we are now at the point that the real questions and truly new paradigms will be in the non-technical disciplines. We understand authentication, we understand assurance, we understand simple intrusion detection, and we understand many other security-related technologies.

The parts we don't understand well seem to be in the non-technical areas. We don't understand the interfaces between the online world and global society. We don't understand all the ramifications of an international communications infrastructure that offers new complexity at the same time it offers completely new capabilities in human interaction.

It seems that any truly new security paradigms will have to be the result of multi-disciplinary work, bringing together the work of the security, legal, and other communities.

Acknowledgements

We would like to thank everyone in attendance a the NSPW 2000 when this discussion topic was presented. Everyone made important contributions and this paper is better for them. We would particularly like to thank Bob Blakley for acting as scribe during the discussion. Mary Ellen Zurko, Marv Schaefer and Brad Wood in particular were very kind to us as NSPW newcomers.

References

[1] Steven J. Greenwald. Discussion Topic: What is the Old Security Paradigm? *In Proceedings of the 1998 New Security Paradigms Workshop*, Charlottesville, VA USA, September 1998.

[2] http://www.cert.org/stats/cert_stats.html "CERT/CC Statistics 1998-2000"

[3] http://www.ietf.org/html.charters/idwg-charter.html "Intrusion Detection Exchange Format (IDWG) Charter"

[4] http://www.distributed.net "Distributed.Net Home Page"

[5] http://www.enropia.com "Entropia Inc. High Performance Internet Computing"

[6] http://philby.ucsd.edu/~sanctuary "The Sanctuary Agent System"

[7]http://www.mitre.org/oubs/showcase/cve "The Mitre Common Vulnerabilities and Exposure Database"

[8]http://www.attrition.org "Forced Attrition Web Site"

[9] Cliff Stoll, The Cuckoo's Egg: tracking a spy through the maze of electronic espionage

New Security Paradigm Workshop 9/00 Ballycotton, Co. Cork, Ireland
© 2001 ACM ISBN 1-58113-260-3/01/0002...$5.00