Fighting Through Cyber Attacks

An Informed Perspective Toward the Future

Partha Pal, Rick Schantz, Michael Atighetchi, Kurt Rohloff
Information and Knowledge Technologies Department
BBN Technologies
Cambridge, USA
{ppal, rschantz, matighet, krohloff}@bbn.com

Nathan Dautenhahn, William Sanders
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, USA
{dautenh1, whs}@illinois.edu

Abstract— **This paper presents an overview of the current state of cyber-security research and development (R&D), and a number of forward looking thoughts focusing on the challenges the community is likely to encounter in the next few years. The research ideas are organized in two categories- the first describes ideas that have already taken roots in the R&D community, whereas the second describes ideas that are more radical and require a significant departure from current practice.**

Keywords-survivability; future directions; challenges

I. WHERE ARE WE COMING FROM? WHERE ARE WE NOW?

The cyber security research and practice is currently in its 3rd generation, focusing on tolerance and survival, following the prevention-focused 1st and detection-focused 2nd generation. Key achievements of the 3rd generation cyber security so far include:

- The concept and validation of survivability architectures as a way to organize security building blocks such as protection measures, detection capabilities and adaptive behavior, and
- Extending the notion of Fault Tolerance to Intrusion Tolerance and development of defense mechanisms aimed at attack avoidance, masking of attack effects, and recovery from attack induced failures.

Demonstration of high-water mark survivability architecture [1] achieved mission length (~12 hrs) survival time, but with graceful degradation—only prolonging an eventual death as attacker actions diminish the pool of resources available for the mission to continue. Ongoing 3rd generation research is developing self-aware mechanisms that attempt to stop and possibly reverse the degradation and resource attrition caused by attacks. Dynamism and adaptation is being applied to security mechanisms that has been traditionally static, leading to "adaptive protection" (e.g., creating and starting new variants of executables with different vulnerability profiles or inserting new filters that prevent once successful attacks from succeeding again) and "adaptive detection" (e.g., detection mechanisms that learn about new attack variants based on past encounters and develop new signatures).

It is expected that the next high-water mark survivable system will incorporate some of these new mechanisms and capabilities. But will these new mechanisms spell the end of cyber insecurity?

The likely answer is "no!" Defense against malicious adversary is inherently hard. The adversary needs to find only one flaw in a system to exploit, whereas the defense needs to identify and address potentially all. Military and civilian information systems are becoming more distributed and interconnected with each other. Software is performing increasingly complex functions and requires complicated configuration settings distributed over multiple nodes. Attack software and attack surfaces are easily accessible. There is abundant motivation for thrill seekers as well as diverging national interests. The technical community has come to accept that there is no "absolute" in security, only "adequate".

So from time to time, we take stock and ponder what lies ahead? Would cyber security keep raising the bar as adversaries keep catching up? Is there something that will stop cyber-defense from being a perpetual arms race? Is technology the solution or the solution lies in regulatory and economic factors in cyber-security (e.g., when the cost of accepting compromises forces system and data owners to find and employ "expensive" solutions, or when "offense is the best defense" becomes the new regulatory norm)?

Based on our successful involvement in cyber-security research, especially in the 3rd generation aspects, our view of the future is a combination of good and bad news. The bad news is that our cyber security problems are going to get worse before the situation gets better. The technology landscape is fast changing—sometimes too fast for security practitioners to keep up. New attack surfaces and exploits are coming on line as computer and information systems become integrated with physical systems. There are many corners in the world where cyber criminals can nest and hide. The adversary can, and will continue to use cyber attacks as a force equalizer against a stronger opponent. The good news is that the community over time has moved in the right direction by shifting the objective from "completely preventing cyber attacks" to "mission continuation or fighting through cyber attacks" and by harnessing new hardware, software and networking technologies in innovative defense mechanisms.

A number of ongoing research efforts are currently focused on basic security techniques and mechanisms that remain valid despite the changing technology and threat landscape (e.g., the need to measure and asses)—they need to continue. At the same time, we also need to explore if there are fundamental flaws in our computing infrastructure that attackers are able to exploit and novel ideas to address them—new R&D investment is needed to focus on these game changers.

The remainder of this paper is organized as follows. In section II we highlight an example of how changing technology landscape brings out new challenges for cyber security. In section III we describe five research ideas organized in two categories. The first category describes ideas that have already taken roots in the R&D community, but additional research is needed to transform the ideas into useful technologies. The second category captures more radical thoughts that require significant departures from the current way of designing and implementing computing systems but attempt to address the lack of cyber security at a fundamental level. Section IV concludes the paper.

II. Change is the Only Certainty

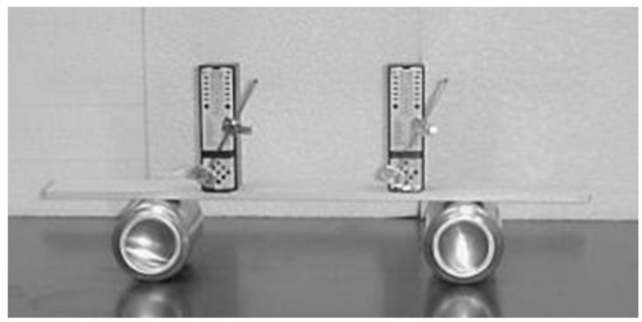
The technology landscape in computing is fast changing. More processing power and memory capacity is packed in increasingly smaller devices. Optical technology offers a huge increase in network capacity and speed. Software construction technology is experiencing Web 2.0, service-oriented architecture (SOA) and semantic web. Advent of new technologies, such as Elliptic Curve Cryptography (ECC) poised to replace RSA, is being complemented by resurgence of classical concepts like functional programming, dynamic time-sharing of resources in newer incarnations like Map-Reduce and cloud computing. It is fair to say that “change” is the only certainty in computing and information technology, and it will remain so in the foreseeable future.

On one hand, the certainty of change makes cyber-defense an infinite problem that can only possibly be solved in increments: security that is adequate today for a given context may not be adequate tomorrow or for a different context. On the other hand, the cyber-security implication of change is frequently overlooked. Let us explain with an example.

Networked computer systems interact with each other in the cyber domain. As it is, not all interdependencies within the cyber domain are well understood or even known. With the emergence of smart grid, there will be a new backplane with significant computing and communication power within the electrical domain. Intelligent computing and communicating devices will be in offices, operation centers and bunkers as part of the power lines and smart storage devices. Globalization will imply that he provenance of the devices or the software they run, or which business or national entities are connected to the power grid may not be known.

Synchronization of coupled systems is a well known phenomenon, first observed by Christian Huygens in 1657 when he found that two pendulum clocks hung from the ceiling were robustly synchronizing with each other despite all attempts to keep them separate and independent. His conclusion was that the wooden truss from which the clocks were hung was transmitting motion even though it was not perceivable by any instrument at that time. Modern day incarnations of the same experiment (see Figure1) show that indeed, the weakest coupling can potentially transmit enough energy to have discernible effect on the coupled system. The underlying theory has been used to model and explain a number of cases observed in biological and electrical systems.

For information systems, the “smart grid” will provide a fairly capable *coupling* that is distinct from and in addition to the couplings that we know and suspect exist in interconnected computing systems. Little attention is being paid to the fact that the grid connects all computing and networking devices (unless the devices are powered off the grid). Most of the current focus in smart grid security seems to worry about grid reliability—how to prevent malicious attacks on the grid; and not on understanding the risks and potential of the smart grid being used as an attack vector on mainstream IT systems. Apart from accidental failures in grid powered computing devices caused by the “synchronization” effect—which for the most part, can be tackled by mechanisms such as surge protectors and uninterrupted power supplies— the power-level computing backplane can be a potent weapon in the hand of an able adversary. For example, the adversary may be able to disrupt, degrade or control the information system by manipulating the level and quality of power delivered to the key computing elements. Unless the smart grid and future computing devices and architectures are built with proper containment and protection, the undiscovered attack surfaces and less understood coupling effects will be exploited by the adversary and cyber security will remain an arms race where the defense starts with a severe handicap.



Two metronomes on a wooden plank over two empty soda cans (from Synchronization of metronomes by James Pantaleone in American Journal of Physics, Vol. 70, No. 10, October 2002, pp. 992-1000)

Figure 1: Huygens' clock experiment using metronomes

Huygens’ clocks and the theory of coupled systems provide a relevant lesson for security in the constantly changing cyber world: intended and unintended coupling is a consequence of change, and coupling will introduce unexpected and sometimes nasty surprises. Let us note that “synchronization” may not be the only undesirable impact in coupled systems and the potential coupling through computationally powerful smart grid is not the only one to worry about: more and more business functions, information systems and networks are getting interconnected at various new levels (e.g., semantic links, social links) introducing different types of coupling.

III. Research Ideas

In this section we outline key research goals that, based on our experience, expertise and insight seem realizable in the next 5 to 10 years. The underlying research directions driving these goals fall into two categories. First we describe three ongoing research directions that are focused on basic security issues that remain important (for example, the need to measure and assess or the need to minimize undue disclosure of information) despite the changing technology and threat landscape. These efforts are at different points in their lifecycle and need to continue. Then, we describe two novel ideas, one aims to repurpose the protection, detection and tolerance focused capabilities developed so far in a new way, and the other aims to eliminate a significant attack vector by means of novel computing primitives. New investment is needed in both categories to further develop the ideas and turn them into usable technologies.

A. Ongoing and Need to Continue

1) Measuring and Evaluating Security

Many a "top ten" list of cyber-security research problems have included the lack of adequate means and metrics to evaluate and assess security of information systems, including a recent one from DHS [2] that posits if we cannot measure, we cannot manage cyber security. Ongoing AFRL funded work [3] offers a potential practical solution by breaking down the larger problem of assessing information assurance into manageable parts. The decomposition is based on the notion of mission stakeholders and their information assurance (IA) requirements. IA requirements are defined in terms of relativistic levels (such as high, medium and low) of individual IA attributes (i.e., Confidentiality, Integrity, Availability etc.) for a given scope (i.e., an end-to-end function, a set of hosts, a sub-network) over time (i.e., during a mission) as required by a specific stakeholder. This decomposition provides a manageable way to continually assess whether the system is delivering the level of IA required by the stakeholder during a mission, which can facilitate timely autonomic mitigation of the deviation.

Automating the response to security incidents is important because today human operators are overwhelmed by the amount of alerts produced by existing intrusion detection systems, and many alerts are false alarms. Identifying the critical alert that needs a response and the information required to mount an appropriate response takes time. On the other hand, every second of delay in response is an opportunity for the attacker to achieve his objective. Deviation from required assurance level can be a reliable trigger for autonomic response. However, additional synthesis and reasoning mechanism is needed to process the observations and reports. A number of techniques such as cognitive reasoning, game theory and probabilistic modeling are being explored by various projects [4, 5].

Advances in assessment and management must be complemented by advanced support for security design. We envision a computer aided design (CAD) tool for security and survivability, along with a library of defenses that takes in a system model/description and performs what if analysis. With this tool, a security engineer will be able to explore different cost/benefit tradeoffs and tailor a specific defense profile to the threat exposure and tolerance requirements. Prior DARPA funded work [4] laid out a foundation for such a tool by demonstrating executable models of systems that include components and behavior of infrastructure, business functions as well as defense mechanisms.

Although more research and engineering is required before developing systems with built in assessment and autonomic response capability becomes routine or the survivability CAD tool becomes a reality, researchers have a clear idea about the end goal and target capability. Work in this area is already underway, and needs to be sustained.

2) Prediction and Avoidance

With increased computing power and storage, high bandwidth communication and social networking—can we be better at anticipating attacks? Can we know and react faster than the spread of attack? Can we dynamically put up speed bumps in the path (adaptive protection) of attack propagation? There are a number of recent and ongoing research activities that look promising. In the DARPA Application Communities (AC) [6] program, researchers are exploring how a monoculture of computing infrastructure and application programs can be leveraged to collaboratively diagnose problems, determine patches and configuration changes that fix the problem, and collaboratively generate awareness and response to the problems encountered as a self-aware organic community. Honey pots, taste testers and sandboxing approaches demonstrated that, for specific contexts (e.g., CORTEX [7]), it is possible to quickly and automatically develop patches and filters to block detected and spreading attacks. Groups like the Internet storm center [8] and CAIDA [9], and projects like the Internet Motion Sensor [10] are watching over the raw traffic pattern as well as the social networking chatter to pick up indications of potential attacks. There has also been early theoretical work on the detection and avoidance mechanisms for worm epidemics [11].

Despite these potential jump-off points, we are still at the early stage of formalizing a realizable end target and sorting through potential directions in this area. A number of research challenges also loom at large. For example, prediction accuracy is still highly context dependent and has room for improvement [12]. The issue of trust, that is, knowing when a report or a patch obtained from a community member is real and not fake or malicious is difficult to address in large general purpose distributed systems. Various organizational and security policies also hinder the ability to collaborate and share internal states, failure and response information with others. In many cases patches and filters are signature based, and are easily subverted by polymorphic attacks or slight variants. If the adversary can determine the prediction and avoidance logic, he can devise targeted attacks to defeat the logic and cause self-inflicted damages on the system. Further research is needed to address these issues.

3) *Minimize Disclosure*

Networked and distributed applications need to disclose some information about them by necessity (for instance, the initial point of contact, the services offered, the signature and return types of remote interactions, and the location and nature of redundancy used in the system). In addition, sometimes defense mechanisms (that are introduced to secure the system) also disclose information—often a denial (successful action from the defense mechanism's point of view) provides useful data points to the adversary. In most cases, much of this information (such as location of services and redundant servers, the OS of the hosts etc.) remain static for a long term. How can we build systems that do not ooze out information unnecessarily? Is there a way to limit the amount or validity period of the disclosed information?

Approaches like single Packet Authentication (SPA) and port knocking attempt to minimize the disclosure about the initial points of interaction. Various dynamic maneuvering, such as restarting hosts and applications, port hopping and service and VM migration attempt to deny or limit the usefulness of information that an adversary may obtain about the system through reconnaissance.

But in most cases, the system architecture in which such techniques are deployed and the protocols that legitimate clients need to use are not "knowledge limiting". For instance, if SPA is used to safeguard access to inside services, an attack on the SPA firewall may crack it open so that it lets specific unauthorized clients get to the inside servers. The architecture needs to ensure that even a corrupt SPA firewall does not disclose any significant information about the system. Redundancy-based techniques require that the redundant entities maintain information about their peers on a continual basis. Therefore, by starting a server for X (assuming he has the authority to do so), or compromising the server at Y, the attacker can find out about all existing service providers. There is some awareness of the need for the research challenges and potential solutions in this area (for instance the AFRL RIKA-08-08 BAA). BBN developed a number of white papers on this topic. For example, in [13] we outlined an approach for knowledge limited architectures and protocols. In [14] we envisioned dynamically changing configuration settings without violating high level policy requirements. But concerted research with this focus has not yet started.

B. *More Radical Thoughts and In Need of Champions*

1) *Assume your Environment is Hostile*

The traditional thought in cyber security has been that the computing environment is generally benign, and we just need to keep a few bad guys out. This belief led to the ideas and approaches that are described as perimeter defense, "intrusion" detection and "intrusion" tolerance. But perhaps it is now time to accept that the computing environment is inherently hostile, and the good guys need to make sure that their interactions can successfully tunnel through the hostile environment without loss of integrity and confidentiality. This requires a fundamental shift in the way systems are architected and software is constructed.

In BBN's submission to the National Cyber Leap Year (NCLY) summit, we envisioned a future where each application a user needs to run resides in a dedicated USB computer (i.e., a small device in the shape of a thumb drive with enough CPU and memory to run Linux and store user data), and when the user needs to use that application he plugs that USB computer into a laptop, desktop or a tablet that simply acts as a chassis to provide display, keyboard and network connectivity. The USB computer relies on Trusted Platform Module (TPM) based attestation to decide whether to trust the chassis. If the TPM based checks are satisfied, the user can then use the application using the keyboard and display, interact with peers using application level security on top of whatever security is accorded by the host and its location (e.g., the laptop may be on SIPRNET, on a security enhanced wi-fi connection, or on an unprotected wi-fi connection). When done, he simply shuts down the USB computer and disconnects from the chassis host. Even if the attacker succeeds to control the chassis and the network, the "good" application will still be a lot safer than it currently is. TPMs and USB sticks with powerful CPU, crypto processor and memory are becoming a reality. The issue of sharing data across applications can be partially addressed by storing encrypted versions in the "cloud". A proof of concept to perform feasibility study looks certainly feasible.

Note that the dedicated USB computer idea described above is an instance of the more general research direction that a future research program should explore: how can we use existing and emerging protection, detection and adaptation capabilities in such way that authorized components can reliably identify each other without any assumption about the environment they are embedded in, and interact only after they are such authenticity is established. Other NCLY participants had thoughts and ideas along similar lines, and there is some precursor work that can be leveraged.

2) *Fundamental Secuity*

In this section we consider research ideas that attempt to get to the root causes of cyber insecurity. Obviously this depends on one's perspective—for example, Microsoft research has been developing a highly-dependable OS [15] where the kernel, device drivers, and applications are all written in managed code, which stems from the point of view that most of the security problems in Windows result from poorly written device drivers.

Another valid view is that social engineering will remain a major attack vector where the attack entry point is established by luring unsuspecting users to download or open a file containing malicious code. Let us explore this issue a bit more.

Defense against malicious code today mostly consists of updates to antivirus and malware detection tools or patches to existing software, a lagging response, because signatures, patches, blacklists are developed by human experts after the malware has caused some amount of damage, although newer methods (e.g., Quorum [16]) are trying to be more proactive. On top of that, signature-based systems are often defeated by polymorphic code. Ongoing and planned (such as DARPA's CyberGenome) research programs are looking at detection, characterization and attribution of malware. However, current attempts to fight malware are missing a crucial flaw in today's processor architecture and operating systems: the file "save" and "open" operations are fundamentally unsafe.

Imagine a future computing universe, where files contain a permanent history and cryptographic material as part of its metadata. The first entry of the permanent history is created when the creator of the file makes an explicit "committed save" differentiating it from the interim saves (ctrl-s or auto-saves) during his work, and is updated at each subsequent "committed save" by any user and also when a program obtains the file from a remote system and stores for the first time in the recipient host's file system. An entry being written into the permanent history will contain information derived from the current file content and creation data (user or program creating or saving the content) tied to the current host environment in a way that cannot be faked. As the file gets saved in different systems, the chain of permanent history gets longer, but the file only needs to maintain a sliding window of such entries consisting of at least the root entry, and the current entry and the immediate precursor. The metadata accompanying the file also includes "diffs" that will enable getting back to the file contents for which the history entries are stored. The idea of including permanent history and other metadata along with stored files is somewhat similar to the notion of "pedigree management". Whereas pedigree management systems such as PMAF [17] focus on the origin (provenance) and lineage (pedigree) of *information* i.e., which is embedded within content of the file, our focus is mostly on the container i.e., the file itself.

In this hypothetical computing universe, before the file is saved in the new environment as well as when a user attempts to open a file, the permanent history, accompanying metadata and file content is checked for consistency and conformance. The user trying to open or the host accepting the file can have different levels of strictness—they may accept files with no permanent history for certain files for certain programs or from certain sources (trusted or white-listed hosts). For others, they may simply accept files with the root entry in the permanent history. For example, a patch from Microsoft hosted at a corporation's internal IT department will be accepted if it can be verified that the file was "commit saved" in a Microsoft environment and has not been tampered with. For others, it may require longer chains of verification. For example, for a file obtained from SourceForge, it may need the root as well the entry for the SourceForge site.

Because in this hypothetical universe all computing hardware and OS conform to the above semantics of file save, commit save and open, this approach will provide accountability of digital artifacts. Malware authors will therefore be forced to try to install files that do not have a root entry or incorrect permanent history chain. If the adversary uses non-conforming hardware or OS, he risks his files being rejected outright by conformant systems. This approach has the potential to strike a fatal blow to the malware scourge.

However, there are many unknowns. The TPM technology is certainly needed to implement the permanent history, but a PKI infrastructure may still be needed in addition. The sliding window algorithms and OS framework need to be fully developed and validated against design flaws and residual issues. The overhead of cryptographic functions in file operations may be a problem (despite projected cost savings due to smaller key lengths in ECC). Finally, the success of this approach critically depends on hardware and OS vendor buy in, which will invariably involve standardization, and possibly government acceptance and requirement.

We envision a future research program that aims to identify fundamental computing primitives that remain unchanged despite the changing technology universe, and augment them with additional security measures (as illustrated in our example with the permanent history and file operations) to gain a multiplicative improvement in cyber defense.

IV. Conclusion

In this paper we outlined a number of research ideas to improve the security of the ever changing cyber universe. The first set of ideas focus on research directions that need to continue and should be reinforced. The second set of ideas focus on game changing strategies that are candidates for future research programs.

References

[1] L. Badger, DARPA Weekly Activity Report, Dec 23 2005.-Jan 6, 2006

[2] D. Maughan, "The need for a national cybersecurity research and development agenda," CACM, vol. 55, no. 2, Feb 2010, pp.29-31.

[3] P. Pal, and P. Hurley, "Assessing and managing quality of information assurance, "NATO RTO Symposium on Cyber Defense and Information Assurance, April 2010, in press.

[4] P. Pal, F.Webber, M. Atighetchi, P. Rubel, and P. Benjamin, "Automating cyber defense management," Proc. of WRAITS 2008, Glasgow.

[5] S. Zonouz, H. Khurana, W. Sanders, and T. Yardley, "RRE: A game-theoretic intrusion response and recovery engin**e, "** DSN 2009, Estoril

[6] http://www.tolerantsystems.org/ac.html, Home page of the DARPA Application Communities program.

[7] http://www.tolerantsystems.org/SRSPIMeeting12/12pi_meeting.html, CORTEX presentation at the SRS PI meeting, Dec 2005.

[8] http://isc.sans.org/, Internet Storm Center Homepage.

[9] http://www.caida.org, Cooperative Association of Internet Data Analysis Homepage.

[10] http://ims.eecs.umich.edu, Internet Motion Sensor project Homepage

[11] K. Rohloff and T. Basar. "Deterministic and stochastic models for the detection of random constant scanning worms. ACM TOMACS, vol. 18, no. 2, April 2008.

[12] K. Rohloff, R. Battle, J. Chatigny, R. Schantz, and V. Asal. "A Trend Pattern Approach to Forecasting Socio-Political Violence." 3rd Int. Conf. on Computational Cultural Dynamics, December 2009.

[13] P. Pal, "Knowledge limited interaction protocols for SOA," BBN Whitepaper, unpublieshed.

[14] M. Atighetchi, "Dynamic and assured configurtion management platform," BBN Whitepaper, unpublished.

[15] http://research.microsoft.com/en-us/projects/singularity/, Microsoft Singularity project Homepage.

[16] http://www.symantec.com/connect/blogs/how-reputation-based-security-transforms-war-malware. Symentec blog on reputation based malware detection.

[17] http://www.oracorp.com/Research/pmaf.html, The PMAF Pedigree Management and Assessment Framework.-