international risk
governance council

WORKSHOP REPORT
EXTERNAL CONTRIBUTIONS

COMPARING METHODS FOR TERRORISM RISK ASSESSMENT WITH METHODS IN CYBER SECURITY

AAAS Washington, DC, USA
28 – 29 May 2015

Question to the community of cyber risk assessment experts and the community of terrorism risk assessment experts:
"What (if anything) can you learn from each other's experiences and tools?"

Formal methods of quantitative risk analysis have a long history. The use of exposure models and dose-response functions became common in the 1960s with the growth of concerns about environmental quality. Similarly, concerns about risks from nuclear power lead to the development of fault trees and failure modes and effects analysis in the 1970's. None of this work was concerned with risks imposed by pernicious intelligent adversaries. Hence, most of the tools that were developed are of limited applicability to the assessment of terrorist or cyber risks. The terrorist attacks in New York and Washington in September 2001 pushed the risk analysis community to extend its tools and methods to consider risks imposed by intelligent adaptive adversaries.

Largely independent of work in the risk analysis community, the computer science community has been working to assess and manage risks from cyber attacks.

Given these two parallel efforts, the workshop brought together a group of leading researchers[1] who have been working on issues of terrorism risk assessment with a group of leading researchers in cyber security. The objective was to explore what (if anything) the two communities can learn from each other and whether there are opportunities to avoid duplication and reinvention. In most conventional risk assessment, there is not an intelligent adversary. Of course, that is not true in the case of terrorism, and it was for this reason that we thought it could be productive to put the two groups together.

Financial support for this workshop was provided by the International Risk Governance Council (IRGC) and its sponsors. In-kind staff support has been provided by the Department of Engineering and Public Policy at Carnegie Mellon University.

[1] The focus in this workshop was on identifying and improving research methods. Discussion with practitioners was not a specific focus.

Contents

Workshop Summary Report

This document includes workshop framing notes written by Prof. Granger Morgan before the workshop, with contribution from several workshop participants[2], as well as notes from presentations and discussions at the workshop. The last section reports a group discussion of a set of questions.

Sessions on terrorism risk and cyber security risk analysis

1. Risks without pernicious intelligent adversaries

Many risks do not involve intelligent adversaries. In the analysis of risks to health safety and the environment, a variety of analytical strategies have been developed to assess and manage both risks that arise from a continuous exposure (e.g., air pollution) and from discrete events (e.g., an explosion). Risks from continuous exposures are typically assessed with methods such as transport and diffusion models and the application of either static or dynamic dose-response models. Risks from discrete events are typically assessed with methods such as fault trees and failure-modes and effects analysis. In both cases, simulation models are often used.

It is common to frame such problems in terms of four stages as shown in Figure 1. Then, depending on the nature of the risk, it may be most appropriately managed by a number of different interventions as shown in Figure 2.

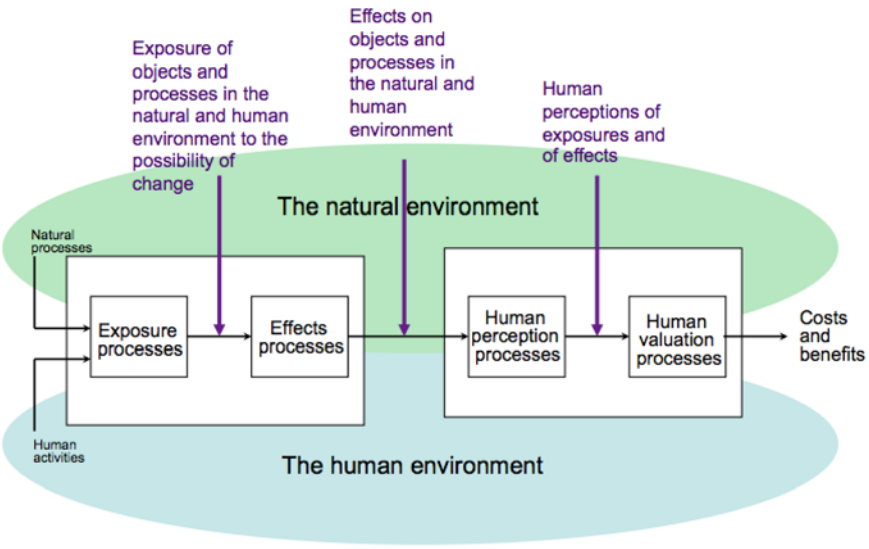

Figure 1: A basic framework for describing the problem of *assessing* physical risks to health, safety and the environment (after Morgan, 1981).

[2] We are in particular grateful for the valuable suggestions provided by Virgil Gligor, CMU; Herbert Lin, NRC/Stanford; Adrian Perrig, ETH-Z; Henry Willis, RAND; and Detlof von Winterfeldt, USC.

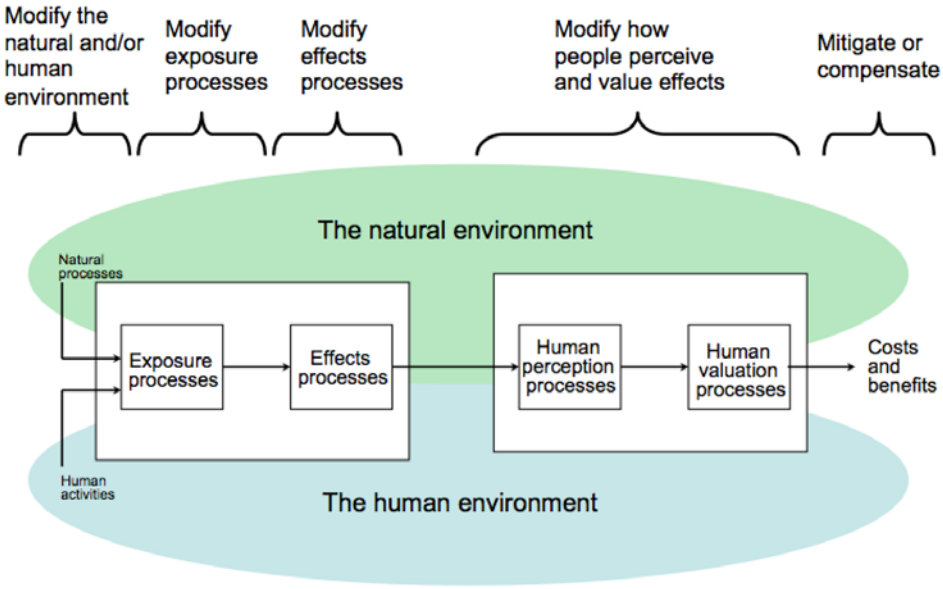

Figure 2: Strategies that can be used in *managing* physical risks to health, safety and the environment. Different strategies are more appropriate for different risks (after Morgan, 1981).

Not all failures in cyber and cyber-physical systems result with the involvement of intelligent adversaries. There may be basic flaws in the logic of the underlying design including unanticipated contingencies, unforeseen combinations of events and dependent failures, there may be errors in how that design is implemented in code, and there may be failures in supporting electronics and hardware.

2. Adding pernicious intelligent adversaries

In the case of physical risks from terrorism, the same analytical methods as those listed above can be used to assess the consequence once an event has, or is projected to, occur. However, there is an almost infinite number of combinations of adversaries, places, and ways events could occur, only one or a few of which may actually occur. Hence, in the case of physical terrorism risk, the major analytical challenges are (a) to assess the probability across this large set of possible attacks, (b) to determine how much to invest over time in protective designs and measures and/or in resilience and rapid recovery, and (c) how to allocate those investments. This involves setting priorities and reducing uncertainty surrounding cyber security investments. The task is complicated by the fact that terrorists can be expected to be adaptive so that defensive measures may not uniformly decrease risk. For example, in response to defensive measures, terrorists may either shift their attention towards less secure targets or continue to target the same locations with more dangerous methods. It is therefore difficult to assess the effectiveness of policies and counter-measures.

Much of the work in terrorism risk analysis has focused on six issues:

1. How to assess the intent and capability of adversaries
2. How to identify and assess the relative attractiveness of different potential targets
3. How to assess the probability that each of these (many potential) targets might become the actual target of an attack
4. How to detect and intervene to prevent an attack
5. How to assess the second and higher order consequences of investing in detection and intervention and in making some targets more secure

6. Given that it may be impossible to avoid all physical attacks on any given class of targets (e.g., a transit or electric power system), how to make those targets more resilient and how to speed recovery after an attack.

There are, of course, a number of cyber-physical systems, such as the examples of transit and electric power systems, which can either be physically attacked (e.g., by destroying tracks, bridges or substation transformers) or attacked via cyber disruption of monitoring and control systems. Indeed, there is also the possibility of combining the two into a physical plus cyberattack. More about cyber-physical systems below in Section 5.

3. Methods used in terrorism risk analysis

Most of the work that has been done in this domain has thus built upon various tools developed and used in the more general field of risk analysis. These tools provide a structured way to describe threats and hazards and a normative approach to decision-making. Methods used in terrorism risk analysis primarily aim to inform deliberation and to optimize decision-making. While conventional methods often fail to assure completeness of threat models and to assess their validity, absent events or well defined adversaries, terrorism risk analysis learns from experiences with past events, including for assessing the cost-effectiveness of investment made (for example, the approach termed "*defender-attacker*" games has proved very useful in identifying the space of potential vulnerabilities and identifying how defensive resources should be allocated), for incorporating behavior in risk models and for understanding the type of rationality that attackers have. One important issue, which is also true in the cyber domain, is the need to invest in understanding how attackers behave, in order to (try to) model their behavior.

Building on methods from decision analysis, Bayesian inference, and game theory, a number of authors (Paté-Cornell, Guikema, Ríos Insua, Rios, Banks and others) have been developing a set of tools and strategies that are now being termed "*adversarial risk analysis*." While considerable progress has been made in the theoretical development of such approaches, practical applications have to date been somewhat modest. However, in their recent book *Adversarial Risk Analysis*, Banks, Rios and Ríos Insua do provide a worked example in which the "client is a railway service that is concerned about fare evaders, pickpockets, and, perhaps, potential terrorist threats."
In contrast to "traditional" risk analysis, where hazard is stochastic (e.g. natural hazards, technical failures) or could be modeled as stochastic, adversarial risk analysis assumes that hazards are adaptive (terrorism, sabotage). The process can be modeled as either exogenous (e.g., optimization) or endogenous (e.g., game theory). *Adversarial risk analysis* could be relevant for cyber risks.

In game theory, it is assumed that the adversary and his behavior are known and presumably rational. Game "*analysis"* can be used when there is no knowledge about the adversary (who it is and what it knows) and no assumption about the behavior or preference of the attacker (what it wants). Of course, cyber systems are most often attacked by individuals who don't want to be visible and whose motives and means are unclear. Defenders often don't know what adversaries want.

4. Cyberattacks and the design of secure cyber systems

After this broad overview, the workshop discussion turned to an examination of a variety of strategies to support the design of more secure cyber systems – largely built around various means of retaining a trustworthy core in a system that must deal with application programs and an outside world that cannot be trusted.

Like in terrorism risk, what characterizes cyber security risk is that the risk results from the deliberate malicious actions of intelligent actors. Also, since almost every computer can be linked to every other computer, the cascading effects of intrusions or attacks can be enormous. A first consequence is the large uncertainties about threats, vulnerabilities and effects. A second consequence is that the threats, the motives, the type and mode of attacks, and the targets change continuously. Threat and consequence assessments, and methods for developing scenarios for that purpose, are largely in their infancy for cyber security risk.

Proposed characterizations and taxonomies of cyber risks reflect the various types and motives of incidents or attacks. For example, Herb Lin[3] has argued that, in very broad terms, the community of attackers in cyber security can be broadly categorized into three groups:

1. The "ankle biters." These include that large community of novice hackers who are forever looking for what they can break into, disrupt or otherwise "hack" but with no specific political, economic or other agenda.
2. The "larcenists." These include individuals and groups who are primarily interested in making money either directly or indirectly. Most of them don't particularly care *what* they break into so long as having broken in they gain access to something they can monetize either themselves or by selling the access or data to others.
3. The "persistent and focused attackers." This group ranges from those with very limited technical skills to those with extraordinarily advanced skills including cyber warfare groups in nation states. They are typically focused on going after specific targets (Iranian centrifuges; naval weapons control systems; the SCADA system of PJM; communications between government officials; etc.).

But this taxonomy involves no clear associated links with consequences.

In the case of many terrorists and of persistent focused cyber attackers, the problem is to figure out which of an almost infinite set of potential targets will be attacked. In the case of all other cyber risks (those perpetrated by "ankle biters" and "larcenists"), it is relatively safe to assume that an attack will be mounted against every system that can be either directly or indirectly accessed over the Internet or other communication channels open to outsiders.

Like in the field of terrorism risk analysis, much of the work for cyber systems has addressed five issues:

1. Assuming that attacks on such systems are inevitable, how to detect them as they begin to happen
2. How to determine the strategies and modes of attack that are being used
3. How to add *patches* and/or make other modifications, such as trying to isolate some of the more critical system elements from (easy) access via the Internet or outside access (via USB flash drives and compromised employees and maintenance personnel), in order to reduce vulnerabilities
4. How to assess the second and higher order consequences of investing in detection and intervention and in making some targets more secure

[3] http://cisac.fsi.stanford.edu/people/herbert_lin

5. Given that it may be impossible to avoid all attacks on any given class of targets (e.g., the electric power or financial systems) how to make those targets more resilient and how to speed recovery after an attack.

Adding *patches* to fix problems when they are identified is commonly used. However, it is very difficult to determine which *patches* are most important to apply, given the hundreds of new vulnerability disclosures each month. Most experts now assume that there is no way to make *commodity software*, such as the MS Windows operating system, completely resistant to cyber attacks. In addition, as new systems are developed, even if they use more robust designs, perform more extensive pre-release checking, and use high quality encryption, sooner or later they can all be expected to be subjected to successful attack, first by advanced hackers, and then by a wider and wider set of attackers as vulnerabilities and attack tools become more widely shared within the hacker community.

Isolating critical elements is also increasingly used. Butler Lampson[4] (*CACM*, Nov 2009) argues that:

> *"Operationally, security is about policy and isolation. Policy is the statement of what behavior is allowed: for example, only particular users can approve expense reports for their direct reports or only certain programs should run. Isolation ensures the policy is always applied. Usability is pretty bad for both."*

Figure 3 illustrates what Lampson terms the standard technical security access control model. The problem, however, is that as the complexity of the objects and resources inside the isolation boundary become more and more complex, it becomes impossible to assure that they can be trusted to only perform as intended, and in no other way.

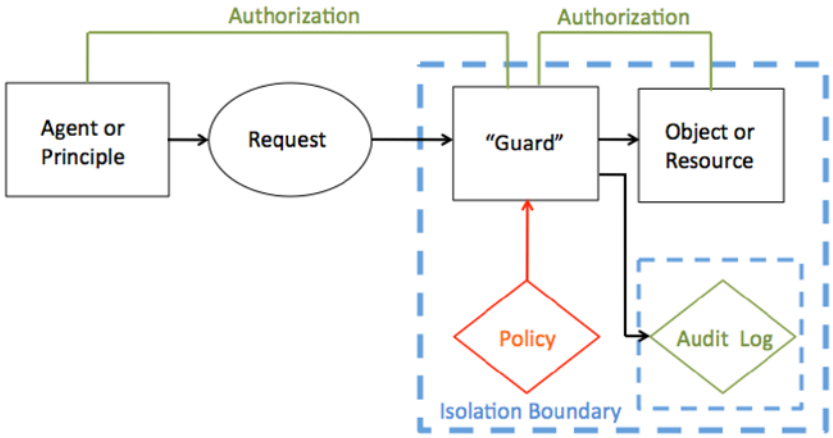

Figure 3: Standard technical security access control model (after Lampson).

The fact that large commodity software will never be fully secure and trustworthy, has led to the idea of *partitioning* systems into small secure sub-elements (a green machine) that works with a much larger set of insecure elements (the red machine). Lampson elaborates:

> *…To reconcile accountability with the freedom to go anywhere on the Internet, you need two (or more) separate machines: a green machine that demands accountability, and a red one that does not.*
>
> *On the green machine you keep important things, such as personal, family and work data, backup files, and so forth. It needs automated management*

[4] http://research.microsoft.com/en-us/um/people/blampson/

to handle the details of accountability for software and Web sites, but you choose the manager and decide how high to set the bar: like your house, or like a bank vault. Of course, the green machine is not perfectly secure—no practical machine can be—but it is far more secure than what you have today.

On the red machine you live wild and free. You don't put anything there that you really care about keeping secret or really don't want to lose. If anything goes wrong, you reset the red machine to some known state.

Virgil Gligor[5] (2014) has further developed these ideas in terms of green "wimps" ("i.e., small software components with rather limited function and high-assurance security properties") and red "giants" ("i.e., large commodity software systems, with low/no assurance of security").

There is, of course, a subset of cyber and cyber-physical systems such as software for avionics or certain robotic control systems where the need to assure security is so high that it is worth adopting advanced methods – proven kernels, complete logical isolation from external communication systems, etc. Such strategies are inherently *very* expensive. Hence, while they can provide a very high level of security, they can only do so at a cost that users consider unaffordable in more general applications.[6]

Effective security that is provided by these features, designs and other analyses is however limited by insufficient knowledge of how users behave in reality and how attackers adapt and exploit the new systems, in response to the strategy and type of response of the defenders. So the field of cyber security overall is marked by a continuous change in many respects. It is unrealistic to believe that a large and complex system can be made fully secure, and perfect security should not be a target.

Since many of the threats are malicious, with the intentional aim to cause damage, classic probabilistic risk assessment is largely seen as inappropriate. However, some analysts who have access to large data sets are working to demonstrate that, for the majority of cyber risks, existing tools and techniques for risk analysis can provide an accurate assessment.

Much of the work in cyber security is focused on the problems of making specific systems more secure and well behaved. That is clearly very important, but Deirdre Mulligan and Fred Schneider[7] (2011) argue that in today's world, it is increasingly not sufficient. Building on a metaphor of public health, they argue that individual strategies (standards, adherence to software engineering good practices, formal methods, red/green machines, filters and firewalls, etc.) are all valuable but as with health, there are aspects of cyber security that are a "*public good*." Just as public health is concerned with the collective health of an entire population, not just the health of specific individuals, this approaches argues that there is a very large element of the "collective good" in the realm of cyber security, and that operational, regulatory, legal and other approaches should all be formulated in a way that addresses these broader issues. Mulligan and Schneider outline a range of strategies, some technical but others behavioral, educational and legal that they believe are needed to create an effective "doctrine for cyber security."

---

5 https://www.cylab.cmu.edu/education/faculty/gligor.html

6 There is, of course, the inevitable issue that what users consider to be too expensive when a system is just beginning to be used may seem very different in retrospect once use has become massive, and a breech with high economic or other losses has occurred.

7 http://www.cs.cornell.edu/fbs/

5. Cyber-physical systems

In the past, many important systems did not involve significant amounts of distributed automated electronic monitoring and control. For example, in the case of the electric power system only the frequency of the AC waveform was needed and used to provide most of the coordination across the system. When the frequency sagged or got too high, automatic generator control (AGC) units would advance or retard the phase angle of generators in order to inject more or less power. However, with restructuring of the power system, the addition of many new players, and growing pressure to operate the system more efficiently and with tighter margins, cyber systems that do both monitoring and control have become increasingly common.

While the motivations and details are somewhat different, the same proliferation of more, and more sophisticated cyber monitoring and control is occurring in a wide variety of other systems such as rail and air traffic control.

For the most part, these systems have grown incrementally, with later additions added on top of existing systems. Especially in the early days, relatively little consideration was given to security. For example, while there were typically chain-link fences around high-voltage sub-stations, wireless control systems (using elements of the MS Windows operating system) were sometimes installed, and in some cases could be accessed from outside the fence-line. Similarly, Some early SCADA systems executed some of their functions over the public Internet.

Today things are improving. Most new systems are adopting at least minimal security standards and efforts are being made to correct some of the most egregious sloppy early implementations. However, it is difficult and expensive to go back and fix things that are already in wide use.

Because cyber-physical systems involve physical components that can also be used as a vector for cyber attack (e.g., physical access to and modification of sensors or monitors and to elements of the communication systems) this is one case in which it appears that the problems faced in terrorism risk assessment and cyber risk assessment may converge. The potential for the emergence of new risks looms ever larger, as do the risks of not doing an adequate fundamental design as pressures grow to implement ever more advanced and capable systems.[8]

[8] For example, Granger Morgan is a member of the NRC/NASA Aeronautics Round Table where he has routinely expressed his concern, and gotten the response that FAA is using "the best state of the art security methods" as they contemplate the development of cyber assisted and fully autonomous aircraft in civilian airspace. However, given that several of the attendees at the workshop probably have the ability to penetrate such systems…this is hardly reassuring!

6. Lists of advice, good practice, attack modes, etc. common in cyber security.

Below are three examples.

Table 1: Twenty top security controls as recommended by the SANS Institute.[9]

1. Inventory of Authorized and Unauthorized Devices
2. Inventory of Authorized and Unauthorized Software
3. Secure Configurations for Hardware and Software on Mobile Devices, Laptops, Workstations and Servers
4. Continuous Vulnerability Assessment and Remediation
5. Malware Defenses
6. Application Software Security
7. Wireless Access Control
8. Data Recovery Capability
9. Security Skills Assessment and Appropriate Training to Fill Gaps
10. Secure Configurations for Network Devices such as Firewalls, Routers and Switches
11. Limitation and Control of Network Ports, Protocols and Services
12. Controlled Use of Administrative Privileges
13. Boundary Defense
14. Maintenance, Monitoring and Analysis of Audit Logs
15. Controlled Access Based on the Need to Know
16. Account Monitoring and Control
17. Data Protection
18. Incident Response and Management
19. Secure Network Engineering
20. Penetration Tests and Red Team Exercises

Table 2: Top ten software security design flaws
A recent IEEE report *Avoiding the Top 10 Software Security Design Flaws*[10] offers the following design advice:

1. Earn or give, but never assume, trust
2. Use an authentication mechanism that cannot be bypassed or tampered with
3. Authorize after you authenticate
4. Strictly separate data and control instructions, and never process control instructions received from untrusted sources
5. Define an approach that ensures all data are explicitly validated
6. Use cryptography correctly
7. Identify sensitive data and how they should be handled
8. Always consider the users
9. Understand how integrating external components changes your attack surface
10. Be flexible when considering future changes to objects and actors

[9] An elaboration of these can be found at www.sans.org, https://www.sans.org/critical-security-controls/
[10] available online at http://cybersecurity.ieee.org/images/files/images/pdf/CybersecurityInitiative-online.pdf

Table 3: GAO classification of categories of cyber attacks[11]

Type of attack
Denial of service A method of attack from a single source that denies system access to legitimate users by overwhelming the target computer with messages and blocking legitimate traffic. It can prevent a system from being able to exchange data with other systems or use the Internet.
Distributed denial of service A variant of the denial-of-service attack that uses a coordinated attack from a distributed system of computers rather than from a single source. It often makes use of worms to spread to multiple computers that can then attack the target.
Exploit tools Publicly available and sophisticated tools that intruders of various skill levels can use to determine vulnerabilities and gain entry into targeted systems.
Logic bombs A form of sabotage in which a programmer inserts code that causes the program to perform a destructive action when some triggering event occurs, such as terminating the programmer's employment.
Phishing The creation and use of e-mails and Web sites—designed to look like those of well-known legitimate businesses, financial institutions, and government agencies—in order to deceive Internet users into disclosing their personal data, such as bank and financial account information and passwords. The phishers then take that information and use it for criminal purposes, such as identity theft and fraud.
Sniffer Synonymous with packet sniffer. A program that intercepts routed data and examines each packet in search of specified information, such as passwords transmitted in clear text.
Trojan horse A computer program that conceals harmful code. A Trojan horse usually masquerades as a useful program that a user would wish to execute.
Virus A program that infects computer files, usually executable programs, by inserting a copy of itself into the file. These copies are usually executed when the infected file is loaded into memory, allowing the virus to infect other files. Unlike the computer worm, a virus requires human involvement (usually unwitting) to propagate.
War dialing Simple programs that dial consecutive telephone numbers looking for modems.
War driving A method of gaining entry into wireless computer networks using a laptop, antennas, and a wireless network adaptor that involves patrolling locations to gain unauthorized access.
Worm An independent computer program that reproduces by copying itself from one system to another across a network. Unlike computer viruses, worms do not require human involvement to propagate.

[11] Available online at http://www.gao.gov/new.items/d05434.pdf

Workshop discussion questions

The final session of the workshop was devoted to a group discussion of a set of questions that had been refined at the end of the previous session. The commentary below combines some summaries of the discussions with some assessments by the workshop organizers.

DISCUSSION QUESTION 1: *How should we identify and strike the appropriate mix/balance between hardening and protection versus resilience and recovery? What if any can be done to support what is clearly a normative (i.e. value-based) choice? What insights can either field draw from the other in this context?*

Some participants suggested that the question is not normative, that one should "simply estimate the cost of each" and opt for the least expensive strategy. To the extent that such estimates could be made, that is a sensible answer. Others noted that while it may be possible to assess the cost of prevention, assessing the costs of not engaging in prevention is *much* more difficult. Indeed, in the area of securing physical infrastructure people routinely assess the costs of different protective measures. The big challenge is assessing the potential costs of an attack, especially given that, in some cases, the direct costs may be dominated by the broader indirect costs. Further, the choice is not binary (protect *versus* facilitate recovery). Rather, both are continuous variables.

Some participants argued that there are pressures to simply "patch" systems rather than back off and adopt a more integrated view. While this is especially true in the cyber realm, the observation applies as well to some issues of protection against terrorism.

There was an extended discussion of "recovery", Participants from both communities argued that recovery was a topic that receives far too little attention, and that there is serious underinvestment. Some participants argued that this is especially true on the cyber side, although there are clearly also compelling examples with respect to physical infrastructure. Suggestions were made that, to better understand resilience and recovery, one might devote more attention to studying past experiences (WWII, hurricanes, ice storms, etc.).

The key idea underlying resilience is to be able to continue to provide critical services over time. The cyber community has worked on the development of systems with such a capability, although there remains a great distance between the capability of research systems and the day-to-day performance of many systems. The same situation applies to the area of physical infrastructure and in cyber-physical systems. People have developed designs for systems that can continue to serve critical needs after disruption, but the implementation of such a capability falls far short. While there are multiple reasons for this, a key factor in many cases it that it is unclear who is responsible for providing resilience, and who should bear the associated incremental costs. Rules that limit liability for failures may also contribute to inaction.

Several examples of efforts on the cyber side were provided, and of organizations such as NASA and DOD, to give the issue of graceful recovery greater attention.

DISCUSSION QUESTION 2: *What are the opportunities to merge or integrate approaches between the two fields? For example, do cyber–physical systems present such an opportunity? Where might cyber (or physical) attack play a role as a "force multiplier?" What are the opportunities for and limitations of the use of scenario analysis, red teaming, etc.?*

Participants suggested that collaboration between the two domains will be essential. One of the first places where such collaboration will be needed is in the area of self-driving vehicles. Addressing risks posed by the growing introduction of autonomous or highly automated vehicles into civilian airspace will also require such collaboration. The power system was also identified as a candidate field in which there is a considerable need for collaboration between those who work on physical risk and those who address cyber risks in monitoring and control.

While at the moment, we tend to divide our thinking about attackers between criminals after money and terrorists after attention and making an ideological point, it was suggested that this distinction has already become blurred and is likely to become even more blurred in the future.

There were mixed views about the **use of scenarios**. One thread in the discussion noted that scenarios are often not sufficiently inventive. For example modern motor vehicles that have extensive automation, GPS and are addressable remotely, hold the potential to be the vector for enormous disruption (e.g. stall them out at critical locations). A particularly worrying example involved an attack on the financial system that goes unnoticed for a while, and has caused very large (perhaps irreversible) damage by the time it is discovered.

There was rather little discussion of the use of red teams, although it is clear that the use of such teams holds considerable potential to identify vulnerabilities in both physical and cyber system – and in the combination of the two.

Several participants noted that, even if inventive scenario analysis and red teaming find potential vulnerabilities, **it is often *very* hard to get decision-makers to take action to increase protection against things that have not happened**. Similarly, system designers often do things in automating systems that are very hard to defend when one adopts a broader, more system-wide view of potential threats.

DISCUSSION QUESTION 3: *The insurance and reinsurance industries are deeply concerned both about issues of terrorism and cyber risk. How can the analytical and research communities best support their needs?*

While there is some overlap, the risks associated with both cyber and physical attacks can be divided into three categories; direct cost, third party costs, and business disruption costs. Many policies in the past were focused only on costs resulting from data breaches and loss of personal or other critical information. However, there is growing awareness of **the risks posed by business and network interruptions**. Today there is a growing line of business in this space.

The **providers of insurance are becoming more proactive**. For example, many firms can simply not get insurance without end-to-end encryption these days.

At the same time, most insurance companies have limited technical expertise. The ISO/IEC 27000 family of standards on Information Security Management Systems are helpful, as are

activities such as training and briefing activities directed at CIOs and others responsible for information security.

However, there are several characteristics of cyber systems that some participants argued make the task of protection much more challenging than the protection of physical systems. It was argued, for example, that one cannot be confident in using induction since some very small change in a cyber system can result in the creation of a very large (and perhaps unrealized) vulnerability. Early detection is important, and there can be great challenges in detecting an intrusion or another form of attack early enough to avoid large damage.

There was some disagreement about whether the data exist, or can be accessed, to perform trend and other needed analysis. Some participants gave an example of how in working with firms this was possible. Others suggested that there are still serious limitations resulting from the reluctance of firms to share data or cooperate. Probably both views are correct. Data collection and sharing is a recurrent topic. There may also be a problem resulting from the fact that the insurance industry is not sufficiently aggressive about asking, or anticipating the longer term consequences of things such as the loss of IP.

Participants noted that some widely used implementations of key cyber systems still contain "massive" security flaws. It was noted that while risks with very long, high tails exist in some cyber and cyber-physical systems, such situations are not unique to these domains.

DISCUSSION QUESTION 4: *Given the various categories of attackers:*

Terrorism	*Cyber*
Petty criminals	*Ankle biters*
Small time and lone terrorists	*Larcenists*
Coordinated individuals or small groups (IEDs)	*Dedicated focused non-state*
Major coordinated attacks (Sep 11)	*Dedicated focused state w/ social, intelligence, or military motivations*

What are the differences in:

- *Data availability and use?*
- *Analytical approaches?*
- *Threat assessment and consequence assessment?*
- *Opportunities for integration/collaboration between our two fields?*
- *Potential for "disproportionate impact" on national confidence (and lifestyle) and how could we assess in advance?*

As the question suggests there is a range of potential attackers and attack modes. Even at the upper end (i.e., big) there is a spectrum of threat activities. While this question stimulated interesting discussion, it did not yield any systematic answers. There was agreement that greater collaboration between investigators from the two fields would be valuable.

It was suggested that it might be beneficial to think more about who the attackers are and to develop additional framings beyond that classic approach of:

threat ➜ vulnerability ➜ consequence

Some suggested that threat and vulnerability may be better understood in the domain of cyber, but that consequence constitutes a gap. To date, the impacts of cyber attacks seem to be more modest than worst case scenarios have suggested.

There was a discussion of vulnerabilities in the form of hidden trap doors that might be built into SCADA systems, air traffic control and similar systems.

One analytical strategy proposed was to **work the problem backward**. In any given domain (either cyber or terrorism) identify a set of a handful of the most serious outcomes against which one would like to guard; then work to identify the many paths that might result in those outcomes. This would not be easy since there is clear evidence of "out of site is out of mind" in the psychological literature. Hence having multiple independent groups do this would be wise.

DISCUSSION QUESTION 5: *In both terrorism risk analysis and cyber security there are too many problems and adversaries and relatively too few good analysts. What are the best strategies to raise the level of analysis and minimize the amount of "snake oil"?*

Other fields, such as environmental impact assessment and conventional risk analysis, have gotten better over time using professional societies, peer review, good publication outlets, strong programs of graduate education, etc. The fields we considered in the workshop should learn from them.

Analysis can be both quantitative and qualitative. **It may be easier to vet the quality of quantitative analysis, but well done, in some setting qualitative analysis can be as or more important**. However, it is harder to assess and enforce the quality of qualitative analysis.

On the cyber side there are tremendous incentives to sell "**snake oil**" because there are so many uninformed decision-makers and groups that know they may have a problem and don't know what to do about it. There has been an insufficient incentive to assess the efficacy of the various programs and get to the truth. However, as with many other fields in the past, with time, and the gradual accumulation of good data, the problem should begin to correct itself. Studies by groups such as the National Academies have been helpful in this process.

There is a need for good and widely shared norms for review of analysis and a need for effective interdisciplinary teams. We are not there yet in the area of terrorism risk assessment, but progress is being made.
Secrecy is a serious constraint on progress. However, there is some progress among key sectors that do share data and best practice. More should be done to promote such groups. At the same time, without wider sharing it may be hard to get wide society buy-in and support. The tone of the discussion often encourages people to turn to "snake oil" as an easy answer.

It was argued that some of the emerging guidelines and tools are actually pretty good for addressing specific issues. Their shortcoming is that they are still insufficient with respect to larger scale questions (see Question 7).

DISCUSSION QUESTION 6: *How effective a framework does the public health analogy provide for both fields to think about and address their respective problems? Can we make any progress in suggesting other frameworks that might apply in at least some contexts?*

During the first day of the workshop Fred Schneider presented an overview of the paper:

- Deirdre K. Mulligan and Fred B. Schneider, "Doctrine for Cybersecurity," *Dædalus*, Fall 2011, *140*(4), pp. 70-92.

After that presentation, the group discussed this framework at some length. It is clear that while risks in both fields affect private individuals and organizations, many also have a large element of affecting the "public good." The public health analogy is designed to focus on that element of cyber security and thus is one strategy to address the problem that present methods are "insufficient with respect to larger scale questions."

DISCUSSION QUESTION 7: *The approach of adversarial risk analysis has great intellectual appeal. How far is it likely to be possible to use it in addressing real physical and cyber risks? What can we say about its potential theoretical and practical limitations?*

There were mixed views. The approach can be very helpful in understanding different groups and what they are trying to attack, and hence for choosing more effective strategies. It was argued that the utility of the approach depends very much on context.

Attachment 1: Background Readings

1. General

- Summaries excerpted from Bruce Schneier, *Beyond Fear: Thinking about Security in an Uncertain World*, Copernicus Books, 295pp., 2003.

2. Papers related to terrorism risk assessment and management

- Jason Merrick and Gregory S. Parnell, "A Comparative Analysis of PRA and Intelligent Adversary Methods for Counterterrorism Risk Management," *Risk Analysis*, *31*(9), pp. 1488-1510, 2011.
- Elisabeth Paté-Cornell and Seth Guikema, "Probabilistic Modeling of Terror Threats: A Systems Analysis Approach to Setting Priorities Among Countermeasures," *Military Operations Research*, *7*(4) pp. 5-20, 2002.
- Insua Rios Insua, Jesus Rios ad David Banks, "Adversarial Risk Analysis," *Journal of the American Statistical Association*, *104*(486), pp. 841-854, 2009.
- David L. Banks, Jesus M. Rios Aliaga, David Rios Insua "Adversarial Risk Analysis", *Chapman and Hall/CRC*, 2015
- Seth Guikema, "Modeling Intelligent Adversaries for Terrorism Risk Assessment: Some Necessary Conditions for Adversary Models," *Risk Analysis*, *32*(7), pp. 1117-1121, 2012.
- Ralph L. Keeney and Detlof von Winterfeldt, "A Value Model for Evaluating Homeland Security Decisions," *Risk Analysis*, *31*(9), pp. 1470-1487, 2011.
- Casey Rothschild, Laura McLay and Seth Guikema, "Adversarial Risk Analysis with Incomplete Information: A Level-*k* Approach," *Risk Analysis*, *32*(7), pp. 1219-1231, 2012.
- Vicki M. Bier, Naraphorn Haphuriwat, Jaime Menoyo, Rae Zimmerman and Alison M. Culpen, "Optimal Resource Allocation for Defense of Targets Based on Differing Measures of Attractiveness," *Risk Analysis*, *28*(3), pp. 763-770, 2008.
- National Research Council, *Review of the Department of Homeland Security's Approach to Risk Analysis*, 160pp., 2010.
- Excerpts from Andrew R. Morral et al., "Modeling Terrorism Risk to the Air Transport System," RAND Report Prepared for the Transportation Security Administration, 95pp., 2012.
- Mark G. Stewart and John Mueller, "Terrorism Risks and Cost-Benefit Analysis of Aviation Security," *Risk Analysis*, *33*(5), pp. 893-908, 2013.
- Naraphorn Haphuriwat, Vicki M. Bier and Henry H. Willis, "Deterring the Smuggling of Nuclear Weapons in Container Freight through Detection and Retaliation," *Decision Analysis*, an INFORMS Journal, *8*(2), pp. 88-102, 2011.

3. Papers related to cyber risk assessment and management

- Deirdre K. Mulligan and Fred B. Schneider, "Doctrine for Cybersecurity," *Dædalus*, Fall 2011, *140*(4), pp. 70-92.
- Butler Lampson, "Privacy and Security - Usable Security: How to get it," *Communications of the ACM*, *52*(11), pp. 25-27, 2009.
- Virgil Gligor, "Dancing with the Adversary: A tale of wimps and giants," pp. 100-115 in B. Christianson et al. (eds.), *Security Protocols*, Springer, 2014.
- Bryan Parno, Jonathan M. McCune and Adrian Perrig, "Bootstrapping Trust in Commodity Computers," *IEEE Symposium on Security and Privacy*, pp. 414-430, 2010.
- Mehran Bozorgi, Lawrence K. Saul, Stefan Savage and Geoffrey M. Voelker, "Beyond

Heuristics: Learning to Classify Vulnerabilities and Predict Exploits," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 105-114, ACM, 2010.

- Pascale Carayon, Sara Kraemer and Vicki Bier, "Ch 3: Human Factors Issues in Computer and E-Business Security," *Handbook of Integrated Risk Management for E-Business: Measuring, Modeling and Managing Risk,* J. Ross Publishing, pp. 63-85, 2005.
- Sjouke Mauw and Martijn Oostdijk, "Foundations of Attack Trees," in *Information Security and Cryptology-ICISC 2005*, Springer, pp. 186-198, 2006.
- Pratyusa K. Manadhata and Jeannette M. Wing, "An Attack Surface Metric," *IEEE Transactions on Software Engineering, 37*(3), pp. 371-386, 2011.
- Marie Vasek, John Wadleigh and Tyler Moore, "Hacking Is Not Random: A case-control study of webserver-compromise risk," *IEEE Transactions on Dependable and Secure Computing*, in press, 2015.
- Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson et al., "Click Trajectories: End-to-end analysis of the spam value chain," in *IEEE Symposium on Security and Privacy (SP),* pp. 431-446, 2011.
- Yilin Mo, Tiffany Hyun-Jin Kim, Kenneth Brancik, Dona Dickinson, Heejo Lee, Adrian Perrig and Bruno Sinopoli, "Cyber-Physical Security of a Smart Grid Infrastructure," *Proceedings of the IEEE*, pp. 195-209, 2011.
- Stephen Papa, William Casper and Tyler Moore, "Securing Wastewater Facilities from Accidental and Intentional Harm: A cost-benefit analysis," *International Journal of Critical Infrastructure Protection, 6*(2), pp. 96-106, 2013.
- Peter G. Neumann, "Inside Risks - Risks and Myths of Cloud Computing and Cloud Storage," *Communications of the ACM*, *57*(10), pp. 25-27, 2014.
- Jonathan Spring, Sarah Kern and Alec Summers, "Global Adversarial Capability Modeling," Tenth Symposium on Electronic Crime Research (eCrime), *IEEE Computer Society's Technical Committee on Security and Privacy*, 2015.

Attachment 2: Workshop Agenda

Thursday, May 28

Introduction: Granger Morgan

Session 1: Tools and strategies for terrorism risk assessment and management:

- Lessons from a decade of using risk analysis to manage terrorism risks – Henry Willis
- Defender-attacker decision tree analysis to counter terrorism – Detlof von Winterfeldt
- Two views on adversarial risk analysis – Seth Guikema and Jesus Rios
- Two (and a half) views on using risk analysis methods in cyber security – Vicki Bier and Elisabeth Paté-Cornell with Marshall Kuypers
- Discussion: What, if any, of this sounds like it might be useful in cyber risk assessment and management?
  - Brief initial observations from Earl Boebert, Doug Sicker, James Larus and Herb Lin, followed by general discussion

Session 2: Tools and strategies for cyber risk assessment and management:

- Risk assessment for cybersecurity: Some challenges and barriers – Herb Lin
- Thinking about cyber security like public health – Fred Schneider
- Lowering aspirations: Some reflections on red and green machines – Butler Lampson
- Wimps, giants, and persistent attackers – Virgil Gligor
- Data-driven vulnerability assessment – Stefan Savage
- Round table on other ideas in cyber security that we should be talking about – Earl Boebert, Tyler Moore, Peter Neumann and Milind Tambe
- Discussion: What, if any, of this sounds like it might be useful in terrorism risk assessment and management?
  - Initial observations from Elisabeth Paté-Cornell, Detlof von Winterfeldt and Seth Guikema followed by general discussion

Friday, May 29

Session 3: Discussion questions

1: *How should we identify and strike the appropriate mix/balance between hardening and protection versus resilience and recovery? What if any can be done to support what is clearly a normative (i.e. value-based) choice? What insights can either field draw from the other in this context?*

2: *What are the opportunities to merge or integrate approaches between the two fields? For example, do cyber–physical systems present such an opportunity? Where might cyber (or physical) attack play a role as a "force multiplier?" What are the opportunities for and limitations of the use of scenario analysis, red teaming, etc.?*

3: *The insurance and reinsurance industries are deeply concerned both about issues of terrorism and cyber risk. How can the analytical and research communities best support their needs?*

4: *Given the various categories of attackers:*

Terrorism	*Cyber*
Petty criminals	*Ankle biters*
Small time and lone terrorists	*Larcenists*
Coordinated individuals or small groups (IEDs)	*Dedicated focused non-state*
Major coordinated attacks (Sep 11)	*Dedicated focused state w/ social, intelligence, or military motivations*

What are the differences in:

- *Data availability and use?*
- *Analytical approaches?*
- *Threat assessment and consequence assessment?*
- *Opportunities for integration/collaboration between our two fields?*

Potential for "disproportionate impact" on national confidence (and lifestyle) and how could we assess in advance?

5: *In both terrorism risk analysis and cyber security there are too many problems and adversaries and relatively too few good analysts. What are the best strategies to raise the level of analysis and minimize the amount of "snake oil"?*

6: *How effective a framework does the public health analogy provide for both fields to think about and address their respective problems? Can we make any progress in suggesting other frameworks that might apply in at least some contexts?*

7: *The approach of adversarial risk analysis has great intellectual appeal. How far is it likely to be possible to use it in addressing real physical and cyber risks? What can we say about its potential theoretical and practical limitations?*

Attachment 3: List of Participants

The workshop was held under the Chatham House rule. Information, views and opinions discussed at the workshop are summarized in the workshop report, but not attributed. IRGC wishes to thank all the participants (listed below) of the workshop for their time and expert contributions, many of whom wrote pieces to prepare discussions at the workshop. While the IRGC endorses the recommendations provided in this summary workshop report, these recommendations do not necessarily represent the views of workshop participants or their employer.

Vicki Bier, Professor, University of Wisconsin-Madison

William (Earl) Boebert, Retired, Sandia National Labs

Robin Dillon-Merrill, Professor, McDonough School of Business, Georgetown University

Virgil Gligor, Professor and co-diretor Cylab, Carnegie Mellon University

Seth Guikema, Associate Professor, Johns Hopkins University

Marshall Kuypers, PhD Student, Stanford University

Butler Lampson, Technical Fellow, Microsoft

James Larus, Professor and Dean, Dept of Computer and Communication Science, EPFL (Ecole Polytechnique Fédérale de Lausanne),

Herb Lin, Senior Research Scholar, Stanford University

Matthew McCabe, Senior Vice President, Marsh

Tyler Moore, Assistant Professor of Computer Science, Southern Methodist University

Granger Morgan, Professor, Carnegie Mellon University

Peter G Neumann, Senior Principal Scientist, SRI International Computer Science Lab,

Elisabeth Paté-Cornell, Professor, Management Science and Engineering, Stanford University

Jesus Rios, Research Staff Member, IBM T.J. Watson Research Center

Stefan Savage, Professor, UC San Diego

Fred Schneider, Professor and Chair Department of Computer Science, Cornell University

Susan Shay, Head Financial Lines, Casualty Reinsurance, Swiss Reinsurance America Corporation

Douglas Sicker, Professor, Carnegie Mellon University

Milind Tambe, Professor, University of Southern California

Detlof von Winterfeldt, Professor, University of Southern California

Sam Weber, Secure Software and Systems Senior Researcher, SEI Arlington, VA office

Henry Willis, Director, RAND Homeland Security and Defence Center; Professor, Pardee RAND Graduate School

Addendum: Short pieces to support and complement contributions made at the workshop

- Risk analysis for the evaluation of cyber threat reduction and counter-terrorism policies
  Elisabeth Paté-Cornell and Marshall Kuypers, Stanford University

- A Holistic View of Terrorism, Cybersecurity, and Risk Assessment
  Peter G. Neumann, SRI International

Risk analysis for the evaluation of cyber threat reduction and counter-terrorism policies - Elisabeth Paté-Cornell and Marshall Kuypers, Stanford University

At this time, there are significant uncertainties about the cost-effectiveness of cyber security investments. Chief Information Security Officers generally do not have an effective framework to compare the risks of different kinds of attack, and therefore, the value of investments in various security safeguards such as encryption technology, employee cyber awareness training, enterprise firewalls or penetration tests. Since there are no clear methods to assess the risk reduction associated with security investments, organizations may purchase ineffective products from various security vendors. The problem of assessing and mitigating the risks of different kinds of cyberattacks is similar in many ways to that of protecting the country, or different organizations, against terrorist attacks. Methods and examples presented in the report describe the risk assessment and game analysis methods that have been proposed.

While many organizations record cyberattack data, few organizations are fully leveraging them. Existing statistical tools address the attack frequencies and the distribution of impacts of basic, frequent cyber incidents. These variables are valuable in providing input to risk models for these kinds of attacks and their effects in the future. Common attacks that occur regularly include phishing, ransomware, and malware infections, and standard probabilistic risk analysis tools and techniques can adequately assess these cyber risks. Based on an actual data set, Stanford researchers are currently developing probabilistic models to incorporate incident data into cyber risk models (Kuypers and Pate-Cornell). However, the 'persistent attackers', often nation states, may require a different set of models because there may not exist sufficient data about their repeated attempts to intrusion. Furthermore, the adaptive nature of these kinds of adversaries and the repetition of attacks (and in some cases, response from the attacked) may require behavioral and game analyses to inform security investment decisions.

These quantitative methods are similar to those used to address terrorism problems but need to be adapted to represent the spectrum of attackers, target organizations, means of penetration and effectiveness of cyber defenses.

A Holistic View of Terrorism, Cybersecurity, and Risk Assessment - Peter G. Neumann, SRI International

Terrorist attacks might have nothing to do with computers (as in the Boston Marathon), or they might affect cyberphysical systems (which are heavily dependent on computer-communication technology), or might be direct cyberattacks on computer systems -- some coordinated combination of these.
An important realization for any risk assessment is that to be effective, it must address the total system rather than just the component pieces. This implies addressing concerns of system flaws, exploitable vulnerabilities, environmental upsets, and sophisticated or end scripted attacks.

It is my contention that we live in a world in which there is essentially only very weak cybersecurity in our computer systems and networks, and indeed in our critical national infrastructures that depend on computer technology -- often connected directly or indirectly to the Internet.
Basically, every system is riddled with potentially exploitable security vulnerabilities.

Although risk assessment and risk management can have some useful but relatively small short-term effects in slightly raising the bar, the most fundamental step forward would be the development of some meaningfully trustworthy total systems that can avoid or dramatically reduce some of the risks. One example of such a computer system is the CHERI system, involving clean-slate hardware-software design and implementation aimed specifically at security, resilience, and dynamic adaptability. CHERI has the ability to run potentially untrustworthy software in a constrained compartment in which it can do no harm, to provide very trustworthy components on which to build trustworthy applications, and in which migration paths exist between compartmented legacy software and high-end software explicitly compiled to understand the security capabilities of the hardware. The system hardware, low-layer software, and compilers have the ability to avoid many of the common design and programming security flaws. In addition, the hardware specifications have a formal basis, and are being subjected to formal analysis. The CHERI system hardware specifications and the software are all open-sourced and available. Tech transfer to the commercial world is in progress, so we have hopes that some of this technology will be available some when in the future.

Some serious problems arise in dual attempts that are evidently somewhat conflicting -- first, efforts to minimize the risks of terrorism, and second, efforts to achieve meaningfully trustworthy computer systems and networks with respect to security, reliability, resilience, and survivability in the presence of external attacks, insider misuse, denials of service, natural disasters, environmental difficulties, and other adversities. Consequential damages have included financial damages, losses of life and accidental injuries due to computer and human errors, privacy violations, and much more. Many risks are considered in my book, Computer-Related Risks (Addison-Wesley, 1995), which also considers what might be necessary to avoid those risks.

In the aftermath of the events of September 11, we have seen a serious potential erosions of personal privacy as a result of extensive surveillance that was hitherto not recognized before the revelations of some of Snowden's information. This has raised a seemingly dichotomous situation between the desire by many governments for ubiquitous surveillance and the need for meaningfully trustworthy computer systems. It would be very interesting to conduct a total-system examination of the risks of each of these would-be goals and also of the risks of

not achieving each of these goals. Of course, such studies would have to transcend national boundaries and recognize the international nature of the problems. Such analyses might demonstrate that the losses of privacy and risks would be negligible compared with the absence of ubiquitous surveillance, or perhaps the converse -- that the risks of having compromisible and seriously flawed systems would seriously undermine personal, corporate, national, and international well-being. However, perhaps more likely, such risk analyses might show that there is no reasonable middle ground between ubiquitous surveillance (with back doors, "carefully" authenticated front doors for law enforcement and intelligence inserted into already compromisible systems and networks) and the consequent increase in disasters and serious disruptions resulting from more easily compromised computer technology.

One other riskful situation deserves noting. I have written an article in my Inside Risks series in the Communiations of the ACM on the risks of not anticipating certain catastrophic events that might render cybersecurity totally ineffective. This article arose from a workshop called Catacrypt, which sought to explore the risks that might result from a breakthrough in factoring large numbers are solving discrete logarithmic equations, both of which could destroy existing public-key cryptography for which there are no practical alternatives ready for prime time (if you will pardon the pun).
This is just one example of something relatively unexpected that could be seriously disruptive.

In all of the above considerations, there is a serious concern that risk analyses might be based on faulty assumptions, that the models might be flawed or intentionally rigged, that the analysis tools were themselves flawed, or perhaps that the entire process was concocted to show a desired result. It is very difficult for such analyses to closely resemble reality when there are so many unkowns and opportunities for errors.

References

Robert N. M. Watson, Jonathan Woodruff, Peter G. Neumann, Simon W. Moore, Jonathan Anderson, David Chisnall, Nirav Dave, Brooks Davis, Khilan Gudka, Ben Laurie, Steven J. Murdoch, Robert Norton, Michael Roe, Stacey Son, and Munraj Vadera: "CHERI: A Hybrid Capability-System Architecture for Scalable Software Compartmentalization", in *Proceedings of the 36th IEEE Symposium on Security and Privacy ("Oakland"), San Jose, California, USA*, May 2015.
http://www.cl.cam.ac.uk/research/security/ctsrd/pdfs/201505-oakland2015-cheri-compartmentalization.pdf
This paper describes our hardware-software compartmentalisation model, as well as implications for operating-system and application design.

David Chisnall, Colin Rothwell, Brooks Davis, Robert N.M. Watson, Jonathan Woodruff, Simon W. Moore, Peter G. Neumann and Michael Roe: "Beyond the PDP-11: Architectural support for a memory-safe C abstract machine", in *Proceedings of Architectural Support for Programming Languages and Operating Systems (ASPLOS 2015), Istanbul, Turkey*, March 2015.
http://www.cl.cam.ac.uk/research/security/ctsrd/pdfs/201503-asplos2015-cheri-cmachine.pdf

Jonathan Woodruff, Robert N. M. Watson, David Chisnall, Simon W. Moore, Jonathan Anderson, Brooks Davis, Ben Laurie, Peter G. Neumann, Robert Norton, and Michael Roe: "The CHERI capability model: Revisiting RISC in an age of risk", in *Proceedings of the 41st International Symposium on Computer Architecture (ISCA 2014), Minneapolis, MN, USA,* June 14--16, 2014.
https://www.cl.cam.ac.uk/research/security/ctsrd/pdfs/201406-isca2014-cheri.pdf
This paper describes the CHERI ISA as of June 2014, and explored architectural and micro-architectural aspects of our work.

International Risk Governance Council

The International Risk Governance Council (IRGC) is an independent foundation based in Switzerland whose purpose is to help improve the understanding and governance of systemic risks that have impacts on human health and safety, on the environment, on the economy and on society at large.

International Risk Governance Council
c/o École Polytechnique Fédérale de Lausanne (EPFL)
CM 1 517
Case Postale 99
CH-1015 Lausanne
Switzerland

Tel +41 21 693 82 90
info@irgc.org

www.irgc.org