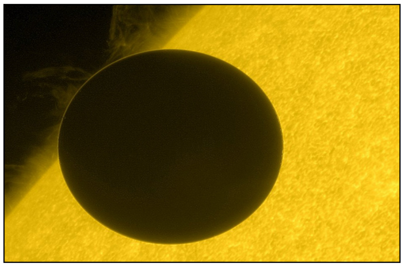
TIM Technology Innovation Management Review

August 2013



Transit of Venus
Image licensed under CC BY JAXA/NASA

Cybersecurity

Welcome to the August 2013 issue of the *Technology Innovation Management Review*. This month's editorial theme is Cybersecurity. We welcome your comments on the articles in this issue as well as suggestions for future article topics and issue themes.

Carleton UNIVERSITY

www.timreview.ca

Publisher

The *Technology Innovation Management Review* is a monthly publication of the Talent First Network.

ISSN

1927-0321

Overview

The *Technology Innovation Management Review* (TIM Review) provides insights about the issues and emerging trends relevant to launching and growing technology businesses. The TIM Review focuses on the theories, strategies, and tools that help small and large technology companies succeed.

Our readers are looking for practical ideas they can apply within their own organizations. The TIM Review brings together diverse viewpoints – from academics, entrepreneurs, companies of all sizes, the public sector, the community sector, and others – to bridge the gap between theory and practice. In particular, we focus on the topics of technology and global entrepreneurship in small and large companies.

We welcome input from readers into upcoming themes. Please visit timreview.ca to suggest themes and nominate authors and guest editors.

Contribute

Contribute to the TIM Review in the following ways:

- Read and comment on past articles and blog posts.
- Review the upcoming themes and tell us what topics you would like to see covered.
- Write an article for a future issue; see the author guidelines and editorial process for details.
- Recommend colleagues as authors or guest editors.
- Give feedback on the website or any other aspect of this publication.
- Sponsor or advertise in the TIM Review.
- Tell a friend or colleague about the TIM Review.

Please contact the Editor if you have any questions or comments: timreview.ca/contact

The PDF version is created with Scribus, an open source desktop publishing program.

Editorial: Cybersecurity

Chris McPhee, Editor-in-Chief

Tony Bailetti, Guest Editor

From the Editor-in-Chief

Welcome to the August 2013 issue of the *Technology Innovation Management Review*. This is the second of two issues covering the editorial theme of Cybersecurity, and I am pleased to introduce our guest editor, **Tony Bailetti**, Director of Carleton University's Technology Innovation Management program (TIM; carleton.ca/tim) in Ottawa, Canada.

In September and October, we will present two issues on Managing Innovation for Tangible Performance, for which the guest editor is **Sorin Cohn**, President of BD *Cohn*sulting Inc. Dr. Cohn also presented the April TIM Lecture on "Enhancing Competitive Position Through Innovation Beyond R&D" (timreview.ca/article/686).

In November, we welcome back **Seppo Leminen**, Principal Lecturer at the Laurea University of Applied Sciences, Finland, and **Mika Westerlund**, Assistant Professor at Carleton University's Sprott School of Business, as guest editors to reprise the theme of Living Labs. Leminen and Westerlund were the guest editors when we covered this theme in our September 2012 issue (timreview.ca/issue/2012/september), and we are looking forward to exploring this theme in even greater depth.

I am also pleased to announce the publication of the TIM program's second ebook: *Business Models for Entrepreneurs and Startups: Best of TIM Review* (tinyurl.com/m3cv88k). This book features 16 of the most insightful, most relevant, and most popular articles published in the TIM Review on the topic of business models. The articles were selected and introduced by **Dr. Steven Muegge**, an Assistant Professor in the Technology Innovation Management Program at Carleton University, and **Claude Haw**, President of Venture Coaches. The foreword was written by **Sir Terence Matthews**, Founder and Chairman of the Board, Mitel Networks Corporation.

We hope you enjoy this issue of the TIM Review and will share your comments online. Please contact us (timreview.ca/contact) with article topics and submissions, suggestions for future themes, and any other feedback.

Chris McPhee
Editor-in-Chief

From the Guest Editor

It is my pleasure to be the guest editor for the July and August issues of the TIM Review, in which we explore the theme of Cybersecurity. These two issues of the journal include 15 contributions from 31 authors, 13 of which are with universities and research institutes; 11 are with industry; and 7 are with the government.

The August issue of the TIM Review includes eight articles. These articles provide: i) an approach to make Canada a global leader in cybersecurity; ii) methods to identify vulnerabilities and countermeasures in networked cyber-physical systems, deliver risk management for enterprises, and analyze all potential pathways of exposure to risk; iii) a research agenda for information system security engineering; iv) overviews of multifactor authentication mechanisms and self-protecting systems; and v) a model to help security providers position their service offers.

Tony Bailetti and **David Hudson** are at Carleton University; **Renaud Levesque** is Director General and **Dan Craigen** and **D'Arcy Walsh** are Science Advisors at the Communications Security Establishment Canada (CSEC); and **Stuart McKeen** is with the Ontario Ministry of Research and Innovation. Their article describes an engine designed to make Canada a global leader in cybersecurity.

Jeff Hughes, President of Tenet 3 and **George Cybenko**, the Dorothy and Walter Gramm Professor of Engineering at Dartmouth College, describe a threat-driven quantitative methodology for identifying vulnerabilities and countermeasures in networked cyber-physical systems. Risk/benefit assessment is performed using a multidisciplinary approach called QuERIES.

Brian Ritchot is a Senior Information Security Consultant with Seccuris Inc. He specializes in the implementation and delivery of intrusion-detection solutions, vulnerability assessment, network analysis, and security architecture. His article provides a business-focused approach to developing and delivering enterprise security architecture for the purpose of providing a sensible and balanced approach to risk management.

Editorial: Cybersecurity

Chris McPhee and Tony Bailetti

Philip O'Neill is Chief Scientist at Deep Logic Solutions Inc. In his article, he presents the strongest-path method of analyzing all potential pathways of exposure to risk – no matter how indirect or circuitous they may be. The network model of infrastructure and operations makes direct use of expert knowledge about entities and dependency relationships without the need for any simulation or any other models.

Rich Goyette and **Yan Robichaud** are Senior Security Architects at Communications Security Establishment Canada and **François Marinie**r is an independent information technology security analyst. They present a research agenda designed to move information system security engineering toward a mature engineering discipline. They propose that a threat model that is actionable from the perspectives of risk management and security engineering and a practical and relevant security-measurement framework be developed as a first step.

Jim Reno, a Distinguished Engineer and Chief Architect for Security at CA Technologies, describes the different mechanisms used to implement multifactor authentication. The article highlights that the selection of a multifactor authentication mechanisms affects both security as well as the overall user experience.

Mahsa Emami-Taba is a doctoral student at the University of Waterloo; **Mehdi Amoui** is a Postdoctoral Fellow working on a joint research project that includes Blackberry Inc. and the University of Waterloo; and **Ladan Tahvildari** is an Associate Professor in the Department of Electrical and Computer Engineering at the University of Waterloo. They provide an overview of self-protecting systems and highlight the importance of creating a holistic decision-making strategy in cybersecurity.

Arto Rajala, a Senior Researcher in the School of Business at Aalto University in Finland; **Mika Westerlund**, an Assistant Professor at Carleton University's Sprott School of Business; **Mervi Murtonen**, a senior scientist at VTT Technical Research Centre of Finland; and **Kim Starck**, a Sales and Security Director at Stanley Security Finland propose a model to help security providers position their service offers. Their 4C model focuses on the conceptualization, calculation, communication, and co-creation of value.

We thank you for reading the journal and urge you to support initiatives to make Canada a leader in cybersecurity worldwide. A nationwide effort to make Canada a global leader in cyberspace offers significant benefits to the users of cyberspace worldwide as well as many opportunities for scholarly inquiry and innovative industrial initiatives.

We hope that you, your colleagues, and your organizations benefit from reading the July and August 2013 issues of the TIM Review.

Tony Bailetti
Guest Editor

About the Editors

Chris McPhee is Editor-in-Chief of the *Technology Innovation Management Review*. Chris holds an MASc degree in Technology Innovation Management from Carleton University in Ottawa and BScH and MSc degrees in Biology from Queen's University in Kingston. He has over 15 years of management, design, and content-development experience in Canada and Scotland, primarily in the science, health, and education sectors. As an advisor and editor, he helps entrepreneurs, executives, and researchers develop and express their ideas.

Tony Bailetti is an Associate Professor in the Sprott School of Business and the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada. Professor Bailetti is the Director of Carleton University's Technology Innovation Management (TIM) program. His research, teaching, and community contributions support technology entrepreneurship, regional economic development, and international co-innovation.

Citation: McPhee, C. and T. Bailetti. 2013. Editorial: Cybersecurity. *Technology Innovation Management Review*. August 2013: 3–4.

Keywords: cybersecurity, cyberthreats, cyberattacks, research, information technology, network security, risk assessment, Canada

Developing an Innovation Engine to Make Canada a Global Leader in Cybersecurity

Tony Bailetti, Dan Craigen, David Hudson,
Renaud Levesque, Stuart McKeen, and D'Arcy Walsh

> *There is one quality which one must possess to win, and that is definiteness of purpose, the knowledge of what one wants, and a burning desire to possess it.*
>
> Napoleon Hill (1883–1970)
> Writer and advisor to President Franklin D. Roosevelt

An engine designed to convert innovation into a country's global leadership position in a specific product market is examined in this article, using Canada and cybersecurity as an example. Five entities are core to the innovation engine: an ecosystem, a project community, an external community, a platform, and a corporation. The ecosystem is the focus of innovation in firm-specific factors that determine outcomes in global competition; the project community is the focus of innovation in research and development; and the external community is the focus of innovation in resources produced and used by economic actors that operate outside of the focal product market. Strategic intent, governance, resource flows, and organizational agreements bind the five entities together. Operating the innovation engine in Canada is expected to improve the level and quality of prosperity, security, and capacity of Canadians, increase the number of Canadian-based companies that successfully compete globally in cybersecurity product markets, and better protect Canada's critical infrastructure. Researchers interested in learning how to create, implement, improve, and grow innovation engines will find this article interesting. The article will also be of interest to senior management teams in industry and government, chief information and technology officers, social and policy analysts, academics, and individual citizens who wish to learn how to secure cyberspace.

Introduction

How can a country become a global leader in a product market and contribute to its own prosperity, security, and capacity? The objective of this article is to examine one response to this research question: the establishment of an engine (i.e., a structure, processes, and values) that converts innovation into system-level results (e.g., prosperity, security, and capacity) that cannot be delivered by a single organization or individual working on its own.

The innovation engine examined in this article cultivates innovation in: i) firm-specific advantages to compete globally; ii) research and development (R&D); and iii) linking with external communities. This engine converts innovation into four system-level results: i) new knowledge jobs; ii) addressed gaps in cybersecurity R&D and in operational limitations; iii) new highly qualified people operating in the cybersecurity space; and iv) sustainable income for the operator of the innovation engine.

We use the authors' experience and knowledge gained designing and growing business ecosystems to offer a generic approach to make a country a global leader in a specific product market. Table 1 list articles published in this journal since 2008, organized on the basis of the nature of their contribution to our understanding of innovation engines and their entities.

Table 1. Contributions that increased our understanding of innovation engines and their key entities

A. Innovation engines	
Engines to convert innovation into desired system-level results	Bailetti and Bot (2013; timreview.ca/article/658)
Models for innovation engines and business ecosystems	Muegge (2011; 495) Muegge (2013; 655) Quinn (2009; 279)
B. Business ecosystems	
Examples of how new technology ventures use business ecosystems to create value	Bailetti (2010; timreview.ca/article/355) Low and Muegge (2013; 703) Rosenblum (2010; 381)
Lessons learned building real-life business ecosystems	Dixon (2011; 441) Milinkovich (2008; 200)
Overview of business ecosystems and identification of distinguishing features	Carbone (2009; 227) Hurley (2009; 276)
Approaches to visualize characteristics of business ecosystems and their evolution	Weiss (2009; 242) Weiss, Sari, and Noori (2013; 683)
C. Projects in business ecosystems	
Examples of projects that co-create value in a business ecosystem	Bailetti and Hudson (2009; timreview.ca/article/308) Westerlund and Leminen (2011; 489) Weiss (2011a; 488) Weiss (2011b; 436)
D. Platforms	
Motivation to use a platform	Makienko (2010; timreview.ca/article/382)
Examples of platform operators and their strategies	Bailetti (2010; 377) Leminen, Westerlund, and Nyström (2012; 602) Mahendran (2008; 114) Majic (2010; 379) Misaka (2013; 684) Muegge and Milev (2009; 245) Noori and Weiss (2013; 647) O'Halloran (2010; 350) Poole (2010; 391) Poole (2011; 446)
E. Foundations	
Examples and characteristics of non-profit organizations that anchor business ecosystems	Prattico (2012; timreview.ca/article/636) Xie (2008; 194) Weiss (2010; 376)

Developing an Innovation Engine to Make Canada a Global Leader in Cybersecurity

Tony Bailetti, Dan Craigen, David Hudson, Renaud Levesque, Stuart McKeen, and D'Arcy Walsh

We use the experience and knowledge gained protecting Canada's critical infrastructures and managing R&D portfolios (Craigen et al., 2013a: timreview.ca/article/704; Craigen et al., 2013b: timreview.ca/article/705) to use Canada and cybersecurity as an example of an application of the innovation engine.

Cyberattacks threaten and limit the benefits that Canadians, as well as citizens of other countries, currently derive from cyberspace. Cyberattacks include, but are not limited to, stealing intellectual property, disrupting critical infrastructure, usurping identity, compromising online bank accounts, creating and distributing viruses, posting confidential information, and encrypting systems to demand ransom. Increasingly, cyberattacks use sophisticated software designed to defeat or bypass security systems. These attacks are criminally or politically motivated, and are executed by very persistent, skilled, and well-funded individuals and organizations.

Cyberattacks that steal intellectual property and disrupt critical infrastructure are particularly damaging. Hard data on the extent of intellectual property theft are difficult to obtain and validate. According to the Canadian Labour Congress, intellectual property theft costs the Canadian economy $22 billion each year (Geist, 2009; tinyurl.com/ptmx2l5). Frontier Economics (2011; tinyurl.com/nauah4a), a research organization based in the United Kingdom, estimates that the theft of intellectual property prevents the world's 20 major economies from collecting €100 billion in tax revenues each year and has "destroyed" 2.5 million legitimate jobs. The Symantec Corporation estimated that companies in the United States lose some $250 billion to intellectual property theft every year; however, this figure has been questioned (Maass and Rajagopalan, 2012; tinyurl.com/c73fp6d).

Critical infrastructure consists of physical and information-technology assets such as energy distribution networks, telecommunications networks, banking systems, manufacturing and transportation systems, and services that support the effective functioning of the private and public sector. Examples of cyberattacks on critical infrastructure include: i) the cyberattacks on Estonia (Ottis, 2013; tinyurl.com/p3juxde) and Georgia (Korns and Kastenberg, 2009; tinyurl.com/oj5ok57); ii) the attack on the Saudi Arabian Oil Company (Bronk and Tikk-Ringas, 2013; tinyurl.com/pegavx8); and iii) brute force attacks on Internet-facing control systems (Industrial Control Systems Cyber Emergency Response Team, 2013; tinyurl.com/q98sqxf).

Cybersecurity will remain a rapidly evolving and significant challenge for the foreseeable future. Protecting cyberspace is a global as well as a domestic priority. There is a sense of urgency for industry, government, academic institutions, not-for-profits, and individuals to work together to ensure that Canadians and citizens of other nations enjoy a secure cyberspace (Auditor General of Canada, 2012; tinyurl.com/otuqxgb). This is easy to say, but very difficult to do. Therein resides the opportunity for Canada to become a leader in cybersecurity.

The global cybersecurity environment presents an increasingly complex set of challenges for Canada (Gendron, 2013; tinyurl.com/p3ela8n). Every adversity, however, has an opportunity couched within. We argue that Canada should act decisively and proactively to become a global leader in cybersecurity. Leadership in this undertaking encompasses the R&D projects; ventures of existing and new companies; content and training; and infrastructures that protect information and information systems.

In this article, we first present the main cybersecurity challenges facing Canada and the ways proposed to improve cybersecurity practice. We then discuss the features of an innovation engine designed to make Canada a global leader in cybersecurity. A unique corporation called the Venus Cybersecurity Corporation anchors the proposed innovation engine. The next section describes the responsibilities and desired results of the corporation. The last section provides the conclusions.

Main Cybersecurity Challenges for Canada

Based on the authors' experience gained protecting electronic information and information infrastructures for the government of Canada, we identify the current challenges faced by those responsible for securing Canada's critical infrastructure. These challenges are:

1. Traditional and ineffective cybersecurity approaches, which focus on prevention, risk management, and deterrence through accountability

2. Uncoordinated approaches between industry, academia, and government

3. Daunting and fractured list of cybersecurity research and development requirements

4. Silo mentality of research disciplines that prevents the development of an interdisciplinary science of cybersecurity

5. Overemphasis on the technical aspects of cybersecurity at the expense of social aspects

6. Chasms between classified and unclassified industry, academia, and government domains

7. Lack of education and training programs in cybersecurity

8. A paucity of Canadian companies operating in the global cybersecurity space

9. An under-investment in cybersecurity-related research and commercialization compared to other jurisdictions

10. Slow and uncoordinated government responses to addressing the root causes of cyberattacks

11. Innovation-stifling contracting processes and procedural requirements of governments (e.g., $25,000 contract limits)

Ways to Improve Cybersecurity Practice

Craigen, Walsh, and Whyte (2013; timreview.ca/article/704) and Craigen, Vandeth, and Walsh (2013; timreview.ca/article/705) offer various suggestions on how to improve the investment in research and experimental development programs in Canada. Their suggestions can be summarized as follows:

1. Establish a healthy ecosystem to incorporate continuously evolving operational concerns into available cybersecurity systems, researchers, and practitioners.

2. Engage social scientists in cybersecurity research.

3. Focus on approaches that: i) are consistent with federal cybersecurity policy; ii) quantitatively assess the cybersecurity risk of complex systems; iii) automate collective action amongst distributed systems to defend individual computers and networks; iv) de-risk emerging technological solutions; v) are ethical and respect privacy concerns; and vi) focus on cyberadversaries, maturity models and standards, “big data”, data scientists, and ways of working and collaborating.

Mulligan and Schneider (2011; tinyurl.com/kt3f3gq) argue that lack of security is the obstacle to success of the information age. Though the problem resides in technologies, the solution requires policies and practices that focus more on the collective than on technology.

Schneier (2008: tinyurl.com/ps78x3y; 2012: tinyurl.com/ousf4cn) argues that understanding the mechanisms of trust is crucial in a connected society. He is a proponent of full disclosure and making security issues public to shed light on the threat as well as encourage its mitigation. According to Schneier, "If researchers don’t go public, things don’t get fixed. Companies don't see it as a security problem; they see it as a public relations problem” (Smith, 2011; tinyurl.com/c34hlbc). Cybersecurity issues as well as their resolutions are community challenges.

The broad set of challenges, the range of stakeholders, and the relationship between the opportunity and national economic well-being suggest that the required response is beyond the capability of any one individual or organization.

Business ecosystems are used to achieve results that no single member can achieve on its own. Business ecosystems provide a networked approach to innovation and commercialization where members act cooperatively for private benefit as well as systemwide benefit (Moore, 2006; tinyurl.com/5rtbj6u). Ecosystems are deeply interlinked. In an ecosystem, a fundamental tension exists between acting in the group's interest and acting in one's own self-interest (Moore, 2006: tinyurl.com/5rtbj6u; Muegge, 2011: timreview.ca/article/495; Schneier 2012: tinyurl.com/oko37dd).

An Engine to Convert Innovation into Desired System-Level Results

We reason that Canada is a country that has the talent, geographical advantage, and political environment to become a global leader in cybersecurity and that an engine that converts innovation into compelling system-level results can be cost-effectively built using the cyclical relationship conceptualization proposed by Muegge (2011; timreview.ca/article/495).

We argue that the innovation engine comprises five key entities (described below), which are linked together by strategic intent, governance, resource flows, and organizational agreements. The innovation engine enhances the firm-specific advantages that determine the outcomes of global competition among firms. These firm-specific advantages include: research and development, size, and managerial capability (Oh and Rugman, 2012; tinyurl.com/o86vnsg), strategic intent of competitors (Hamel and Prahalad, 1989; tinyurl.com/o9evsdh), and capability to use distribution and brand positions to leverage revenue generated in one market to subsidize market-share battles in other markets and increase sales volume (Hamel and Prahalad, 2013; tinyurl.com/p4w6xs9).

Key entities

The five key entities of the innovation engine that is core to the strategy designed to make Canada a global leader in cybersecurity are:

1. The Venus Cybersecurity Ecosystem (hereafter "Venus Cyber Ecosystem")

2. The Venus Cybersecurity Project Community (hereafter "Project Community")

3. The External Community

4. The Venus Cybersecurity Platform (hereafter "Platform")

5. The Venus Cybersecurity Corporation

Figure 1 illustrates the key entities in the proposed innovation engine and the system-level results that are desired by 2017. The entities in Figure 1 exist at different levels of abstraction, the higher the level, the lower the detail presented. The five entities in Figure 1 are interdependent, and each entity relies on the other entities for the innovation engine to achieve the desired system-level results.

The Venus Cyber Ecosystem is the entity at the highest level of abstraction. Ecosystem members include: i) users, buyers, suppliers, partners, and channels of cybersecurity research, products, services, infrastructure, and solutions; ii) new ventures; and iii) the organizations and individuals who serve them (e.g., legal, accounting, intellectual property, economic development organizations) and provide them with requisite inputs (e.g., technology, capital).

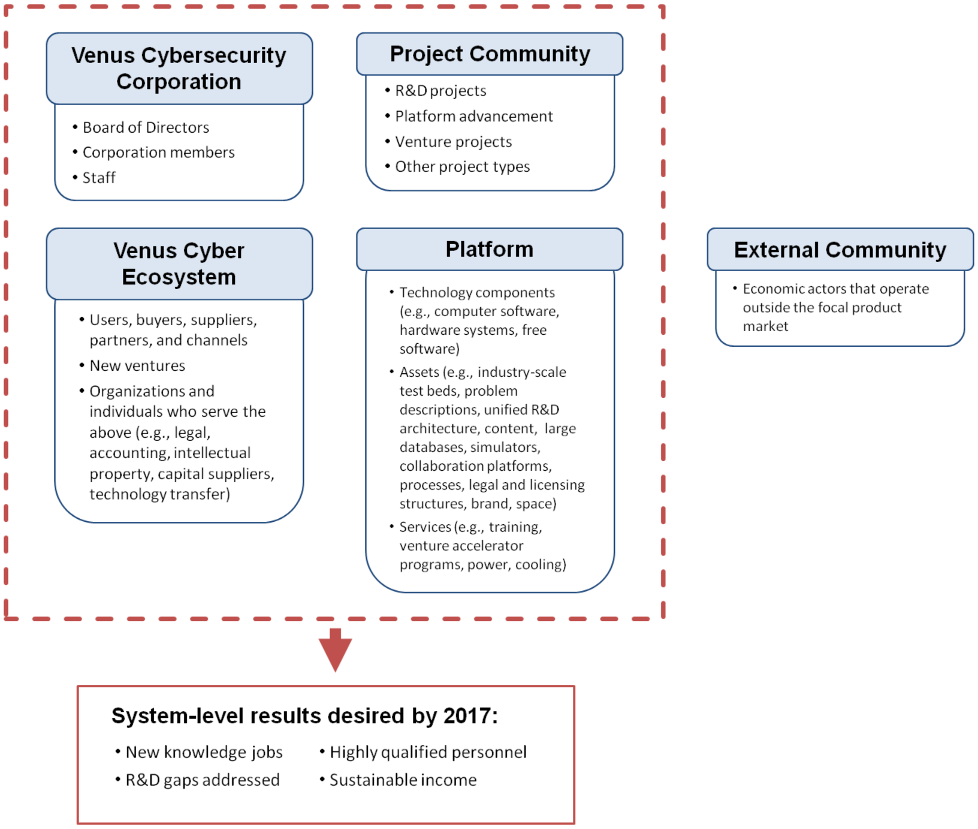

Venus Cybersecurity Corporation
• Board of Directors
• Corporation members
• Staff
Project Community
• R&D projects
• Platform advancement
• Venture projects
• Other project types
Venus Cyber Ecosystem
• Users, buyers, suppliers, partners, and channels
• New ventures
• Organizations and individuals who serve the above (e.g., legal, accounting, intellectual property, capital suppliers, technology transfer)
Platform
• Technology components (e.g., computer software, hardware systems, free software)
• Assets (e.g., industry-scale test beds, problem descriptions, unified R&D architecture, content, large databases, simulators, collaboration platforms, processes, legal and licensing structures, brand, space)
• Services (e.g., training, venture accelerator programs, power, cooling)
External Community
• Economic actors that operate outside the focal product market
System-level results desired by 2017:
• New knowledge jobs
• R&D gaps addressed
• Highly qualified personnel
• Sustainable income

Figure 1. The innovation engine that is core to the strategy designed to make Canada a global leader in cybersecurity

The Project Community comprises the individuals working within a project portfolio sanctioned by the Venus Cybersecurity Corporation. Projects are defined and organized by their desired cybersecurity knowledge, technology, and business outcomes. The project portfolio includes R&D projects to reduce gaps and operational limitations, platform advancement projects, venture projects, and so on. Membership in the Project Community provides rights to engage in one or more of the projects launched by members of the Venus Cybersecurity Corporation.

The External Community refers to people who collaborate outside of and with the Venus Cyber Ecosystem and the Project Community. They may contribute to the Platform. People in the External Community can operate inside and outside Canada. The External Community is a source of human capability, technology, relationships, and other resources. The Venus Cyber Ecosystem and the External Community will exchange resources through the identification of important technology and business opportunities, acceleration of members' businesses, open source developments, standards activities, training seminars, and the like.

The Platform comprises a set of technology components (e.g., computer software, hardware systems, free software), infrastructure (e.g., industry-scale test beds, large databases, simulators, and systems to distribute assets, manage contributions, communicate between members, and coordinate work), assets (e.g., descriptions of industry problems, unified R&D architecture, courseware, validation requirements, legal and intellectual property licensing structures, brand) and services (e.g., training, venture accelerator programs). Members of the Venus Cybersecurity Corporation will be able to use and consume these technology components, infrastructure, assets, and services to develop their market offers and carry out R&D projects as well as other projects.

The Venus Cybersecurity Corporation is an organization that: i) supports and structures the collaboration of organizations and individuals in the Venus Cyber Ecosystem; ii) sustains the strategic intent of making Canada a global leader in cybersecurity over the long term; and iii) advances and operates the Platform. The Venus Cybersecurity Corporation comprises the Board of Directors, Members of the Corporation, and employees.

The Venus Cybersecurity Corporation is a not-for-profit, member-supported corporation. Membership in the Venus Cybersecurity Corporation provides rights to engage in the governance of the corporation to the extent allowed by the various membership levels. Strategic members of the Venus Cybersecurity Corporation pay the highest cash fees and thus will have a significant influence over the direction and strategic intent of the Corporation. Other membership levels can influence the direction of the Corporation through their representation on the Board and through participation at the annual General Meeting.

There are differences between members of the Venus Cybersecurity Corporation and members of the Project Community. For example, members of the Venus Cybersecurity Corporation pay annual (cash) membership fees for which they receive the right to vote on governance matters. Voting rights allow corporation members the ability to shape how the corporation operates and what it achieves relative to its strategic intent. Further, the ability to decide on, and launch, cybersecurity-related projects is the purview of corporate membership. Project Community members can only participate in the specific projects to which they make in-kind contributions or provide cash.

Relationships among the five key entities

Figure 2 illustrates the relationships among the Venus Cyber Ecosystem, the Venus Cybersecurity Corporation, the Platform, the Project Community, and the External Community that produce the desired system-level results. The inner triangle in Figure 2 (shown in heavy red arrows) highlights that the resource cycle of the proposed innovation engine will move from the Platform, to the Venus Cyber Ecosystem, to the Project Community, and back to the Platform.

The Project Community is the focal point of innovation in R&D. Projects leverage their access to the Platform to transform resources received from the Venus Cyber Ecosystem and External Community into technology components, assets, and services that increase the relevance of the Platform.

The Venus Cyber Ecosystem is the focal point of innovation in the factors that determine the outcomes in global competition for Canadian firms and new ventures. Organizations and individuals in this ecosystem leverage the technology components, assets, and services of the Platform to create competitive advantages in the global markets where they operate for their own economic gain and to secure Canada's critical infrastructure.

Developing an Innovation Engine to Make Canada a Global Leader in Cybersecurity

Tony Bailetti, Dan Craigen, David Hudson, Renaud Levesque, Stuart McKeen, and D'Arcy Walsh

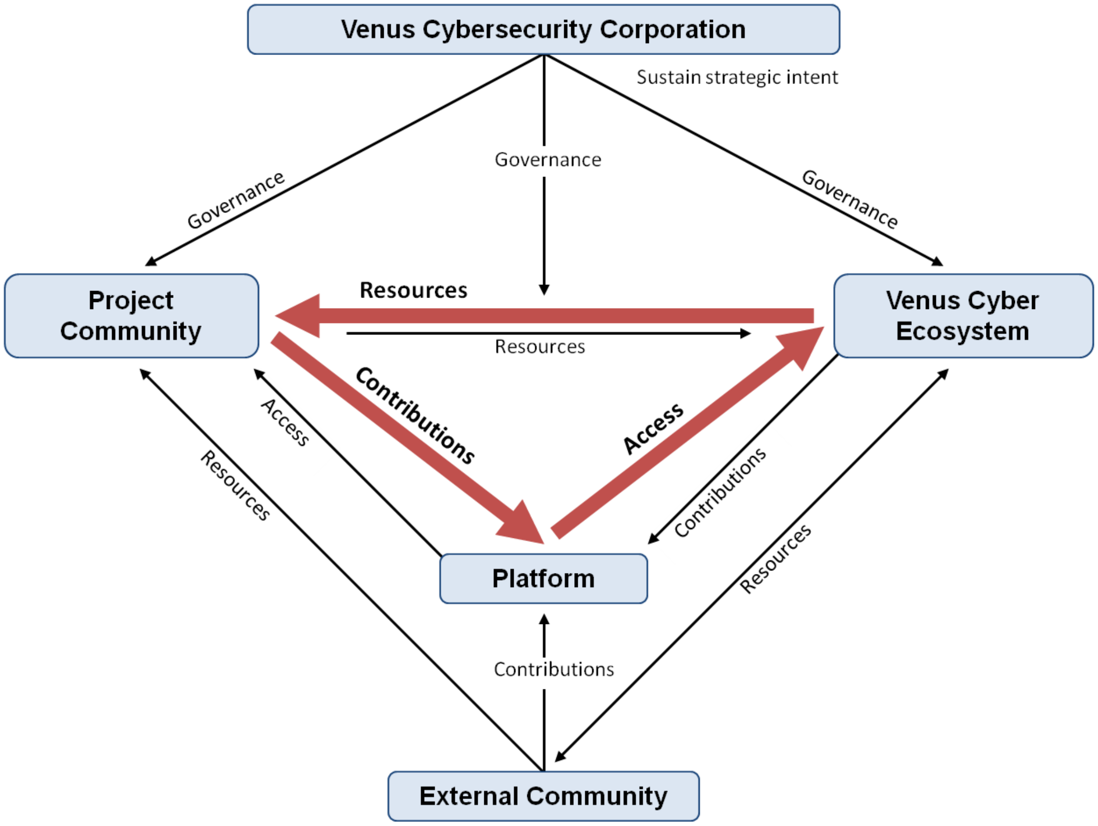


Figure 2. Relationships among the five key entities in the innovation engine

Organizational agreements will enable the following activities:

1. Members of the Venus Cyber Ecosystem will be able to use, extend, and commercialize the assets of the Platform to create and capture economic value.

2. The organizations and individuals in the Venus Cyber Ecosystem and External Community will be able to make the resources required to carry out projects.

3. The Project Community will be able to contribute new technology components and assets to the Platform thereby increasing the Platform's value.

4. Members of the Venus Cyber Ecosystem will be able to contribute technology components and assets acquired from other communities to the Platform.

5. The Project Community will be able to contribute resources such as information, customer leads, and skills to the Venus Cyber Ecosystem.

Venus Cybersecurity Corporation

The Venus Cybersecurity Corporation is the organization that anchors the innovation engine illustrated in Figures 1 and 2. The Venus Cybersecurity Corporation has five important responsibilities:

1. Sustain the strategic intent of the innovation engine over the long term. Strategic intent is an obsession created to attain the desired leadership position and to develop a process that sustains this obsession over the long term. Strategic intent is a vivid picture that captures the essence of winning in cybersecurity and that is stable over time to keep the ecosystem focused. The strategic intent is sufficiently detailed to set targets that deserve personal effort and commitment from members who drive cybersecurity technology and business innovation. Finally, the strategic intent creates a sense of urgency to keep an aggressive pace of ecosystem work and ensures consistency in resource allocation over the long-term (Hamel and Parahalad, 1989; tinyurl.com/o9evsdh).

2. Lead the Venus Cyber Ecosystem and the Project Community. By leadership, we mean organizing groups of people to achieve a common goal.

3. Govern. The corporation will make decisions that define expectations, grant power, allocate resources, and verify performance.

4. Increase the relevance of the Platform. The corporation will use contributions from the Project Community, the Venus Cyber Ecosystem, and External Community to advance the Platform. The Project Community contributions may include up front in-kind contributions as well as project outcomes.

5. Provide access to the Platform. The corporation will provide the Project Community and the Venus Cyber Ecosystem with access to a state of-the-art platform.

Iansiti and Levien (2004a: tinyurl.com/7t4xgvn; 2004b: tinyurl.com/nmfpyms) refer to the organization that anchors a business ecosystem as the "keystone." The responsibilities of the Venus Cybersecurity Corporation include the responsibilities that Iansiti and Levien attributed to a keystone plus an additional one: the leadership role described as the second responsibility above.

Not-for-profit versus for profit
In the Canada/cybersecurity example described in this article, the innovation engine is anchored around a not-for-profit corporation. This decision was made to reduce the time required to make and execute decisions; to increase information and resource exchange among industry, government, and academia; to reduce overhead; and to establish strong links with cybersecurity centres in allied countries.

In Canada, a group of private sector firms should lead the proposed not-for-profit organization. There is not one firm that can lead. Government agencies, universities, and other not-for-profits can join as members.

Desired system-level results
Table 2 identifies the four system-level results that differentiate the Venus Cybersecurity Corporation and that will motivate organizations and individuals to become members in the first four years. These system-level results require a business ecosystem, platform, project, and external communities because they are not attainable by any organization or individual working alone. Table 2 also shows the dimensions that will be used to assess the success of the corporation as of December 31, 2017.

Table 2. System-level results and success dimensions of the Venus Cybersecurity Corporation

Desired system-level results	**Dimensions used to assess success**
1. New knowledge jobs in Canada	• # of jobs created • # companies launched • # and revenue of products delivered to Canadian and international markets
2. R&D gaps and operational limitations in cybersecurity addressed	• # of R&D gaps addressed • # of government programs influenced • # of large industrial programs influenced
3. New highly qualified people operating in the cybersecurity space	• # of academic initiatives generated • # of advanced graduates produced • # of new cyber-related academic disciplines established
4. Income for the corporation	• # of decision makers engaged in the corporation's projects and initiatives • # of jointly sponsored projects in cybersecurity • # of public campaigns recognizing leadership roles of members in the Venus Cybersecurity Corporation

Developing an Innovation Engine to Make Canada a Global Leader in Cybersecurity

Tony Bailetti, Dan Craigen, David Hudson, Renaud Levesque, Stuart McKeen, and D'Arcy Walsh

Conclusion

In this article, we offer a generic approach to making a country a global leader in a specific product market and use Canada and cybersecurity as an example of its application. The success of the proposed innovation engine relies on properly structuring the collaboration among organizations and individuals in an ecosystem, project community, platform, and a corporation, and creating links to communities external to their sphere of activity.

The cybersecurity opportunity is not exclusive to Canada as a country. Other countries are just as well positioned as Canada to become global leaders in cybersecurity. We use Canada as an example because it is the focus of our work. Our "definiteness of purpose" to make Canada a global leader in cybersecurity may encourage our allies to work towards making their countries global leaders as well. We would welcome this outcome. If Canada and its allies commit to attaining global leadership positions in cybersecurity, the rising tide will lift all boats and the networked world will benefit as a result.

Implementation of the innovation engine to make Canada a global leader in cybersecurity through putting the five entities in place is expected to: i) accelerate and strengthen the process of participation through which organizations and individuals work together to achieve results not possible by any entity working on its own; ii) enable continuous improvement and rapid adjustment to environmental changes; iii) increase the positive impact of the results attained; iv) accelerate learning; and v) identify the salient factors that determine a sustainable global leadership position in cybersecurity.

The cybersecurity challenge transcends the abilities of any single organization or individual to address alone. Consequently, academic, private, and public sector participants must unify their efforts when identifying the relevant issues, finding solutions, informing choices, and educating society in direct response to domain-specific requirements for the protection of information technology. This article contributes a way to unify these efforts.

To implement the approach proposed in this article, a task group has been formed. The task group has assumed responsibility for the embryonic development of the proposed innovation engine, including the launch of the Venus Cybersecurity Corporation. This not-for-profit corporation will be launched by March 31, 2014.

Developing an Innovation Engine to Make Canada a Global Leader in Cybersecurity

Tony Bailetti, Dan Craigen, David Hudson, Renaud Levesque, Stuart McKeen, and D’Arcy Walsh

About the Authors

Tony Bailetti is an Associate Professor in the Sprott School of Business and the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada. Professor Bailetti is the Director of Carleton University's Technology Innovation Management (TIM) program. His research, teaching, and community contributions support technology entrepreneurship, regional economic development, and international co-innovation.

Dan Craigen is a Science Advisor at the Communications Security Establishment Canada (CSEC). Previously, he was President of ORA Canada, a company that focused on High Assurance/Formal Methods and distributed its technology to over 60 countries. His research interests include formal methods, the science of cybersecurity, and technology transfer. He was the chair of two NATO research task groups pertaining to validation, verification, and certification of embedded systems and high-assurance technologies. He received his BScH in Math and his MSc in Math from Carleton University in Ottawa, Canada.

David Hudson has recently completed his doctoral studies at Carleton University’s Sprott School of Business in Ottawa, Canada. He is a lecturer in information technology innovation in the MBA program at Sprott, a Director of the Lead to Win entrepreneurship program, and Chair of the Ontario Centres of Excellence advisory board for the Information, Communication, and Digital Media sector. David also consults with Fortune 500 firms on innovation management. Previously, he was the Vice President for advanced research and development at a large technology firm and has had an extensive career in technology development and product line management. David received Bachelor's and Master's degrees in Systems Design Engineering from the University of Waterloo, Canada.

Renaud Levesque is the Director General of Core Systems at the Communications Security Establishment Canada (CSEC), where he is responsible for R&D and systems development. He has significant experience in the delivery of capability and organizational change in highly technical environments. His career began at CSEC in 1986 as a Systems Engineer, responsible for the development and deployment of numerous systems, including the CSEC IP corporate network in 1991. In 2000 Renaud went to work in the private sector as Head of Speech Technologies at Locus Dialogue, and later at Infospace Inc., where he became Director of Speech Solutions Engineering. He rejoined CSEC in 2003, where he assumed the lead role in the IT R&D section. Subsequently, as a Director General, he focused efforts towards the emergence of CSEC's Joint Research Office and The Tutte Institute for Mathematics and Computing. Renaud holds a Bachelor of Engineering from l’École Polytechnique, Université de Montréal, Canada.

Stuart McKeen works for the Ontario Ministry of Research and Innovation (MRI), where he just finished serving a three-year secondment with the Federal Economic Development Agency for Southern Ontario (FedDev). At FedDev, he was both the Agency’s Manager of Innovation and the Manager of Entrepreneurship, Internship, and Youth Programs. He has worked in six different ministries of the Ontario Government over the past 30 years. In 2008, he was awarded the Amethyst Award, the Province of Ontario’s highest employee recognition award for his pioneering work on prospecting and developing large-scale international research consortiums that have brought jobs and investment to Ontario. Stuart holds a BScH degree in Zoology from the University of Western Ontario, Canada and a BA degree in Economics from the University of Toronto, Canada.

D’Arcy Walsh is a Science Advisor at the Communications Security Establishment Canada (CSEC). His research interests include software-engineering methods and techniques that support the development and deployment of dynamic systems, including dynamic languages, dynamic configuration, context-aware systems, and autonomic and autonomous systems. He received his BAH from Queen’s University in Kingston, Canada, and he received his BCS, his MCS, and his PhD in Computer Science from Carleton University in Ottawa, Canada.

Citation: Bailetti, T., D. Craigen, D. Hudson, R. Levesque, S. McKeen, and D. Walsh. 2013. Developiong an Innovation Engine to Make Canada a Global Leader in Cybersecurity. *Technology Innovation Management Review.* August 2013: 5–14.

Keywords: cybersecurity, innovation engine, business ecosystem, innovation in research and development, innovation in commercialization

Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity

Jeff Hughes and George Cybenko

> “*Risk comes from not knowing what you're doing.*”
>
> Warren Buffett
> Business magnate, investor, and philanthropist

Progress in operational cybersecurity has been difficult to demonstrate. In spite of the considerable research and development investments made for more than 30 years, many government, industrial, financial, and consumer information systems continue to be successfully attacked and exploited on a routine basis. One of the main reasons that progress has been so meagre is that most technical cybersecurity solutions that have been proposed to-date have been point solutions that fail to address operational tradeoffs, implementation costs, and consequent adversary adaptations across the full spectrum of vulnerabilities. Furthermore, sound prescriptive security principles previously established, such as the Orange Book, have been difficult to apply given current system complexity and acquisition approaches. To address these issues, the authors have developed threat-based descriptive methodologies to more completely identify system vulnerabilities, to quantify the effectiveness of possible protections against those vulnerabilities, and to evaluate operational consequences and tradeoffs of possible protections.

This article begins with a discussion of the tradeoffs among seemingly different system security properties such as confidentiality, integrity, and availability. We develop a quantitative framework for understanding these tradeoffs and the issues that arise when those security properties are all in play within an organization. Once security goals and candidate protections are identified, risk/benefit assessments can be performed using a novel multidisciplinary approach, called “QuERIES.” The article ends with a threat-driven quantitative methodology, called “The Three Tenets”, for identifying vulnerabilities and countermeasures in networked cyber-physical systems. The goal of this article is to offer operational guidance, based on the techniques presented here, for informed decision making about cyber-physical system security.

Introduction

Cyberattacks are increasing in frequency and severity. Prolexic Technologies (2013; tinyurl.com/n66algm) reports that the average packet-per-second rate of distributed denial-of-service (DDOS) attacks reached 47.4 million packets per second and the corresponding average bandwidth reached 49.24 Gbps in the second quarter of 2013. These are increases of 1,655% and 925% respectively over 2012.

Although DDOS attacks are relatively brutish cyberweapons, the so-called “advanced persistent threat” (APT) refers to sophisticated attackers who operate more subtly against specific targets with specific goals. For example, Operation Aurora deployed a zero-day web-browser exploit to extract detailed intellectual property from high-tech companies (McAfee Inc, 2010; tinyurl.com/np89339).

Whether done with blunt objects (DDOS) or scalpels (APT), cyberattacks continue to be effective. In fact, enterprise IT security managers believe their networks are becoming less secure. A survey of 671 IT security practitioners conducted by the Ponemon Institute (2012; tinyurl.com/afk94px) found that only 33% believed their IT networks were more secure in 2012 than in 2011. In spite of such concerns, a recent Oracle study (2013; tinyurl.com/l76858h) found that, even with increased overall IT security spending, enterprises are still not protecting the right assets.

Combining all these facts and findings, it is evident that the growing cyberthreat environment is becoming more complex and more targeted while our ability to respond with appropriate defences at the appropriate investment levels is becoming more difficult.

The cybersecurity research, development, and vendor communities have not been helping matters. Most researchers and vendors promote their specific point solutions at the expense of seeing the bigger security picture. For example, on the one hand, the "build security in" community advocates redesigning and rebuilding IT systems from scratch to be more secure from the start (e.g., U.S. Department of Homeland Security: tinyurl.com/mh4a2e3; Darpa: tinyurl.com/6nf5yp3; McGraw, 2013: tinyurl.com/mu4oz24). On the other hand, "big data" security technologies promote extensive IT instrumentation, logging, and analysis for whatever application and network infrastructure that has already been deployed (e.g., Hewlett Packard: tinyurl.com/kdsrvuj; Splunk: tinyurl.com/kj3ujkp).

These extremes beg the key question of what combination of cyberdefensives are appropriate for securing an enterprise from the spectrum of threats that it realistically faces. Government efforts at articulating security best practices and risk assessments (e.g., National Institute of Standards and Technology, 2013: tinyurl.com/c8vukj7) are comprehensive and noble but too generic to be operationally prescriptive for such purposes.

New ideas are needed for enterprise-level cybersecurity assessment and investment. The novel approach proposed in this article is based on the authors' 30 years of combined experience in securing complex cyber-physical systems in government and private sector environments. The approach consists of three ingredients that will be outlined in detail below:

1. Confidentiality, integrity, and availability requirements

2. Quantification and assessment of cybersecurity defence investments

3. Identification of cybersecurity threats and vulnerabilities

This article is organized around these ingredients, as follows. The second section argues that tradeoffs between confidentiality, integrity, and availability are intrinsically unavoidable in typical enterprise operations and proposes an analytic framework for managing those tradeoffs. The third section describes a methodology for quantifying the impact of vulnerabilities and defences that are used to mitigate them, namely "QuERIES". The fourth section presents the underlying cybersecurity model, called "The Three Tenets" of cybersecurity-vulnerability assessment and mitigation. Finally, the fifth section provides a summary and a discussion of ways forward based on these results.

Confidentiality, Integrity, and Availability Requirements

Security considerations and metrics are not the only criteria enterprise IT managers use to make decisions. Revenue (or service in the case of a non-profit or government entity) is a result of providing users access to networked services and information and so it is often a primary driver when trading off security against access.

In practice, decision makers must constantly balance availability (i.e., the ability of end users to derive benefit from the system), confidentiality (i.e., the protection of information from access by unauthorized users), and integrity (i.e., the protection of information from unauthorized modification). This task involves complex, typically enterprise- and system-specific, tradeoffs that require an appropriate balance between properties that are not entirely consistent with each other.

In order to make such tradeoff decisions more rigorous and quantitative, we have started to develop a model and corresponding framework for confidentiality, integrity, and availability (CIA) risk management. Here, we briefly introduce our work on the specific issue of introducing "diversity" into an enterprise IT environment for the purpose of increasing "security". Information-system diversity, as opposed to "monoculture", has often been praised as a mechanism for building more resilient and secure systems, ones in which the compromise of one system does not immediately translate into the subsequent compromise of all similar systems.

Diversity can be introduced into an IT system by deploying hardware and software from different vendors or by mechanisms such as randomizing address layouts or compiler generation of executable code (Jajodia et al., 2011; tinyurl.com/mz5d8fn). Further details of the model and associated results can be found in a forthcoming technical paper devoted to this issue (Cybenko and Hughes, *in press*; tinyurl.com/m3jexfv).

Our basic model assumes a network of nodes that comprise an asynchronous distributed system that an enterprise operates. These nodes could be mirrored web or database servers, clients, routers, or other replicated devices or services in an information system. The designer has a choice of making the components the same (i.e., homogeneous or a monoculture) or making the components different in some way (i.e., diverse, moving targets, heterogeneous, or some other approach).

A compromise of a node (component or device) means that an attacker has control of that node, such as root or administrator privileges in an operating system, for example. The goals of a compromise are often summarized as violating one or more of the confidentiality, integrity, and availability properties (Smith and Marchesini, 2008; tinyurl.com/l8jx7op). We interpret these goals in the context of a networked system of functionally redundant components. In this article, we define these concepts as follows:

- *Availability* means that at least one of the nodes has not been compromised and is therefore functioning properly. Stated otherwise, not all of the nodes in the system have been compromised and so at least one is still functioning in a reliability theory sense.

- *Confidentiality* means that none of the nodes have been compromised. This definition is based on the assumption that all clients, servers, or other nodes under consideration contain or have access to critical, possibly the same, information. Therefore, if one node is compromised, that critical information is available to the attacker and so confidentiality of the overall system has been breached.

- *Integrity* means that a majority of the nodes (components) have not been compromised so that, if we request information from the components and compare results, at least one half of the results will match. Once an attacker has compromised more than one half of the components, we no longer have any confidence that the information being provided by a majority is correct. Byzantine failures (Lamport et al., 1982; tinyurl.com/klewe3x) can also be modelled in this framework whereby at least one third, a different but constant fraction, of the components need to be compromised for an integrity attack.

The time-to-compromise, t_i, of the ith node is a random variable distributed according to a probability density function, $f_i(t)$. The concept of time-to-compromise, discussed in more detail below, is based on the premise that any node is ultimately compromisable and the time when an attacker achieves the compromise is a random variable (which can include the attacker's skill level, resources, choice of attack strategies, and so on).

For example, the time to achieve success in a brute-force attack on a password would be distributed according to a uniform density between time 0 (when the attack begins) and time N/M where there are N possible passwords and M random passwords are tried per time unit. Techniques for estimating f_i and t_i for more complex computing systems have been developed and evaluated by the authors (Carin et al., 2008; tinyurl.com/mfyxu9r). Moreover, estimates of the time-to-compromise density allow us to estimate the cost-to-compromise of the ith component as well as the overall system or mission.

For simplicity, we assume that each density has the same form for each component and define α to be the lower bound on the support of f, β to be the upper bound on the support of f, μ to be the mean, and m to be the median. Moreover, we let n denote the degree of diversity (i.e., the number of distinct versions, for example, where clearly $n = 1$ represents a monoculture) as well as the number of parallel attackers such as would occur in a coordinated nation-state or organized crime attack.

Table 1 summarizes several analyses we have performed. The columns labelled Attackers and Diversity are as described; the entries in the columns for C, I, and A are the expected times to compromise confidentiality, integrity, and availability respectively. A graphical depiction of this analysis for the last line in the table is shown in Figure 1 to illustrate the wide variability on time-to-

Table 1. Expected times for an attacker to achieve one of the CIA goals

Attackers	Diversity	*C*	*I*	*A*
1	1	μ	μ	μ
1	n	μ	$n\mu/2$	$n\mu$
n	1	α	α	α
n	n	α	m	β

Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity

Jeff Hughes and George Cybenko

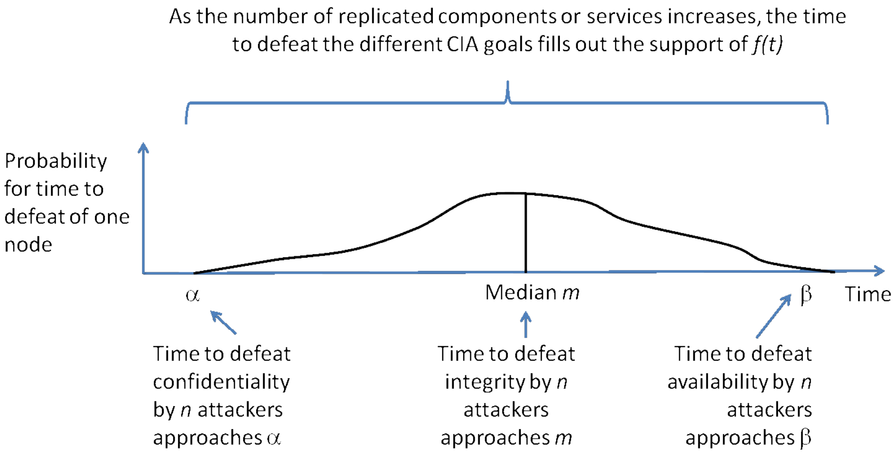


Figure 1. Expected times to achieve CIA security goals with n-fold diversity and n parallel attackers

compromise under these different scenarios. In this situation, where we have n-fold diversity and n parallel attackers, the expected times to achieve the various CIA security goals varies significantly. Decision makers must understand which security properties are most important to their organization's missions and invest accordingly.

This work quantitatively develops the trade space between confidentiality, integrity, and availability as a function of network diversity and time-to-compromise. In any such trade space, the IT manager must determine the "operating point of the design" or the balance between security properties and other important system properties such as "maintainability" and "mission utility".

It is informative to craft a simple opportunity-cost comparison based on this trade space. For instance, "cost-to-disrupt" is a cost to the adversary to compromise the enterprise that is directly estimated from the time-to-compromise scenarios provided above. This cost can be contrasted to the "cost-of-mission-disruption", which is a cost to the IT manager when considering the three types of security objectives (i.e., CIA) and the compromise of which disrupts the enterprise’s mission (e.g., a disruption cost can be proportional to the number of users affected). Hence, analytically describing the trade space enables a richer strategic analysis regarding various IT enterprise objectives. This type of analysis is more explicitly described using the methodology in the next section.

Quantifying Cybersecurity Risk: The QuERIES Technique

The discussion above provides a framework and methodology for identifying various security goals and understanding the possible tradeoffs between them. Moving forward, if we are given a collection of identified security vulnerabilities impeding the achievement of the goals and possible defences or responses to those vulnerabilities, then we would next like to have some sense of how effective the proposed defences are in terms of performance metrics that go beyond a simple checklist.

In physical security, the time-to-compromise of a system is an accepted and measurable performance metric. Consider for example the case of the Overly Door Company (tinyurl.com/kdfdjm2), a supplier of US government General Services Administration approved Class 5 security vault doors suitable for storing national security information. These doors must provide protection against unauthorized entry for the following periods of time:

- 20 man-hours surreptitious entry
- 30 man-minutes covert entry
- 10 man-minutes forced entry
- 20 man-hours against manipulation of the lock
- 20 man-hours against radiological attack

Note that different times are specified for different types of attacks. Surreptitious entry means a method of entry that would not be detectable during normal use or during inspection by a qualified person. Covert entry means a method of entry that would leave evidence, but would not be detectable by a user during normal use of the door and would only be detectable during inspection by a qualified person. Forced entry means a method of entry that would leave evidence of the attack and would be readily discernible in the normal use of the door; the attacker has no concern over leaving evidence that the vault door has been penetrated. Manipulation of the lock is defined as the opening of the combination lock without alteration of the physical structure or disarranging of parts. Ordinarily, manipulation would be accomplished by movement of the lock dial. Entry by radiological attack means the use of radioactive isotopes and other sources judged to be effective in determining the locks combination. Any entry made under these conditions within 20 man-hours shall be considered a failure of the vault door.

Physical security is relatively mature with much operational experience, so measures such as these have emerged as accepted standards within that community. Cyber-physical system security is still relatively new, so such performance measures are not yet standardized.

The authors have developed a technique, called QuERIES, for quantifying cybersecurity risk using ideas from a variety of disciplines and have demonstrated those techniques in software protection scenarios.

An example result of using the QuERIES methodology is depicted in Figure 2, which shows the distribution of times for completing a successful attack against a protected software system. The horizontal axis is time (equatable to cost in man-hours) and the vertical axis is the percentage of attempts requiring the corresponding time. The empirical probability density function was estimated using the QuERIES methodology and is depicted by the vertical bars. The specific compromise that was modeled in this example was an attack against a protected software system by an adversary whose goal is to extract specific parameter values from the protected code.

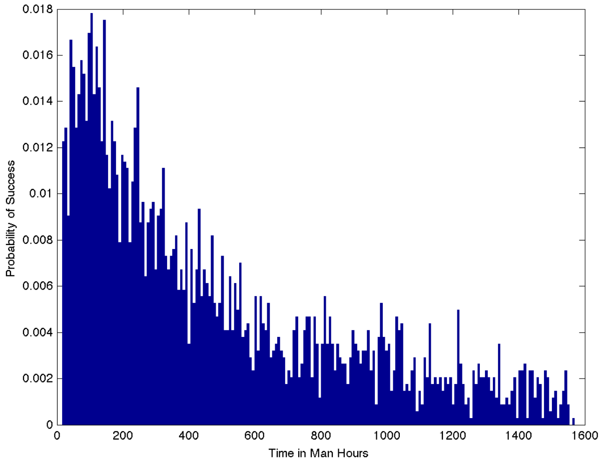


Figure 2. An example time-to-compromise density function for a software-protection defence developed by the authors in previous work on QuERIES

The plot in Figure 2 shows a probability density function for the time required by an attacker to compromise the protections. We model the time-to-compromise as a random variable in QuERIES because it depends on the skill level and approach an attacker takes. It might also depend on luck. Consider, for example, that we do in fact model brute-force password attacks in this way already – an attacker can be lucky and very quickly guess correctly but with very small probability. For a brute-force password attack, the corresponding plot would be a horizontal line at a very small probability going very far into the future.

The QuERIES methodology consists of a number of steps and has been successfully applied in a variety of cyber security situations. All seven steps in the QuERIES methodology are depicted on the right side of Figure 3; the four major ingredients of the methodology are shown on the left side of Figure 3 and are described below:

1. **Model the Problem**: Obtain objective quantities such as the economic value of the intellectual property (IP)(i.e., the protected software asset) to the IP owner; the cost of developing the IP by an adversary; the cost of obtaining the IP through other possible means; and a map of the specific protections applied to the IP asset.

2. **Model the Attacks:** Use the protection map and knowledge of reverse-engineering methodologies to build an attack graph represented as a partially observable Markov decision process (POMDP) (Russell and Norvig, 2002; tinyurl.com/lcpmldm).

3. **Quantify the Models:** Perform a controlled red-team attack against the protected IP and use another red- or black-hat team to conduct an information market for estimating the parameters of the POMDP.

4. **Use the Results:** The resulting estimates can be used to decide if the proposed protections are appropriate for the specific vulnerabilities in terms of various possible cost-benefit analyses.

Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity

Jeff Hughes and George Cybenko

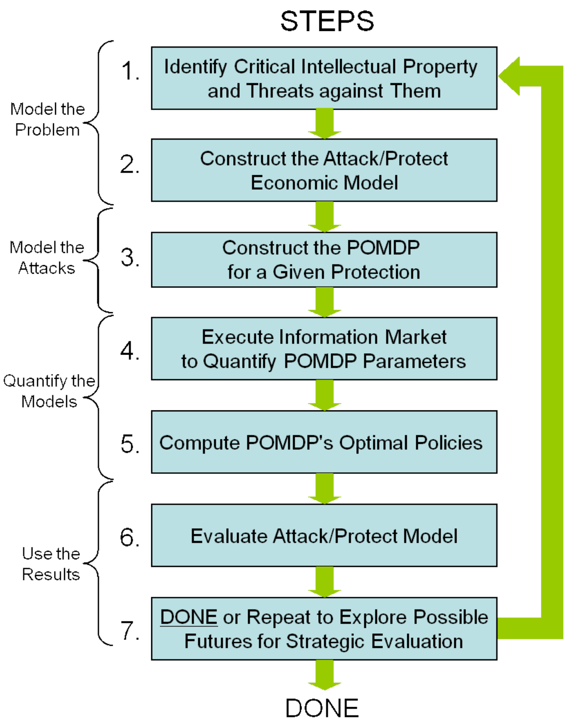

STEPS
Model the Problem
Model the Attacks
Quantify the Models
Use the Results
1. Identify Critical Intellectual Property and Threats against Them
2. Construct the Attack/Protect Economic Model
3. Construct the POMDP for a Given Protection
4. Execute Information Market to Quantify POMDP Parameters
5. Compute POMDP's Optimal Policies
6. Evaluate Attack/Protect Model
7. DONE or Repeat to Explore Possible Futures for Strategic Evaluation
DONE

Figure 3. The seven steps of the QuERIES methodology

To illustrate such a cost-benefit analysis, consider the plot shown in Figure 4, which compares two approaches for attackers to decide when to stop an attack. The red line is the difference between the cost of the attack up to the corresponding time and the benefit of a successful attack; the blue line is the “cost-to-go” value of continuing the attack given that it has failed up to that time. The cost-to-go value is computed using dynamic programming based on the probabilities shown in Figure 2 and is similar to the techniques used for American Options pricing.

Figure 4 illustrates that a binary metric (i.e., true or false) is not suitable for determining whether or not a cyber-physical system can be compromised. Any system requiring a password can be compromised by a lucky guess and so would be considered insecure if that were the metric. Instead, we argue that the right kind of metric is, for example, the expected cost of a successful attack. If that cost is high enough, an attacker would not undertake the attack in the first place. This is the basis for all state-of-the-art encryption schemes, so our position on this is entirely consistent with existing practice and experience in that realm.

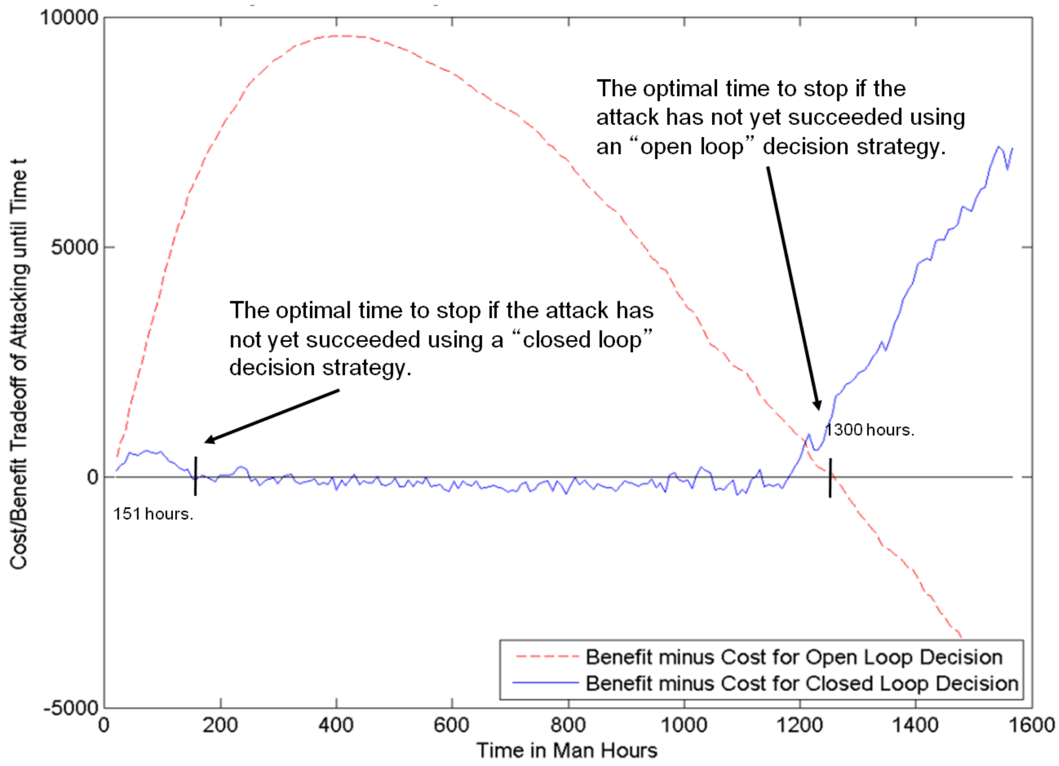


Figure 4. A comparison of two approaches to determine the optimal time for an attacker to stop an attack

Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity

Jeff Hughes and George Cybenko

The Three Tenets of Cybersecurity

Early in our work on threat and vulnerability analysis, we sought to identify simple – but still operationally meaningful, necessary, and sufficient – conditions for cyber-physical system vulnerabilities to exist. Once such conditions are identified, specific mitigations could be identified and evaluated.

This led us to identify three elements as being necessary and sufficient for successful attacks to occur:

1. The existence of inherent system susceptibilities

2. The threat's access to the susceptibility

3. The threat's capability to exploit the susceptibility

It is evident that, when these three elements are present in a system, an actual vulnerability exists.

Murphy's Law – "Anything that can go wrong, will go wrong." (Bell, 1989; tinyurl.com/llaps5q) – suggests that a system with vulnerabilities will be exploited given the appropriate operational environment. Moreover, a threat model that supports reasoning about whether an inherent system weakness rises to the level of a vulnerability is essential for cost-effective system-security engineering. This aspect is important because security for security's sake is neither affordable nor desirable, and so vulnerabilities must be quantified and only mitigated to the degree necessary to prosecute the enterprise's business processes or other missions.

We briefly describe these three elements below:

1. **System Susceptibility:** Absolute system confidentiality, integrity, and availability cannot be simultaneously achieved. Therefore, all systems will have design trade-offs resulting in inherent weaknesses. Such weaknesses will be manifest as software errors/bugs, protocol flaws, misconfigurations, or physical implementation constraints, and can be organized into the following eight categories of susceptibilities (National Vulnerabilities Database; nvd.nist.gov):

   a. *Input Validation Error (IVE):* includes failure to verify the incorrect input and read/write involving an invalid memory address. This category of susceptibility is also known as a boundary condition error (BCE) or buffer overflow (BOF).

   b. *Access Validation Error (AVE)*: causes failure in enforcing the correct privilege for a user.

   c. *Exceptional Condition Error Handling (ECEH):* arises due to failures in responding to unexpected data or conditions.

   d. *Environmental Error (EE):* triggered by specific conditions of the computational environment.

   e. *Configuration Error (CE):* results from improper system settings.

   f. *Race Condition Error (RC):* caused by the improper serialization of the sequences of processes.

   g. *Design Error (DE):* caused by improper design of the software structure.

   h. *Others:* includes susceptibilities that do not belong to the types listed above. This category of susceptibility is sometimes referred to as nonstandard.

2. **Threat Accessibility:** A threat will probe and analyze a system in order to discover which susceptibilities are accessible and how, with the goal of subsequent exploitation. Generally, the threat will use access points or services offered by a system to legitimate users as the original point of entry. Threat access is typically a superset of legitimate user access, because some access points may be undocumented or not of interest to legitimate users. Possible access points include wireless networks, legacy dialup lines, maintenance/service ports, automatic updates, and so on. Moreover, commercial and open source systems are accessible by the attacker for testing and exploit validation prior to launching a real attack. Any access offered an attacker provides a learning opportunity.

3. **Threat Capability:** After thorough surveillance (either via remote observations or in situ instrumentation) of the system design and its operation, an attacker will attempt to gain control, tamper with, or steal detailed system-design knowledge or other valuable data. Such attempts are often made using either a known or zero-day exploit determined after additional system reverse engineering. Skilled attackers typically employ a methodical approach to reverse engineering during which they expect to observe certain system behaviours. These system behaviours serve as exploitation guideposts and significantly aid

> the attacker. The degree to which the attacker is successful will depend on their level of system knowledge, their ability and resources to develop and use specialized tools given the system's functionality and operating environment, and their overall level of reverse-engineering experience.

This form of threat model has deep roots in the Electronic Warfare (EW; tinyurl.com/kzrbp) test and evaluation community. That community shares a similar adversarial framework (i.e., measure-countermeasure) with cyber-physical system security. A version of this threat model was suggested for EW vulnerability analysis in 1978, and is called data link vulnerability analysis (DVAL) (Guzie, 2000; tinyurl.com/n4ge8w7). DVAL has four components in its vulnerability definition: susceptibility, interceptibility, accessibility, and feasibility. However, in contrast to DVAL, The Three Tenets threat model assumes that "feasibility" and "interceptibility" are effectively merged into what we call "capability." In today's complex cyber-physical systems based on commercial-off-the-shelf technologies, attackers can rehearse for almost any given operating environment and develop an exploitation capability. Such rehearsals are even possible with specialized computer-controlled systems, as demonstrated, for example, by Stuxnet (tinyurl.com/3vol5nk).

Thus, The Three Tenets threat model posits that three ingredients are necessary and sufficient for cyber-physical vulnerabilities to exist: i) a system susceptibility, ii) threat accessibility, and iii) threat capability. The three threat-model elements are illustrated in Figure 5. This figure depicts the co-occurrence of those ingredients as the space of vulnerabilities and therefore successful attacks.

Additional evidence that this vulnerability model is suitable comes from so-called routine activity theory (RAT) (Cohen and Felson, 1979; tinyurl.com/pml7vcq) that is used in criminology. This theory posits that crimes occur when three elements coincide: i) there is a motivated offender, ii) there is a lack of guardianship, and iii) there is a suitable target. The elements of RAT and the threat-model elements listed above have a clear correspondence: capability = motivated offender, accessibility = lack of guardians, susceptibility = suitable target. Our threat model is also related to "means, motive, and opportunity" arguments for convincing a jury of a suspect's guilt. The point is that previous work in criminology is relevant and consistent with our approach to cyberthreat modeling.

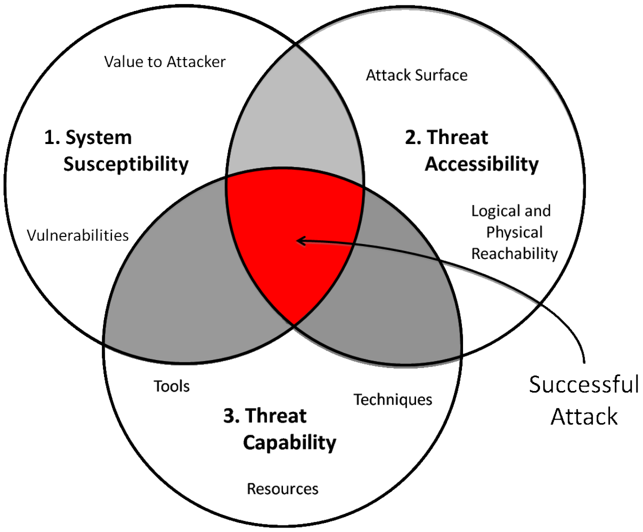


Figure 5. The three ingredients necessary and sufficient for cyber-physical vulnerabilities to exist

With these three necessary and sufficient conditions for cyber-physical vulnerabilities to exist, we can develop mitigation techniques and metrics for each condition. These mitigation techniques are called The Three Tenets and correspond to each condition outlined above. Collectively, The Three Tenets comprise a system security engineering approach consisting of both a secure design methodology and an assessment tool for security evaluation. The Three Tenets are introduced and described below:

1. **Focus on What is Critical:** The first Tenet instructs the designer to consciously and methodically focus on including only those system functions that are essential to the mission. This is an acknowledgement of Occam's razor (tinyurl.com/gxvu2) by the system designer. Adherence to this Tenet reduces the number of potential susceptibilities, and therefore, the paths between the attackers' starting state (i.e., the system access points) and goal states in which mission-essential functions, critical security controls, or critical data are compromised. This Tenet eliminates those access points and susceptibilities associated with unneeded functionality.

2. **Move Key Assets Out-of-Band:** The second Tenet instructs the designer to consciously differentiate between user access and attacker access for a given system's mission. This Tenet modifies system availability and is accomplished by moving the data/processes used by mission-essential functions, their

security controls, and associated access points out-of-band of the attacker either logically, physically, or both. By "out-of-band" we mean not accessible by the attacker through their preferred or available access methods. Adherence to this Tenet reduces threat access for a given mission (i.e., use case) and may enable unalterable observations of system state by a security control sensor. The extent and strength of access differentiation between the user and attacker is greatly influenced by the type of out-of-band mechanism employed and whether it is done in software or hardware.

3. **Detect, React, Adapt:** The third Tenet instructs the designer to employ dynamic sensing and response technologies (e.g., a security control sensor or reference monitor) that mitigate the threat's capabilities and exploitation attempts through automated (preferably autonomic) system behaviour. Adherence to this Tenet confounds the attacker's capabilities by making the system's defences unpredictable (i.e., nonstationary) and adaptive (i.e., with penalties) instead of merely being passive.

Just as each ingredient of the threat model has grounding in EW and classical criminology theory, each of The Three Tenets has been advocated and practiced in one form or another by computer security researchers and developers in the past. Further details and a more comprehensive treatment of The Three Tenets is available in a longer and more technical article (Hughes and Cybenko, 2013; tinyurl.com/l5wl5nt).

The Three Tenets provide a quantitative basis for the following security metrics, which are merely illustrative of more comprehensive and quantitative metrics that are possible:

1. **System Susceptibility Metric:** In its simplest instance, this system-construction metric instructs us to minimize the number of functionalities and services that act as access points to system-critical functions. This “reachability” metric is a direct consequence of the first Tenet: to identify, implement, and protect only what is mission critical.

2. **Access Point Metric:** Minimize the amount of input/output and system processes visible to an attacker. This metric is a direct consequence of the second Tenet: to move critical data “out-of-band.” Enumeration of “in-band” versus “out-of-band” access points is one way to measure application of the second Tenet.

3. **Threat Capability Metric:** Minimize useful insight into system operations in the sense that data observed at one time may or may not be similar or consistent with data observed at another time or on another system by the attacker. This “evidence variability” metric is a direct consequence of the third Tenet: to detect, react, and adapt. It is referred to by cybersecurity practitioners as a “moving target defence.”

These metrics can be readily measured by an enumeration of access points and data input/output or process observations together with determination of system functional behaviour.

Moreover, there are clear economic and effectiveness tradeoffs between, for example, implementing Tenet 3 (detect, react and adapt) and Tenet 1 (implementing only mission-critical functionality). These tradeoffs can be addressed through a QuERIES-type methodology and are the subject of ongoing work.

Conclusion

In this article, we have presented threat-driven, descriptive security methodologies that enable reasoning about cyber-physical system design in a strategic fashion. We feel that this approach is a clear alternative to traditional prescriptive approaches to cybersecurity that provide little insight into the comparative value of security solutions given the entirety of the system security trade-space. Underpinning our methodologies is the concept of “time-to-compromise.” We suggest that this is a fundamental metric associated with any adversarial environment and that cyber-physical system security is no different than physical security in this respect. Concrete metrics are described that are functionally related to and expand upon time-to-compromise. These metrics serve as informative and quantitative guides in secure system design. Future work will describe the mathematical underpinnings of The Three Tenets and provide a more complete derivation of the resultant quantitative security metrics. Additionally, the benefits of analyzing complex system security by employing probabilistic formulations such as QuERIES and the CIA analysis will be illustrated via reduction to practice for varying use cases. Finally, we intend to develop a more explicit coupling of these methodologies to a lifecycle security analysis for cyber-physical systems.

Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity

Jeff Hughes and George Cybenko

About the Authors

Jeff A. Hughes is President of Tenet 3, LLC. Tenet 3 is a cybertechnology company with a focus on autonomous cyber-physical systems, analyzing their trustworthiness, and evaluating economical ways to demonstrably mitigate security risks. Previously, Jeff held various positions in the US Air Force Research Laboratory (AFRL), where he led research into advanced techniques for developing and screening trustworthy microelectronic components and performing complex system vulnerability and risk analysis for cyber-physical systems. Jeff has an MS in Electrical Engineering from the Ohio State University and has completed graduate work towards a PhD at the Air Force Institute of Technology in Ohio, United States.

George Cybenko is the Dorothy and Walter Gramm Professor of Engineering at Dartmouth College in New Hampshire, United States. Professor Cybenko has made multiple research contributions in signal processing, neural computing, information security, and computational behavioural analysis. He was the Founding Editor-in-Chief of both *IEEE/AIP Computing in Science and Engineering* and *IEEE Security & Privacy*. He has served on the Defense Science Board (2008-2009), on the US Air Force Scientific Advisory Board (2012-2015), and on review and advisory panels for DARPA, IDA, and Lawrence Livermore National Laboratory. Professor Cybenko is a Fellow of the IEEE and received his BS (Toronto) and PhD (Princeton) degrees in Mathematics.

Citation: Hughes, J. and G. Cybenko. 2013. Quantitative Metrics and Risk Assessment: The Three Tenets Model of Cybersecurity. *Technology Innovation Management Review.* August 2013: 15–24.

Keywords: quantitative cybersecurity, risk assessment, vulnerabilities, confidentiality, integrity, availability

An Enterprise Security Program and Architecture to Support Business Drivers

Brian Ritchot

> *We will bankrupt ourselves in the vain search for absolute security.*
>
> Dwight D. Eisenhower (1890–1969)
> 34th President of the United States

This article presents a business-focused approach to developing and delivering enterprise security architecture that is focused on enabling business objectives while providing a sensible and balanced approach to risk management. A balanced approach to enterprise security architecture can create the important linkages between the goals and objectives of a business, and it provides appropriate measures to protect the most critical assets within an organization while accepting risk where appropriate. Through a discussion of information assurance, this article makes a case for leveraging enterprise security architectures to meet an organizations' need for information assurance. The approach is derived from the Sherwood Applied Business Security Architecture (SABSA) methodology, as put into practice by Seccuris Inc., an information assurance integrator. An understanding of Seccuris' approach will illustrate the importance of aligning security activities with high-level business objectives while creating increased awareness of the duality of risk. This business-driven approach to enterprise security architecture can help organizations change the perception of IT security, positioning it as a tool to enable and assure business success, rather than be perceived as an obstacle to be avoided.

Introduction

Many organizations find that their existing security controls are preventing them from getting something done or are reducing their effectiveness. Conversely, an organization may question if there is sufficient protection for information that is to be shared with a new business partner, customer, or the general public. If a critical system is compromised, what will be the *business* impact?

In order for a security program to be effective, it must demonstrate value to the business while avoiding the traditional pitfalls associated with the perception of security being an inconvenience and an obstacle to effective business operations. Security practitioners are challenged to consider security in the context of the business and understand the duality of risk: some risks represent business opportunities and should therefore be accepted. However, risk avoidance is a common practice within IT organizations, where security expenditures, policies, procedures, and technologies are not proportional to the risk appetite of the business. When security controls become overly intrusive to employees of a business, and in fact impede business operations, individuals will seek the means to bypass these controls. This desire to avoid security is due to the perception that security is an obstacle. As a result of this security avoidance, new risks are introduced, however are not known to the security team and cannot be monitored and managed.

The situation describes a common struggle faced by most organizations. Organizations strive to achieve the appropriate balance between security controls to protect business information, while also allowing their employees to be productive and share information easily. Achieving this balance requires information assurance.

An Enterprise Security Program and Architecture to Support Business Drivers

Brian Ritchot

This article will provide an initial understanding of information assurance and present the case for leveraging enterprise security architectures to meet an organization's need for information assurance. The approach is derived from the Sherwood Applied Business Security Architecture (SABSA; tinyurl.com/9hg3se2) methodology, as put into practice by Seccuris Inc. Seccuris (seccuris.com) is a Canadian information assurance integrator that helps organizations achieve their business goals through effective management of information risk. To help customers effectively manage risk and capitalize on business opportunities, Seccuris has come to rely on business-driven enterprise security architectures. Seccuris adopted the SABSA methodology for enterprise security architectures to provide organizations with the often-missing critical link in effective cyberthreat mitigation. This missing link is an appropriate understanding of business goals and a structured repeatable process with which to identify assets of critical value to the organization and deliver appropriate safeguards within an established risk appetite.

An understanding of Seccuris' approach will illustrate the importance of aligning security activities with high-level business objectives while creating increased awareness of the duality of risk. The business-driven approach to enterprise security architecture can help organizations change the perception of IT security, positioning it as a tool to enable and assure business success, rather than an obstacle to be avoided.

The article is intended for senior executives within an organization who are trying to rationalize an appropriate balance between the protection and availability of information that supports the business. The article will also help security practitioners, in particular security architects, understand how to align security initiatives with business goals to deliver an effective security program.

Information Assurance

Information assurance relates to the management of risk and security related to the use, processing, storage, and transmission of data. It is part of a broader category, known as information security, that is predominantly focused on IT security controls and processes. IT security deals primarily with the confidentiality, integrity, and availability of information and provides mechanisms to protect these aspects. When information is compromised, the result is a change in state of one of these aspects.

1. **Confidentiality:** ensures that privileged or sensitive information is accessible only to those individuals with a valid requirement to view and access the information. It is particularly important when concerning personal information, intellectual property, and classified or sensitive information in a government context.

2. **Integrity:** refers to a lack of corruption in data or overall consistency. When integrity of information is compromised, it creates a lack of trust wherein data may have been manipulated, changed, or deleted.

3. **Availability:** relates to having access to authorized information when it is required. Should information be affected so it cannot be accessed when needed and authorized, then availability has been compromised.

Information risk arises when the confidentiality, availability, or integrity of data can be compromised. To mitigate risk, controls can be developed and implemented to provide increased assurance of information. A control is a safeguard or countermeasure designed to avoid, minimize, or counteract risk.

The practice of information assurance relies on the identification of risk and the application of appropriate controls. However, over time, this practice has come to be categorized as "risk adverse" and to be seen as an impediment to business effectiveness. Information security professionals are labelled as obstacles to successful implementation and delivery of IT solutions. The resulting business culture is reluctant to involve and solicit input from IT security teams, because the input can create business risk for a project and delay implementation. This view comes from the practice of creating a risk-avoidance approach to information security, based solely on technical threats, identification of risk, and use of as many controls as possible to mitigate risk. The end result of this practice is a security program that fails in its effectiveness, given reluctance at an organizational level to involve security in the early stages of projects and planning. True information-assurance practices must also recognize the value and importance of making information available and establishing safe information-sharing practices.

In the first 6 months of 2010, McAfee (2010; tinyurl.com/n7rykk6) discovered over 10 million new pieces of malware. According to the US Intellectual Property Commission Report (IP Commission, 2013; tinyurl.com/pnyjnod), hundreds of billions of dollars are lost each year to the theft of intellectual property. Increasingly, this theft is the result of cyberattacks against United States' electronic infrastructure. Sophisticated samples of malware have been discovered in recent years, with demonstrated capability to attack SCADA control networks (tinyurl.com/jcrlz) and negatively impact critical infrastructure. The two most notable examples are "Stuxnet" and "Flame" (Klochender, 2013; tinyurl.com/l6vb6ja). Together, all of these examples illustrate a failing in existing information assurance practices and a rise in the sophistication and capabilities of cyber-adversaries. When faced with these challenges, many organizations may default to implementing more security controls to minimize vulnerabilities and deny cyber-adversaries access to systems. An approach focused solely on controls will ultimately prove unsuccessful given the resourcefulness and capabilities of malicious threat actors, as demonstrated in the examples above. Through a history of assisting customers across multiple sectors and varying levels of government, Seccuris has identified a common thread: while most organizations are security conscious and acknowledge the need for good IT security practices, they lack the necessary knowledge needed to build effective security programs.

Security programs should be designed to provide appropriate protection for information and information assets. This protection should be tailored to the environment, which requires the identification of what information is most vital to an organization. Without a clear understanding of priorities for information, and information security, organizations are incapable of prioritizing control improvements. A lack of a structured security architecture impacts all aspects of an organization's IT security program, including threat monitoring, vulnerability management, identity management, and incident response just to name a few. Effective security operations require effective security architecture.

Sherwood Applied Business Security Architecture (SABSA)

When security is found to be cumbersome or intrusive into business practices, loopholes and shortcuts are taken to bypass the implemented controls, creating increased risk that is not accounted for and is easily capitalized upon by threats seeking to compromise the confidentiality, integrity, and availability of data. The Sherwood Applied Business Security Architecture (SABSA) methodology for an enterprise security architecture and program can be leveraged to address this shortcoming (Sherwood, et al., 2009; tinyurl.com/mkggknj).

In essence, the SABSA approach is centered on making security a business enabler rather than an obstacle and avoidable inconvenience. The SABSA approach creates an understanding of an organization’s business objectives and provides a structured approach to designing a security program that supports these objectives. Security does not hinder business objectives, but instead provides assurances around operational risk that could negatively impact the business and, in fact, enables the organization to take on new strategic opportunities.

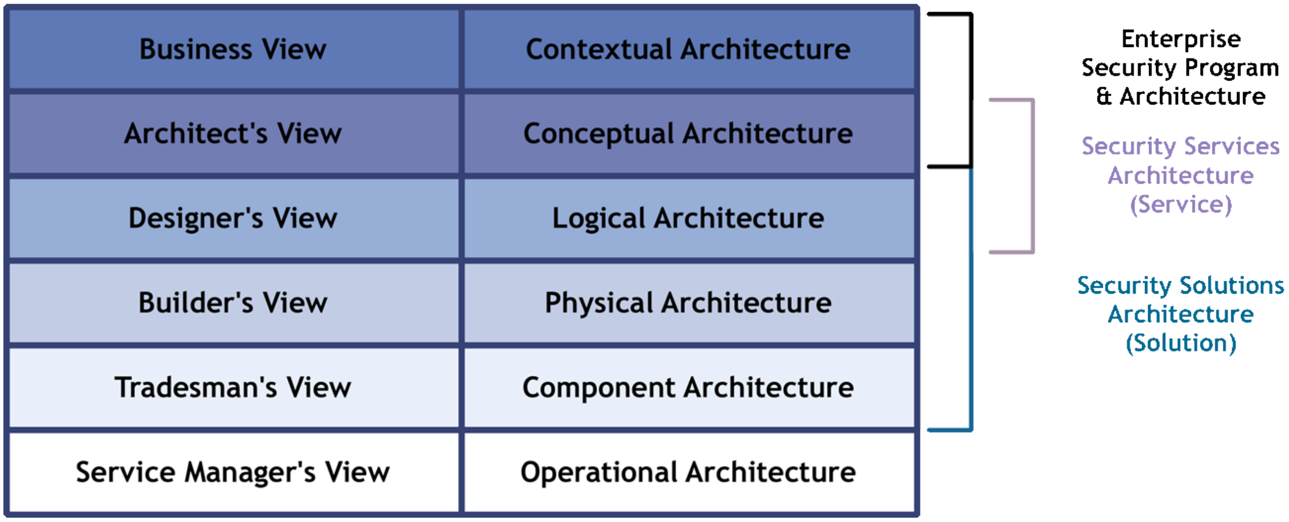
SABSA is a unique approach to information assurance because it seeks to align security programs with an organization's fundamental business objectives and drivers. In doing so, the SABSA approach treats risk as something that can not only hinder a business, but can also enable new opportunities. It is necessary for organizations to accept risk in order to do business and be effective. Embracing the right type of risk has the potential of leading to good fortune for a business (Card, 2013; timreview.ca/article/696) .

An important element to consider when selecting an architectural framework for security, is that many organizations already have an established IT architecture program to facilitate delivery of IT projects. Some existing architectural frameworks include the Open Group Architecture Framework (TOGAF; tinyurl.com/yj72xcf) and the Zachman Framework (tinyurl.com/4axvn2e); however, such frameworks do not traditionally address security requirements. Furthermore, many organizations implement service-management programs to manage and operate IT systems and services. The Information Technology Infrastructure Library (ITIL; tinyurl.com/mukhg) is an example of such a service-management framework. SABSA is unique among architectural frameworks in that it does not seek to replace or interfere with these existing frameworks and practices, but instead integrates with them and provides the necessary tools to align and support existing architectural programs. This ease of alignment with existing frameworks and the focus on leveraging security as a business enabler are key criteria that affected Seccuris’ decision to leverage SABSA as a framework for delivering enterprise security architectures.

To make security relevant to all stakeholders within an organization, the SABSA framework introduces a layered approach to architecture. Each layer corresponds to a different player's view within the organization as it relates to specifying, designing, constructing, and operating a security architecture, as shown in Figure 1.

An Enterprise Security Program and Architecture to Support Business Drivers

Brian Ritchot



Business View	Contextual Architecture
Architect's View	Conceptual Architecture
Designer's View	Logical Architecture
Builder's View	Physical Architecture
Tradesman's View	Component Architecture
Service Manager's View	Operational Architecture

Enterprise Security Program & Architecture
Security Services Architecture (Service)
Security Solutions Architecture (Solution)

Figure 1. The SABSA model for security architecture

For each of the architectural views in Figure 1, SABSA encourages security architecture to consider the following key questions:

1. What are you trying to protect at each layer? (Assets)
2. Why are you protecting these assets? (Motivation)
3. How will you achieve your objective? (Process)
4. Who is involved in applying security? (People)
5. Where are you applying security? (Location)
6. When are you applying security?(Time)

SABSA provides the theory and background knowledge essential to delivering business-driven security architectures; however. the practice of delivering the architecture is left up to the security practitioner. As described in the next section, Seccuris has developed a repeatable process and procedures that answers the above questions for each layer of the architectural model.

Enterprise Security Architecture: Establishing the Business Context

A business-driven approach to enterprise security architecture means that security is about enabling the objective of an organization by controlling operational risk. This business-driven approach becomes a key differentiator to existing security practices that are focused solely on identifying threats to an enterprise and technical vulnerabilities in IT infrastructure, and subsequently implementing controls to mitigate the risks introduced. A purely threat-based approach to risk management fails to enable effective security and business operations. The term *security* will carry very different meanings to different organizations. For example, consider security as it relates to a military organization and security related to an online retailer that processes credit card information. The business models for these two organizations will be very different and, as a result, the security programs should be unique and relevant to their underlying businesses. A military organization may determine that the most critical asset to protect is the life of its soldiers as they are engaged in military operations. To provide assurance as to the safety of a soldier, complex security architectures are needed to protect information and information systems that could impact the soldiers' safety. Solutions could range from ensuring that logistic systems that manage the delivery of supplies, food, and ammunitions remain available and that data integrity is protected to protecting confidentiality of mission plans and military intelligence that, if compromised, could cause considerable harm to war fighters. Conversely, an online retailer is likely most concerned with compliance with standards set by the payment card industry. These standards are tailored to protect the confidentiality of personal information and the integrity of transactions. An online retailer may have lower thresholds for availability then a military logistics system. The needs for confidentiality, availability, and integrity of data must be balanced and appropriate to the business activity.

An Enterprise Security Program and Architecture to Support Business Drivers

Brian Ritchot

Developing a security architecture begins with an understanding of the business, which is achieved by defining business drivers and attributes. A business driver is related to the organization's strategies, operational plans, and key elements considered critical to success. A business attribute is a key property of the strategic objectives that needs to be enabled or protected by the enterprise security program. An organization's senior executives, who set the long-term strategy and direction of the business, can typically provide knowledge regarding business drivers. The drivers are often reflected in an organization's mission and vision statement. Consider our military organization, which may have a strategic objective of "operational excellence". This business driver can be distilled into relevant attributes that require assurance to satisfy the overarching business driver. Conversely, the online retailer may have a strategic objective of being "customer focused", as expressed in their vision statement to provide a superior online shopping experience.

Business attributes can generally be identified through an understanding of the business drivers that are set by the top levels of an organization. Security architects will often conduct structured interviews with senior management in order to identify business attributes by determining the essence of what is conveyed by high-level business drivers. In the example of the business driver labelled "operational excellence", the executives might be referring to the availability, reliability, and safety of their operations and resources. In this case, the business attributes defined are "available", "safe", and "reliable". Each attribute is then linked to the business driver they support. This pairing of a business driver and attribute results in the creation of a proxy asset. Again, building on our example, a sample proxy asset is "operational excellence" with the attribute of "available". Each proxy asset is owned by the organization and is assessed as having value to them. The fact that the proxy asset has value sets the requirement that it should be protected. The value of these proxy assets is difficult to define given that they are often intangible and exist at a very high level. Despite being unable to assign a monetary value to a proxy asset, it is still possible to identify risks that may act against the asset. Our online retailer may have attributes of "confidential", "reputable", and "error-free".

An inventory of proxy assets can be maintained by the security architect and will be considered as key assets to the organization. This is later used to conduct a business threat and risk assessment to identify risks to the business. It is through a business threat and risk assessment that the sometimes-competing aspects of confidentiality, integrity, and availability can be reconciled. When the overall objective and needs of a business are understood, through proxy assets, then impact can be understood as it relates to confidentiality, integrity, and availability. Understanding of the business helps prioritize which of these elements is most important, and which aspects of the business are most in need of protection.

Identification of Risk

Traditional threat-based risks are those that can be mitigated via a control because they would result in the loss of value. Many organizations rely on threat risk assessment to create an inventory of threats that may have a negative impact on the business. These threats are then mapped to vulnerabilities that, when exploited, result in a compromise to the organization's business. Managing this risk often relies on the deployment of security controls that offer a form of mitigation. Consider, for example, an external website that has a technical vulnerability that could result in a threat that exploits this vulnerability, such as a denial-of-service attack, which would render the website inaccessible. Threat-based risk analysis would identify the threat and provide recommendations on control improvements to decrease the risk. This may include the deployment of a web-application firewall to provide intrusion-prevention capabilities, increased network monitoring, the deployment of additional firewalls, and software upgrades to reduce the vulnerability.

The traditional approach described above successfully identifies risk based on an analysis of possible threats and provides the means of mitigating that risk. The relation of the risk to business impact, however, is missing in this approach. Perhaps the server is used for testing with a small number of business partners, and denial of service is not a risk of importance. The implementation of controls sometimes does not offer protection for what is truly important to the business. Controls can increase complexity within the network and incur costs that could have been avoided. When identifying risk using only threat-based approaches, key information is likely missing, which could have informed a security architect on where to prioritize control improvements. Another problem with threat-based approaches is that they do not consider the potential opportunity that can be realized when embracing risk.

Consider the non-traditional idea that risks can also be categorized as opportunity-based – this perspective is lacking from traditional information assurance practices. An opportunity-based risk can increase the value of an asset. This approach enables us to understand the duality of risk. Some risks should be mitigated, while others might be accepted as something that cannot necessarily be avoided; businesses always operate with some level of inherent risk.

When a business and security architects understand the duality of risk, they can also focus on risk acceptance rather than just risk avoidance. To this end, the organization develops key performance indicators (KPIs) and key risk indicators (KRIs). KPIs are measures of the value and performance of business attributes in the context of the business driver. KRIs are measures of risk, and they establish risk thresholds to provide early warning when a risk will exceed an organization's risk tolerance.

In our military example, we identified "operational excellence" as a business driver and "available" as an attribute. Developing a KPI around the availability of operational systems would allow the organization to measure the availability and uptime of a key business application. The performance of this application could be tracked over time to ensure that it remained available and supported the business driver of "operational excellence". Conversely, the same example could be considered from a different angle and result in the development of a KRI. A KRI differs from a KPI in that it establishes a threshold or condition that creates a warning state. Monitoring of KRIs provides an indication when risk is about to exceed established tolerance levels. In our example, a KRI might enable ongoing measurement of the time that an application is inaccessible. The system may be inaccessible due to technical circumstances such as network failure, software updates, or other IT incidents. Measuring the time that an application is inaccessible will facilitate identification of an established threshold that sets unacceptable behaviour. When the key business application approaches this threshold, alerts can be generated to warn that the established risk tolerance is soon to be exceeded.

In our online retailer example, the attribute "error free" can apply to the processing of financial transactions. In order to provide a consistent experience for users and to maintain customer confidence, the online retailer wants to ensure that any transactions are without error. A KRI can be created to capture any time a financial transaction is disputed due to a potential error. It is likely that a very low threshold would be established and any errors would trigger appropriate response to investigate and remediate the cause of the error.

Regardless of the indicator selected (i.e., KPIs or KRIs), when considered in the context of business drivers and attributes, it becomes possible for security to have sufficient information to consider risk within the context of business objectives. This understanding of risk will contribute to an overarching model of business risk, which is needed for a security architect to successfully develop security architecture at an enterprise level.

Creating a Model of Business Risk

A model of business risk provides a mechanism for quantifying risk and ensuring that it remains relevant to business drivers and attributes (as described above). The business-risk model builds on the understanding of risk, centered on the established proxy assets as well as KRIs and KPIs. In addition to the proxy assets and KPIs and KRIs, there are other models that also must be developed and understood to complete the business-risk model: i) trust models and business relationships, ii) threats operating against the business, and iii) safeguards that have been implemented. This approach is similar to a threat risk assessment; however, the difference is that a threat risk assessment measures the risk to a system or IT environment. An enterprise security architecture measures the risk to the proxy assets that represent the organization's business.

Trust models

Trust needs to be considered in the context of the overall business as a business attribute, not a technical one. Whenever two or more entities are required to interact and exchange information, trust must first be established between the two entities. Trust can be established through registration of the entity by the other. The registering entity will then trust the entity that has been registered, based on assurance mechanisms. The levels of assurance required to establish this trust vary, based on the degree of risk involved.

As an example, consider a shopkeeper selling lottery tickets to a customer. For the shopkeeper to trust the customer, they require a valid form of identification, such as a driver's licence, to validate the customer's age. The decision to require the driver's licence is based on a risk decision taken by the shopkeeper on whether the individual appears of age or not. Conversely, the customer must trust the shopkeeper as a valid merchant authorized to sell lottery tickets. This trust is established through the clear display of the lottery licence issued and validated by the government. Once these assurances are validated, the shopkeeper and customer establish a two-way trust relationship to complete the transaction.

Understanding business risk requires that relationships internal and external to a business are properly understood. Evaluation is required to gather information related to the criticality of a relationship, the sensitivity of information shared, and existing methods of assuring trust between the parties involved in a relationship. A security architect can explore the relationships both internal and external to an organization to identify the points where information is exchanged. This information can be used by a security architect to layer logical relationships to the physical IT environment to prioritize the placement of controls to assure trust and protect information as it is exchanged as part of business relationships.

Threat models

Threats are entities or things, operating against a business, that cause damage or harm to an organization. Developing an understanding of threats is an important step in developing awareness of business risk. Typically, safeguards are implemented to mitigate damage caused by a threat that is exploiting a vulnerability. A threat may be a deliberate action taken by an entity or it may be accidental, based on an unintentional realization of a scenario that creates risk. Natural hazards, such as fire and flood, are threats that may impact a business. Understanding these threats, the likelihood of them operating against a business, and the gravity of consequences should they succeed, helps in creating a prioritization of threats based on the impact and likelihood of being realized.

In the case of our military organization a threat to availability may materialize from a hostile government agency or military that would seek to disrupt the availability of key systems essential to effective military operations. The online retailer would not necessarily be concerned with military threats, and would instead consider organized crime as a threat that would seek to abuse the technical application in the retailer environment for the purposes of financial gain. This threat may include gaining access to confidential customer data for identity theft, or exploiting application vulnerabilities to order receive goods without paying full price. The threat model can be made relevant to the organization by considering threats that affect that proxy assets defined earlier in the process.

Safeguards

The final model required to establish the business risk model is an understanding of existing safeguards and any gaps that exist in mitigating risk associated with threat activity. Industries follow various standards of best practice and frameworks to assess their relative maturity regarding security controls and safeguards implemented to mitigate risk.

A popular general purpose control framework is ISO 27002 (tinyurl.com/lcz75nz), which is provided by the International Organization for Standardization (ISO). This standard provides industry best practices for information security management. Other control frameworks specific to industries include the North American Electrical Reliability Corporation (NERC) Critical Infrastructure Protection (CIP; tinyurl.com/l22ggon) standard and the International Society of Automation (ISA; isa.org) standards. Selecting the appropriate framework is an important step in conducting a review and identifying relevant gaps. Each business will have a unique set of controls relevant to their industry and the regulated protection required. Once a control framework is selected, a review of the controls can be undertaken along with assessment of the maturity level for each control and the extent to which it is implemented in the business.

Once all data around risks, relationships and trust, threats, and control effectiveness have been gathered, the inherent risk to a business can be described and represented. An initial assessment is undertaken where the business impact of a risk being realized is considered, irrespective of any existing controls, to quantify the severity of risk. Based on this understanding of the inherent risk, a security architect can identify areas for improvement that will manage risk and establish acceptable thresholds. Inherent risk is a measure of the risk to the enterprise prior to any controls being implemented.

Risk can be classified on a scale tailored to meet an organization’s need; however, fundamentally, risk will be quantified as negligible, acceptable, significant, or severe. In the case of negligible risk, no action is required, whereas acceptable risk requires monitoring to ensure it remains at an acceptable level. Significant and severe risks require action to establish an appropriate risk threshold in which the business is comfortable operating, to maximize opportunities. The resulting risk score is known as residual risk, which is the risk that remains after security controls and improvements are selected, approved, and implemented in an environment.

Improvements to controls can be simulated and new risk scores can be calculated to develop risk-reduction strategies. This process shows how security improvements can affect enterprise-wide risk. Because various improvements carry different complexities and cost, multiple models can be constructed to show options and potential benefits over time as control improvements are made. This approach allows an understanding of risk appetite and provides the overall business-risk model.

Using a Business-Risk Model to Drive Security Architecture

Once a business risk model has been completed, a security architect can leverage this information to create logical security services. A logical service is specified independent of any physical mechanism that might be used to deliver the service. Most importantly, a logical service is driven from the business attributes and the business-risk model. A security service is therefore a combination of security controls that work together to support the delivery of a valued business service. A number of security services can be developed to provide "defence in depth" and increase the overall assurance of information that is important to the business. Sample security services that can be created may include threat management, vulnerability management, and network access. These services will form part of a security-services catalogue and, as a result, services can be selected and implemented based on the needs of a specific business initiative. Unlike traditional security controls, these services are derived from business drivers and attributes, and they provide traceability to the business objectives.

The services to be designed will draw on the information gathered in the business-risk model, particularly the understanding of relationships expressed as physical boundaries where information is shared, threats that are operating against the environment, and opportunities for control enhancements to mitigate the risks introduced by threats. The added benefit of developing security services based on the business-risk model is that full traceability can be maintained, thereby demonstrating that security initiatives are tied to supporting an organization's business and effectively managing both threat- and opportunity-based risk.

Relevance to Cybersecurity

Although the concept of enterprise security architecture does not provide a concrete technical tool with which to counter advanced persistent threats or zero-day attacks, it provides a critical tool for identifying and assessing assets of value to an organization. This step is often missing when developing security programs, which can lead to the deployment of controls without a proper understanding of how they can hinder or support the overarching business objectives.

Consider our two examples: the military organization and the online retailer. Both are likely to have an Internet-accessible website. The purpose of these websites would be very different in both cases. For the military, it likely provides general information at the unclassified level to the general public; it would not likely be used for operational missions and would be in a separate network segment from critical operational systems. The online retailer, however, would use the website as a front-end to their e-business, where the website provides customers with the ability to browse and purchase merchandise. If both organizations have an enterprise security architecture, they are better equipped to respond to, and deal with, threat activity when it materializes against the website.

In the event of a zero-day attack against their webserver, the military would realize that the attack did not affect any critical information and was not related to a key relationship that was foundational to success of the organization. During incident response, the effort and tools used to respond would be appropriate, and control improvement would be balanced based on cost, impact, and risk tolerance. Although the reputation of the organization may be damaged should the information become public, the overall business impact would be minimal. The online retailer, however, would need a much different response based on their business needs. Furthermore, the logical security services that we alluded to would be geared towards protecting confidentiality and ensuring error-free processing. Although the discussion of logical-service design falls outside of the scope of this article, consider briefly the concept of vulnerability management. The online retailer would likely perform regular web-application assessments and vulnerability testing against its website and web application to appropriately protect customer data. The military organization, also implementing a vulnerability-management program, would most likely scan the website on a much less frequent basis and focus efforts on securing mission-critical systems, requiring high availability.

An enterprise security architecture helps organizations identify assets of critical importance to their organization. Attempting to counter emerging cyberthreats without a clear understanding of the business needs of an organization will result in ineffective security controls and practices. When security is considered in the context of the enterprise, as both an enabler and means of assuring business success, control improvements can be tailored to the environment to address more sophisticated and complex threat scenarios. In the cases of 10 million new malwares every six months, intellectual property theft in the billions of dollars, and sophisticated intrusions such as Stuxnet, organizations are often incapable of prioritizing security initiatives, and they default to technical solutions without proper identification of critical assets and information. The cyber-challenge requires that organizations be better equipped than the threats acting against them, and an enterprise security architecture provides this capability.

Conclusion

For information assurance to effectively strike the appropriate balance between the protection of information and making the correct information easily assessable to authorized parties, an enterprise security architectures needs to be developed with a focus on business goals. An enterprise security architecture should provide a means of mitigating risk while also supporting the business to pursue new opportunities.

The approach described in this article provides security practitioners and senior executives with the knowledge and foundational information needed to connect security performance to business performance. As the IT threats facing an organization continue to grow in volume and variety, a reasonable approach is needed to address threats while continuing to support the operations of the business. Organizations that focus solely on the eradication of threats and vulnerabilities when developing security architecture risk creating an environment where security becomes an obstacle to operations. A business-driven approach to security architecture helps organizations prioritize where controls are needed to protect critical information, and it helps them define what level of risk is acceptable.

About the Author

Brian Ritchot is a Senior Information Security Consultant with Seccuris Inc, specializing in the implementation and delivery of intrusion-detection solutions, vulnerability assessment, network analysis, and security architecture. With 11 years of prior experience in the federal government, Brian has developed skills and expertise to support the detection, discovery, and mitigation of cyberthreat activity. Brian has led and managed several high-profile teams and projects to deliver operational security solutions that monitor and protect systems of importance to the Government of Canada. Brian now focuses his time in the private sector, helping a variety of customers across the critical infrastructure sector with their IT security needs. These activities span enterprise security architecture development, incident response and handling, vulnerability assessments, forensic investigations, and specialized IT security expertise to mitigate sophisticated cyberintrusions.

Citation: Ritchot, B. 2013. An Enterprise Security Program and Architecture to Support Business Drivers. *Technology Innovation Management Review.* August 2013: 25–33.

Keywords: security architecture, risk, cybersecurity, information assurance, cyberthreats, information risk, information security

Protecting Critical Infrastructure by Identifying Pathways of Exposure to Risk

Philip O'Neill

> “*If you can look into the seeds of time, and say which grains will grow and which will not, speak then to me who neither beg nor fear your favour nor your hate."*”
>
> William Shakespeare (~1564–1616)
> Poet and playwright

Increasingly, our critical infrastructure is managed and controlled by computers and the information networks that connect them. Cyber-terrorists and other malicious actors understand the economic and social impact that a successful attack on these systems could have. While it is imperative that we defend against such attacks, it is equally imperative that we realize how best to react to them. This article presents the strongest-path method of analyzing all potential pathways of exposure to risk – no matter how indirect or circuitous they may be – in a network model of infrastructure and operations. The method makes direct use of expert knowledge about entities and dependency relationships without the need for any simulation or other models. By using path analysis in a directed graph model of critical infrastructure, planners can model and assess the effects of a potential attack and develop resilient responses.

Introduction

The complex connectedness of infrastructure, processes, commodities, and services in our society gives rise to risk. Failure of any particular system or service can cause far-reaching harm propagated through networks of other systems. A striking example is the electrical power failure that crippled much of Ontario and the northeastern United States in August 2003 (tinyurl.com/6qa5ode). The triggering event was a software bug in a control room system in Ohio, and it allowed a disastrous power surge to cascade through the power distribution grid. In addition to the usual direct effects of a power failure, unanticipated *indirect* effects were also experienced in the telecommunications, food, and transportation sectors. For example, in the Detroit area, residents lost water pressure because of failed pumps in the water supply system. However, the lack of pressure in the system resulted in potential contamination of the drinking water, which resulted in a “boil water advisory” after the pressure was restored.

In order to safeguard society, we need to connect with our connectedness. Models are needed to prepare for all recognized risks so that actions can be taken to make our communities, businesses, governments and environments as resilient as possible.

Typically, risk analysis of systems with complex dependency relationships is carried out by means of simulation. (A constructive simulation is a computer program in which software components mimic the behavior of infrastructure entities. Systems dynamics (tinyurl.com/yrqbyx) provides a framework and tools for building a constructive simulation. However, such models are governed by mathematical equations that are difficult to calibrate against the real world. Extensive data gathering and complex computer programming are required to create a useful model. Consequently, simulation models are time-consuming and costly to develop.

This article presents a technique, called the strongest-path method, which has been evolving since preparations began to solve the Year 2000 problem, or "Y2K" (tinyurl.com/2z675x). The paradigm, together with calculation tools and graphical output features, has been implemented under the trade name RiskOutLook (riskoutlook.com) as java-based software available exclusively from Deep Logic Solutions Inc (deeplogicsolutionsinc.com).

The strongest-path method fathoms all potential pathways of exposure to risk – no matter how indirect or circuitous they may be – in a network model of infrastructure and operations. The method makes direct use of expert knowledge about entities and dependency relationships without the need for any simulation or any other models. It can, however, incorporate results from simulation and other models if any are available.

This article will present the strongest-path method as a modelling paradigm, founded on risk analysis, that makes use of path analysis in a network representation of the entities and relationships in the environment of interest. The article starts with a description of the fundamental ideas in the paradigm. Next, it provides background information on risk analysis and path analysis in representations of networks in order to develop the tools for risk analysis used in the strongest-path method. Next, an example problem is used to demonstrate a practical use of the method and tools. Finally, the implications of this approach for planners and managers are discussed and conclusions are provided.

The Modelling Paradigm

Modern society can be viewed as a collection of networks that overlap and interact with each other. There are transportation networks, communications networks, energy networks, supply chains, distribution networks, social networks, cyber networks, and so on. In both private enterprise and public service, from the national level down to the local community level, planners typically divide their planning domain into coherent subsets called sectors. For example, at the provincial emergency planning level in Ontario, Canada, the following sectors are defined:

- Food
- Water
- Electricity
- Communications
- Healthcare
- Finance
- Natural Gas
- Oil
- Transportation
- Government
- Public Safety and Security

Each sector takes *inputs* from other sectors and, by means of its own *actors* and *activities*, produces *outputs* that are in turn taken as *inputs* by other sectors. As well, there are internal and external controls and regulations that govern the activities of any sector along with monitoring and verification agents who oversee the activities and processes.

Planners describe their domains in terms of the actors, actions, controls, agents, inputs, and outputs that exist in their sector. For purposes of risk analysis, distinct and significant actors that exist in a sector will be referred to as *entities*. An interaction between two entities will be referred to as a *relationship*. By using a mathematical object known as a *directed graph*, RiskOutLook creates a network model of entities and relationships that embodies the planner's domain. Entities will be modelled as *nodes* in the graph and relationships will be modelled as links in the graph, which are called *edges*. Because all of the edges have a *direction* from one entity to another, the graph is a *directed graph*.

A *dependency relationship* is a special kind of relationship in which there is a transaction between two entities. The *transaction* can be physical (e.g., electricity or water) or non-physical (e.g., data or instructions). The risks in our society that result from connectedness can be characterized as stemming from dependency relationships. The strength of a relationship is measured by a weight called the *degree of dependence*.

Direct dependency relationships are well understood by domain experts; indirect dependencies are more complex. Modelling and analysis is needed to verify or correct intuition and to synthesize expert knowledge into a comprehensive assessment of the effects of direct and indirect dependencies and their associated risks.

Protecting Critical Infrastructure by Identifying Pathways of Exposure to Risk

Philip O'Neill

Risk and Dependency Analysis

For purposes of infrastructure analysis, risk is the possibility of loss (Rowe, 1988; tinyurl.com/kmo7gd2). The strongest-path method describes loss using two dimensions: i) degree of impact and ii) likelihood of occurrence. The method then proceeds to aggregate assessments of cumulative impact resulting from multiple pathways of exposure to loss through pathways in the network model.

Estimates of likelihood of occurrence can be made in terms of "degree of belief" or "expert judgment". For the purposes of modelling infrastructure, experience has shown that a scale of high, medium, and low is sufficient. High-likelihood events are deemed to have more than an 80% likelihood of occurrence and low-likelihood events are deemed to have less than 20% likelihood of occurrence during the time interval under consideration; all other events therefore have a medium likelihood of occurrence. More complex models can be created for situations that evolve over multiple time intervals.

An infrastructure is said to "fail" if it falls below a threshold level of its expected or required outputs. From this definition, we develop the following criteria for *degree of direct dependence*:

1. If failure of entity x inevitably leads to failure of entity y, then y has a *high direct dependency* on x, and conversely x has a *high direct impact* on y.

2. If failure of entity x leads to degradation of entity y to the extent that y must enact a contingency plan or resort to alternate operating procedures in order to stay above the expected threshold, then y has a *medium direct dependency* on x and conversely x has a *medium direct impact* on y.

3. If failure of entity x leads to significant degradation of entity y, but y can stay above its expected threshold without significantly changing its operating procedures, then y has a *low direct dependency* on x and conversely x has a *low direct impact* on y.

4. If failure of entity x does not lead to significant degradation of entity y, then y has *zero direct dependency* on x, and conversely x has *zero direct impact* on y.

A *directed path* in a *directed graph* is a sequence of nodes with the property that each node in the sequence is connected to its successor by an edge. For example, in Figure 1, $\{s \rightarrow w \rightarrow x \rightarrow y\}$ is a *directed path*, whereas $\{s \rightarrow w \rightarrow x \rightarrow v\}$ is not, because $\{x \rightarrow v\}$ does not exist.

In order to derive a method for estimating the impact of every node in a graph on all nodes in the graph, path analysis is used. In particular, the analysis will be used to identify the *paths of strongest impact*, from any node x to any node y (including x itself). The paths of strongest impact, will be referred to as *strongest paths*. To visually indicate the strength of impact in a directed graph (e.g., Figure 1), the edges are coded as follows: red/solid for high impact, orange/dashed for medium impact, and yellow/dotted for low impact.

We have described the effect of a *high* direct impact event on a direct dependent. However, we also need to estimate the effect of a *medium* direct impact event and a *low* direct impact event on a direct dependent. There are two dimensions for this estimate: i) the degree of the triggering impact event and ii) the degree of the direct dependency relationship.

It is reasonable to expect that a strong triggering event will have little impact if the degree of dependence is low, whereas even a relatively weak triggering event will be felt if the degree of direct dependence is high. Thus, we estimate that the propagated impact can be no higher than the lesser of the triggering degree of impact and the degree of dependence. For example, according to

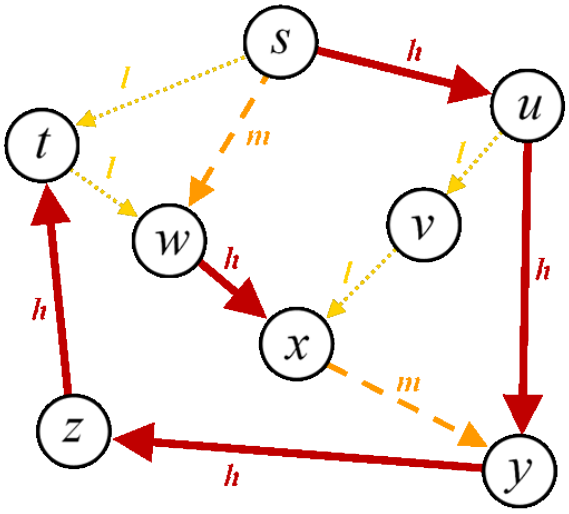


Figure 1. A directed graph with weighted edges (red/solid = high impact, orange/dashed = medium impact, yellow/dotted = low impact)

this principle, a medium-degree triggering impact acting over a low-degree of direct dependence will cause a low impact because the degree of direct dependence is low, whereas, a medium-impact trigger acting over a high degree of direct dependence can cause a medium impact because of the high degree of direct dependence.

As we proceed along a path in Figure 1, our propagation rule compares the lowest-degree edge we have yet encountered with the degree of the next edge on the path and sets the triggering degree of impact to the lower value. Therefore, the indirect dependence of any node on any other node along a selected path is driven by the lowest-degree edge along that path. For example, consider the graph in Figure 1 and all paths connecting *s* to *t*. These four paths are shown in Figure 2 and can be described as follows:

- Path 1 = $\{s \rightarrow t\}$
- Path 2 = $\{s \rightarrow w \rightarrow x \rightarrow y \rightarrow z \rightarrow t\}$
- Path 3 = $\{s \rightarrow u \rightarrow v \rightarrow x \rightarrow y \rightarrow z \rightarrow t\}$
- Path 4 = $\{s \rightarrow u \rightarrow y \rightarrow z \rightarrow t\}$

Using the propagation rule, we find that the impact of *s* on *t* from Path 1 is low by virtue of *(s, t)*, the impact of *s* on *t* from path 2 is medium by virtue of *(s, w)* and *(x, y)*, the impact of *s* on *t* from Path 3 is low by virtue of *(u, v)* and *(v, x)*, and the impact of *s* on *t* is high from Path 4 by virtue of all of its edges being high degree. Therefore, the indirect dependence of *t* on *s* is high and the strongest path is Path 4.

Proceeding from the direct impact and likelihood of failure of each node coupled with the ability to measure the strongest path from one node to any other, we can calculate other useful metrics. For any nodes *x* and *z* in any directed graph we can calculate:

1. **Strongest-path impact of *x* on *z*:** This is the strongest-path degree of dependence of *z* on *x* multiplied by the direct impact of *x*.

2. **Cumulative impact of *x* on *z*:** This includes a term for every pathway that exists from *x* to *z*. Similar to the binomial probability function, this metric compounds the effects of all of the terms.

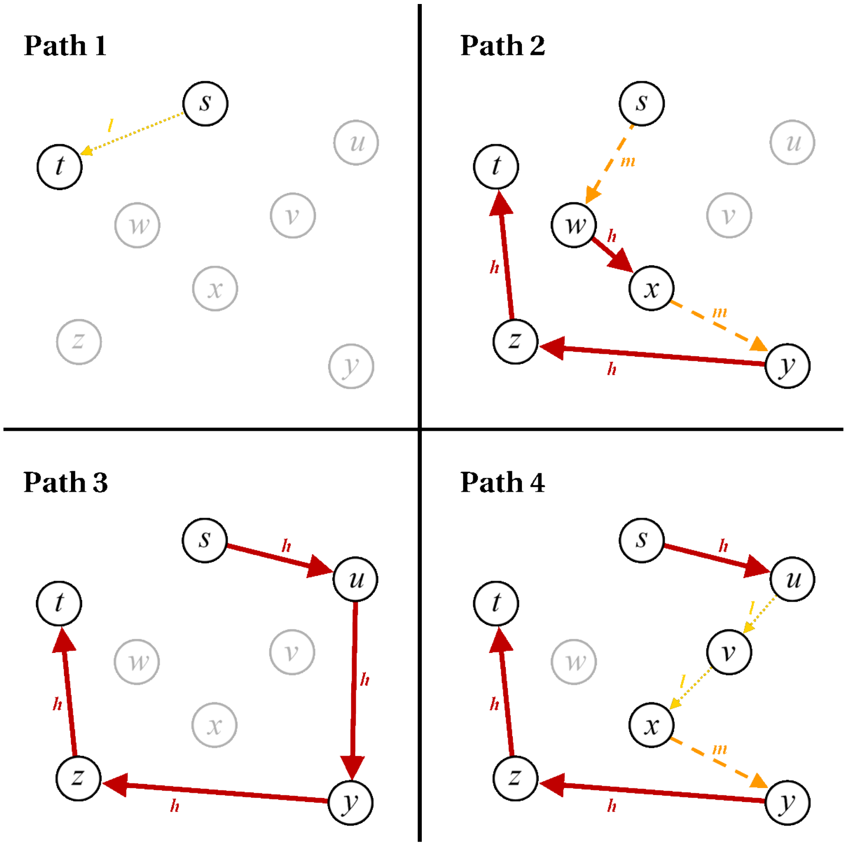


Figure 2. All paths connecting *s* to *t* in the graph shown in Figure 1

3. **Global impact of *x* on the entire graph:** By "global", we mean the impact of any node *x* on the entire graph. This metric is calculated using the impact *x* on every node *z* in the graph and summing the terms.

4. **Global vulnerability of *x* from the entire graph:** The cumulative vulnerability of *x* from all nodes in the model is the binomial probability that a failure event of any node will cause *x* to fail.

5. **Risk index of *x*:** The risk index of an entity is the product of the global impact of an entity times its global vulnerability. This metric provides a single score for comparing risk among all of the entities in a model.

Finally, we describe how to find the strongest-path degree of dependence between all pairs of nodes. By adapting any "shortest path" algorithm we can find a strongest path from any node *x* to any node *z* as follows:

1. Within the graph, remove all edges except for those of high degree.

2. If the shortest path from *x* to *z* exists, then it is a strongest path and *z* has *high* dependence on *x*.

3. Otherwise, put the medium-degree edges back into the graph.

4. If the shortest path from *x* to *z* exists, then it is a strongest path and *z* has *medium* dependence on *x*.

5. Otherwise, put the low-degree edges back into the graph.

6. If the shortest path from *x* to *z* exists, then it is a strongest path and *z* has *low* dependence on *x*.

7. Otherwise, *z* has *zero* dependence on *x*.

A Practical Application of the Method

In this section, we use an example network model to illustrate the practical use of the strongest-path method. The example model, shown in Figure 3, is a small infrastructure model with 10 entities: Drinking Water, Local Electrical Distribution, Natural Gas Storage and Transport, Ambulance Services, Local Food Outlets, Local Food Distribution, Farm Food Production, Health Canada/Food Inspection, Hospitals & Clinics, and Cyber Networks.

For each of these entities, the degree of impact has been assessed, as indicated by the number on the left side of each node in Figure 3: high (score = 7, dark orange), medium (score = 5, orange), or low (score = 3, yellow). As well, the likelihood of failure has been assessed for each entity, as indicated by the number on the right side of each node: high (score = 7, dark orange border), medium (score = 5, orange border), or low (score = 3, yellow border).

There are 30 direct-dependency relationships that have been scored high (score = 9, red edges), medium (score

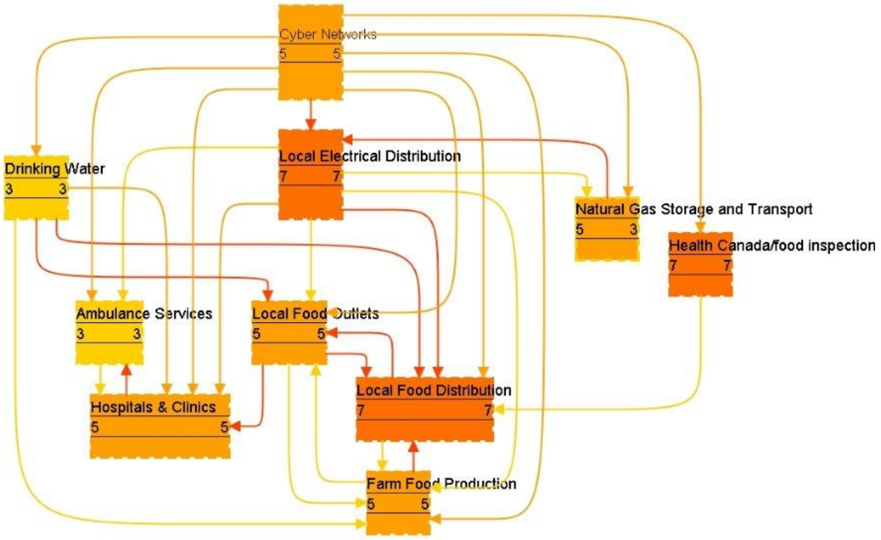


Figure 3. A small infrastructure model

= 5, orange edges), and low (score = 3, yellow edges). All entities except Local Electrical Distribution have been assessed as having medium dependence on Cyber Networks. Local Electrical Distribution, however, has been scored as having high dependence on Cyber Networks.

After path analysis is carried out and scores for global impact and global vulnerability are computed, a histogram of the risk indices can be created, as shown in Figure 4. The bars represent risk-index scores arranged from the lowest score to highest score, from left to right. The highest-risk entity in the model is Local Electrical Distribution. The second-highest entity is Local Food Distribution. At medium risk are three entities: Health Canada/Food Inspection, Local Food Outlets, and Cyber Networks. The remaining entities are of relatively low risk.

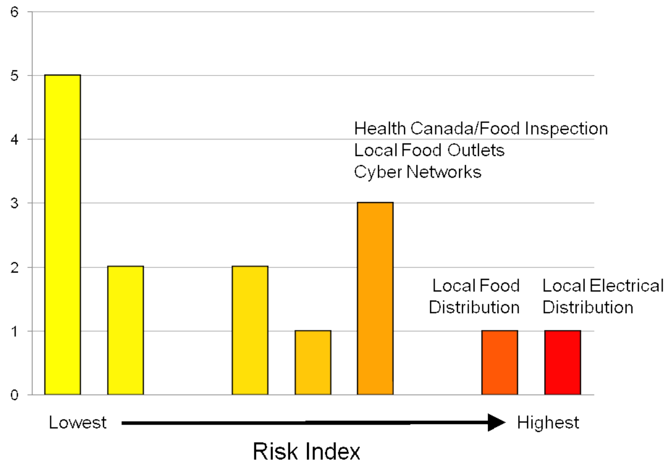


Figure 4. Histogram of the risk indices for the model shown in Figure 3

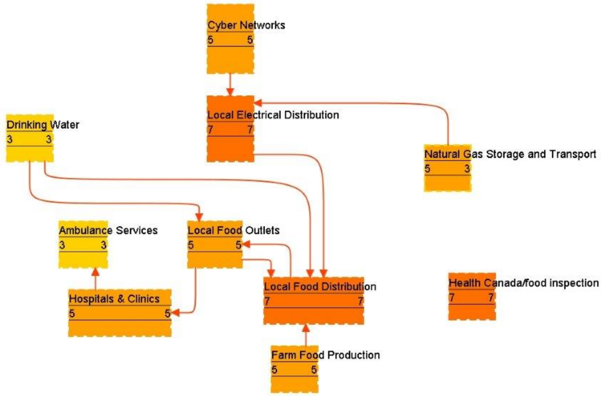


Figure 5. High-dependency relationships for the example model shown in Figure 3

Even though Cyber Networks is assessed as medium risk in the model, we can still assess all consequent impact were it to fail. Figure 5 shows the sub-network of high-dependency relationships in the model.

From this sub-network, we can isolate the paths of high impact emanating out of the Cyber Networks entity, as depicted in Figure 6.

Although only Local Electrical Distribution has a high direct dependence on Cyber Networks, Figure 6 shows that Local Food Distribution, Local Food Outlets, Hospitals & Clinics, and Ambulance Services would all fail if Cyber Networks were to fail.

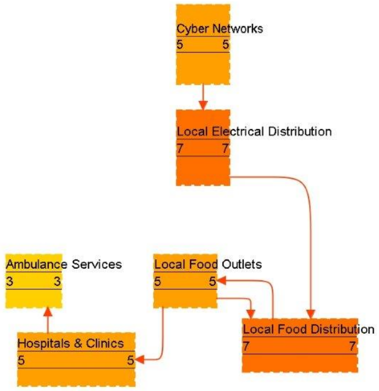


Figure 6. Paths of high impact from the Cyber Networks entity for the model shown in Figure 3

Discussion

The strongest-path method provides a tool for assessing and prioritizing risk. The risk index provides a global measure of risk for every entity in a model. This depiction of risk is of strategic value to decision makers in that it gives them a strategic prioritization of every entity.

Moreover, with the strongest-path method, the global impact and global vulnerability of every entity in a model is assessed so that a separate prioritization – based on either impact or vulnerability – is available to decision makers. Thus, any scenario of high impact can be identified, no matter how unlikely it is to occur. Conversely, possible situations of triggering events for unlikely scenarios can be identified from the path analysis.

With the strongest-path method, the influence of every node on every other node in a model is assessed. Consequently, decision makers can make plans based on identified pathways of exposure to risk. Risk-mitigation plans and contingency plans can take into consideration chains of events that might otherwise have gone unnoticed.

A model based on the strongest-path method can be made as detailed as required for the decision-making requirements of the planners or managers who will use it. Entities should embody the level of detail that is significant in the environment of interest. For example, a cyber network might be modelled at the level of individual servers and computers together with the direct connections that link them together. This model might also include entities and relationships that model power sources and the distribution of power.

In situations where entities and relationships change over time, a time-oriented model can be built using time-intervals that represent periods when changes do not occur among the entities and relationships. For example, consider a hospital that has a backup power generator with 36 hours of fuel to sustain it during a power failure. At t=0, the hospital has medium dependence on local power distribution because of its backup system. Once the power failure starts, it has high dependence on its backup generator until t=36 hours. By that time, it must have acquired more fuel or it must shut down, and it is deemed to have failed if there is low likelihood of having fuel delivered by that time. Any entities with high dependence on that hospital will also fail at t=36 hours.

Conclusion

The strongest-path method is a paradigm for modelling infrastructure risk using a directed graph. Models are constructed from entities that are assessed with a degree of impact and a likelihood of failure together with dependency relationships between the entities that are scored for degree of dependence according to well defined criteria.

The paradigm allows the knowledge of experts to be used for infrastructure risk analysis. Results from other analytical models, such as simulations, can also be included in a model. As a result of performing the path analysis, such models reveal the potential consequences of the failure of any entity on all of the others. This enables contingency planners to anticipate all outcomes in any infrastructure damage scenario.

The strongest-path method and the RiskOutLook software are currently being used by Emergency Management Ontario to manage risks in critical infrastructure. The provinces of New Brunswick and Saskatchewan will soon begin projects to build similar infrastructure models.

About the Author

Philip O'Neill is Chief Scientist at Deep Logic Solutions Inc. He holds a PhD in Combinatorics and Optimization from the University of Waterloo, Canada. He is a specialist in operational research and risk analysis, and has additional expertise in mathematical modelling, quantitative analysis, algorithms, and decision support. His career has included 17 years of practice in the Operational Research Division of the Department of National Defence (DND); he has served as chairman of the NATO Panel 7 Specialist Team on the Evaluation of Readiness and Sustainment Policy; and he was chosen by the DND to model dependency relationships among infrastructures in Canada as part of risk analysis for the millennium turnover. Since 2001, he has designed and managed the software development of RiskOutLook, an analytical tool for risk analysis that identifies and quantifies risks that result from dependency relationships.

Citation: O'Neill, P. 2013. Protecting Critical Infrastructure by Identifying Pathways of Exposure to Risk. *Technology Innovation Management Review.* August 2013: 34–40.

Keywords: cybersecurity, risk analysis, strongest-path method, directed graph, path analysis, critical infrastructure, simulation, modelling

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

> “*We need to establish security engineering as a valid profession in the minds of the public and policy makers. This is less about certifications and (heaven forbid) licensing, and more about perception – and cultivating a security mindset. Amateurs produce amateur security, which costs more in dollars, time, liberty, and dignity while giving us less – or even no – security.*”
>
> Bruce Schneier
> Cryptographer and computer security specialist

Despite nearly 30 years of research and application, the practice of information system security engineering has not yet begun to exhibit the traits of a rigorous scientific discipline. As cyberadversaries have become more mature, sophisticated, and disciplined in their tradecraft, the science of security engineering has not kept pace. The evidence of the erosion of our digital security – upon which society is increasingly dependent – appears in the news almost daily.

In this article, we outline a research agenda designed to begin addressing this deficit and to move information system security engineering toward a mature engineering discipline. Our experience suggests that there are two key areas in which this movement should begin. First, a threat model that is actionable from the perspectives of risk management and security engineering should be developed. Second, a practical and relevant security-measurement framework should be developed to adequately inform security-engineering and risk-management processes. Advances in these areas will particularly benefit business/government risk assessors as well as security engineers performing security design work, leading to more accurate, meaningful, and quantitative risk analyses and more consistent and coherent security design decisions.

Threat modelling and security measurement are challenging activities to get right – especially when they need to be applied in a general context. However, these are decisive starting points because they constitute the foundation of a scientific security-engineering practice. Addressing these challenges will require stronger and more coherent integration between the sub-disciplines of risk assessment and security engineering, including new tools to facilitate that integration. More generally, changes will be required in the way security engineering is both taught and practiced to take into account the holistic approach necessary from a mature, scientific discipline.

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

Introduction

Despite nearly 30 years of research and application, the practice of information systems security engineering has not yet begun to exhibit the traits of a rigorous scientific discipline (Cybenko and Landwehr, 2012; tinyurl.com/kc3nm7p). As a result, it is still not possible to examine an information system and answer the question "How secure is it?" in a scientifically meaningful way. This is a significant problem because, increasingly, the economic and physical well-being of our society *depends* on the secure design and operation of business, government, and critical-infrastructure information systems. They appear in almost every facet of our daily lives but we actually know very little about how they stand up when it comes to security (Viega, 2012; tinyurl.com/mnwqd8c). It would be truly alarming to ask the question "How safe is it?" with respect to an aircraft only to discover that neither the engineers nor the certifiers really understood the answer. Yet, this is precisely the situation in which the information-technology security community finds itself today.

Although most of the concepts and ideas found in this article are applicable to security engineering at large, here we use term "security engineering" with specific reference to the security of information systems. Thus, in the context of this article, we define security engineering as "the art and science of discovering users' information protection needs and then designing and making information systems, with economy and elegance, so they can safely resist the forces to which they may be subjected" (National Security Agency, 2002; tinyurl.com/kcx5y4u). This definition has an analog in the physical sciences where the "forces" are natural and safety is an absence or avoidance of physical injury. As we shall see, this natural analog can inform us when answering questions such as "How secure is it?" in a more precise and consistent fashion.

Challenges

We see two significant challenges holding back the science of security engineering. First, it is unlike other engineering fields in the respect that the majority of the "forces" to be modelled are caused by human threat actors with *deliberate intent*, as opposed to forces due to natural and accidental causes. Thus, the first major hurdle facing security engineering is to define and maintain a threat model that can be used to calculate or bound these "forces" in a way that results in consistent engineering outcomes. This does not mean that threat models do not exist. In fact, like the nascent stages of any young scientific discipline, there are many models which, unfortunately, can lead to inconsistency and duplication of effort. For example, in many methodologies for assessing threats and risks, such as the Harmonized Threat and Risk Assessment Methodology developed by Communications Security Establishment Canada and the Royal Canadian Mounted Police (CSEC/RCMP, 2007; tinyurl.com/kfrjgv8) and the Guide for Conducting Risk Assessments developed by the National Institute for Standards and Technology (NIST, 2012; tinyurl.com/6srqlug), assessors are coached on *developing* a threat model. To be sure, threat analysts are not likely to rebuild their threat models each time they perform an assessment. Rather, the models are developed incrementally over time and are based substantially on individual knowledge and experience. However, while there may be commonality between models created by different assessors, there is certainly no guarantee that this is the case. This inconsistency (along with variations in categorization schemas, methodologies, definitions, and terminology) makes it challenging to validate and reuse results that would eventually drive the community to a small set of the most successful models. This convergence, which is a hallmark of a mature science, has not yet occurred within the security community.

Thus, we argue that a common threat model should be a primary goal of the security engineering community. This model should define the threat environment and the "forces" involved in a way that can be validated and built upon over time through repeatable qualitative or quantitative analyses. Such a model would also be "actionable" in the sense that threat-assessment results would point naturally to design options for security engineering that, at the outset, may be its primary measure of success. Such an undertaking would, of course, require a concerted research and development agenda to lay a common foundation upon which validation and refinement can begin to occur.

A second and potentially more challenging problem is the need for a useful framework for *security measurement*. Currently, there is no practical, relevant way of measuring the *absolute* security of an information system. In fact, there is no clear understanding of what absolute security means (e.g., Pfleeger, 2012: tinyurl.com/mt2xnsw; Böhme, 2010: tinyurl.com/kn99d4q; Davidson, 2009: tinyurl.com/l7475p3; Houmb et al., 2010: tinyurl.com/lw9ffqz; Savola, 2007: tinyurl.com/la87e84; Pfleeger, 2007: tinyurl.com/k3rjb6z; McHugh, 2002: tinyurl.com/m4z9nuu).

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

Absolute security is different from the common practice of “measuring” a system’s compliance against arbitrary security requirements. Although a system may be 100% compliant with a set of security requirements, in most cases, there is little direct evidence that those requirements actually result in a more “secure” system. Clearly, security measurement will be assisted to a large extent by a common threat model, but the two approaches are co-dependent because threat modelling will eventually require quantitative measurement in order to demonstrate success.

We examine both of these challenges in the following sections and describe our vision for a security community-driven research program. For both challenges, we contend that the best approach is to take cues from established disciplines such as civil, mechanical, or electrical engineering and to draw analogies wherever possible. We feel that the closer we draw these parallels, the clearer will be our understanding of where we need to proceed next.

An Actionable Threat Model

A generally accepted model of deliberate threats is central to the advancement of the security engineering discipline. The most important aspect of such a model is that it be *actionable* from an engineering perspective. That is, when defining security requirements and undertaking risk-based design, the model would help to consistently and coherently identify a suite of design options – along with the associated security controls and their required level of implementation assurance – that could meet the goals identified by the system owner in terms of cost, operational utility, and risk tolerance. This is, of course, analogous to using standard engineering models during design activities (e.g., high-frequency antenna design).

Typical methodologies for assessing threats and risks, such as the Harmonized Threat and Risk Assessment Methodology (CSEC/RCMP, 2007; tinyurl.com/kfrjgv8) and the NIST's Guide for Conducting Risk Assessments (2012; tinyurl.com/6srqlug), generally focus on the generation of information related to risk decisions. In these assessments, a potentially long list of threat actors or threat scenarios is generated to estimate threat attributes and calculate the potential for risk. In other words, motivation is assessed for the purposes of determining the likelihood of an attack. Few, if any, threat attributes are identified and assessed in a way that purposely helps with the selection of design options or security controls, or that helps determine levels of implementation assurance.

In order to achieve an actionable threat model, there are two fundamental changes that must be made to the way threats are assessed. First, the act of performing a *threat assessment* must be divorced from the act of performing a *risk assessment*. Existing threat and risk-assessment frameworks make little distinction between these activities. Although they have elements in common, each activity requires a different skill set and targets different audiences. Second, threats should be assessed based (at least initially) on the *capabilities* that a threat actor *could* wield rather than on attributes that are specific to threat actors themselves (e.g., motivation, intent, risk aversion, willingness to invest time). Actor-specific attributes are more appropriately addressed during risk assessment. We explore a capabilities-based approach in the following sections.

Threat assessment based on threat actor attributes

Typically, threat-assessment methodologies begin by asking which threat actors are likely to attack an information system. In some cases, threat actors may be dropped from consideration if the likelihood of an attack is deemed to be very remote. More often, this likelihood is used to condition the potential risk from the attacker downwards (the injury from an attack does not change, just the magnitude of the outstanding risk). This approach places bounds on the costs of security, both in terms of money and constraints on operational freedom, and focuses limited resources where significant injury is *expected* to occur.

The likelihood that a threat actor will attack is often determined by examining certain attributes such as the actor's capabilities (what kinds of attacks they are capable of doing), motivations and intents, aversion to risk, willingness to invest time and effort, degree of access, etc. A difficulty with this approach is that many of these attributes cannot be accurately modelled or assessed because they can change frequently over time or they are based on complex mental states or behavioural patterns. As a consequence, assessments based on these attributes have large uncertainties, which makes the expectation of where significant injury will occur less accurate.

A more significant challenge with this approach is that an analyst must develop an exhaustive list of credible threat actors and their attributes in order to ensure that all threat scenarios are addressed. However, it is difficult to reason about the completeness of this list as

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

demonstrated by such events as the Oklahoma City bombing (tinyurl.com/gesg2), the September 11 attacks (tinyurl.com/nhx7m), the Fukushima Daiichi nuclear disaster (tinyurl.com/44hjgf3), and the Lac Mégantic derailment (tinyurl.com/q2fh9nk).

In the course of compiling this list of credible threat actors, the analyst must also think about, enumerate, and assess the threat actor's *capabilities* – the ways in which each threat actor might attack the system. As illustrated in Figure 1, the challenge with this approach is that capabilities can be missed if, for example, the assessor fails to consider every possible threat actor scenario or if threat actors are not adequately considered because they are viewed as being unlikely to attack. Some of these unidentified capabilities, if exercised, could be crippling to an organization. This revelation will not necessarily be made obvious by thinking only about threat actors. It is also natural to assume that even those threat actors that *have* been considered will evolve over time and that some may come into possession of more sophisticated, or even as-yet unidentified capabilities. Only a well-disciplined, frequent refresh of the assessment of threat actors will be able to track this evolution. In the next section, we argue that a better approach is to base a threat assessment on capabilities instead of threat actors themselves.

Threat assessment based on threat actor capabilities
Instead of modelling the characteristics of threat actors such as their motivation and intent, resources, and tolerance to risk, we propose that the threat assessment should be focused on the *capabilities* that can be employed to attack a system. Using this approach, it is possible (although potentially challenging) to: i) develop a more exhaustive survey of the threat *potential*, ii) reason about the completeness of the analysis, and iii) identify potential gaps in our knowledge. A capabilities-based approach also lends itself to a community effort because system-specific information need not be divulged. Given that capabilities can be assessed in a general context, the material will be highly reusable. A final benefit of this approach is that it provides a common interface between attacks by threat actors and the controls necessary to effectively counter them, as illustrated in Figure 2.

But what do we mean by threat actor capability? In our view, a *capability* is composed of a *vector* and a *sophistication level*. In simple terms, a *capability vector* defines a coarse taxonomy of attacks on an information system (Figure 3). A capability vector identifies *where*, *how*, and *what*. These fields of a capability vector can include:

1. **Access Mode:** This field identifies the means by which access to the target system is obtained. *Direct modes* include physical, personnel, logical, and electromagnetic access. *Indirect modes* involve direct modes that are applied to lifecycle elements of the system (e.g., system development, software patches, replacement hardware, and support-system operations). Indirect modes are potentially recursive and generally relate to the supply chain of a system.

2. **Target Layer:** A capability vector will be designed to act against one or more architectural "layers" within a system, such as the application, data, operating system, virtualization, network, firmware, and hardware.

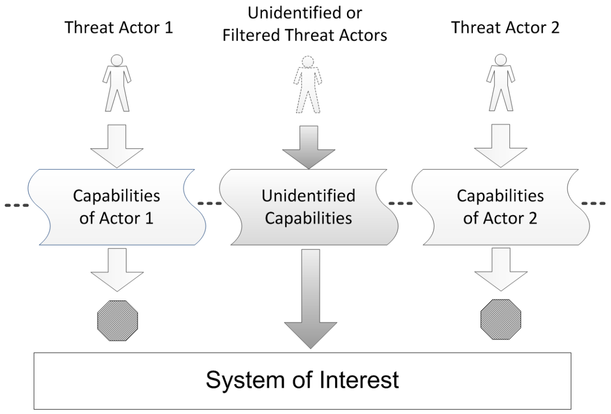


Figure 1. Capabilities from unidentified threat actors may be overlooked

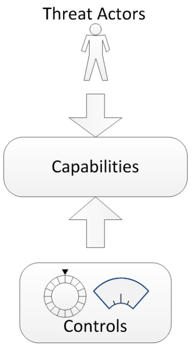


Figure 2. Threat actors are related to controls through their capabilities

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

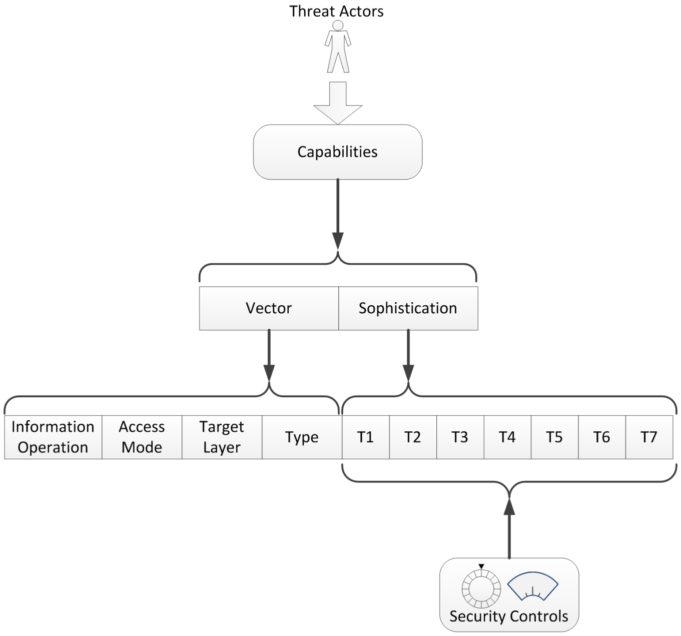


Figure 3. The threat capability model

3. **Type and Sub-type:** These fields identify a common name or grouping for a specific capability in order to maintain semantic compatibility with common terminology for various attacks. The Type field may be followed by a number of Sub-type fields to further distinguish capabilities. An example of the material that could be found in the Type field could include top-level categories found in the Common Attack Pattern Enumeration and Classification (CAPEC; capec.mitre.org).

4. **Information Operation:** This field is optional but can be helpful as a way of grouping or indexing attack vectors that have similar effects. The Information Operation field hints at *why* a given capability might be exercised (i.e., its intent). A capability vector may belong to more than one type of information operation. Potential categories include deny, exploit (infiltrate, exfiltrate), reconnoiter, deceive, etc.

It should be noted that this breakdown is only a recommended starting point. More, fewer, or different categories may be needed as the framework evolves.

The second element of a capability is defined by its *sophistication*. For example, consider a denial-of-service capability. A distributed version of this capability can be performed using software downloaded from the Internet and executed from a dozen computers. The same capability can be launched from 5000 computers distributed all over the world using code that exploits a previously unknown vulnerability. The differences between these capabilities are: i) the level of sophistication required to set up and execute them and ii) the set of controls required to prevent or limit the successful use of the capabilities against a system (as well as the rigour with which they are designed, implemented, and operated).

Thus, simply identifying a capability vector is not enough. To fill out the threat assessment, security assessors must also determine if there are distinguishing features (or attributes) that make the same capability vector harder to detect or prevent and then identify what options exist to address these more sophisticated variants (i.e., controls, architecture changes). In Figure 3, we divide sophistication into seven distinct levels according to those originally proposed in the National Security Agency's Information Assurance Technical Framework (NSA, 2002; tinyurl.com/kcx5y4u). However, we have not yet determined what would constitute a worthwhile set of distinguishing sophistication attributes, although we suspect that they may be somewhat dependent on features of each individual capability vector.

It is important to emphasize that identifying graduated levels of sophistication leads to the selection of security controls and design options that are generally more expensive or more operationally constraining as one moves up the scale of sophistication. This approach provides risk assessors with more explicit information regarding the tradeoffs between threat mitigation and costs.

As a final note, an important feature of the capabilities-based approach is that it has some predictive potential. That is, if we expanded every combination of the first attribute with the second and then third attribute, we obtain the universe of possible capability combinations. Some combinations will not make sense and can be disregarded while others will have ample evidence to show that they are in active use by threat actors at various levels of sophistication (e.g., logical access, operating system, Trojan, infiltration). Other combinations will appear strangely unfamiliar, either because they have never been exercised or they have been exercised but have never been publicly observed (e.g., electromagnetic, hardware, audio covert channel, exfiltration). Thus, capability vectors should tell us where we need to be looking for evidence of attack and, as a corollary, where threat actors might look in order to find new opportunities to expand their capabilities.

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

Figure 4 illustrates a fictitious set of capabilities related to a denial operation. For each capability vector, there are seven cells in which information about sophistication can be captured. Exactly what information should be contained in each cell and how it is determined is a fundamental research problem. To satisfy risk assessors' need to quantify the "potential" for an attack using a given capability, we propose three general categories as follows:

1. The capability has been observed to be in use by at least one threat actor at the given level of sophistication (i.e., the black boxes marked with an "O" in Figure 4).

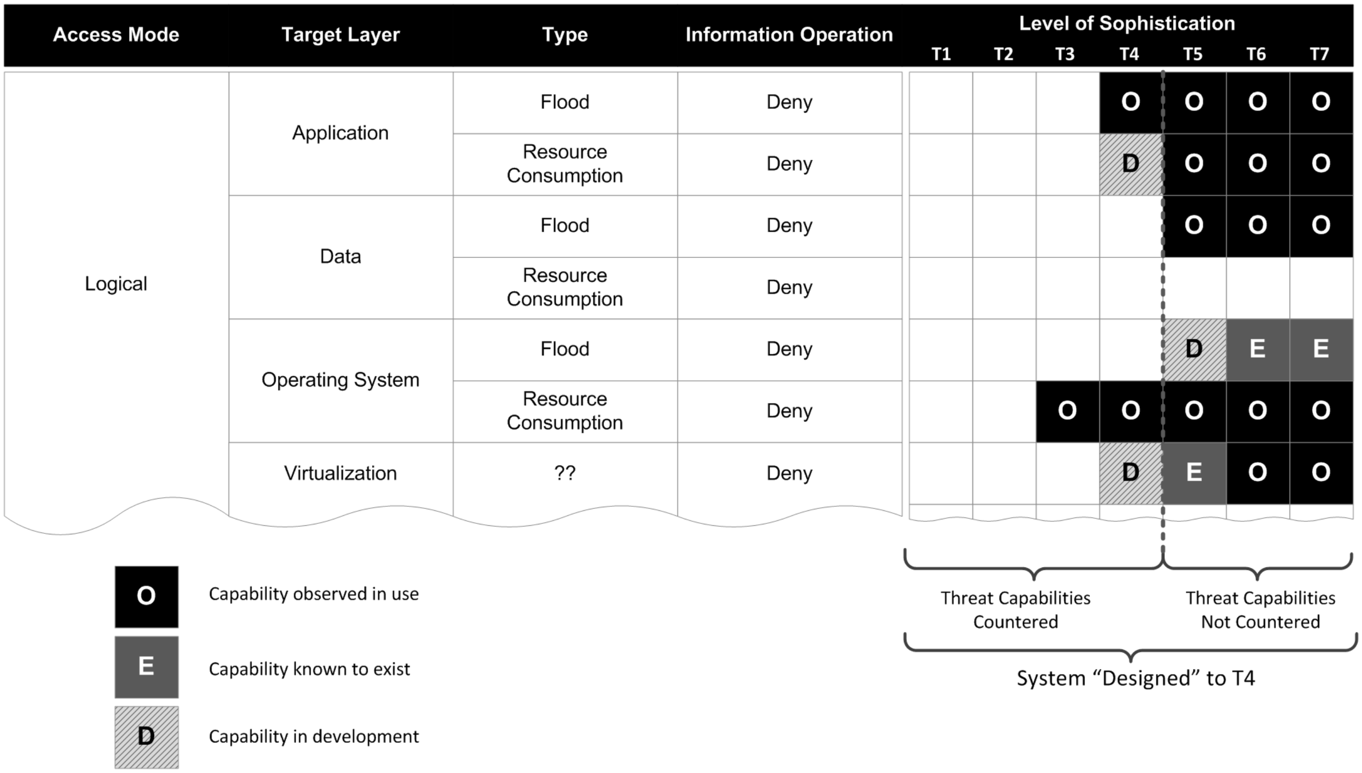
2. The capability has been demonstrated at a conference such as DEF CON (defcon.org), but has not yet been observed "in the wild" (i.e., the hashed grey boxes marked with a "D" in Figure 4).

3. The capability is known to exist at a given level of sophistication but has not been observed (i.e., the dark grey boxes marked with an "E" in Figure 4); an example would be a nuclear-generated electromagnetic pulse.

Many capabilities follow a "commoditization" lifecycle in which they are generated at high levels of sophistication but are subsequently made easier to implement and become more widely available at lower levels of sophistication. This can be represented in Figure 4 as a heat map and could provide valuable information for risk assessors when considering the need for security controls over a long period of time.

From an engineering perspective, each cell should map to a set of security controls, mechanisms, strategies, security design patterns, and implementation-assurance requirements that have been shown (preferably through quantitative analysis) to effectively counter the threat capability at the specified level of sophistication.

Regardless of the type of information included with each cell, the basic research problem is as follows: given a post-incident analysis of a threat event (or an analysis based on vulnerability research work), how do we determine what sophistication level is represented? We see this as a difficult challenge because it will be a necessarily subjective exercise (at least at the outset). However, there are both theoretical and practical approaches that can help to reduce variation and uncertainty introduced by this subjectivity.



Access Mode	Target Layer	Type	Information Operation	Level of Sophistication						
				T1	T2	T3	T4	T5	T6	T7
Logical	Application	Flood	Deny				O	O	O	O
		Resource Consumption	Deny				D	O	O	O
	Data	Flood	Deny					O	O	O
		Resource Consumption	Deny							
	Operating System	Flood	Deny					D	E	E
		Resource Consumption	Deny			O	O	O	O	O
	Virtualization	??	Deny				D	E	O	O

Threat Capabilities Countered
Threat Capabilities Not Countered
System "Designed" to T4

O	Capability observed in use
E	Capability known to exist
D	Capability in development

Figure 4. A threat-capability example for a fictitious set of vectors

Security Measurement

A second key challenge facing information-system security engineers is the difficulty involved in actually *measuring* security in a *practical* and *relevant* way (Pfleeger, 2007; tinyurl.com/k3rjb6z). There is a significant body of research on the subject but all studies seem to fail one or both tests of practicality or relevance (e.g., Lundin et al., 2006: tinyurl.com/mez2k8k; Shin et al., 2011: tinyurl.com/k2nvk53). If this were not the case, we would have a working solution by now. Being able to measure security in a useful way is absolutely critical to the advancement of security engineering as a discipline, because measurement is the bedrock of the scientific approach.

Why measurement is difficult

Security measurement is challenging for a number of reasons. First and foremost is the problem of concisely defining what is meant by the term *security*. Krautsevich, Martinelli, and Yautsiukhin (2010; tinyurl.com/kpuek8j) note that "we do not have a widely-accepted and unambiguous definition" that enables us to indentify one system as more secure than another. However, the definition of security often depends on perspective and context; it means different things to different people in different roles. Thus, we believe that there are a multitude of definitions that may all be equally useful within their own contexts. For example, contradicting security objectives may arise when considering an organization's need to monitor and control what happens on their systems versus an employee's need for legal privacy protection. Nevertheless, because the definition of security will have a significant impact on the way it is measured, it is critical to ensure that it is chosen appropriately and used consistently.

A second difficulty is that, regardless of how they are defined, security properties must actually be *measurable* and those measurements must be *practical* to obtain. There are at least three types of *security measurement* that information-system security should be concerned with:

1. **Engineering measurement:** These measurements are used by engineers to build models that "provide a formal representation (e.g., sets of equations) that corresponds well to security for systems under consideration" (Verendel, 2010; tinyurl.com/lgsxnrl). These are the same kind of measurements that one would expect from, for example, stress and strain analysis of various materials in civil engineering.

2. **Compliance measurement:** These measurements establish the degree to which an information system meets a set of specifications derived from security functionality and assurance requirements. Compliance measurement is normally performed throughout the process of system development. Examples of these types of measurements are described in the "Overview of IT Security Risk Management: A Lifecycle Approach" (CSEC ITSG-33: Annex 2, 2012; tinyurl.com/kf5ejyu) and the Common Criteria (commoncriteriaportal.org).

3. **Operational measurement:** These types of measurement provide metrics to reflect the operational security performance of an information system. Examples include patch-management coverage, mean time to mitigate, etc. Related resources include the ISO/IEC 27004:2009 standard for measurement techniques in information security (tinyurl.com/ln92xe3) and the metrics used by the Center for Internet Security (cisecurity.org).

In this article, we are concerned primarily with engineering measurement because it is a prerequisite for advancing the science of security engineering. Unfortunately, these kinds of measurement also appear to be the most difficult to obtain in a *quantitative way* (Wang, 2005: tinyurl.com/mgj3mj3; Verendel, 2010: tinyurl.com/lgsxnrl). They require a common, objective scale and a measuring device, and both must be accepted broadly across the security-engineering community in order to gain traction (Zalewski et al., 2011; tinyurl.com/lqw6865). Security engineering lacks these quantitative standards, primarily because security is often expressed in abstract terms.

In the absence of quantitative measurements, qualitative assessments have been used to derive security metrics. Qualitative assessments may be the best that security engineering can achieve until appropriate quantitative measures become available. Unfortunately, subjectivity implies inconsistency, which is unacceptable in a science-based discipline. Although it may not be possible to eradicate subjectivity altogether, there are certainly ways to minimize it. In some respects, we are advocating the same approach (but on a much larger scale) that the National Institute of Standards and Technology (NIST) has taken with cryptography. NIST arranges encryption algorithms by key size according to number of "bits of security" that they provide. The scale is *nominally* objective but an algorithm's placement on the scale is the result of expert judgment by one or more

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

cryptographic mathematicians who estimate the "amount of work" necessary in order to crack a ciphertext using an algorithm with the given key length (NIST 800-57: Revision 3, 2012; tinyurl.com/n5dk85u). Here, the measurement is a subjective assignment on a constructed scale and clients use that as an engineering measurement to compose their designs.

A third challenge facing security measurement is the notion of *assurance*. In the cryptography example above, we looked at measurement from the perspective of answering "How strong is it?" but we have not asked the question "How well does it work?" Without addressing assurance, efforts in addressing security are wasted. For example, although an encryption algorithm may have strong *conceptual* security (i.e., theoretical strength), if the algorithm is implemented incorrectly, then its *actual* security (i.e., robustness) is weak.

Assurance measurement is not as widely considered as *strength* within the security research community. Notable exceptions include cryptographic evaluations following the Federal Information Processing Standard (FIPS 140; tinyurl.com/pn9mb4) and Common Criteria assurance requirements (commoncriteriaportal.org/cc/). This lack of attention to assurance measurement is probably due to the fact that it appears to be even more abstract and (in most cases) more subjective than measurements of security strength. Generating and communicating assurance information in a "standardized" way would serve to reduce subjectivity, and this is the focus of at least one object-management group's specification (Alexander et al., 2011; tinyurl.com/mtjuufn). However, combining information about assurance and strength into a composite measure of security should be subjected to further analysis and validation. Some work has been accomplished in this direction with the notion of "robustness" in the CSEC's IT Security Guidance (CSEC ITSG-33: Annex 2, 2012; tinyurl.com/kf5ejyu), the Information Assurance Technical Framework (National Security Agency, 2002; tinyurl.com/kcx5y4u) and some Common Criteria protection profiles (commoncriteriaportal.org/pps/).

Going forward

On the issue of pursuing a research agenda that addresses practical and relevant security measurement, we propose a straightforward approach: consider security engineering in relation to other mature engineering disciplines and draw as many analogies as possible. This approach has been advocated before by Zalewski and colleagues (2011; tinyurl.com/lqw6865), who perceive that moving ahead requires "a closer alignment of security assessment with concepts developed in measurement science and physics". Where an analog in security engineering does not exist or does not translate easily, then we have an item to add to the research agenda. We suspect this might help identify the form that measurements *should* take.

For example, civil engineers build structures that are designed to exist in a certain threat context. "Loads" (in terms of forces) are applied in three-dimensional space – often downward with the pull of gravity but sometimes in other directions due to other natural or man-made forces. The scale and magnitude of these loads is directly proportional to defined levels within each threat event; for example, an "F2 tornado" (tinyurl.com/2frdj2) or a "Cat3 hurricane" (tinyurl.com/kl5ukgo). Minimal load levels for certain threat events are specified by regulatory bodies and have an effect on the way the structure is architected and designed (e.g., minimal spacing between load bearing members). Other, unregulated threat events may be of specific concern to certain clients and the forces to be countered by these threats may be specified as additional design requirements (e.g., bollards in front of federal buildings).

A natural security analog to "loads" is provided by the spectrum of threat sophistication that we proposed earlier. However, although we imply that the "load" imposed by a threat capability vector at sophistication level 7 is "greater" than level 6, we do not have a clear understanding of what "forces" are applied by threat agents against the information-system infrastructure and what effects these may have. Addressing these gaps in our understanding may help us develop the engineering metrics that are needed to advance the science of security engineering.

Putting It All Together

In the following sections, we outline a few specific areas where improvement can be expected as a result of taking on the research agenda proposed in this article.

Composite security

The "holy grail" of security engineering is to be able to answer the *composition problem* (Irvine and Rao, 2011: tinyurl.com/n626rgo; Datta et al., 2011: tinyurl.com/kukg98y). That is, given an information-system architecture or design made up of discrete security and non-security components, solving the composition problem would allow us to determine the overall security of the information system. The composition problem is a common lament in the information-security domain; "We simply have no theoretical basis for judging the security of a system as a whole" (McHugh, 2002; tinyurl.com/m4z9nuu). However, the composition problem cannot be solved without measurement, and measurement cannot be performed without a generally accepted threat model.

Development of mandatory security requirements
Having a science-based threat model and security-measurement framework would allow the security community to influence the development of security standards that are based on sound engineering principles. In civil engineering (and certainly other engineering disciplines), threat events that may pose a risk to safety are incorporated over time into standards, codes, and regulations. This information is gleaned from engineering measurement and, in some cases, spectacular failures such as the Tacoma Narrows Bridge (tinyurl.com/c77rpw). A hallmark of a mature engineering science is the ability to investigate and learn from these failures and recycle that information into curricula, codes, and regulations.

Security engineering appears to have few close equivalents to requirements specified in codes and regulations – anti-virus and access control mechanisms seem to be a standard requirement found in most system specifications, although these are by no means mandated. In order to begin embedding security controls in security standards (especially if they are very expensive), it is necessary to thoroughly understand those controls from an engineering-science perspective.

Security-engineering curriculum
Finally, we note the fact that many curricula being proposed for security engineering in a college or university setting are simply computer engineering or computer science degrees that have been sprinkled with topics in security, assurance, and, unfortunately, risk assessment or risk management (e.g., Hjelmås and Wolthusen, 2006: tinyurl.com/kncwfek; Older and Chin, 2012: tinyurl.com/l5bbtah; Irvine and Nguyen, 2010: tinyurl.com/mvzj4xa). As far as we know, there is no curriculum that seeks to build (or build upon) a set of mathematical (or at least more formal) models that allow the composite security of an information system to be determined in a repeatable, meaningful manner. We suspect this is due to a lack of understanding of where exactly to begin.

Conclusion

In this article, we broadly outlined a research agenda that, with sufficient effort, would help begin the process of placing security engineering for information systems on foundations equivalent to other mature engineering disciplines. Two significant areas requiring attention were identified: threat modelling and engineering-security measurement. We argued that these areas are critical starting points because they affect almost all other aspects of security engineering, and more generally, the field of IT security. In addition, we believe that in order to be successful, these areas of research should be performed by a multi-disciplinary team of subject-matter experts. In taking on this research agenda, there is considerable opportunity to affect a significant change in the security posture of existing and future information systems. And, in doing so, the security and privacy of Canadians and the trust that they invest in the information systems of businesses, governments, and critical-infrastructure information systems will also be positively affected.

Acknowledgements

We would like to thank Larry Stoddard for his contribution and insights into this work.

A Research Agenda for Security Engineering

Rich Goyette, Yan Robichaud, and François Marinier

About the Authors

Richard Goyette is Senior Security Architect at Communications Security Establishment Canada. Richard has a BEng and MEng in Electrical Engineering, both from the Royal Military College of Canada in Kingston, Canada. Richard spent 22 years as a Signals officer in the Canadian Forces, where he was involved with a multitude of projects in the areas of intelligence, security, and command and control. He is currently employed in the area of architecture and technology assurance developing security guidance for the wider Government of Canada.

Yan Robichaud is a Senior Security Architect at Communications Security Establishment Canada. Yan has a BASc degree in Computer Engineering and MSc degree in Electrical Engineering, both from Université Laval, Québec City, Canada. He provides advice and guidance related to security architecture and engineering, threat assessment, and risk management to Government of Canada departments and agencies. He is involved in key government IT initiatives, such as large IT consolidation projects, enterprise security architecture, and the security of space-based systems. Yan is also involved in the development of IT security courses and leads the production of publications about IT-security guidance, such as "ITSG-33 IT Security Risk Management: A Lifecycle Approach".

François Marinier is an independent IT security analyst with experience in all facets of IT-security risk management. François started his career working in computer operations and mainframe application support. He eventually migrated to IT security, where he acquired knowledge and experience in the development and application of processes for IT-security risk management. He has also worked as an analyst, supporting large IT-infrastructure initiatives, in both the public and private sectors. For the last three years, François has dedicated his work almost exclusively to the development of ITSG-33, the next generation of guidelines for IT security risk management for the Government of Canada.

Citation: Goyette, R., Y. Robichaud, and F. Marinier. 2013. A Research Agenda for Security Engineering. *Technology Innovation Management Review.* August 2013: 41–50.

Keywords: cybersecurity, information system security engineering, research, risk management, security engineering, threat modelling, security measurement

Multifactor Authentication: Its Time Has Come

Jim Reno

> “ *'What does it mean by* speak, friend, and enter*?' asked Merry.* ”
>
> *'That is plain enough,' said Gimli. 'If you are a friend, speak the password, and the doors will open, and you can enter.'*
>
> *'Yes,' said Gandalf, 'these doors are probably governed by words. Some dwarf-gates will open only at special times, or for particular persons; and some have locks and keys that are still needed when all necessary times and words are known.'*
>
> *The Fellowship of the Ring*
> J.R.R. Tolkien

Transactions of any value must be authenticated to help prevent online crime. Even seemingly innocent interactions, such as social media postings, can have serious consequences if used fraudulently. A key problem in modern online interactions is establishing the identity of the user without alienating the user. Historically, almost all online authentications have been implemented using simple passwords, but increasingly these methods are under attack. Multifactor authentication requires the presentation of two or more of the three authentication factor types: “What you know”, “What you have”, and “What you are”. After presentation, each factor must be validated by the other party for authentication to occur. Multifactor authentication is a potential solution to the authentication problem, and it is beginning to be implemented at websites operated by well-known companies. This article surveys the different mechanisms used to implement multifactor authentication. How a site chooses to implement multifactor authentication affects security as well as the overall user experience.

Introduction

The last year has brought news of a number of prominent security breaches centered on authentication, with, in some cases, severe consequences. A not uncommon pattern is a revelation that some server has been hacked and a large number of account passwords have been potentially exposed. *Potentially* because while we know files containing things such as password hashes have been copied, there is often no subsequent information on actual fraudulent use of the data or real damage done. An example of a security breach where damage actually resulted is the attack on the Associated Press Twitter account of April 2013. A bogus tweet about explosions at the White House caused a brief, but serious, disruption to the financial markets (Selyukh, 2013; tinyurl.com/d6zozam).

The industry is slowly reacting to password attacks and is starting to try to find better ways to prevent them. Media attention is growing. In particular, each publicized password attack is usually followed by a series of articles decrying the “end of the password” and calling for implementation of multifactor authentication (MFA). An online site using MFA is harder to attack – to “break into” – than a site authenticating users with only a single factor such as a password. The widespread adoption of MFA would improve online security and help reduce fraud.

MFA is not a new idea. Consider a Roman soldier guarding the Senate door and requiring senators to show a ring and speak a password. This is an example of two-factor authentication. MFA has been implemented in online systems for many years. Until recently, however, MFA has rarely been deployed successfully in very-large-scale websites intended for communities such as consumers. In the light of the increasing password attacks, practices are beginning to change.

In this article, the next three sections describe the types of authentication factors, examine the authentication solutions users want, and introduce emerging authentication systems. Then, examples of authentication implementations used in websites of well-known companies are reviewed. The last section includes the conclusions.

Types of Authentication Factors

Authentication factors can be categorized as: “What you know”, “What you have”, and “What you are”. What-you-know factors include passwords or answers to secret questions, and are by far the most commonly used of the three types. What-you-have factors are things you physically carry and must have in your possession in order to authenticate. What-you-are factors measure characteristics of your person, such as fingerprints.

Within a given type, a factor can be more or less secure, such as a password that is more or less easily guessed. But, the real increase in security comes from requiring more than one factor of different types. Two factors of the same type are not enough; the reason is that different types require an attacker to mount separate and unique attacks. Consider the case of “phishing” – a general term for emails, text messages, and websites fabricated and sent by criminals. These messages are designed to look like they come from well-known and trusted senders in an attempt to collect personal, financial, and sensitive information (Royal Canadian Mounted Police, 2010; tinyurl.com/mjqpt78). A phishing email might get your password (i.e., what you know) but cannot get your hardware token (i.e., what you have); conversely, a pickpocket might steal your token (i.e., what you have) but will not get your password (i.e., what you know).

What-you-know factors

Passwords are the most common of the what-you-know factors and are the target of much criticism. But, the death of the password has been greatly exaggerated. Even if everyone moves to MFA, a what-you-know factor in the form of a password will almost certainly be one of the factors. Moreover, even though technologists think of passwords as “old technology”, in broader consumer terms, they are not. Most consumers really only started becoming comfortable with passwords as a result of the adoption of email and online services (e.g., home banking), going back perhaps 15 years. After passwords, the next most common what-you-know factors are answers to “secret questions”, sometimes called knowledge-based authentication.

Password systems have a number of problems. Today, most users access too many distinct systems requiring passwords, leading to poor security practices such as password reuse or passwords being written down. Knowledge-based authentication suffers when the secret is not-so-secret because it is based on information about the user that is available from public sources.

The rise of social media has aggravated the knowledge-based authentication problem because facts about users that previously might have been known only to a few close friends are now online and widely shared. As a result, what-you-know systems are subject to different attack vectors (i.e., paths or means by which a hacker accesses a computer or network server in order to commit fraud). Attack vectors enable hackers to exploit system vulnerabilities, including the human element. Attack vectors that target what-you-know systems include phishing and spearphishing (Associated Press, 2013; tinyurl.com/ahjw9bd). Phishing and spearphishing messages, usually emails, appear to come from a trusted source. Phishing messages often appear to come from a large and well-known company or website with a broad membership base. In the case of spearphishing, however, the apparent source of the email is likely to be an individual within the recipient's own company, often someone in a position of authority.

Other attack vectors that target what-you-know systems include: attacks on password recovery and reset systems (Honan, 2012; tinyurl.com/c2ao8ur); malware; and server-side attacks (Ku, 2012; tinyurl.com/kh55qkb).

What-you-have factors

The most common what-you-have factors are hardware one-time-password tokens and smart cards. One-time-password tokens are small devices with a display that generate a periodically changing code. Authentication requires entry of that code (usually along with a password), so the user must be in possession of the token. A more recent variant is an application for a mobile device that replicates the function of the token, which has the advantage of using something the user is already carrying. Smart cards are credit cards with an embedded microprocessor that securely store secrets such as cryptographic keys. Authentication involves the card communicating with some other system, such as the user's personal computer or a point-of-sale system, and executing some authentication protocol. In addition to authentication, both of these choices can perform other functions such as digitally signing a transaction.

What-you-have factors are costly and inconvenient. Tokens must be purchased, inventoried, distributed, and managed. Users must remember to carry them; they can be lost, stolen, or broken. Also, backup systems for forgotten tokens are an issue. Often, these systems fall back to knowledge-based authentication, which then becomes an attack vector that bypasses the what-you-have factor.

Application variants are decreasing the cost and increasing the convenience of what-you-have factors. Tokens, however, are popular solely in enterprise deployments. Smart cards have had success in government situations that require high security or where their use can be mandated. The largest consumer smart card deployment has been the EMV credit card (tinyurl.com/3k8puz) or "Chip and PIN" card. EMV stands for Europay, MasterCard, and Visa, a global standard for authenticating credit card and debit card transactions that is widely used outside the United States. The wide adoption of EMV took many years: the first EMV standard was set in 1995. There have been a few attempts to use EMV online, however, it is almost entirely used at point-of-sale terminals. So far, there has not been a successful consumer deployment of smart cards used for online authentication.

Token theft is one possible attack for what-you-have factors. There have been some server-side attacks, such as the breach of RSA Security's keys (Rashid, 2011; tinyurl.com/kub4l8a). Targeted malware can also attack tokens and smart cards, by intercepting the one-time-password, session hijacking, or by causing the card to sign data other than what the user intended.

What-you-are factors

What-you-are factors, or biometrics, include: fingerprints, handprints, face or eye geometry, voice prints, typing patterns, and behavioural analysis. Many of these factors require some sort of sensor to measure a physical characteristic, adding to the cost and complexity of the solution. Enabling things such as facial recognition using hardware that is already in the user's hands (e.g., cellphone cameras) is one way to lower both cost and complexity.

Biometrics is very different from other authentication factor types due to false positives and false negatives. Although a password check is a binary test (i.e., it either matches or it does not), the outcome of a biometric authentication event has only a probability of correctness. There is an explicit tradeoff. Systems that are more secure will also reject more legitimate users; conversely systems that reject few legitimate users will be less secure. Some biometric products allow this tradeoff to be explicitly tuned, giving implementers the ability to set their own policy.

Possible attack vectors for what-you-are factors include replicating the physical characteristic and fooling the sensor. Although this is a common theme in movies, it is difficult to implement in real life. But it is possible. There have been demonstrations of successful attacks in popular media, such as the television show "MythBusters" (tinyurl.com/kekbbj9), in which the presenters successfully duped a thumbprint scanner. As with other factors, server-side attacks on the stored characteristic data are possible, as well as malware on the user system.

Authentication factors in online systems

In the physical world, the factors types identified above are very distinct. Imagine a door with a guard. To open the door, you must have the proper key, be recognized by the guard, and speak the correct password: three-factor security. For online systems, however, the types overlap and their distinction is somewhat fuzzy. This is because they all end up represented as data inside a computer – usually the user's personal computer and eventually some server.

One what-you-have mechanism used by some organizations is the "bingo card", which is a card printed with a matrix of short codes. During authentication, the server asks the user to enter the code from, say, row 3 and column 4. If the user memorizes the entire card, is it still a what-you-have factor? Or does it become what-you-know? Used alongside a password, is that really MFA? Attack types also can overlap: one-time-password tokens can be attacked by phishing. Approaches that use more complex protocols, such as public-key infrastructure-based smart cards, can avoid these attacks.

A real-life attack that demonstrates this overlap is what is commonly called “ATM skimming”. In a skimming attack, a device is placed over the card reader slot on an Automated Teller Machine (ATM). The device is built to appear as if it is part of the ATM, so the user does not notice its presence. As the card is inserted into the ATM, it passes through the device, which reads the magnetic stripe on the card. The device also usually includes a tiny camera, focused on the ATM keypad, to capture the user’s personal information number (PIN). The captured data might be saved within the device in memory, and retrieved later by the attacker, or it may be transmitted wirelessly to the attacker who lurks nearby. Using the captured magstripe data, the attacker can create a duplicate of the ATM card using almost any other card as a “blank” – even, for example, a hotel key card. The attacker then has a duplicate of the what-you-have factor (the card) and, with the PIN, can withdraw funds from the user’s account.

This attack is possible partially because the what-you-have factor in this case simply holds a bit of data that is read by the ATM. The ATM has no way to distinguish whether that data came from the legitimate card belonging to the user or from a copy. Data is data; inside the ATM, both factor types – what you have and what you know – look the same.

Malware

Malware on the user’s system is the bane of all factor types. It can target the authentication system directly by intercepting the data entered by the user or read by a sensor. Even for systems using cryptographic protocols, sufficiently targeted malware can hijack a session after authentication or can cause the data presented to the user, and the actual transaction being executed, to be different. In this context, "transaction" refers to any user action, including the act of authenticating or communicating to exchange an asset for payment.

The financial industry understood this problem many years ago and solved it through hardware mechanisms. Point-of-sale systems that accept credit cards and debit cards typically use a tamper-proof, integrated pad. This single device reads the card, displays the transaction information, reads the user's PIN, and contains cryptographic keys to encrypt information before it leaves the device. For security, the device depends on its physical tamper-resistance and the inability of an attacker to insert code into it. That approach will not work for general-purpose computers, although there are efforts to put secure hardware components, such as the Trusted Platform Module (tinyurl.com/on9vqcj), into personal computers.

What Do Users Want?

Given that there is a wealth of authentication mechanisms available, it is worth considering the needs and preferences of users, which highlight the tradeoff between security and convenience. Users want security, however, their willingness to accept inconvenience depends on their perception of the immediate threat. Consider the case of people who live in a neighborhood they perceive to be “safe”. They may tend to leave doors unlocked – until they hear of a nearby break-in. Then they are careful, and lock up when leaving – until time passes and complacency sets in. Even though identity theft receives a reasonable amount of attention from the press, for online systems, the threats are more esoteric and harder for non-technologists to understand. To the majority of users, technology is supposed to be convenient and “just available” – such as television, where you do not have to log in to use it.

A number of user behaviour patterns have emerged. One is for users to share credentials across many sites. By using a single password in many places, the user (even if unconsciously) is opting for convenience over security. Similarly, the selection of weak passwords is also the result of users opting for convenience over security.

Another popular pattern supported by many online services is to leave the user logged-in semi-permanently. For example, a website might require re-authentication periodically or when the user attempts a sensitive operation such as changing the password. The overall user experience is smoother because the user is required to authenticate less frequently.

To businesses, the security versus convenience tradeoff directly affects their success. Greater inconvenience risks alienating users and driving them to competitors; yet, weaker security can lead to direct monetary loss. This tradeoff is commonly resolved based on the real or perceived value of the assets the business controls. Financial websites, such as those for home banking, deal with high-value assets where real monetary loss is possible. Often, they also typically have regulatory responsibilities, so high security is important. On the other hand, businesses want an easy-to-use experience for their customers. As a result, they do not use the “always-logged-in” model and require authentication for every session. Session lifetime is limited to a short period, usually measured in minutes. However, few businesses have opted for MFA, and they typically only use it for accounts with very high value. For example, tokens may be used to provide access to corporate accounts or brokerage systems that move or trade large amounts of money.

Websites with lower-value assets opt for approaches that decrease inconvenience for the user, typically by requiring authentication only occasionally. Most browser-based email systems operate this way. A cookie set on the user’s system establishes the session when the user logs in. The cookie can be thought of as a what-you-have factor, and the act of logging in exchanges a what-you-know factor (i.e., the password) for the cookie. Social media websites have used this pattern often. However, recent incidents are changing their perception of “low-value”, and some websites are starting to implement stronger authentication.

Emerging Authentication Mechanisms

Emerging authentication mechanisms include risk analysis and use of an alternate channel. These mechanisms are helping organizations address the problem of increasing security while minimizing user inconvenience. The use of risk analysis during authentication, or when the user attempts a sensitive or high-value transaction, is one of these mechanisms.

Risk analysis focuses on the characteristics of the event – independently of the actual authentication – by searching for suspicious patterns. Comparisons can be made against historical data for the user as well as common patterns for fraudulent access. Examples of questions that drive a risk analysis include:

- What device is being used? Has this user used this device in the past? Has this device been used to commit fraud?

- Where is the user located? What time is it? Are these patterns consistent with past usage?

- Has the user moved physically in an impossible way (e.g., logged in from San Francisco, then from New York only moments later?)

- Is the transaction typical for the user? Is the user executing an unusual number of transactions?

Risk analysis is popular because it layers with other authentication mechanisms and is invisible to the user. The result of the risk analysis must be acted on, according to organizational policy. For example, transactions scored as "very risky" might be blocked. Moderate risk might trigger additional authentication, such as asking the user a security question.

Another emerging mechanism is the use of an alternate channel during authentication. This mechanism is receiving the most amount of attention because of the widespread adoption of mobile computing devices. Alternate channel involves establishing some communication between the user and the server over a path that is different than the one being used to log in. Most often, the alternate channel is the user’s mobile phone. For example, if a user logs in using a personal computer, the server might send a code using Short Message Service (SMS) to the user’s phone. SMS is a text-messaging service component of phone, web, or mobile communication systems that allows the exchange of short text messages between fixed line or mobile phone devices. To complete the login, the user must enter the code at the user’s personal computer in addition to providing a password. SMS, voice calls, push notifications, and emails are among the possible channels. The interaction can be simple or may involve a more complex sequence with the user. Transaction details might be sent to the alternate device for the user to review and approve. Quick response codes or bar codes might be used and read by the phone’s camera. Moreover, cryptographic keys and protocols can be involved.

From the perspective of factor types, this kind of authentication is difficult to characterize. Ostensibly it is what-you-have authentication because the user must be in possession of the phone. However, it really is based on ownership of the phone *number*, not the device itself. Therefore, the security of the approach actually depends on how well the phone carrier has secured the network. Similarly, email as an alternate channel depends on the security of the email account, which often depends on just a password, and so it is arguably a what-you-know factor. Alternate channels can help with the malware problem. It is possible to devise an alternate-channel system that would require the malware author to attack both devices. For example, with a single device, the malware can always take over the session after the user has authenticated, regardless of the authentication technology being used or the number of factors. Such malware might subsequently submit fraudulent transactions using that session, or modify transactions entered by the user. With an alternate channel, however, the server can send transaction details to the second channel – say, the phone – where the user could verify them. Because the malware is on only one device, the user is protected. However, that protection is lost if there is no second device – such as when the user is originating the transaction from the phone itself, as opposed to a personal computer and phone. If the malware is sufficiently “smart”, it can target whatever authentication mechanisms are being used or attack the user’s session after authentication.

Implementation

This section provides examples of the authentication mechanisms used by well-known organizations, including large organizations with large user communities – sometimes with hundreds of millions of users. Many of these organizations are seen as industry leaders, especially in terms of user experience. These examples are worth examining to understand how these organizations have tried to add authentication factors and balance the convenience–security tradeoff. Other organizations are likely to follow their lead, and their success or failure will likely have a big impact on future implementations of MFA.

The mechanism names vary – “two step” instead of “two factor” or “verification” instead of “authentication” – but, effectively, all of these examples describe forms of MFA. Also, the specific time of usage varies. For example, some organizations use MFA at every login, whereas others use it only occasionally or in special circumstances.

Financial institutions
Card associations, such as Visa and MasterCard, have a long history of security innovation. The EMV smart cards were a major advancement in physical card security and required significant investment over many years. More recently, financial institutions have addressed online fraud using systems such as 3-D Secure (tinyurl.com/38qjke), a protocol designed to be an additional security layer for online credit card and debit card transactions. The protocol ties the financial authorization process to online authentication based on a three-domain model:

1. **Acquirer domain:** the merchant and the bank to which money is being paid

2. **Issuer domain:** the bank which issued the card that is being used

3. **Interoperability domain:** the Internet or Message Passing Interface (tinyurl.com/qxwe2)

For online access, such as for home banking, many banks have implemented risk analysis systems, often in response to regulatory pressure. These systems are layered with simple passwords. Fallback systems, used when the user forgets a password or when an account is locked, often use knowledge-based authentication. Banks have an advantage over many purely online sites in that they have a physical presence (branches) and call centres that can be used for fallback. The costs of servicing users this way, however, are significant.

Google (tinyurl.com/d27xnr7)
Google implemented a system called “two-step verification” using alternate-channel authentication. In addition to a password, the user could receive a text or phone call. They also support the alternative of using one-time-password applications. Computers can be designated as trusted by the user, such that two-step verification is not required when logging in from those systems. There are multiple fallback approaches. More than one phone number can be registered. During enrollment, the user can print and save a set of backup codes to use in the event of a lost phone. Finally, if all else fails, an account recovery form can be sent to Google.

Google also has a mechanism for handling account access from mobile devices. A common problem with MFA is that users access their accounts from many devices, some of which might not support the MFA technology very well. For example, a fingerprint reader might be present on a user’s personal computer, where a driver could be loaded and the reader could be used when logging in. But, if the account requires access from an application on a phone, there may be no reader available; plus, it is unlikely that the application will support more than simple password authentication. Google allows the user to generate, on a personal computer, “application passwords" that can be used specifically by the mobile applications. Because these passwords are long-lived, this method arguably reduces the overall solution to a single-factor. However, these passwords are phishing-resistant (unlike user passwords), because they are not used regularly and are not required to be memorized by the user.

Apple (tinyurl.com/czwun9b)
Apple also uses the term “two-step verification” for their approach. Their system is based on three elements: i) a password; ii) an alternate channel with SMS and push notifications; and iii) a 14-character recovery key that is generated at setup. MFA is not used at every authentication; it is only used when the user wants to perform sensitive operations such as account management or changing a password. This method solves the problem of application access because normal authentication involves only the password. Resetting any of the three elements requires the user to have two elements. For example, to reset a forgotten password, the user must have the recovery key and be able to receive a code through the alternate channel.

One concern in the use of this two-step verification approach is that there appears to be no other fallback mechanism. If two of the factors are lost – say the user forgets the password and has lost the recovery key – Apple suggests that the user should create a new AppleID. Given that purchases are tied to the AppleID, presumably this means that the user loses access to them.

LinkedIn (tinyurl.com/k3cwqcv)
LinkedIn also calls their approach “two-step verification”. As the alternate channel, they use a code sent via SMS. Applications are handled by appending the code to a regular password and giving that to the application. This approach depends on the application remaining logged in for a long time. Fallback mechanisms seem unclear. The website’s help page has an “ask us” form that can be submitted in the event of a problem.

Twitter (tinyurl.com/paya4rj)
Twitter has implemented “login verification” in two successive steps. In the spring of 2013, they implemented alternate-channel authentication via SMS messages, with some limitations. Only one phone number was allowed per account, and only one account was allowed per phone number. Applications could be handled by generating a temporary password with a one-hour lifetime, so, as with LinkedIn, the application is usually expected to remain logged in continuously. The fallback mechanism was to contact support. There was no apparent provision for multiple users on the same account, which was a problem for corporate accounts that handle tweets from multiple employees.

During the summer of 2013, Twitter has added an additional mechanism that involves a cryptographic key that is stored on the user's phone. When logging in at a personal computer, a notification is sent to the phone. The user must approve the login using the Twitter application on the phone. The application communicates with the server using the key and a cryptographic protocol, and the login proceeds. The new mechanism also provides backup codes generated at the phone that can be used for fallback. The multiple-user problem is addressed by allowing the phone application to support multiple simultaneous accounts. Therefore, a user can be logged in to both the user’s personal and corporate accounts at the same time. Multiple users of the corporate account can be logged in, each user using his or her own phone.

Facebook (tinyurl.com/3ocrlc3)
Facebook uses a mechanism referred to as “login approvals”. Alternate-channel authentication via SMS is supported, as well as one-time-password generation in the Facebook application or via third-party applications. MFA is used only if the login device is not recognized. Fallback is supported by reset codes that the user can print in advance or by contacting support. Applications are handled by one-time application passwords that can be generated by the user.

Conclusion

Solving the online authentication problem – improving security without alienating users – is a critical and growing need. Authentication attacks are increasing every year and attackers are becoming more sophisticated. MFA will be one important tool, but it is a complex and evolving concept. Although the history of MFA goes back many years, for many online sites it is only now being applied. However, a rethinking of authentication is happening across the industry. The future of MFA will depend on how well popular sites – such as those mentioned above – implement it, and on how well users like it. No data is available yet on adoption rates. The common trend of using an alternate channel, particularly mobile devices, is likely to continue given its selection by well-known companies.

There are steps everyone can take. Businesses with online sites should implement some form of MFA. User education is also important. The adoption rate of MFA can increase by helping users understand why they need more than a simple password. Partnerships between industry, academia, and governments can help fund research into new authentication technologies and the effectiveness of existing authentication technologies.

Individual users should examine the options presented by the sites they frequent and consider enabling MFA, particularly for those services where high-value assets are involved. If MFA is not available, users should reach out and try to influence those organizations to use MFA. Often, businesses will not move to adopt MFA until after an attack; however, they can be influenced by customer demand. Given the increasing frequency of highly publicized attacks, it is better to proactively prevent them than to reactively respond.

About the Author

Jim Reno is a Distinguished Engineer and Chief Architect for Security at CA Technologies. He joined CA with the Arcot acquisition in October 2010. At Arcot, Jim led the development of strong authentication and risk management systems. He has more than 30 years' experience in software development, working on operating systems, databases, networking, systems management, and security. Jim is one of the inventors of the 3-D Secure protocol used in the Verified by Visa and MasterCard SecureCode programs. He holds multiple patents in the area of credit card verification and authentication. At CA he guides the overall architecture of CA's security products as well as security aspects of the entire CA portfolio.

Citation: Reno, J. 2013. Multifactor Authentication: Its Time Has Come. *Technology Innovation Management Review.* August 2013: 51–58.

Keywords: multifactor authentication, authentication mechanisms, online security, authentication attacks

On the Road to Holistic Decision Making in Adaptive Security

Mahsa Emami-Taba, Mehdi Amoui, and Ladan Tahvildari

> *“When you have to make a choice and don't make it, that is in itself a choice.”*
>
> William James (1842–1910)
> Philosopher and psychologist

Security is a critical concern in today's software systems. Besides the interconnectivity and dynamic nature of network systems, the increasing complexity in modern software systems amplifies the complexity of IT security. This fact leaves attackers one step ahead in exploiting vulnerabilities and introducing new cyberattacks. The demand for new methodologies in addressing cybersecurity is emphasized by both private and national corporations. A practical solution to dynamically manage the high complexity of IT security is adaptive security, which facilitates analysis of the system's behaviour and hence the prevention of malicious attacks in complex systems. Systems that feature adaptive security detect and mitigate security threats at runtime with little or no administrator involvement. In these systems, decisions at runtime are balanced according to quality and performance goals. This article describes the necessity of holistic decision making in such systems and paves the road to future research.

Introduction

Cybersecurity threats, such as Internet worms (tinyurl.com/lg2wghw), can spread too quickly for humans to respond and pose a genuine risk to users and systems. In March 2013, a computer scam fooled some Canadian Internet users by picking up their location and making it appear as though the Royal Canadian Mounted Police had frozen their screens; pop-ups demanded that users must pay a $100 fine to have their computer unlocked (CBC, 2013; tinyurl.com/lhuwq82). In the same month, a computer virus paralyzed computer networks of broadcasters and banks in a network attack in South Korea (BBC, 2013; tinyurl.com/cgustwk). The economic and national security consequences of these types of attacks are severe. The official website of the United States Department of Homeland Security (DHS; tinyurl.com/kttv9qo) indicates that the Secret Services Cyber Intelligence Section has directly contributed to the arrest of transnational cybercriminals who were responsible for the theft of hundreds of millions of credit card numbers and the loss of approximately $600 million to financial and retail institutions. The same resource indicates that, in 2011, the DHS prevented $1.5 billion in potential losses through cybercrime investigations. The distributed architecture of networks results not only in faster propagation of cyberattacks, but it also affects a greater number of vulnerable cyberdevices. For example, in 2003, the Slammer worm infected more than 90% of vulnerable hosts in 10 minutes (Moore et al., 2003; tinyurl.com/koweuj5). Traditional security models are not able to keep up with the security attacks that propagate at machine speed.

McConnell (2011; tinyurl.com/65udd87) explored the technical options to enhance cybersecurity through three major building blocks: automation, interoperability, and authentication. These building blocks provide the means to limit the spread of attacks and thus minimize consequences. McConnell introduced the concept of automated courses of action (ACOA), which encapsulates many of the complex decisions and activities involved in defending cybersystems. The concept of ACOAs is a novel step toward enabling the collective action required to protect against evolving cyberthreats. Novel decision-making approaches will enhance these courses of actions in response to cybersituations.

Automation accelerates the analysis of monitored data and perhaps increases the number of symptoms that can be detected in order to prevent a threat. Moreover, automation helps to speed up the decision-making process at the time of attack. An immediate, suboptimal response can sometimes be more effective than a later, optimal response. These timely actions prevent the spread of attack and therefore minimize the consequences of the attack. In recent years, interest in building software systems that are adaptive to their security goals has increased. Self-adaptive software (SAS) systems address automation in response to changes in the requirement and environment. SAS *monitors* itself and its context, *detects* significant changes, *decides* how to react, and *executes* such decisions (Salehie and Tahvildari, 2009; tinyurl.com/lffu25g). *Adaptive security* refers to solutions that aim to adapt their defence mechanisms at runtime. This class of SAS is called self-protecting software (SPS). SPS systems have the ability to detect security attacks and trigger countermeasures. These systems not only defend against the malicious attack but also are capable of anticipating problems and taking steps to avoid them or moderate their effects (Salehie and Tahvildari, 2009; tinyurl.com/lffu25g). In this article, we focus on the role of *automation* in cybersecurity. First, we raise awareness of the importance of addressing adaptive security from a holistic view of the system. Second, we show how game theory can contribute to decision making in adaptive security.

The rest of this article is organized as follows. The next section provides an overview of the active work on self-protecting systems. Then, we highlight the importance of creating a holistic decision-making strategy in cybersecurity, after which we discuss the use of game theory in the network and application architecture layers of the system. Finally, we conclude by describing the steps required to achieve a holistic decision-making strategy.

SPS Tools and Techniques

Projects in both academia and industry have addressed adaptivity in software systems. Table 1 lists recent research and development achievements in self-protecting software systems.

A revealing insight from this overview of tools and techniques is the absence of adaptation decision-making that captures all the possible knowledge from the software system and incorporates that knowledge in making effective adaptive decisions. In both academia and industry, SPS is still in its early years.

Holistic Decision Making in Adaptive Security

The fundamental relationship between security and decision making is highlighted by Alpcan and Başar (2010; tinyurl.com/mfvae39). Making systematic decisions, such as allocating resources while balancing risks, can benefit the system with efficient protection against both known and unknown attacks. The dynamic nature of network security requires dynamic analysis and decision making based on the monitored data. Dynamic measurements of the system metrics and states manifest dynamic changes both in the system itself and in the environment.

Figure 1 illustrates the process of acquiring data from different layers of the software's architecture through sensors. The adaptable software may contain one or more layers than are shown in this figure. Here, the rest of the layers that are not included in the software *itself* are considered as the *environment*. A *holistic decision-making strategy* considers knowledge from different layers of the system in its decision-making process. The monitored data is gathered from the sensors of the system itself and its environment. Depending on the system, some layers may not provide access for the sensors or effectors in that layer. The data gathered by sensors is transmitted through event buses to the adaptation manager, which contains the four main adaptation processes: monitor, analyze, plan, and execute. The planning process encapsulates the decision-making engine. The knowledge of the system itself and its environment is shared among the adaptation processes. Correspondingly, adaptation action is applied through effectors in various layers of the software system. The decision-making technique must embody the gathered knowledge from various sources and find the effective alternate action in the most appropriate layer of the software system. The set of adaptive security actions can be applied in more than one layer of the software system. The effectors that are responsible for performing adaptation actions reside in the layers of the system itself and its environment based on the access permission to different architecture layers.

Table 1. Notable examples of research from academia and industry relating to self-protecting systems and adaptive security

Academic Research

Author (Year)	Project/Approach	Description
Hashii et al. (2000) tinyurl.com/lkbsdbw	An extensible security infrastructure that supports fine-grained security policies	Accommodates adaptive security by dynamically modifying the policies based on the mobile code environment.
Feiertag et al. (2000) tinyurl.com/kdqqdu8)	Intrusion detection inter-component adaptive negotiation (IDIAN)	Allows intrusion-detection systems to dynamically cooperate and evolve based on the changes in the environment. The negotiation among intrusion-detection systems is facilitated by a negotiation protocol.
Scott and Davidson (2001a) tinyurl.com/n3fofdb (2001b) tinyurl.com/m3vb5ez	Strata project	Uses software dynamic translation (SDT) technology to alter code at the instruction-level. Strata can be exploited to provide adaptive security by defining dynamic and adaptive security policies.
Knight et al. (2002) tinyurl.com/kyyye9q	Willow architecture	Provides adaptive security by reconfiguration.
English et al. (2006) tinyurl.com/lwyqmbn	Trust management	Provides adaptive security by reconfiguration.
Claudel et al. (2006) tinyurl.com/kcu5veo	Application of JADE	Benefits from component-based software engineering to protect distributed systems.
Al-Nashif et al. (2008) tinyurl.com/lrfk6uh)	Multi level intrusion detection system (ML-IDS)	Detects network attacks by inspecting and analyzing the traffic using several levels of granularity.
Blount et al. (2011) tinyurl.com/k4c43r2	Adaptive rule-based malware detection	Leverages learning classifier systems to improve the accuracy of intrusion detection in detecting unknown attacks.
Pasquale et al. (2012) tinyurl.com/kgovcan	SecuriTAS	Enables software designers to model security goals and requirements of a system at the design time. The model is used at runtime to analyze and plan processes of adaptation.

Industry Research

Author (Year)	Project/Approach	Description
Burns et al. (2001) tinyurl.com/lqjp8w5	Automatic management of security policies in dynamic networks	Validates policies by models of network elements and services.
Ryutov et al. (2005) tinyurl.com/l3r6r7d	Adaptive trust negotiation and access control (ATNAC)	Uses a framework that provides adaptive access control.
Costa et al. (2005) tinyurl.com/me5ch4u	Vigilante	Provides automatic worm containment. The advantage of Vigilante is that it is not limited to network-level information about the worms.
He and Lacoste (2008) tinyurl.com/m8xxhqv)	Component-based software paradigm	Provides adaptive security in ubiquitous systems.

On the Road to Holistic Decision Making in Adaptive Security

Mahsa Emami-Taba, Mehdi Amoui, and Ladan Tahvildari

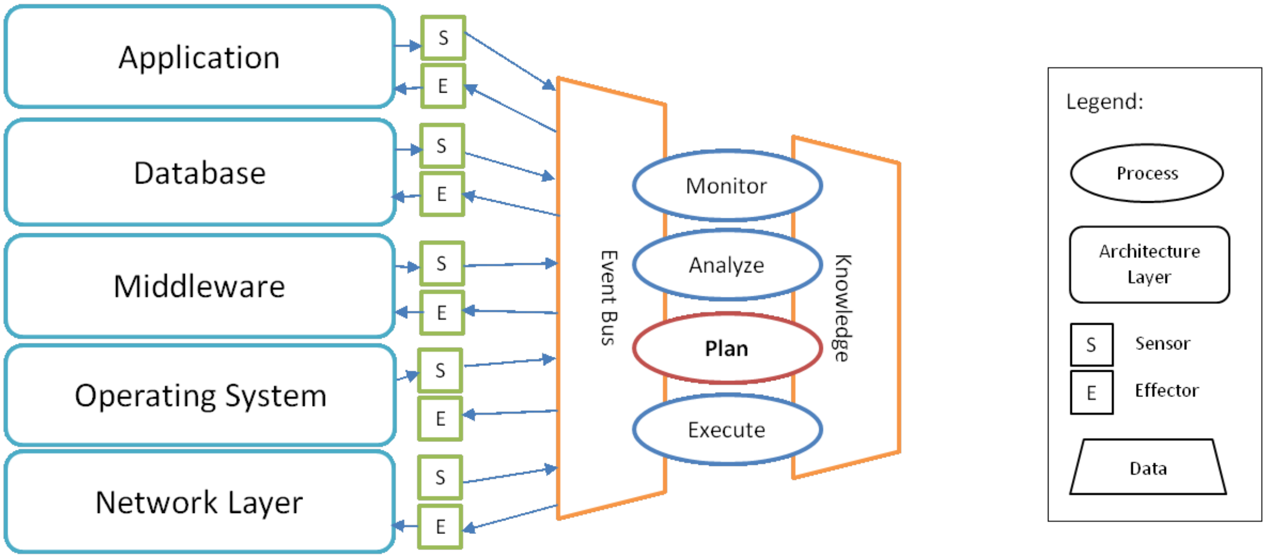


Figure 1. Decision making based on holistic shared knowledge of system layers

The need for holistic management in disciplines such as management science is explored through enterprise integration (Dalal et al., 2004; tinyurl.com/kkwvapu). Besides, vulnerability and risk management can benefit from a holistic methodology by assessing the non-linear relations of contextual parameters and the complexity and dynamics of social systems (Cardona, 2003; tinyurl.com/l3m6zdl). Recently, the idea of delivering a holistic approach to addressing cybersecurity has received greater attention. Bencomo, Belaggoun, and Issarny (2013; tinyurl.com/kr6sc56) provide a holistic view to tackle self-adaptation under uncertainty. They use the mathematical model of dynamic decision networks (DDNs) to support decision making under uncertainty for self-adaptation. An architecture-based approach in SPS systems was recently proposed by Yuan and colleagues (2013; tinyurl.com/n6ydvn7); their approach benefits from the holistic view of the systems that is provided by the software architecture.

A holistic view of the application and its environment can be completed through feedback loops. Feedback loops help to combine the result of adaptation with theoretical formulation of the problem. Developing a decision solely based on the mathematical model does not reflect the actual consequences of the decisions made. Incorporating a feedback loop in the decision-making engine helps to repeatedly observe the result of the actions made and consider its effectiveness in future alternative actions.

In summary, to achieve a holistic decision making strategy: i) security goals must be defined at each architecture layer of the system; ii) appropriate decision-making models and techniques should be applied to reduce conflicts and increase the decision quality; and iii) adaptation should not be limited to detecting and preventing attacks – adaptation must also stop the spread of the attack after it happens.

From game theory to adaptive security

A variety of mathematical theories can be used to model and analyze cybersecurity. Resource-allocation problems in network security can be formulated as *optimization problems* (Alpcan and Başar, 2010; tinyurl.com/mfvae39). In dynamic systems, *control theory* is beneficial in formulating the dynamic behaviour of the systems. In contrast, *game theory* provides rich mathematical tools and techniques to express security problems. Security games allow players (the defender and the attacker) to develop a systematic strategy based on formalized methods. In security games, players do not have access to each other’s payoffs; therefore, they observe the opponent’s behaviour and estimate the result of their action. Security games can be modelled as *non-cooperative games* in which players make decisions independently.

Due to limited resources in software systems, a practical approach is to utilize the resources and protect them against malicious attacks. Critical assets such as personal or sensitive information also require protection. Game theory provides a formal approach to maximize the effectiveness of resources against cyberthreats (Tambe, 2011; tinyurl.com/m6nwedq). From simple deterministic games to more complex stochastic games, security games can be used to model security in intrusion-detection systems and social, wireless, and vehicular networks (Alpcan and Başar, 2010; tinyurl.com/mfvae39).

The analytical foundation of game theory can be applied to security problems at various architecture layers of the system. For example, intrusion detection is a defence mechanism at the network layer. Intrusion-detection systems can take adaptive actions such as intensifying monitoring efforts when malicious behaviour is detected. In the remainder of this section, we look at the applicability of game theory in two architecture levels: the network layer and the application layer.

Security games at the network level
Network security is a strategic game between the malicious attacker and the administrator (Alpcan and Başar, 2010; tinyurl.com/mfvae39). In a simple intrusion-detection game, the attacker chooses between the alternative actions of attacking or non-attacking. Due to limited resources by the systems and the fact that monitoring and analyzing the monitored data adds overhead to the system, the system has the option to continue the default monitoring or to intensify monitoring. This simple formulation can be extended in complex cases such as stochastic games or games with limited information, which are discussed in greater detail by Alpcan and Başar (2010; tinyurl.com/mfvae39). After distinguishing the alternative actions by each player, the next step is to associate the payoff for each action. Based on the decision strategy, players select the alternative that yields a better payoff. Similar modelling can be applied to intrusion-prevention systems and efforts to prevent denial-of-service attacks. In the latter case, the alternative actions of the attacker could be changing the rate of data generation in the network. Meanwhile, the system's alternative actions are: i) checking the rate of congestion and ii) modifying the refresh interval. After identifying the main components of the game theory (i.e., players, the set of alternative actions, and the payoffs), the more appropriate type of game can be selected based on the availability of data. For example, if complete knowledge of the adversary payoffs is available, *repeated complete-information games* can be exploited in modelling.

Security games at the application level
Existing cybersecurity approaches based on game theory are mostly focused on providing security at the network level. The mathematical foundation of game theory can also be applicable to security at a variety of architecture levels such as the database or operating system. Here, we discuss the applicability of game theory in providing security at the application level. Depending on the architecture layer, the source of the data to be monitored is different. To detect a cyberattack at the network level, the data to be monitored can be packet data, network traffic, etc. At the application level, a cyberattack can be detected from various data sources. For example, the system can monitor the number of transactions by a specific user or the access rights of a user over a specific window of time. Even though the nature of the monitored data may vary, the problem can still be modelled as a non-cooperative game. The alternative set of actions includes more high-level actions that should align with the system's specified policies. As an example, the dynamic change to the access rights of a user should satisfy the pre and post conditions specified in the IT policy.

Previous approaches, such as those used by Alpcan and Başar (2010; tinyurl.com/mfvae39), only apply game theory at one layer of the system. To provide a holistic approach in making decisions at runtime using game theory, defining the set of alternative actions that can be taken by both players should not be limited to actions in only one layer of the systems. The same requirement applies to the data gathered by sensors in various architecture layers.

Conclusion

This article presents a brief overview on adaptive security and existing tools and techniques for SPS, and it introduces a visionary approach in holistic decision making to achieve adaptivity in cybersecurity. It provides insights into the use of game theory as a decision-making strategy that can be applied in different architecture levels. A proper decision-making strategy not only helps to model security goals and actions at runtime, but it also enables systematic decision making after the attack happens and it consequently limits the spread of attack in distributed systems.

On the Road to Holistic Decision Making in Adaptive Security

Mahsa Emami-Taba, Mehdi Amoui, and Ladan Tahvildari

Acknowledgements

The authors would like to express their appreciation for the generous financial support received from the Communications Security Establishment Canada (CSEC) on this project. Special thanks to Sharon Liff, D'Arcy J. Walsh and Daniel H. Craigen for their tremendous technical support. We also thank the reviewers for their valuable feedback.

About the Authors

Mahsa Emami-Taba received her BEng degree in Computer Engineering from Shahid Beheshty University, Iran, in 2005. She received her MMath degree in Computer Science from the University of Waterloo, Canada, in 2009. After completing her studies, she worked as a software designer and developer. She is currently working toward a PhD degree in the Department of Electrical and Computer Engineering at the University of Waterloo. Her research interests include self-adaptive software systems, adaptive security, and nature-inspired adaptive software.

Mehdi Amoui is a Postdoctoral Fellow at the University of Waterloo, Canada. He currently works as a researcher/consultant on a joint research project with the Software Verification and Validation team at Blackberry Inc., Canada. In 2002, he received his PhD from the University of Waterloo on the topic of an evolving software system for self-adaptation, and in 2006, he received an MASc degree in Artificial Intelligence and Robotics from the University of Tehran. His main research interests include self-adaptive software systems, search-based software engineering, software evolution, and software quality.

Ladan Tahvildari is an Associate Professor in the Department of Electrical and Computer Engineering at the University of Waterloo, Canada, and she is the founder of the Software Technologies Applied Research (STAR) Laboratory. Together with her research team, she investigates methods, models, architectures, and techniques to develop higher-quality software systems in a cost-effective manner. Her research accomplishments have been recognized by various awards, including the prestigious Ontario Early Researcher Award, which recognized her work in self-adaptive software. She is a Senior Member of the IEEE, a member of the ACM, and a Professional Engineer (PEng).

Citation: Emami-Tabla, M., M. Amoui., and L. Tahvildari. 2013. On the Road to Holistic Decision Making in Adaptive Security. *Technology Innovation Management Review.* August 2013: 59–64.

Keywords: cybersecurity, cyberattacks, adaptive security, holistic decision making, self-adaptive software, self-protecting software, automation, game theory, architecture

Servitization in a Security Business: Changing the Logic of Value Creation

Arto Rajala, Mika Westerlund, Mervi Murtonen, and Kim Starck

> *Security technology and security services – can we separate these two? Can you have one without the other? I cannot figure out how.*
>
> A manager in our case company

How can a firm change its value-creation logic from providing technology to selling technology-based services? This is a question many security companies face today when trying to apply a solutions-based business model in response to recent macro- and microeconomic trends. The fact that customers increasingly demand security as a service, rather than technical equipment, challenges the basis of a security firm's value provision and alters the logic of its operation. In this article, we investigate a technology- and product-oriented security business that is now rapidly transforming into a service business. We use data from a case study to propose a 4C model (conceptualization, calculation, communication, and co-creation of value) that can help security providers to objectify their service offerings and succeed in the servitization of their security businesses.

Introduction

We are living in an era that is characterized by the increasing importance of the service economy, as predicted by Vandermerwe and Rada (1988; tinyurl.com/n4fjfn5) 25 years ago. Accordingly, more and more companies are confronting the challenge of shifting from selling products to providing services (Grönroos and Ravald, 2011; tinyurl.com/l8b59lt). Vargo and Lusch (2004; tinyurl.com/cuzndc) describe this shift as moving away from the goods-dominant logic to the service-dominant logic. The shift is also known as "servitization" (Vandermerwe and Rada, 1988; tinyurl.com/n4fjfn5), which means that a physical product is no longer the basis of exchange, and the process of value creation that translates business strategies into value to customers and suppliers is changing dramatically (Fischer et al., 2012; tinyurl.com/kx9qkm7). With services, the customer is seen as the creator of value and the supplier helps them to achieve the desired outcome in the value-creation process (Grönroos, 2011; tinyurl.com/mct9mcu). Servitization has been the trend in manufacturing industries that face increasing pressures to renew business practices, but now even sectors with service traditions are striving to better understand how to define and conceptualize the value that customers perceive.

Servitization is also increasingly occurring in the private security sector, where the rapid development of technology had previously encouraged companies to focus on security products and technologies (cf. Lucintel, 2013; tinyurl.com/lrnc9gs). However, selling security equipment such as digital security products offers little room for specialization and differentiation in today’s market. Many security providers are responding to this challenge by developing new service-based business models, but this change may not be straightforward. Servitization suggests that an increasing focus on services rather than products requires new approaches, skills, and mindsets that were previously unknown to many security providers. Security companies need a better understanding of the new service-business logic, including the formation of customer value and the relevance of security services to the customer. Thus, among service practitioners and researchers, there is a growing interest in the topic of customer value (Smith and Colgate, 2007; tinyurl.com/k479dc7).

Servitization in a Security Business: Changing the Logic of Value Creation

Arto Rajala, Mika Westerlund, Mervi Murtonen, and Kim Starck

In this article, we investigate how servitization is manifested in business-to-business security services. In particular, we discuss how a security system and solutions company can use objectification to provide technology-based services to its customers and how customers perceive the benefits of these services. We provide some recommendations for service providers to better comply with the service-oriented mindset and implement objectification into their business models. Our empirical study is based on a research project that took place in the Finnish Security Sector from 2009 to 2012. Here, we discuss servitization in Niscayah's security business based on an analysis of interviews with 10 managers, an investigation of the company's marketing material, and interviews with five of their long-term customers.

Servitization and Customer Value Creation

Servitization brings the concept of customer value to the forefront. Traditionally, value in business-to-business exchange refers to monetary or non-monetary benefits and sacrifices perceived by customers in terms of their expectations, needs, and desires (Lapierre, 2000; tinyurl.com/nxrd27w). However, the service perspective means that value can only exist when an offering is used (i.e., value-in-use), and the experience and perception of use are essential for the customer (Vargo and Lusch, 2008; tinyurl.com/myn8efl). In other words, value from using the service comes from the ability to act in a manner that is beneficial to the user. Value is subjective and always determined by the beneficiary that is the co-creator of value (Lusch et al., 2007; tinyurl.com/blazss).

Customer value creation in services is not like product-based customer value creation. Therefore, companies need to reform their mindsets concerning the value-creation logic when providing services (Heinonen et al., 2010; tinyurl.com/jwq224j). First, they need to recognize customers as co-producers and maximize customer involvement in the service development. These service providers can then expand the markets by assisting customers in focusing on each customer's core business. Tangible goods may only serve as platforms for service provision, thus providers can retain the ownership of goods and earn by charging a fee based on the extent of use (Vargo and Lusch, 2004; tinyurl.com/cuzndc). Given that servitization is driven by the changing customer needs, providers need to carefully analyze what benefits customers are looking for to better understand the value perceptions of customers (Matthyssens and Vandenbempt, 2008; tinyurl.com/m4xjq3u). This understanding is even more challenging in business markets where the ultimate customer value can only be improved by increasing the value of the market offerings of intermediaries (Ulaga, 2003; tinyurl.com/c77vpud).

Providers of services need to recognize whether they are supporting their customer's core business or merely taking care of the customer's outsourced routine activities when conceptualizing offerings. Thus, a service business comprises both services that support products and services that support customer actions (Mathieu, 2001; tinyurl.com/kxzbcfs). Actually, value for the customer emerges from the whole spectrum of provider–customer interactions that support the use of core resources rather than from one source (Grönroos, 2011; tinyurl.com/mct9mcu). For example, product-lifecycle services, such as inspection of an automated teller machine, facilitate the customer's access to the provider's product and ensure its proper functioning over every stage of the lifecycle. In contrast, asset-efficiency services, such as pre-emptive maintenance and remote monitoring of manufacturing gear, strive to achieve productivity gains from assets invested by customers. Moreover, process-support services such as security consulting assist customers in improving their own core business processes. Finally, process-delegation services, such as cybersecurity incident response, carry out processes on behalf of the customers (Ulaga and Reinartz, 2011; tinyurl.com/murteex).

Servitization needs to be complemented by objectification. Whereas servitization means the customization of offerings, objectification concerns packaging and making services more tangible (Lindberg and Nordin, 2008; tinyurl.com/kypbmpw). At best, these two logics exist simultaneously and successful firms combine them by delineating distinct products, services, and processes (Sundbo, 2002; tinyurl.com/lpcl5kx). We refer to Ulaga and Reinartz's (2011; tinyurl.com/murteex) notion that hybrid offerings can help companies manage the balance between the servitization and the objectification. Based on these views, we conclude that the objectification of technology-based services presupposes a change towards a service-related mindset. However, there is limited understanding of what objectification of services really means and how companies can use it to respond to the challenges of servitization. In the following section, we introduce a case study designed to improve our understanding of objectification and to help security providers objectify their service offerings.

Methods

This empirical study is based on qualitative research and comprises multiple data sets. We interviewed 10 managers at Niscayah, a security service provider in Finland, to gather their views and perceptions on customer value, as well as their intentions to respond to servitization in the security markets. Now being part of Stanley Security Solutions (stanleysecuritysolutions.com), a division of Stanley Black & Decker Corporation, Niscayah has a strong foothold in the global security service market. It is an integrator and supplier of access control, intruder alarms, fire alarms, and video surveillance solutions. In Finland, the company operates in multiple locations and has over 250 employees, positioning itself as a market leader. At the time of the data gathering (2009–2012), its annual sales exceeded 35 million euros. Its main customer segments are retail, healthcare, transportation and logistics, insurance and finance, energy industry, and manufacturing. Niscayah pursues global market reach, strong customer orientation, comprehensiveness in offerings, and extensive field experience.

The selection of the interviewed managers was based on referral sampling, where the contact person at Niscayah identified the suitable managers for the interviews. The main selection criteria were involvement and experience in the development and delivery of security services. We examined the company's marketing material (e.g., brochures, leaflets, customer magazines, and web pages) to analyze how it communicates the value Niscayah is providing to the customers. Finally, we interviewed five of Niscayah's long-time customers to examine perceived value and benefits of acquiring security services. The interviewed customers were nominated by the contact person at Niscayah and included large Finnish enterprises representing pharmaceutical, diagnostics, telecommunications, forestry, and metal industries. We interviewed the managers at Niscayah in 2009 and 2010, and the interviews with customers took place at the end of 2011. With this schedule, we were able to examine Niscayah's intentions to provide value and serve customers through security services and then evaluate their customers' perceptions of how well the company succeeded in doing so.

Empirical Findings

The interviews with Niscayah's managers and customers clustered around four themes that reveal how the servitization and objectification are addressed in business-to-business security services. These four themes are conceptualization, calculation, communication, and co-creation of value. In the following subsections, these four themes are discussed in detail both from the managers' and the customers' perspectives.

Conceptualization

The interviewed managers emphasized the importance of service conceptualization, meaning that the benefits from using the company's services should be objectified as concrete and usable offerings. All interviewees perceived that the company is a forerunner in the development of technology-based security services, and that service concepts are the way forward. According to the managers, technology is at the core of services that are actually designed on the basis of products: *"We are able to combine technical security and the national maintenance network in a way that we can help the customer throughout the whole lifetime of the system. For this purpose we have developed the so-called 'one-stop shop' principle where we can maintain, use, and monitor security costs efficiently. This also includes remote control 24 hours 7 days a week. None of our competitors is able to offer this kind of service."*

Clearly defined and packaged services are seen as a cutting edge in the highly competitive security market and are prerequisites for a market-oriented security service offering. Niscayah's extensive product and service repertoire forms a solid basis for novel offerings, all of which include technology-assisted services. Consequently, their new services can be based on customer segmentation, where specific customer needs and requests are identified and a purposeful product/service mix is selected for each segment. Managers also felt that service conceptualization will lead to better service quality; it involves a guarantee of full-functioning security systems with fixed costs, periodic reviews, and feedback.

At the time of the interviews, Niscayah's service conceptualization was at an early stage. Therefore, managers frequently discussed priorities and the interfaces between products and services, as well as the balance between customer-tailored services and industry-specific services. One of the managers said: *"This means we also have high-quality products, albeit I should not be talking about the products at all… but we are selling products anyway."* They also expressed concerns that, as a result of the service conceptualization process, they will have fewer personal interactions with customers. This is because conceptualization requires the alignment of business units and service activities, as well as ensuring a uniform quality in the different geographical locations.

From the customers' perspective, conceptualization in security services is reflected by how well the offered services are perceived to fulfill the customers' security needs. However, these needs often relate to basic types of security such as access control systems. The interviews revealed that many of the customers are not aware or interested in more sophisticated or comprehensive security services. Some customers were considering buying or outsourcing more security services and concentrating those purchases with preferred providers: *"These [security] markets are changing all the time, but we could buy other services from Niscayah as well … for example, we have some 600 cameras out there, and why not, when renewing them, buy the whole system from Niscayah?"*

Customers were also aware of the value they receive for the money they invested in services. In many cases, they only achieved the minimum level of security required by the law or regulations. However, the interviewed customers were predominantly happy about what they were provided. As one customer put it: *"When a company spends a certain amount of euros [on security services], it receives an equivalent quality of services… however, what I have ordered has worked well."* On the other hand, another customer said: *"If the money was not a bottleneck, we would make things differently, for example, [we would] co-develop more sophisticated security systems based on RFID [radio-frequency identification] technology together with the service provider."*

Calculation
Niscayah managers considered the components of customer value from different perspectives and identified several mechanisms through which they are able to create value for the customer. These included releasing customers from the security control activities and responsibilities, enabling customer's core business functions, and cutting operational costs and crime-related costs. The interviewed managers highlighted that they need to understand the customer' core business and know the stakeholders and the business environment to be able to identify the right value drivers for each customer. These value drivers are industry- and company-specific, and therefore are difficult to identify. The managers also said that the customers are often not aware of their security needs and what the provider's security services are worth.

Although the customers' value drivers are acknowledged among the security-service managers we interviewed, calculations related to the benefits of using the company's security services are still lacking. The managers said that they need to illustrate the value of their services in monetary terms, but by the same token admitted that there are many aspects of the security service that cannot be quantified. In a security context, in which uncertainties and unforeseeable events are particularly inherent contingencies, services may be associated with a variety of negative consequences, and security as the content of a service is perceived subjectively. Therefore, reliable value estimates are difficult to calculate, even though some quantifiable measures can assure customers of the value of security services. Many of these measures are related to service quality and include security-system availability rates and response times to calls and alarms.

For the customers, value calculations refer to the price of the security services. Customers anticipate that striving for lower service prices means narrowing the scope and lowering the quality of the security they will receive. However, most of the interviewed customers perceived that some security services had become less expensive due to the technological development. One customer commented that *"Niscayah is more expensive than its competitors, but we will not change the service provider just because of the price, because proper security services cannot be provided on the cheap."* In contrast, most interviewees pointed out that security services provided by a professional security firm are more reliable, safeguard the continuity of customer's business, and increase the customer's credibility in the eyes of its customers. In addition, training in security was appreciated as a benefit. Surprisingly, we found that some of the customers did not calculate lifetime costs of the service (e.g., maintenance costs) when making purchase decisions for security services.

Security services also play a critical role from point of view of the customer's business. One of the customers said: *"Security systems are a part of our quality system, and thus support our business operations. In fact they are a kind of a concealed benefit whose value is realized when something happens."* Moreover, in some industry sectors, such as pharmacy, security procedures are highly regulated (e.g., access-control requirements). Therefore, external security professionals are needed. One of the interviewees said: *"When we started to export our products to the U.S., their border control and customs detachment visited us and inspected our security systems."* Security-service providers were perceived to provide invaluable help and up-to-date information in such cases.

Communication
The managers noted that Niscayah is a service company whose main business mantras include focusing on customer relationships. As one manager noted: *"Vis-*

iting our customers means we will nurture that particular customer relationship, not maintain equipment and devices." All the interviewed managers emphasized close, long-term partnerships with customers and open communication towards customers. Open communication is achieved through continuous and frequent customer encounters and good interpersonal relationships. Nevertheless, the supplier-customer dialogue focuses on relationship management instead of customer value, and several managers explained how difficult it is for them to start a proper discussion with their customers on the value of security services.

In its marketing communications, Niscayah primarily emphasizes the provision of security in general, and only secondarily explains about its security services. In practice, this means that the company sells its solutions by describing the security benefits customers gain when using security services, particularly those solutions provided by Niscayah. Consequently, customers can focus on their core business, save resources, and reduce costs. Niscayah's marketing material suggests that the company openly communicates its mission, vision, and values. Communication seems to be rather consistent throughout all channels. This consistent communication supports the company's aims to create a unified corporate image, bring them closer to customers, and assure customers about their intentions.

Customers were mostly satisfied with Niscayah's communications. Most of the customers had long relationships with Niscayah, which affected the way communication was carried out and perceived. Common methods included phone calls and emails, but we identified two broad types of communication. First, contact at the operative level takes place when something happens or there is a need for professional help and problem solving. Second, another type of communication comprises keeping in touch with the contact person(s) at Niscayah to get information about new security-related issues and possible re-evaluation or changes in the service provision. This communication is related to customer relationship management on a regular basis.

Interestingly, one of the interviewees hoped that the service provider would not contact them proactively. *"The security manager easily gets the information (s)he needs about the security service providers – even too much information… sometimes I have to say them: No, don't contact me, I'll be in contact with you if needed."* This might reflect that some service providers' representatives are too keen to be in contact with their customers. On the other hand, many interviewees expressed that the service provider's representative should visit them personally at least once a year and inform customers about new security products and services and whether the customer should update their security systems.

Co-creation

According to Niscayah's marketing material, the company is branded as a system integrator that provides total solutions. The material suggests that customers require a more proactive approach and better understanding of suitable business security strategies from their security suppliers. Consequently, the interviewed managers explained that they work closely with customers to solve their problems using Niscayah's accumulated expertise and doing whatever is required to find a solution. In addition, managers highly appreciate long-term customer relationships and strive toward partnerships with their key customers. In other words, the managers displayed strong customer orientation and clear intentions to co-create value with the customers. At the same time, the Niscayah managers viewed their role as providing external expertise to the customer company, not working *with* the customer but rather working *for* the customer to resolve the security issues.

Based on the interviews with customers, co-creation in security services is not yet extensive. Although customers highly value long-lasting relationships with their service provider, service development still lacks deep collaboration. As one of the interviewees put it: *"Niscayah is a service provider but I would not talk about partnership, because we know what we want and they will deliver it to us."* Another customer explained that *"Co-creation requires a lot of resources and is risky; failure would be horrific for us."* However, one customer mentioned that the relationship with Niscayah has developed remarkably towards a true partnership. They have had a myriad of different security systems developed by Niscayah *"as a system supplier, and this mode has deepened throughout the collaboration. Currently, it is truly reciprocal and mutual."*

The 4C Model of Objectification

Figure 1 summarizes the empirical findings of our study. Both managers' intentions and customers' perceptions of objectification centre around four main themes: conceptualization, calculation, communication, and co-creation of value. The views of both man-

Servitization in a Security Business: Changing the Logic of Value Creation

Arto Rajala, Mika Westerlund, Mervi Murtonen, and Kim Starck

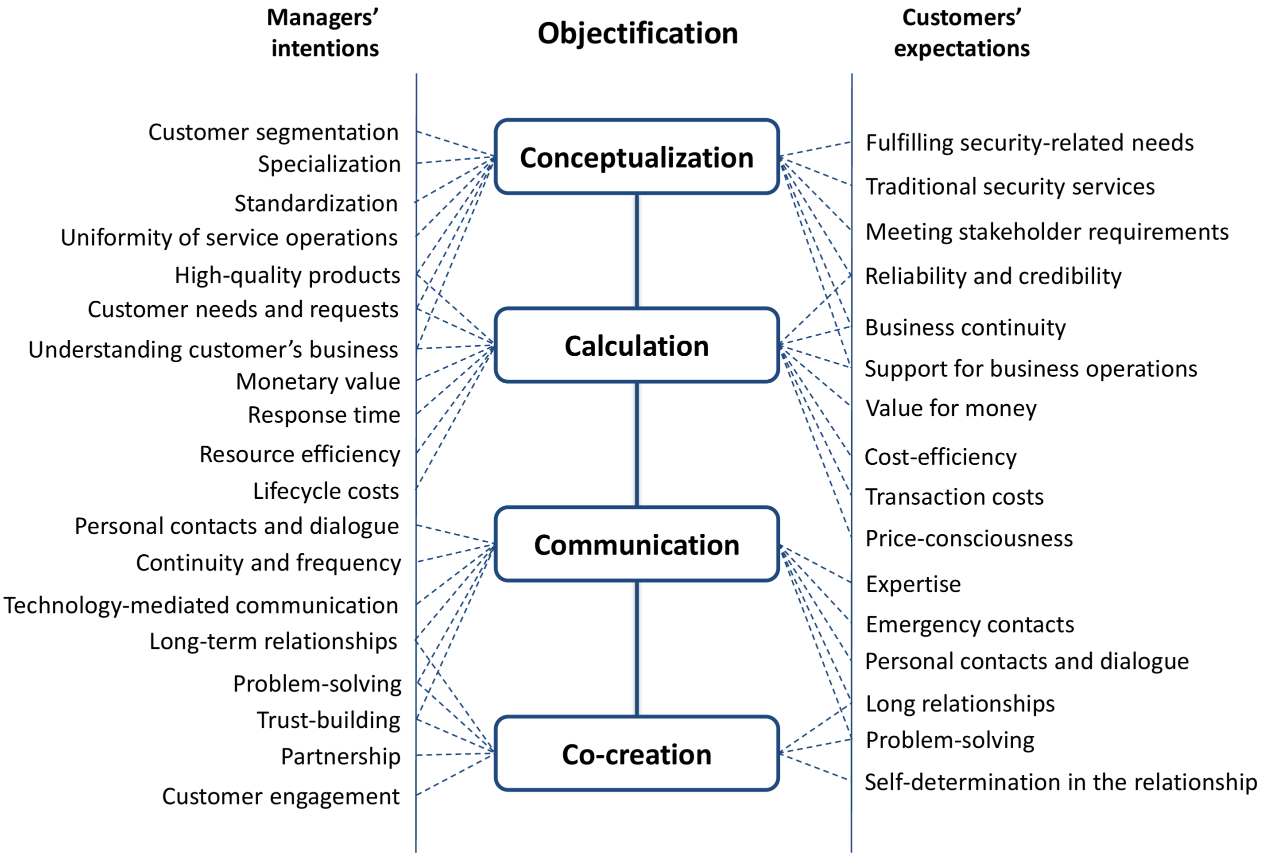


Figure 1. The 4C model of objectification, illustrating service providers' intentions and customers' expectations concerning objectification in business-to-business security services

agers and customers are well aligned with regard to service concepts, value estimates, and value communication, but are contradictory with regard to value co-creation. Managers indicated their intentions towards closer and more co-operative relationships with customers, but there is little evidence of successful co-creation. Customer interviews support this notion by suggesting that value co-creation in security services is still limited. Therefore, security suppliers need to consider how to motivate customers for more co-operative service delivery. Customers should consider how to better utilize the resources and competences of security suppliers. Grönroos (2011; tinyurl.com/kzv22gf) argues that value creation and co-creation are distinct processes, and that a customer creates value for itself, whereas the service supplier can only facilitate the customer's value-creation process.

Conclusion

This article focused on the changing logic of value creation in servitization. In the private security sector, this suggests a movement towards more specialized, more customized, and increasingly technology-based security services, such as the design of complex yet interoperable alarm and surveillance systems or security training. Our case study of Niscayah, a security-service provider in Finland, illustrated that, even if a security firm has a fair understanding of their customers' needs, and despite that they are able to communicate key benefits of security solutions in their marketing communications, customers do not understand the costs and benefits of the total security solution, and may fail to see the value of deep provider-customer collaboration.

Therefore, security-service providers should adopt a mindset that promotes deeper relationships with customers, and they should focus on techniques that help them to objectify their service offerings to make the value and benefits more tangible. After reviewing our findings, we proposed the 4C (conceptualization, calculation, communication, and co-creation of value) model of objectification that illustrates the alignment or mismatch of manager’s intentions and customer’s perceptions on provider’s security services. The model can help security providers to objectify their service offerings and succeed in the ongoing servitization of their security businesses. Furthermore, our interviews with customers and the analysis of marketing material brought about some practical suggestions for Niscayah and other security providers to support their service-objectification efforts:

1. **Assign a personal contact to each customer**: Customers value personal service and continuity; therefore, the service provider should assign a representative to each customer – preferably one that does not change roles too often. Should problems arise, customers perceive that they will more quickly receive help if they have a named, personal contact in the firm that provides their security service. A personal contact knows the customer account and, consequently, has all the relevant background information required to quickly solve a problem.

2. **Become a more proactive partner:** There is a demand for more comprehensive security services; however a customer's budgetary constraints and strict focus on their core businesses may limit their view of potential new ideas and may prevent such services from emerging. Many customers are interested in strengthening their relationship with the security-service provider and look for all-inclusive services, but they expect the service provider to be the initiator and assign dedicated people to initiate and coordinate such projects.

3. **Shorten response times:** For a customer, solving an acute problem that affects their business is of utmost importance, and this is where the capability of a security provider is measured. Customers value service providers based on this capability and perceive that security service providers should respond immediately when customers face security problems or in the event of false alarms.

4. **Put the fundamentals in place:** Several customers suggested ideas on how to avoid delays in implementing security-service systems. Customers indicated that they are unable to discuss individual objectives and needs until standard features and issues are solved. Therefore, security-service providers should ensure that their fundamentals are in good shape before promising anything about the implementation schedule or service features.

5. **Develop superb marketing materials:** Security companies need to take marketing communications seriously. They have to ensure that marketing materials clearly communicate the value of using security solutions, focus on enhancing the customer’s business instead of focusing on product attributes, provide a consistent description of the provider’s security-service offerings, and sharpen the positioning of the security company against its competitors.

Servitization in a Security Business: Changing the Logic of Value Creation

Arto Rajala, Mika Westerlund, Mervi Murtonen, and Kim Starck

About the Authors

Arto Rajala, D.Sc. (Econ.) is a Senior Researcher in the School of Business at Aalto University in Finland. He earned his doctoral degree in Marketing from the Helsinki School of Economics. Arto's current research interests include business networks, business marketing, business-to-business service development, and innovation ecosystems.

Mika Westerlund, D. Sc. (Econ.) is an Assistant Professor at Carleton University's Sprott School of Business in Ottawa, Canada. He previously held positions as a Postdoctoral Scholar in the Haas School of Business at the University of California Berkeley and in the School of Economics at Aalto University. Mika earned his doctoral degree in Marketing from the Helsinki School of Economics. His current research interests include open innovation, business strategy, and management models in high-tech and service-intensive industries.

Mervi Murtonen is a senior scientist at VTT Technical Research Centre of Finland. Her research interests include risk assessment practices, security management systems and contracted security services. Mervi holds an MSc degree in Electrical Engineering from Tampere University of Technology, Finland. Currently, she is finalizing her doctoral thesis on supplier-perceived customer value in business-to-business security services.

Kim Starck is a Sales and Security Director at Stanley Security Finland. He has strong experience in sales, sales management, as well as security and quality management. Kim has broad understanding of business operations and operations management, and he holds a Professional Master of Security (MBA) degree from Aalto University, Finland. He has been actively involved in process and solution development at Stanley Security.

Citation: Rajala, A., M. Westerlund, M. Murtonen, and K. Starck. 2013. Servitization in a Security Business: Changing the Logic of Value Creation. *Technology Innovation Management Review*. August 2013: 65–72.

Keywords: servitization, security, products, services, value provision, co-creation, customer value creation, objectification

Author Guidelines

These guidelines should assist in the process of translating your expertise into a focused article that adds to the knowledge resources available through the *Technology Innovation Management Review*. Prior to writing an article, we recommend that you contact the Editor to discuss your article topic, the author guidelines, upcoming editorial themes, and the submission process: timreview.ca/contact

Topic

Start by asking yourself:

- Does my research or experience provide any new insights or perspectives?
- Do I often find myself having to explain this topic when I meet people as they are unaware of its relevance?
- Do I believe that I could have saved myself time, money, and frustration if someone had explained to me the issues surrounding this topic?
- Am I constantly correcting misconceptions regarding this topic?
- Am I considered to be an expert in this field? For example, do I present my research or experience at conferences?

If your answer is "yes" to any of these questions, your topic is likely of interest to readers of the TIM Review.

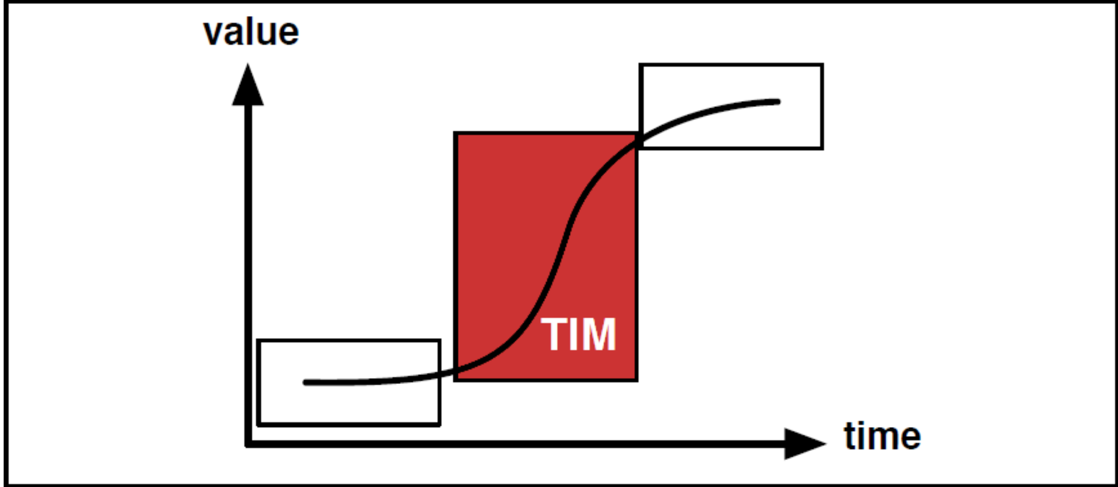
When writing your article, keep the following points in mind:

- Emphasize the practical application of your insights or research.
- Thoroughly examine the topic; don't leave the reader wishing for more.
- Know your central theme and stick to it.
- Demonstrate your depth of understanding for the topic, and that you have considered its benefits, possible outcomes, and applicability.
- Write in a formal, analytical style. Third-person voice is recommended; first-person voice may also be acceptable depending on the perspective of your article.

Format

1. Use an article template: .doc .odt
2. Indicate if your submission has been previously published elsewhere. This is to ensure that we don’t infringe upon another publisher's copyright policy.
3. Do not send articles shorter than 1500 words or longer than 3000 words.
4. Begin with a thought-provoking quotation that matches the spirit of the article. Research the source of your quotation in order to provide proper attribution.
5. Include a 2-3 paragraph abstract that provides the key messages you will be presenting in the article.
6. Only the essential references should be included. The URL to an online reference is preferred; where no online reference exists, include the name of the person and the full title of the article or book containing the referenced text. If the reference is from a personal communication, ensure that you have permission to use the quote and include a comment to that effect.
7. Provide a 2-3 paragraph conclusion that summarizes the article's main points and leaves the reader with the most important messages.
8. Include a 75-150 word biography.
9. If there are any additional texts that would be of interest to readers, include their full title and location URL.
10. Include 5 keywords for the article's metadata to assist search engines in finding your article.
11. Include any figures at the appropriate locations in the article, but also send separate graphic files at maximum resolution available for each figure.

Issue Sponsor



TIM is a unique Master's program for innovative engineers that focuses on creating wealth at the early stages of company or opportunity life cycles. It is offered by Carleton University's Institute for Technology Entrepreneurship and Commercialization. The program provides benefits to aspiring entrepreneurs, employees seeking more senior leadership roles in their companies, and engineers building credentials and expertise for their next career move.

Carleton
UNIVERSITY