Doctrine for Cybersecurity*

Deirdre K. Mulligan† Fred B. Schneider‡

May 15, 2011

Abstract

A succession of doctrines have been advocated in the past for enhancing cybersecurity: prevention, risk management, and deterrence through accountability. None has proved effective, and their failings are discussed. Proposals are now being made to view cybersecurity as a public good or to adopt mechanisms inspired by those used for public health. This landscape is surveyed through the lens that a new doctrine—public cybersecurity—provides.

1 Introduction

Governments, business, and individuals are growing increasingly concerned about the security of networked computing systems. This concern is justified. Press reports of successful attacks grow ever more frequent: cross-site scripting used to pilfer consumers' passwords, large-scale breaches of corporate customers' personal information, massive distributed denial of service attacks on web sites, cyber espionage aimed at classified documents, and all manner of attacks on civil critical infrastructures.

Computer Scientists and those who fund them are, consequently, investing heavily in technological means for improving cybersecurity. But technological solutions are useless if they are not deployed or if operating

*Supported in part by AFOSR grant F9550-06-0019, National Science Foundation grants 0430161, 0964409, CNS-0524745 (ACCURATE), and CCF-0424422 (TRUST), ONR grants N00014-01-1-0968 and N00014-09-1-0652, and a grant from Microsoft. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of these organizations or the U.S. Government.

†School of Information, University of California, Berkeley 102 South Hall, Berkeley, CA 94720–4600.

‡Computer Science Department, Cornell University, Ithaca, New York 14852.

practices allow them to be circumvented by attackers. Policy must create incentives for system developers, operators, and users to act in ways that enhance rather than weaken system security. Moreover, neither technologists nor policy makers have the luxury of starting with a clean slate. All must labor in the shadows of legacy networks and end-systems that are not secure (nor easily made so) and in the context of extant policy that reflects societal values from a time when dependence on networked information systems was minimal.

Enhanced levels of cybersecurity can entail tensions involving cost, function, convenience, and societal values such as openness, privacy, freedom of expression, and innovation. Absent some widely accepted doctrine, evaluation of proposals for improvement is difficult and debate about their adoption can be neither compelling nor conclusive. The utility of a doctrine is thus determined by the extent to which it offers a framework for resolving these tensions but does not impose, ignore, or rule out possible technical or policy solutions.

We thus conclude that a prerequisite for achieving enhanced cybersecurity is articulating a *cybersecurity doctrine*, which specifies *goals* and *means*.

- **Goals.** These define some agreed upon kinds and levels of cybersecurity, characterizing who is to be secured, at what costs (monetary, technical, convenience, and societal values), and against what kinds of threats. The goals might be absolute or they might specify a range of permissible trade-offs. In allowing trade-offs, we acknowledge the political nature of cybersecurity and the need for conversations among those affected when goals are set.

- **Means.** These might involve technological, educational, and/or regulatory *measures*. We should expect a means to include policy that creates incentives—which might range from market-based to coercive—that foster adoption and/or deployment of the measures it proposes.

Through incentives provided as part of its means, a cybersecurity doctrine can address barriers to market production of cybersecurity that others[1] have aptly noted reflect a lack of will rather than a lack of ability. Incentives also can prompt continued improvement to address the constantly emerging landscape of threats and the new needs that arise as a growing range of applications are being migrated to networked information systems.

[1] R. Anderson and T. Moore. The Economics of Information Security, *Science 314* (5799), 610-613, October 27, 2006.

One candidate for such a cybersecurity doctrine is the subject of this paper. Our doctrine of *public cybersecurity* is rooted in the thesis that cybersecurity is a public good. The focus of this doctrine, then, is on the collective rather than on any single individual's or entity's computer, network, or assets. Comparisons with public health—another public good—are apt; these are made below.

The next section (§2) analyzes the limitations of various cybersecurity doctrines that have been proposed to date. Then §3 discusses in detail our new doctrine of public cybersecurity. The production of cybersecurity by building systems that have fewer vulnerabilities is the subject of §4. Subsequent sections discuss approaches for managing insecurity: diversity (§5), surveillance (§6), installation of patches (§7), isolation (§8), and the role of intermediaries (§9). Finally, §10 puts this work into perspective and §11 offers some conclusions.

2 Cybersecurity Doctrines: Past and Present

The advent of time-sharing in the 1960's meant that computations on behalf of multiple users were interleaved on a single computer. Each user's computation and data thus had to be protected from misbehavior by programs being run by other users. Confronted by a problem born of technology, engineers of early time-sharing systems sought solutions in technology. Therefore, the focus of early cybersecurity doctrine was on developing new technology. Societal values could be and were ignored, because users of these early computing systems had shared values which meant those values would be respected without explicit discussion.

Technological solutions for creating the needed isolation were beyond our capabilities, especially when users could be motivated and capable adversaries bent on disrupting another user's computation or stealing information. Improved technology, however, is not the only way to solve problems that technology has created, and subsequent cybersecurity doctrines focused on policy to leverage those technological solutions that were at hand. These doctrines too were unsuccessful. But had they succeeded, they ultimately would have been inadequate because the problem was changing.

Computer systems were becoming pervasive, which had two broad consequences. First, the information technology sector became a significant economic force. Concerns about freedom to innovate and success in the marketplace come with this first consequence. The second consequence was that computer systems increasingly touched the lives of ordinary people, as

citizens whose records were stored electronically, and as workers who used information technology to be more efficient. Eventually, computer networks and the web were created; they changed how people shopped, communicated, socialized and engaged in politics. As a consequence, privacy and other societal values grew in importance. The requirement for a cybersecurity doctrine to take account of societal values became crucial.

Because their effectiveness has been limited, it is instructive to review the three doctrines that have dominated cybersecurity thinking for the past 50 years—the doctrines of prevention, risk management, and deterrence through accountability. In particular, by analyzing the measures these doctrines propose, we might hope to gain insights into properties that cause measures to be effective.

Doctrine of Prevention. The goal of this doctrine is systems that are completely free of vulnerabilities. Absent vulnerabilities, attacks are not possible, so the resulting system is secure.

Such *absolute cybersecurity* is worthwhile but unlikely ever to be achieved. For systems that incorporate humans as users and operators, we would need some way to prevent social engineering and intentional insider-malfeasance. Prevention, here, requires overcoming the frailty of humans, which is likely to involve more than technology.

If we ignore humans in a system then the problem is different but not easier. Software systems today are just too large and complicated to be verified using formal logics. Researchers, assisted by computers, have been able to devise formal proofs for small systems[2] (i.e., under 10,000 lines of code), and software producers regularly employ automated checking for relatively simple properties of code and for analyzing specifications. Techniques are also being explored that reduce the size of code bases for certain key systems, because a smaller code base is more amenable to formal verification.[3] But revolutionary advances are needed before formal verification could be used to verify the entire code base found today on a desktop system or server, yet that would be necessary for implementing the Doctrine of Prevention.

[2]Klien et al. seL4: Formal Verification of an OS Kernel. *Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles*, SOSP '09, (Big Sky, Montana, USA), ACM, 207–220.

[3]For example, "prerendering" can reduce the code required to generate the user interface in a voting system, thereby simplifying the vote-entry software and making it more amenable to verification. Yee, Ka-Ping. *Building Reliable Voting Machine Software.* Ph.D. Dissertation, Computer Science Department, University of California at Berkeley, Fall 2007.

System testing is the obvious alternative for gaining assurance that a system has no vulnerabilities. Tests, however, only can reveal the presence of vulnerabilities—not their absence. Exhaustive testing would be necessary for demonstrating the absence of vulnerabilities, and the amount of work involved here is prohibitive even for small components, much less for large systems.

Formal proofs and testing are necessarily relative to some expectations about what the system must do and the environments in which the system will operate. So we must discharge the Doctrine of Prevention relative to some expectations, which is tantamount to establishing the absence of vulnerabilities only if certain assumptions hold. Unfortunately, assumptions about the environment that hold one day might subsequently be invalidated. Attacks evolve in sophistication in response to better defenses. And threats emerge to exploit new opportunities for disruption that are created when cyberspace provides access to new forms of value. So a system that was once considered secure might not remain so for very long

In light of this dynamic, expectations about the environment must be periodically revisited and, if necessary, revised. Thus the Doctrine of Prevention involves a recurring expense. That recurring expense is inconsistent with the business model employed by many of today's software providers, which favors reuse and extension of existing hardware and software in order to lower the cost of producing new systems.

The adoption of mandatory standards can be seen as a way to support the Doctrine of Prevention by increasing the chances that what is built and/or deployed has fewer vulnerabilities. Some standards directly concern what functions an artifact must or must not support, some govern its internal structure, while others concern the process by which the artifact is constructed or maintained, and yet others govern qualifications of the personnel who are involved in creating the artifact. Examples include the TCSEC[4] (also known as the Orange Book), its successor the Common Cri-

[4]DoD Computer Security Center. *Department of Defense Trusted Computer System Evaluation Criteria*, CSC-STD-001-83, August 1983.

teria[5] as well as security provisions in information privacy laws[6] [7] [8], the Federal Information Security Management Act[9], and the Voluntary Voting System Guidelines[10]. Current market activity suggests that such mandates show value in some areas. However, a correlation between the absence of vulnerabilities and standards compliance has not yet been documented. So the stated goal for the Doctrine of Prevention is unlikely to be achieved through these measures.

Doctrine of Risk Management. Absolute cybersecurity is not affordable but, fortunately, it is also not needed for most systems. The Doctrine of Risk Management responds by stipulating a more modest goal—that investments in security reduce expected losses from attacks. To adopt this doctrine is to admit that all vulnerabilities are not equal; one should focus only on vulnerabilities whose exploitation is (i) sufficiently likely by the perceived threats and (ii) could enable system compromises that are expensive (by some cost measure). In contrast to the Doctrine of Prevention, the more modest goal of defending a smaller body of code against a more restricted set of threats is likely within our capabilities. Moreover, to maintain that steady state, fewer assumptions about the environment would have to be revisited periodically.

In theory, the Doctrine of Risk Management seems quite sensible. But the lack of information about vulnerabilities, incidents, and attendant losses makes actual risk calculations difficult. Companies and individuals do not know how to value (i) confidentiality of information, (ii) integrity of information, or (iii) the pain of dealing with recovery from an attack's effects (e.g., bad credit ratings). And these costs seem to vary tremendously. Also,

[5] US National Institute of Standards. *Common Criteria for IT Security Evaluation*, 1999. ISO Standard 15408. http://csrc.nist.gov/cc/.

[6] Health Insurance Portability and Accountability Act of 1996 (HIPAA), Pub. L. No.104-191, 110 Stat. 1936 (1996) (regulating the use and disclosure of "Protected Health Information").

[7] Title V of Gramm-Leach-Bliley Act (GLBA), Pub. L. No. 106-102, 113 Stat. 1338 (1999) (codified at 15 U.S.C. §§6801-6827 (2006)), 15 U.S.C. §§6801, §§6805 (empowering various agencies to promulgate data-security regulations for financial institutions).

[8] Children's Online Privacy Protection Act, 15 U.S.C. §6501 et seq. (prohibiting the collection of personally identifiable information from young children without their parents' consent).

[9] Federal Information Security Management Act of 2002 ("FISMA", 44 U.S.C. §3541, et seq.).

[10] Election Assistance Commission. *Draft Voluntary Voting System Guidelines*, Version 1.1, May 27, 2009. Available at www.eac.gov/testing_and_certification/voluntary_voting_system_guidelines.aspx

people have difficulty reasoning about extremely low-probability events. Finally, investment incentives are missing when costs are borne by third parties, materialize only well after the breach occurs, and causation is difficult to discern much less prove.

Accurate information about threats and attacks may not be publicly available because those with that knowledge fear tarnishing their reputations or compromising their intelligence methods and/or sources. Even were that information available, deployment of replacement systems and upgrades changes the uses to which systems are being put and alters the set of vulnerabilities, which in turn can lead to the creation of new attacks. These differences mean that the past is not a good predictor of the future. As a consequence, actuarial models cannot be constructed, so insurance to transfer risk is impossible to price in a way that ensures profits to the policy underwriter.[11]

Were there some way to analyze a system mechanically and obtain a quantity that indicates just how secure that system is, then we could have a basis for assessing what is gained from specific investments made in support of cybersecurity. But today such cybersecurity metrics do not exist. Quantities derived entirely from empirical observations also don't work for justifying investments. The absence of detected system compromises could indicate that investments in defenses worked, attacks haven't been attempted, or the compromise escaped notice (e.g., theft of confidential information). So whether or not prior security investments were well targeted is impossible to know, leaving security professionals to justify investments based solely on non-events.

Approaches based on risk management are further confounded by externalities that derive from the emergent nature of cybersecurity in networks. Individuals and entities here can neither fully reap the benefit of their security investment nor entirely control their vulnerability through investments.[12] For example, a single compromised system anywhere in a network can serve as a launching point for attacks on other systems connected to that network. So local investment in defenses not only provides local benefits but also benefits others; and under-investment in defenses elsewhere in the network could facilitate local harms. Absent coordination, the only

[11] Rainer Boehme and Galina Schwartz. Modeling Cyber-Insurance: Towards A Unifying Framework. Working paper presented at Workshop on Economics of Information Security, Harvard University, June 2010. http://weis2010.econinfosec.org/papers/session5/weis2010_boehme.pdf

[12] R. Anderson and T. Moore. The Economics of Information Security, *Science 314* (5799), 610-613, October 27, 2006.

logical strategy would be to invest in insurance (were it available), because here an entity can reap the entire benefit of that investment.[13] However, this strategy does nothing to improve security and, as noted above, viable long-term business models for insurance do not exist today.

The situation for risk management is not completely bleak, though. In the policy arena, state security breach notification laws[14] can be viewed as a risk management intervention. Significant costs are incurred to notify individuals and to manage the adverse publicity surrounding reportable breaches. These potential costs act as a proxy for the costs of security failures to customers, forcing companies to internalize previously externalized risks of security failures. The price tag on breaches also means that these laws have created a set of data to use in risk and return on investment calculations. However, the laws focus on only a narrow set of breaches and, as a result, might artificially skew the focus of investments.

Doctrine of Deterrence Through Accountability. This doctrine is concerned with treating attacks as crimes, and it therefore focuses on infrastructure to perform forensics, identify perpetrators, and prosecute them. We deter attacks by increasing the chances that perpetrators will be found and prosecuted.[15] Implementations of this doctrine require strong authentication technologies and surveillance of network activity. Robust forms of user identity allow us to overcome the loose binding that exists today between individuals and machines.[16]

Absent an effective means for retribution, this doctrine has no teeth and fails. And efforts to punish perpetrators of cyber-attacks are not always feasible in today's global environment. Attribution of actions by machines to individuals is complicated, agreement about illegality is illusive, and cross-

[13] J. Grossklags, N. Christin, J. Chuang. A game-theoretic analysis of information security games. *Proceedings of the 17th International World Wide Web Conference* WWW 2008, (Beijing, China, April 2008).

[14] Security breach notification statutes that require companies to notify individuals when certain personal data has been accessed or disclosed without authorization are in place in 46 states, the District of Columbia, Puerto Rico and the Virgin Islands. The first was California's Notice of Security Breach Law Cal.Civil Code §1798.29 (2002).

[15] Butler W. Lampson. Computer security in the real world. *IEEE Computer* 37(6), June 2004, 37–46.

[16] South Korea currently requires internet users to attach their real names and resident ID numbers when they post messages on the Internet. To accomplish this, Web sites that allow posting must collect and confirm names and resident IDs with a government server. Se Jung Park, Yon Soo Lim, Steven Sams, Sang Me Nam, and Han Woo Park, "Networked politics on Cyworld: The text and sentiment of Korean political profiles," *Social Science Computer Review*, September 21, 2010.

border enforcement requires more cooperation than is likely to emerge between important constellations of nations. Recent attacks against U.S. and other systems suggest that we cannot ignore non-nation state actors that engage in terrorism and large-scale financial crimes. The very features that make the internet a profitable environment for criminals—world-wide reach, connectedness, neutral treatment of packets, and weak binding of machines to individuals—make it difficult for law enforcement to identify and catch them, while other features of the international landscape make it difficult to bring them to justice.[17]

There are also conceptual obstacles that limit the effectiveness of the doctrine. First, this doctrine is punitive. Like most criminal law, it is aimed primarily at punishing with expectations of producing both general and specific deterrence. This does little to keep networks up and running when they are under siege nor does it prompt proactive security investments. Second, the doctrine could require individuals to sacrifice privacy and, in the extreme, abandon the possibility of anonymity and the protections for freedom of speech and association it affords.

Nevertheless, many attacks are indeed criminals plying their trade, and it makes good sense that criminal activity in cyberspace face the risk of retribution that we employ to deter criminal activity in the physical world. The Doctrine of Deterrence Through Accountability thus brings value here. But in cyberspace, unlike in the physical world, terrorists or state actors are difficult to distinguish from common criminals.[18] Deterrence through accountability is not necessarily effective against these trans-national threats. Other doctrine is also required.

3 A New Doctrine: Public Cybersecurity

Cybersecurity is non-rivalrous and non-excludable so, by definition, it is a *public good*. It is non-rivalrous because one user benefiting from the security of a networked system does not diminish the ability of any other user to benefit from the security of that system. And it is non-excludable because users of a secure system cannot be easily excluded from benefits security brings. Doctrines to foster the production of public goods thus constitute a sensible starting point in our search for doctrines that promote cybersecurity.

[17] J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of internet miscreants. *Proceedings of the 14th ACM Conference on Computer and Communications Security* (CCS '07). ACM, New York, NY, USA, 375-388.

[18] William J. Lynn III. Defending a new domain. The Pentagon's Cyberstrategy. *Foreign Affairs*, September/October 2010.

Economists define a *common good* as one that is rivalrous and non-excludable. The sea, outer space, and air are examples. In so far as common goods (by definition) are inherently different from public goods, doctrines for common goods are likely to be unsuitable for fostering cybersecurity. Indeed, we can see this irrelevance in laws for protecting common goods, which typically aim to ensure rights of equal use, and in the mechanisms these laws introduce, which are intended to manage the depletion and inequitable consumption by first comers or more sophisticated users. The production of cybersecurity has little to do with any of these issues.

3.1 Public Health as an Exemplar

Public health—the prevention of disease and promotion of good health in populations writ large—is a public good. It is non-rivalrous because having the population healthy implies a lower prevalence of disease, which in turn decreases the chances any member can fall ill. And it is non-excludable because nobody can limit an individual's ability to profit from the health benefits that living amongst a healthy population brings.

The essential characteristics of public health law are a focus on the health of the population as a whole and the singular responsibility of government in that enterprise.[19] To discharge these responsibilities, various agencies are mandated to engage in a broad set of activities, which include the following.

- Public education about the cause and effects of disease, as well as methods of prevention, since that education empowers individuals to act in ways that optimize their own health, which in turn furthers public health.

- Creation and use of methods for the prevention and treatment of specific diseases. This could involve (i) providing subsidies to procure care needed by those who could not otherwise afford it[20] or (ii) imposing specific health standards as eligibility requirements for various societal

[19]The mission of public health was defined by an influential Institute of Medicine Committee as "fulfilling society's interest in assuring conditions in which people can be healthy". *The Future of Public Health*, Institute of Medicine, 1988 p. 7. A rich description of the legal framework is set out in Lawrence O. Gostin, *Public Health Law: Power, Duty, Restraint*, University of California Press, January 2001.

[20]For example, after several outbreaks of measles among primarily unvaccinated children, a federal law was passed to provide free vaccines to certain groups of children and to provide funds to states to support efforts to enhance vaccination levels.

benefits.[21]

- Identification and management of disease and infected individuals though surveillance, information gathering, and analysis—including mandatory reporting requirements for certain diseases and conditions, mandatory testing or screening for others, symptom surveillance that seeks to identify non-obvious public health threats in masses of routine records, and mandatory treatment.

While interests of an individual and the public often align, it is the points of conflict that public health law speaks to—by setting out the rights of individuals as sovereigns over their physical bodies versus the obligation of the state to protect the population as a whole and by offering frameworks to mediate conflicts between the two.

As an example, public health mandates that children be vaccinated because in a generally healthy population such vaccinations cannot be justified based on the benefit to the individual. In fact, the optimal choice for any given child might be to avoid vaccination and thus avoid the risk of side-effects. What mandatory vaccination creates is *herd immunity*, which benefits the collective by reducing the total number of hosts available to carry a disease, thereby decreasing the risk to individuals who have not been vaccinated. However, if too many individuals act in their own self-interest and eschew vaccinations, then the herd immunity that allowed a non-vaccinator to freeload may disappear. This is a "tragedy of the commons" where individuals acting rationally leave everyone worse off.

Every state in the U.S. conditions a child's attendance at school on satisfying some specified regimen of vaccinations. And, in addition, vaccine manufacturers are indemnified from liability for side-effects users might experience. Specifically, the Vaccine Injury Compensation Program (VICP) provides certain payouts from public coffers to children injured as a result of a vaccine. VICP also provides a compensation mechanism outside seeking damages from the vaccine manufacturers, thereby providing an environment conducive to both the production and willing use of vaccines. While this cannot fully compensate for negative health consequences from vaccinations, it is an important component of the overall public health strategy.

Public health is a logical outgrowth of having disease detection and prevention mechanisms, which transformed societal perception of health from a primarily private concern to a concern of the collective. Ultimately, this

[21] As discussed later, children in the United States are eligible to attend public school only after receiving a mandated set of vaccinations.

led to public health being seen as a public good that the government should enable.[22] With a focus on the collective, health now becomes intertwined with societal values. For example, we impinge on societal values when we introduce mandatory reporting and surveillance systems that (i) alert individuals at specific risk so they can be tested and treated, and (ii) allow isolation, quarantine, and even mandatory treatment to be imposed. At the same time, public health interventions aim to minimize their intrusiveness due to the chilling effect they may have on access to healthcare. Thus, we pursue programs such as anonymous HIV testing and needle exchange.

This same public health framework (*viz*, laws, agencies, and measures) applies equally well to weaponized pathogens. This is not to suggest that motive is irrelevant in considering public health strategies. For example, weaponized pathogens may change more quickly than those that evolve in nature, and certainly the transmission vectors may differ when pathogens are used as weapons. But the basic tools of public health still apply: public education (to minimize exposure and facilitate early detection), investments to create means for prevention and treatment (antidotes and vaccines), and surveillance and analysis (facilitating isolation and quarantine as defenses).

3.2 From Public Health to Public Cybersecurity

Public health and cybersecurity both aim to achieve a positive state (health or security) in a loosely affiliated but highly interdependent network. With one, it is a network comprised primarily of people existing in an environment over which they have some limited control; with the other, the network is comprised of people, software, and hardware (for communications, storage, and processing). And because this positive state is ultimately unachievable, both struggle with how to manage in its absence as well as with how to prompt its production. Success ultimately depends not only on technical progress but on reaching a political agreement about (i) the relative value of some public good in comparison to other societal values and (ii) the institutions granted authority to resolve conflicts (and the methods they use).

We define a *doctrine of public cybersecurity* to be any cybersecurity doctrine whose goals are (i) producing cybersecurity and (ii) managing insecurity[23] that remains, where political agreement balances individual rights and

[22] *The Future of Public Health*, Institute of Medicine, 1988 p. 3.

[23] Systems that employ technical means to enable continued operation in the face of attacks are sometimes called *intrusion tolerant*. A sampling of specific techniques for achieving intrusion tolerance is discussed in *Foundations of Intrusion Tolerant Systems*

public welfare. There is no single doctrine of public cybersecurity. Different definitions for what is meant by "cybersecurity" and "insecurity" as goals[24] lead to different doctrines of public cybersecurity. Also, different choices of measures and incentives that together constitute the means result in different doctrines of public cybersecurity. But none of the doctrines discussed in §2 has all of the elements we require in a doctrine of public cybersecurity.[25]

The analogy with public health inspires cybersecurity measures like prevention, containment, mitigation, and recovery—strategies that direct resources toward production and preservation of cybersecurity. But modern public health doctrine does not compensate victims of disease, so by analogy, a doctrine of public cybersecurity would not focus on restitution. Indeed, restitution is economically efficient only when attacks are infrequent, and that assumption is not realistic today.

Modern public health also does not punish victims of disease, though there is some nuance. Quarantine, in response to disease, benefits the collective by depriving an individual of certain freedoms. Such a response could be seen as harsh consequences, which is one definition of "punishment". By analogy, a doctrine of public cybersecurity could dictate responses that deprive individuals of actions, but only if those responses benefit the collective. Punishments solely for retribution could not be part of a public cybersecurity doctrine (since retribution does not benefit public welfare), though nothing precludes implementing a doctrine of public cybersecurity alongside a cybersecurity doctrine that does admit retribution. Finally, the parallel with public health also suggests that prevention be preferred to recovery.

With regard to incentives, ensuring that actors contribute to public cybersecurity requires interventions to overcome positive and negative externalities that lead rational individuals to under-invest, just like for public health. And when incentives are insufficient to prompt private provisioning, the public interest requires making value-ridden choices to interfere with the rights and interests of individuals and organizations. Those choices would be embodied in goals that reflect political agreement about the good in

(Jaynarayan Lala, editor), IEEE Computer Society, Order Number PR02057, ISBN 0-7695-2057-X, 2003. Uses of the term "managing insecurity" in this paper are intended to denote something broader, admitting non-technical means as well as intrusion tolerance techniques.

[24]H. Nissenbaum, Where Computer Security Meets National Security. *Ethics and Information Technology*, Vol. 7, No. 2, June 2005, 61-73.

[25]Doctrine of Prevention is not concerned with managing insecurity; Doctrine of Risk Management and Doctrine of Deterrence Through Accountability are not concerned with producing cybersecurity. And none concern trade-offs of individual rights for public welfare.

question (its definition), the socially desirable level given competing priorities and values, and provisions for determining when the individual's desires yield to the collectives' need. For example, an agreement might stipulate that state coercion is permitted only when certain incursions into the rights and interests of individuals are tightly circumscribed.

Public health solutions don't always translate into sensible support for public cybersecurity, but we have been pleasantly surprised at how often the former inspires strategies for the latter. Below, we explore some examples. These also illustrate how doctrines of public cybersecurity can be useful for evaluating current cybersecurity proposals. Our choice of examples should not, however, be seen as an endorsement for any particular proposed set of interventions.

4 Producing Security during Development

Underutilized approaches (e.g., formal methods, testing, and improved software engineering processes and standards) that were developed, in part, to serve the Doctrine of Prevention are effective in producing cybersecurity (by reducing the number of vulnerabilities present in a system), even if they cannot produce absolute cybersecurity. Thus, methods do exist that could serve as a means for a doctrine of public cybersecurity, just like disease prevention through vaccination and the monitoring of our food and water supplies fosters public health. The question is: What incentive structures would ensure that these methods are used?

Education could play a key role in defect reduction. Knowledgeable developers are less likely to build systems that have vulnerabilities. They are also better able, thus more likely, to embrace leading-edge preventions and mitigations. There is, however, no agreement about what should be taught. To reach that agreement would require a dialog among universities and with practitioners.

Once agreement has been reached about a body of knowledge for cybersecurity practitioners, we could require that mastery of this material be certified as a condition for practice. But the details of how certification is handled can be subtle. Possession of a certificate does not by itself compel the use of best practices, and it is easy to imagine certified system-builders who cut corners by choice (out of laziness, for example) or by mandate (because management is trying to reduce costs). Moreover, unless the certification process imposes a continuing education requirement to ensure that certificate holders will stay current with new developments, it might impede

rather than promote the spread of innovation. And even when continuing education is mandated, old habits die hard; physicians, for example, who have been shown new methods, empirically demonstrated to be superior, nevertheless exhibit a tendency to stick with familiar practices.[26]

Employing practices to reduce defects during system development and employing better-educated practitioners will mean systems become more expensive to produce. Today's software-procurement market does not provide incentives that would prompt developers to incur those additional expenses. Moreover, purchasers are unable to predict the costs of a system's vulnerability to attack and, without ways to measure a system's security, have no way to rationalize paying higher prices—the same underlying reasons that the Doctrine of Risk Management failed.

Law could force system producers and/or purchasers to make the necessary investments. Software distributors currently disclaim liability beyond the purchase price for damages caused by their product. This probably reduces the time and energy that developers devote to eliminating defects, as evidenced by the number of buffer overruns and other exploitable coding errors still being discovered and exploited by attackers. The law could, for example, be revised to disallow limits on damages flowing from attacks that exploit poor coding practices that lead to buffer overflows and other easily exploited vulnerabilities. Limits on liability could depend on the use of formal methods, type safe languages, or specific forms of testing (e.g., fuzz testing[27]). Creation of a class of certified security professionals also could provide the basis for a professional duty of care supporting liability for shoddy security.

Law also could require that software developers adhere to security standards. Or safe harbor provisions could be created to protect software developers against future findings of liability for those systems built according to specified standards. In fact, the law arguably already mandates that companies follow certain standards regarding personally identifiable information. Through a set of settlement agreements, the Federal Trade Commission established a de facto standard that requires a company collecting and handling personal information of consumers (i) to establish reasonable security processes and (ii) to mitigate system vulnerabilities that are known in the marketplace and for which mitigations exist. A first step in deter-

[26] Deborah G. Mayo and Racehelle D Hollander. *Acceptable Evidence: Science and Values in Risk Management*. Oxford University Press, 1994

[27] In *fuzz testing*, a system is exposed to random inputs of unexpected kinds. This form of testing reveals inadequacies in the input validation routines of a system. Several classes of attacks are blocked by implementing stringent input validation.

mining whether law should more broadly mandate the adoption of security standards might be research that identifies connections between security development processes and good security outcomes.[28]

5 Managing Insecurity: Diversity

Monocultures in nature risk extinction from pathogens and have less chance of adapting to changing conditions. Diversity—of the individuals within each species and by virtue of many species co-existing within an ecosystem—is one way nature creates a resilient ecosystem. Public health benefits from individuals in a population having different inherent resistance to pathogens and, by virtue of the different exposures[29] to diseases that each experiences, having different immunities.

Although nature abhors monocultures, cyberspace seems to favor them. A collection of identical computing platforms is easier, hence cheaper, to manage because we need master only one interface and make one set of configuration decisions. In addition, user training costs are reduced when job transfers do not have the overhead of learning yet another operating system and suite of applications; investments in education about how to use or manage a system also now can be amortized over a larger user base in a monoculture. Finally, interoperability of a few different kinds of systems is far easier to orchestrate than integrating a diverse collection, standards not withstanding. So networking is usually easier to support with a monoculture. Mindful of these advantages, the public and private sectors both tend to adopt procurement policies that foster creating computer monocultures.[30]

[28] For a case-study comparison of four vulnerability reduction techniques see K. Buyens, B. De Win, W. Joosen: Empirical and statistical analysis of risk analysis-driven techniques for threat management. ARES 2007: 1034-1041, *Proceedings First International Workshop on Secure Software Engineering* (SecSE 2007). For a theoretical comparison of two high-profile development processes (Microsoft SDL and the Open Web Application Security Project's Comprehensive, Lightweight Application Security Process (CLASP) see Johan Gregoire, Koen Buyens, Bart De Win, Riccardo Scandariato, and Wouter Joosen: On the secure software engineering process: CLASP and SDL compared. In *SESS'07: Proceedings of the 3rd International Workshop on Software Engineering for Secure Systems*, Minneapolis, MN, USA, May 2007.

[29] Vaccination works by causing exposure to a relatively benign form of the disease against which protection is being sought.

[30] A notable example is U.S. Office of Management and Budget policy memorandum M-07-11 "Implementation of Commonly Accepted Security Configurations for Windows Operating Systems" which lists the permitted versions of Windows that should be used by certain civilian federal agencies. See www.cio.gov/documents/FDCC_memo.pdf.

Methods do exist, however, to artificially and automatically create diversity in software systems—and do so without sacrificing the advantages given above for having a monoculture. These methods involve tools that randomly transform code and/or stored information in semantics-preserving ways. Due to the *artificial diversity*, internal details of an individual system are no longer predictable. So an attack that exploits knowledge of internal details is more likely to cause a system crash after a small number of instructions than to cause an attacker to receive control. In many settings, having a system crash is preferable to it being controlled by attackers. Moreover, a platform that crashes in response to an attack cannot then help propagate that attack to other platforms. The crash also (implicitly) signals system operators that something is wrong, creating an opportunity for initiating other means to block an attack's spread.

As with the diversity found in nature, artificial diversity is inherently a probabilistic defense, because an attack against any given individual component might not be derailed by the specific random transformations that were made to that component. Also, by converting some attacks into crashes, artificial diversity can adversely affect a system's availability.

Despite these limitations, artificial diversity does facilitate public cybersecurity because it provides a way to cope with residual vulnerabilities, thereby providing a way to manage insecurity. Today, artificial diversity is often used in operating systems but less so in applications[31] (even though increasingly it is applications that attackers target). However, the various legal approaches discussed above in §4 for incentivizing defect reduction during development are equally well suited for incentivizing system producers to support artificial diversity. So a rich space of incentives is at hand for encouraging broader adoption of the measure.

6 Managing Insecurity: Surveillance

Surveillance is used extensively in support of public health, where data is collected through a variety of means. The data enables disease containment and mitigation through

- the dissemination of information that facilitates individual actions to avoid exposure to infection,

[31] D. Ladd et al. The SDL Progress Report. Microsoft Corporation, 2001. page 23–24. Available at http://www.microsoft.com/downloads/en/details.aspx?FamilyID=918179A7-61C9-487A-A2E2-8DA73FB9EADE&displaylang=en.

- isolation and quarantine, which limit the interaction of affected individuals with the rest of the population, and
- mandatory treatment to reduce danger to the public.[32]

Data collection for public health occurs at many levels. At the lowest level is the self-serving inclination of individuals themselves to assess their well-being. Education equips individuals with a basic level of knowledge about health indicators—normal body temperature, pulse, blood pressure, and respiratory rate—as well as with simple precautions to limit infection and the spread of disease (e.g., frequent hand washing). Other data is collected by primary care providers in conjunction with annual check-ups and, when symptoms require further analysis, at hospitals and other more-advanced diagnostic facilities. Each successively higher level is concerned with the overall health for a larger population and, thus, provides a natural venue for constructing and analyzing larger data aggregations.

By minimizing the disclosure of information about an individual's health, public health law strives to reduce one potential source of reluctance about seeking health care—fear an individual might have about being shunned because of a publicized health condition. Generally, identifying information flows away from primary health care providers only in instances where aggregation and/or analysis is necessary to identify significant trends. And when information does flow, efforts are undertaken to protect each individual's privacy.

Surveillance in Networked Systems. In contrast to public health, coordinated surveillance is not used extensively today in support of cybersecurity. Yet it would be feasible and advantageous to do so. Low-level indicators about the basic "health" of a computer can be made available by running built-in checking software (e.g., virus scanners and intrusion detection systems). And each of the Internet Service Providers (ISPs) that constitutes the network has an infrastructure that facilitates monitoring of events internal to its network as well as interactions with other networks.

Surveillance of network traffic (volume, distribution over time and destinations, etc.) could be a powerful potential source of information about certain attacks and vulnerabilities. Denial of service attacks, for example, have an obvious manifestation and a natural mitigation based on traffic filtering by ISPs. However, the source(s) of such attack packets, the target(s),

[32]The general rule allows individuals to refuse treatments. Some states mandate treatments for communicable diseases, such as tuberculosis, that pose a danger to the public.

and the intermediaries are likely to span multiple ISPs, which would have to share data and coordinate for mitigation. Data sharing among ISPs is, unfortunately, inhibited today by competition and, in some cases, varied interpretations of privacy law.[33] So ISPs are less able to have situational awareness that could enable them to suppress packets delivering attacks. Widespread sharing of information, however, can introduce a risk if chances are now increased that attackers learn about vulnerabilities for specific sites.

Just as there are privacy issues with collecting data about an individual's health, network traffic surveillance raises privacy concerns. The extent to which collecting packets actually does impinge on privacy depends on what information is recorded, how long it is stored, how it is used, who can have access to the information, etc. For example, realtime responses to protect networks can be accomplished by authenticating machines, a far less politically fraught solution than proposals for "internet drivers licenses" and other tight bindings between machines and individuals.[34]

ISP cooperation and information sharing is less likely to raise privacy concerns than collection of information by centralized government organizations. Yet the packets themselves can be invaluable to a government seeking situational awareness about threats in cyberspace, and defense of its citizens is a clear responsibility of government. Unfortunately, packet inspection is also easily abused by a government for spying on citizens; critics cite this fear (among others) when discussing the Einstein[35] systems recently deployed by the U.S. Government for monitoring all federal civilian agency connections to the internet. As with public health, political agreement must be reached based on the expected benefits (backed by sound research and field experience) of surveillance and the risk it poses to other values.

An understanding of the kinds vulnerabilities found in systems is a form of situational awareness of potentially great value to system builders. In the absence of mandatory reporting requirements for cybersecurity incidents, a diverse collection of public and private reporting mechanisms have evolved. The U.S. Department of Homeland Security (US-CERT) and the NIST Computer Security Division maintain databases of common vulnerabilities.

[33] M. Van Eeten and J. M. Bauer. Economics of malware: security decisions, incentives and externalities. OECD STI Working Paper 2008/1. OECD. Available at www.oecd.org/dataoecd/53/17/40722462.pdf.

[34] See, D. D. Clark and S. Landau. Untangling Attribution. *National Security Journal* Vol. 2, Issue 2 (2011). http://harvardnsj.com/2011/03/untangling-attribution-2/.

[35] Center for Democracy and Technology, Einstein Intrusion Detection System: Questions That Should Be Addressed, July 28, 2009. http://cdt.org/security/20090728_einstein_rpt.pdf.

Many organizations contribute, but these are not the only such databases and none provides anything close to a complete view. Some vulnerabilities do not get reported to any. For example, a private sector community of "security researchers" search systems for vulnerabilities but report their findings to middlemen, who offer it for sale to companies that build and sell anti-malware or intrusion prevention/detection products.[36] The ad hoc Conficker Working Group[37] is an example of a rather successful coordinated private-sector activity formed to defend against a particular attack.

7 Managing Insecurity: Patching

A *patch* is an update that can be applied to an installed system in order to eliminate one or more previously identified vulnerabilities. Exploitation of an unpatched vulnerability on a computer could target that machine and assets of the individual and, therefore, be fully internalized. Or the exploitation could target the machines and systems of others, producing a negative externality.[38] This uncertainty about consequences means that self-interest of machine owners is not a strong incentive to apply patches.

Various policy interventions could raise patch rates, though. Choosing among them requires additional information about why people and businesses delay or outright fail to apply patches.

- Research might conclude that low patch rates in the consumer market are caused by an under-appreciation of the risks. Public education that applying patches improves cybersecurity might then dramatically increase patching.

- We might find that individuals lack awareness of vulnerabilities present on their machines. Here, built-in software to check whether all current patches have been applied might suffice for triggering consumers to be more attentive to downloading patches for their machines.

- Feedback about what others do, thereby creating new norms of behavior, might lead to better patching practices. Researchers in other areas

[36] The risk that such vulnerabilities might be sold or disclosed to irresponsible or hostile parties is cause for concern about this kind of market, though the structure flourishes in the absence of alternatives.

[37] See www.confickerworkinggroup.org.

[38] For example, unpatched machines can be co-opted by attackers into a botnet. Such collections of remotely-controlled machines are used today for a variety of illicit activities, including generation of spam email and distributed denial of service (DDoS) attacks.

have found that showing individuals how their behavior compares to others taps into a competitive and/or social consciousness. So simple statements that a great percentage of others have patched their machines and are doing their part for cybersecurity just might be enough to push the laggards into applying patches.

- Research might find that individuals or enterprises hesitate to install patches for fear of destabilizing their other software. Reluctance by individuals might be overcome if software vendors were more transparent about what specific configurations and applications they tested. And enterprises that depend on home-grown software could have have their fears somewhat assuaged if test suites could be provided for use by patch developers.[39] As a final safety-net, we might require that all software contain mechanisms whereby a patch that has been applied can easily be removed, and the system and data restored to the pre-patch state.

- If the impediment to installing security patches is the time or expertise needed then an obvious mitigation is to have vendors configure defaults that automatically download and apply security patches.

- Consumers being charged for Internet access in proportion to the amount of bandwidth they use will incur lower costs by not downloading patches, which could be an excuse to forgo installing security patches. One solution, here, is to subsidize the bandwidth required for such downloads; tariffs that distinguish between different kinds of traffic is another solution.

- Those who run pirated software might hesitate to install patches for fear that the installation process would disable the illegal software or detect and report it. Regulations could address this by prohibiting security-patch installation from implementing functionality in support of license-enforcement or any other form of intellectual property protection.

Incentives to apply patches could also have a useful indirect effect. If patch installations are frequent and disruptive then consumers have reason to prefer products that have fewer security vulnerabilities from the start. This, in turn, pressures software producers to build and deploy more secure products.

[39] However, some enterprises regard their software and data as proprietary. They might not be comfortable providing test suites to patch developers.

Mandates to apply patches raise concerns about subsidizing the installation of patches[40] and compensating injured parties when patches cause harm. Losses from applying mandated patches, particularly where unacknowledged and uncompensated, will breed suspicion and resistance to patching efforts. Thus, it seems advisable to consider back-stop measures, analogous to what VICP provides to incentivize the use and production of vaccines and to the process used at the Food and Drug Administration to ensure vaccine efficacy.

8 Managing Insecurity: Isolation

Geological features, like mountains and oceans, have proved valuable in protecting individuals and populations. And when natural boundaries are absent, we build our own—fences to surround buildings and nations, often with guards to control who is allowed to transit the border. Such boundaries protect activities on one side from activities occurring on the other. A boundary might limit travel in one direction or in both directions, and it might be completely impervious or it might be selective about what may pass. Neither is a panacea. An impervious boundary could block the good with the bad; a selective boundary must employ some kind of *filter*, and that filter might block things it shouldn't or pass things it should.

Firewalls, so-called network guards, intrusion detection/prevention systems, and "air gaps" are examples of mechanisms that implement boundaries in networked systems.[41] Data collected through surveillance can serve as the basis for *signatures*, which are then used to define filters, effectively creating dynamic boundaries. Surveillance thus can lead to automatically-imposed quarantines. And since attacks in networks propagate rapidly, automatic response is especially attractive.

Ideally, we would deploy selective boundaries that block attacks but nothing else. In practice, though, filters will be far from perfect.

- Filters that inspect packet payloads (known as *deep packet inspection*) in addition to checking packet headers are ineffective when packet payloads are encrypted or otherwise obfuscated. Encryption is not

[40] Richard Clayton. Might governments clean up malware? *Proceedings Ninth Annual Workshop on Economics and Information Security* WEIS10, (Cambridge, MA, June 2010).

[41] The term "air gap" originally referred to isolation caused when no wires are connected to a given component. With the advent of wireless networking, the term's meaning has broadened to denote isolation caused when the Laws of Physics ensure no signal can reach the component.

today used extensively in networks, but that could easily change. And encryption is often used by attackers to evade detection. Moreover, it is malware variants that often are being spread, where each variant is obfuscated by applying a different random set of semantics-preserving transformations. It is quite difficult and often impossible to construct a signature that matches all variants by generalizing from a few.

- A filter might be designed either to (i) block packets and protocols corresponding to known attacks or (ii) pass packets from protocols or conveying content that is known to be normal. Filters that implement (i) are fooled by new attacks (in addition to suffering limitations described in the previous bullet). Filters that implement (ii) could block previously unseen protocols and kinds of packets, which stifles innovation.

- Whether a packet is part of an attack could depend only on sender intent. Consider a large number of request packets being sent to a web server. Are many people trying to access the same particularly topical content or is a denial of service attack is in progress? Sender intent is the sole differentiator.

There is also a human element to consider. Boundaries and filters must be installed, configured, and managed by human operators, and people do make mistakes. Moreover, when such a mistake allows unimpeded flow, then the error might be difficult to detect until after it is too late.

Network providers are understandably reluctant to publicize details of defenses, because revealing that information could help attackers. Yet we do see example defenses in today's commercial networks, which create and reference so-called "black lists" of sites whose communications will be ignored and "white lists" of sites that are known to be trustworthy. Some ISPs create a competitive advantage by offering a service to their customers whereby suspicious inbound-traffic spikes directed at that customer's site will automatically prompt upstream filtering to block those suspicious packets. Denial of service attacks in such networks are harder to undertake, as a result. Other ISPs monitor each endpoint, disconnecting it if outgoing traffic suggests that endpoint is compromised.[42]

A boundary might be deployed around a system (be it a single computer or a network) that must be protected from attacks or it might be deployed

[42]OECD. Malicious software (malware): A security threat to the internet economy. Ministerial Background Report DSTI/ICCP/REG(2007)5/FINAL, June 2008. Available at http://www.oecd.org/dataoecd/53/34/40724457.pdf.

around a system that is likely to harbor attackers. Different incentives are effective in each case. One natural scenario for direct government investment occurs when security boundaries coincide with national ones. Systems in different countries are subject to different laws, typically reflecting different societal values. A government might therefore justify installing a boundary whenever systems subject to its laws are being connected to systems located in a jurisdiction that allows system behavior the first considers an attack.

Boundaries are more likely to accepted and effective when instigated by the collective rather than by individuals. First, an individual is unlikely to have the necessary authority for mandating changes to defenses on all of the remote systems that could be involved in creating a quarantine. Second, the possibility of freeloading means externalities limit the incentives that might prompt owners or operators of networks or individual systems to make the investments to support enforced isolation. Finally, an agent of the collective would have the fuller picture that enables better signatures to be defined for filters.

We see an example of such boundaries in recent proposals for deterrence through accountability. Some have suggested that the internet be partitioned into national or multi-national enclaves. Those enclaves that serve the population we are trying to protect (i) run protocols that enable packet-sender tracing and (ii) do not carry traffic from enclaves where packet-sender tracing is not supported or cannot be trusted. Accountability of attack packets back to an individual machine can now be supported in enclaves that serve populations of concern.

With powerful enough filters, boundaries have the potential to intrude on societal values. One concern arises when the defining filters not only block packets that contain attacks but can be configured to block other kinds of packets. Such a filter could be used to prevent data from leaving an enclave, which makes it well suited for protecting confidential information against theft. But content filters also could allow governments to implement censorship, as illustrated by the firewalls China has installed to protect that nation's computing systems from receiving information in violation of local laws about allowed speech. Deployed in the reverse direction, a content filter could block someone from sharing information with others, thereby stifling debate.

So there may be trade-offs: societal values as well as potential benefits for the collective versus constraints on activities by individuals and business. Moreover, no criteria for deciding where a system should be segregated will be infallible. The result is a complex risk-management decision procedure that society must prescribe, with imperfect information and unknowable

consequences.

9 Managing Insecurity: ISPs as Intermediaries

The public health system leverages health professionals and other institutions to influence individuals' behavior. For example, health professionals educate individuals about benefits of vaccinations, schools demand conformance with vaccination schedules, and airports screen passengers for symptoms during some infectious disease outbreaks. Intermediaries clearly play an important role in public health strategies.

Intermediaries also have an important role to play in fostering cybersecurity. For example, in many cases today, network operators, such as employers and universities, require that all machines on their networks run virus detectors or malware detectors (with up-to-date signature files). These intermediaries could require that all machines are up-to-date on security patches. Similarly, some ISPs have chosen to notify subscribers when a computer seems to be infected.[43] At least one ISP restricts web surfing until the machine is cleaned-up, while another ISP reportedly quarantines any compromised machine until it is clean.[44]

ISPs are well positioned to facilitate patching and, by monitoring traffic, to enforce isolation of machines harbouring certain malware. Yet ISPs today have little incentive to engage in such behaviors, because in doing so they incur the bulk of the costs while any costs from the infected machines is more widely dispersed. Moreover, an ISP that disables or limits a machine's access to the internet will likely bear the burden of assisting that customer as she attempts the necessary repairs. Analysis[45] suggest that the cost incurred by an ISP in fielding a customer's tech-support call approaches the ISP's annual revenue from that customer. So by making this sort of monitoring and clean-up a mandatory obligation for ISPs, we would not

[43] Elinor Mills, "Comcast pop-ups alert customers to PC infections" Cnet News, October 8, 2009 (discussing Comcast October 2009 announcement of a trial of an automated service that will "warn broadband customers of possible virus infections" called Comcast Constant Guard the pop-up notice of possible infection directs customers to resources to rid their machine of infection, a similar service by Qwest, Customer internet Protection Program that displayed a Web page warning to customers provided ways to remove the infection for free before allowing them to surf the Web; and a similar older, now discontinued service of SBC that quarantined computers until they were cleaned)

[44] Id

[45] Richard Clayton. Might governments clean up malware? *Proceedings Ninth Annual Workshop on Economics and Information Security* WEIS10, (Cambridge, MA, June 2010), page 5, footnote 2.

only force action by the ISPs but we would also prevent consumers from shopping around for ISPs with weaker security requirements.

More daunting are the potential costs an ISP might incur from making an incorrect decision to disconnect a customer.[46] To limit spam email, for example, an ISP might block all bulk sending of email. But missives sent by a political organization might then be blocked, resulting in unwanted attention from advocacy groups and the press.[47] And while the law is evolving to provide ISPs who take steps designed to protect security with immunity from suits brought by providers of malware, those users who suffer losses after installing required patches or system upgrades or suffer due to isolation might also file legal complaints. In summary, the costs of ISP intervention present a formidable barrier to such action, but the law could remove these disincentives.

A recent proposal[48] by Lichtman and Posner argues that expanding ISPs liability "for violations of cyber-security" would improve cybersecurity because (i) individual attackers are often either beyond the reach of the law or are judgment-proof and (ii) because ISPs "can detect, deter, or otherwise influence the bad acts in question." The proposers speculate that ISPs could detect cybersecurity violations by building profiles of their users and looking for traffic anomalies, in a manner analogous to the monitoring of cardholder purchases done by credit card companies. But anomaly detection with usable levels of fidelity has eluded cybersecurity researchers for decades (as Lichtman and Posner openly admit), so implementing the proposal is not at present feasible.

Still, a policy holding ISPs liable for the damage caused by infected machines running on their networks might encourage more diligence in monitoring and fixing their subscriber's machines. The details are subtle and depend on the standard for liability—strict, knowledge based, etc. One (undesirable) outcome would be if the ISP undertakes less monitoring as a way to avoid its duty to intervene.

Alternatively, we could provide indirect or direct subsidies to foster cybersecurity-preserving activities by ISPs. For example, the government

[46]For a discussion of some of the issues raised by an ISP decision to cut off broadband access due to infection see, George Ou, "Comcast heading the right direction on cybersecurity" Digital Society Web site, October 9, 2009 http://www.digitalsociety.org/2009/10/comcast-heading-th-right-direction-on-cybersecurity/.

[47]http://www.eff.org/wp/noncommercial-email-lists-collateral-damage-fight-against-spam.

[48]See generally Doug Lichtman and Eric Posner, Holding Internet Service Providers Accountable, 14 S. Ct. Econ. Rev. 221, 233-34 (2006).

could create a centralized service for hosting patches or it could subsidize bandwidth to all endpoints in order to ensure that cost or delay to download a patch cannot become an impediment to installing that patch.

Given the decentralized and private provisioning of network resources in this country and many others, figuring out the role of intermediaries in driving cybersecurity is essential. As in other areas, such as copyright, the challenge is to establish policies that incentivize the desirable behavior while minimizing impact on other values.

10 Discussion

A biological basis for cybersecurity has been discussed by computer scientists for at least two decades. Their thrust has been to understand whether computer networks can benefit from implementing defenses like those that have worked so well for protecting living things. So we see intrusion detection systems that mimic pathogen detection by immune systems[49] and software defenses based on artificial diversity.[50] A recent white paper from DHS[51] describes how a human immune system's response mechanisms might serve as the blueprint for software that defends individual computers and networks against cyber-attacks; much research remains to be done, though, before those ideas are reduced to running code.

In contrast to this biological metaphor, which focuses on technical measures for blocking cyberattacks, the analogy between public health and cybersecurity is primarily concerned with new policy and new institutions. The notion of a Cyber-CDC, for example, has been attracting considerable interest.[52] Inspired by the existing Centers for Disease Control and Prevention, the proposed Cyber-CDC would be a government institution that provides leadership in organizing public and private sector strategies to enhance cybersecurity. It would also undertake data collection about threats and attacks, analyze and disseminate that information (perhaps in partner-

[49] S. Forrest, A.S. Perelson, L. Allen, and R Cherukuri. Self-nonself discrimination in a computer. *Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy* (Oakland, CA, May 1994).

[50] S. Forrest, A. Somayaji, and D. Ackley. Building diverse computer systems. *Proceedings of the Sixth Workshop on Hot Topics in Operating Systems* (Cape Cod, Mass, 1997).

[51] Department of Homeland Security. Enabling distributed security in cyberspace. http://www.dhs.gov/xlibrary/assets/nppd-cyber-ecosystem-white-paper-03-23-2011.pdf.

[52] National Cyber Leap Year Summit 2009, Co-chairs Report. http://www.nitrd.gov/NCLYSummit.aspx

ship with the private sector), serve as a repository for technical remedies, and educate the public about best practices, defenses, and remedies.

An IBM white paper[53] broadens the analogy. It proposes not only borrowing from public health but also from public safety. The broader analogy leads to additionally recommending that a Cyber Federal Emergency Management Agency be established and that a Cyber National Response Framework be devised. Independently, Microsoft's Scott Charney has advocated "device health" measurement and using a device "health certificate" as a basis for controlling access by devices to network resources.[54] Hunker, instead of focusing on institutions, suggests that public health could be a model for norms of behavior.[55] Individuals are expected to satisfy those norms; government institutions should focus on supporting the norms.

None of the aforementioned work includes a compelling argument for why the analogy to public health is a suitable starting point for cybersecurity. Public health informs people's behaviors (seemingly an obvious route to enhanced cybersecurity) but so does religion (which is not being advocated as a cybersecurity solution). We argued the analogy between public health and cybersecurity by using economic theory to highlight their shared status as public goods, since economics provides a vocabulary for talking about externalities and incentives. So while viewing cybersecurity as a public good is not new[56] [57] [58] we seem to be the first to use economics insights to justify the analogy between public health and cybersecurity.

Our public cybersecurity doctrine also goes beyond prior work that explores cybersecurity counterparts for institutions and policies that have served public health well. Public cybersecurity derives from identifying cybersecurity counterparts to the *goals* of public health—not the *institutions* of public health. First, through public health law, we gain and have

[53] Daniel B. Preito and Steven Bucci. Meeting the cybersecurity challenge: Empowering stakeholders and ensuring coordination. White Paper, IBM U.S. Federal Division, Feb 2010.

[54] Scott Charney. Collective Defense. Applying Public Health Models to the Internet. White paper available at http://www.microsoft.com/security/internethealth

[55] Jeffrey Hunker. U.S. International policy for cybersecurity: Five issues that won't go away. *Journal of National Security Law & Policy* 4, 197–216.

[56] Benjamin Powell. Is cybersecurity a public good? Evidence from the financial services industry. Working Paper Number 57, The Independent Institute, March 15, 2001.

[57] Bruce H. Kobayashi. An economic analysis of the private and social costs of the provision of cybersecurity and other public security goals. Working paper 26, George Mason University School of Law, 2005. Available at http://law.bepress.com/gmulwps/gmule/art26.

[58] Brent Rowe and Michale Gallagher. Private sector cyber security investment strategies: An empirical analysis. RTI International, March 2006.

exploited a powerful framework for balancing collective versus individual interests. Second, just as managing disease is an important goal of public health, managing insecurity is an important goal of public cybersecurity. We thus advocate that the siren call for the production of "secure" systems and networks be augmented with a mandate to manage the inevitable insecurity that comes from the constant vulnerabilities and adversaries we face.

The goals of public cybersecurity focus on the collective. Individual high-consequence systems, such as those that control critical infrastructures, are not singled out. Why not instead just focus our efforts on a seemingly smaller problem—making only the high-consequence systems secure? In public health terms, why not focus only on keeping important [sic] people healthy? Isolation is not a realistic proposition for either. Public health teaches that it is easier to keep specific individuals healthy when everyone is healthy. And so it will be with cybersecurity. If we foster the production of cybersecurity generally and our networks are able to manage insecurity then it will be easier to ensure that our high-consequence systems are secure.

Cybersecurity, like security in so many other contexts, involves trade-offs with other values.[59] So conflicts will have to be resolved between public cybersecurity and other values or interests of specific individuals, entities, and society at large. And a cybersecurity doctrine is obliged to provide principles and processes to negotiate and resolve these. Public health already offers principles to resolve such conflicts for a public good.

First, an individual's decision is most drastically prescribed by the state where that decision might directly impact the health of others, and the state generally lacks the ability to coerce, when the health impact is primarily on the individual. Change "health" to "security" and we obtain sensible guidelines for public cybersecurity.

Second, public health guidance applied to how public cybersecurity should deal with externalities suggests the following.

- The state's obligations and abilities to shape and override private choices should turn on the extent to which they have a direct impact on the security of the public broadly rather than the security of an individual or entity.

- Where their decision may be poor due to a lack of information or a misapprehension of the risks, the state's role should be aimed at facilitating better individual decisions through information or gentle

[59] H. Nissenbaum, Where Computer Security Meets National Security. *Ethics and Information Technology*, Vol. 7, No. 2, June 2005, 61-73.

interventions that influence the perception of risk, not supplanting them.

- Where security choices of the individual will impact the security of others, then the state can use a wider array of tools to alter behavior.
- Even where state action is permissible, impact on other societal values must be considered in choosing among solutions.
- Whenever possible, opt for minimal interventions implemented in a decentralized manner, so as to limit the negative impact they may have on willingness to participate.

11 Conclusions

Cybersecurity is the obstacle to success of the information age. Though the problem resides in technologies, the solution requires policies. It requires intervention in the private choices of individuals, hard trade-offs, and political agreements that could span nations.

We believe that a doctrine of public cybersecurity can be the basis for those policies. The doctrine establishes a framework for state incentives and coercion that we believe is rational, defensible, and legitimate. It alters the focus of cybersecurity to be on the collective rather than on the individual; and it advocates that systems be built with fewer vulnerabilities but acknowledges that systems will have vulnerabilities and must still be resilient in the face of attacks. If adopted, public cybersecurity will reorient public policy and discourse toward the proper goals of encouraging collective action to produce the public good of cybersecurity and managing the insecurity that remains.

Acknowledgments. We benefited from comments on early drafts of this paper and discussions with Marjory Blumenthal, Aaron Burstein, Scott Charney, John Chuang, Dave Clark, Craig Fields, Kelly Garrett, Jens Grossklags, Joshua Gruenspecht, Joseph Lorenzo Hall, Carl Landwehr, Susan Landau, Steve Lipner, Helen Nissenbaum, Shari Pfleeger, Audrey Plonk, Ashkan Soltani, participants at the 2009 Workshop on the Economics of Securing the Information Infrastructure, and attendees at several Daedalus meetings.