Comments on Cybersecurity Research and Development Strategic Plan

David D. Clark (MIT/CSAIL) and kc claffy (UCSD/CAIDA)

July 16, 2015

Comment in response to Request for Information (RFI)-Federal Cybersecurity R&D Strategic Plan, posted by the National Science Foundation on 4/27/2015: `https://www.federalregister.gov/articles/2015/-04/27/2015-09697/request-for-information-rfi-federal-cybersecurity-randd-strategic-plan`. *Background on authors: David Clark (MIT Computer Science and Artificial Intelligence Laboratory) has led network architecture and security research efforts for almost 30 years, and has recently turned his attention toward non-technical (including policy) obstacles to progress in cybersecurity through a new effort at MIT funded by the Hewlett Foundation. kc claffy (UC San Diego's Center for Applied Internet Data Analysis (CAIDA)) leads Internet research and data analysis efforts aimed at informing network science, architecture, security, and public policy. CAIDA is funded by the U.S. National Science Foundation, Department of Homeland Security, and CAIDA members. This comment reflects our views and not necessarily those of the agencies sponsoring our research.*

1 Introduction

The RFI asks *"What innovative, transformational technologies have the potential to enhance the security, reliability, resiliency, and trustworthiness of the digital infrastructure, and to protect consumer privacy?"*

We believe that it would be beneficial to reframe and broaden the scope of this question. *The security problems that we face today are not new, and do not persist because of a lack of a technical breakthrough.* Rather, they arise in large part in the larger context within which the technology sits, a space defined by misaligned economic incentives that exacerbate coordination problems, lack of clear leadership, regulatory and legal barriers, and the intrinsic complications of a globally connected ecosystem with radically distributed ownership of constituent parts of the infrastructure. Worse, although the public and private sectors have both made enormous investments in cybersecurity technologies over the last decade, we lack relevant data that can characterize the nature and extent of specific cybersecurity problems, or assess the effectiveness of technological or other measures intended to address them.

We first examine two inherently disconnected views of cybersecurity, the *correct-operation* view and the *harm* view. These two views do not always align. Attacks on specific components, while disrupting correct operation, may not map to a specific and quantifiable harm. Classes of harms do not always derive from a specific attack on a component; there may be many stages of attack activity that result in harm. Technologists tend to think about assuring correct operation while users, businesses, and policy makers tend to think about preventing classes of harms. Discussions of public policy including research and development funding strategies must bridge this gap.

We then provide two case studies to illustrate our point, and emphasize the importance of developing ways to measure the return on federal investment in cybersecurity R&D.

2 Disconnected views of cybersecurity

One fundamental obstacle to progress is the different conceptions of cybersecurity used by computer scientists vs. policymakers (and users). Computer scientists tend to define security in terms of the correct operation of a system: a *secure system* functions as designed (or *specified*), even when it is under attack. An old saying among security experts goes: "A system without a specification cannot fail; it can only surprise." A more user-centric framing is that a secure system *detects and prevents harms*, which implies that what matters in the use of a system is the outcomes, not behavior of components. Juxtaposing these two framings reveals a tension: "correct operation" of components does not guarantee that harm will not result, and preventing harms does not require that all parts of the system operate correctly. It is not clear that we can start from a harms-based conception of security and derive specifications of components that make up the overall system, especially a system of interdependent technology, infrastructure, and content, composed of and specified by many different groups with different incentives and economic constraints.

- **Overall recommendation: We believe the federal government can increase the effectiveness of its investment in cybersecurity research and technology development by expanding its focus to include multi-disciplinary conceptions of cybersecurity, including mapping and characterizing the roles of mis-aligned economic incentives, externalities, and informed and responsive policy that can promote and measure progress.**

2.1 An architecture-based (layered) view of Internet cybersecurity

The TCP/IP Internet is a *layered platform*[1], meaning it is composed of layers, each of which serves as a platform to provide services to the layer above. These layers are implemented by protocols; and the set of protocol layers is known as the *TCP/IP architecture*. The *network layer* of this architecture moves (IP) packets of data. On top of this layer, people have built general protocols at the *application layer* – the most famous of which is the hyper-text transfer protocol (HTTP) which connects the world wide web (WWW). Many other platform layers exist on top of this layer, including proprietary platforms such as Facebook. In turn, there are specific apps built for use on the Facebook platform. The Internet architecture is a stack of platform layers, each with its own specification. Security functionality in a protocol specification generally means that the layer fulfills its service commitment even while under attack.

We can use this layered framework to guide an analysis of cybersecurity threats, whether they are vulnerabilities of hosts, applications, the communication channel, and network infrastructure. This framing makes it clear that the resilience and robustness of the system relies not on precise specification but on pragmatic balance of effort and investment at each layer, which are parameters that may continually adjust based on the political economy in which infrastructure components are embedded.

The **physical layer** of the Internet is made up of hardware equipment, e.g., links, routers, servers. Routers and servers are computers, and can be placed in physically secure facilities but are still susceptible to remote software-based attacks; links are mostly susceptible to physical attack, e.g., using cutters and explosives. The functional specification of this layer of the Internet is weak: components do what they are designed to do except when they fail. Higher layer (e.g., routing, transport) protocols include functionality to adapt to failures at the physical layer.

The **network layer** is characterized by routers forwarding units of data (called *packets*) from an entry point to an exit point. The exit point is defined by an address specified in the packet header. This weak functional specification is famously known as "best effort". The network may fail to forward a packet, or deliver packets out of order, multiple times, after inexplicable delays, etc.

The **transport layer** provides services that compensate for failures at the network layer. Its protocols include a software module at the two endpoints of a connection that implements the *Transmission Control Protocol (TCP)*, which copes with packet loss in the Internet. TCP numbers the packets, tracks which ones are received, resends lost ones, and passes them up to the application layer in the correct order.

[1] k. claffy and David Clark, "Platform Models for Sustainable Internet Regulation", Journal of Information Policy, vol. 4, pp. 463–488, Sep 2014. http://www.caida.org/publications/papers/2014/

Because higher layers are designed (specifications) to accommodate failures at lower layers, characterizing the *security* of the network infrastructure requires looking beyond the loose specifications of these layers.

Over time, as engineers have fixed vulnerabilities of lower protocol layers, attacks have moved up the layers of the TCP/IP stack. Many security problems on the Internet today arise at higher layers: naming, routing, and the application layer. The **application layer** describes the software through which most of us experience the Internet: the Web, email, computer games, audio and video conferencing, streaming video. Application-layer security concerns include spam, email phishing, malware downloaded by an unsuspecting user from a malicious web site, etc. Applications that perform such mischief may perform exactly according to its technical specification, but induce harms to user, i.e., they are "insecure by design". Why, for example, does our browser download and execute software applications that can cause us harm? Such vulnerabilities are not accidental, they are a consequence of features added to web software that greatly enhance the functionality (and ease of development) of web pages and services when used correctly. These features are too appealing in terms of functionality and convenience, so we cannot reasonably posit or propose their removal. But research funding agencies can prioritize research that acknowledges their pervasive usage and reconceptualizes application design around this reality.

More precisely, application designers operate in a four-dimensional space where they balance security, features, cost and usability against each other. But there is no reason to believe that current application design practices are anywhere near the intrinsic frontier in this tradeoff space. Analyzing applications requirements and tradeoffs and trying to generalize from them may reveal new concepts for more secure applications. One example is linking the use of risky features to the level of trust between communicating parties, e.g., opening email attachments only from senders known to the receiver.

- **Recommendation: The software engineering and application research community should be challenged to explore new concepts of application design that improve security without degrading other design dimensions.** One positive result of these efforts would be a set of *application design patterns* that reduce the need to resolve these tradeoffs from scratch for each new application.

As application code becomes more secure, the attacks evolve to directly target the user, e.g., exploiting risky modalities in applications to fool users into executing an unsafe action such as downloading and executing code or going to an malicious website and entering information. There has been progress in making host operating systems more robust as well as more helpful in protecting applications running on the host. *Sandboxing* places application code in a confined environment before it interacts with the network, and discards this environment at the end of the interaction, thus discarding in passing any malware or other modifications that may have resulted from the interaction.

- **Recommendation: Research programs should emphasize the importance of safe execution environments, and encourage operating system designers to work with application designers to provide execution environments that protect against harmful effects of running insecure applications.**

This layered approach also informs consideration of security against attacks on the communication channel itself, often characterized as having three sub-objectives: confidentiality, integrity and availability (CIA). Information should not be disclosed except to parties authorized to see it, information should not be corrupted, and it should be available. Although the IP packet forwarding layer does not provide confidentiality or integrity (routers may copy or modify packets), applications can use encryption to achieve these two objectives. But effective use of encryption assumes robust key management, which is a persistent challenge today, both at the local level (protection of private keys in individual systems) and at the system level (PKI systems that can function in a world lacking global trust).[2] The current problems with today's Certificate Authority system illustrate the challenge of a using a competitive industry structure to build a global trust ecosystem

[2]Crypto-currencies bring an increased urgency to this line of research. A user of BitCoin must protect his private key, since loss of the private key allows theft of the coins. We should thus expect regular attacks on systems of users with significant crypto-currency wealth.

composed of mutually untrusting (and variably trustworthy) actors. The controversy over surveillance to support national security, which includes traffic analysis capabilities at multiple layers of the architecture, provides another example of the tension between rights of the citizen and the rights of the state in trying to achieve a tolerable balance of varied conflicting interests in pursuit of a secure and trustworthy cyberspace.

There is a related tension between protecting the communication among trusting parties and protecting a user from attack by another. Cryptography assumes mutually trusting actors at the endpoints, which is not a valid assumption for many Internet transactions, e.g., email or web browsing. By analogy, if trusting parties want to send a private letter, they want assurances that the letter is not opened in transit. But if recipients think they may get a letter full of anthrax, then their security objective reverses–they want that letter opened and inspected by a trained, trustworthy (and well-protected) intermediary. An encrypted exchange with an untrustworthy party is like meeting them in a dark alley–there are no witnesses and no protections. Crypto systems are an excellent example of research and technology outrunning our collective ability to effectively apply them to pervasive and persistent cybersecurity problems.

- **Recommendation: Research should prioritize trust (and key) management frameworks and systems for use in different trust contexts**. This goal requires moving beyond pursuit of isolated security objectives, and mitigating tensions between conflicting objectives in the overall landscape of cybersecurity.

Encryption schemes to safeguard confidentiality and integrity have the consequence of mapping a wide range of attacks on those objectives into attacks on *availability*, since they halt progress if an attack is detected.[3] If the best we can do using encryption is to turn a range of attacks by untrustworthy actors into attacks on availability, we should explicitly conceptualize an approach to availability.

There are several ways to deal with untrustworthy actors: constrain them, eject them, or avoid them. Constraint is tricky in the context of today's *insecure by design* applications. Ejection sometimes works; an ISP might drop a customer known to send spam. Unfortunately, malicious individuals show great resilience to constraint and expulsion, especially across jurisdictional boundaries. Thus, the more common approach is to formally avoid untrustworthy actors. Organizations coordinate to share "blacklists" of networks that demonstrate evidence of misbehavior, so security-conscious networks can take action to selectively block traffic (e.g., spam) from misbehaving networks.

This shift in mindset – assume the presence of untrustworthy actors but devise mechanisms to detect, localize, and avoid them – can form the basis of a theory of availability. A layered analysis can inform development of this theory: can a layer detect attacks at that layer but targeting a higher layer? For example, routers that drop, add, or change bits in packets do not break the forwarding layer, just the end-to-end communication. Should routers track and share packet counts, or recompute encryption functions (which would require appropriate keys), to detect manipulation of a data flow? The performance cost is daunting, which is why the Internet architecture leaves such responsibility to the endpoints. But in today's Internet, endpoints do not control network routing. If endpoints had such capabilities, the capabilities might themselves become attack vectors. Fundamentally improving availability of the Internet faces this basic conundrum: in general only endpoints can detect attacks on availability, but if they are allowed to reconfigure the network they would be another attack vector. Working around this conundrum is a challenging design problem that would require creation of control structures and functions that build on a diversity of trustworthy elements, tools to manage dependencies among them, and mechanisms to use constraints to compensate for lack of trust. In the meantime, the research community does not have good methods and models to relate different sorts of redundancy to different availability outcomes under different failure scenarios. The range of dynamic and evolving adaptation algorithms across all regions and layers of the Internet renders such methods and models elusive. What we have instead today is a pragmatic engineering approach, which has prevented major disruptions to Internet service, even in the face of major jolts, but the research community has no methodical way to reason about availability.

[3]This observation may explain why so many users deal with dialog boxes warning about potential hazards by clicking the "proceed anyway" option–they want to make progress. Another reason is the often inexplicable content of those warnings..

- **Recommendation: Network architects and security experts should collaborate toward a "theory of availability" relevant to cybersecurity concerns**, rather than just simple failures.

2.2 A harms-based (use-centric) view of cybersecurity

Another way to examine the cybersecurity landscape, more user-centric than component-centric, is based on potential harms to different actors in the system. A harms-based view asks whether and how the network, operating systems, and applications can help detect and prevent resulting harms. Harm prevention is not always possible: evidence may only be observable in retrospect, or the harm may materialize only with some future event, such as when legitimately accessed data is improperly combined with other (perhaps also legitimately obtained) data.

Harms to the users of the system: The individual (or user, or consumer or citizen) has a well-known set of security concerns related to theft or "kidnaped" data (where a malicious actor compromises a user's machine and encrypts data, demanding a ransom to decrypt it), and subsequent unauthorized use of that data. Malware on a home computer can capture and steal passwords, facilitating access to personal information. Loss of privacy is another harm, which can be due to a failure of security, but also due to behavior by legitimate actors who have interests misaligned with those of users, or misuse of data originally obtained for a legitimate purpose.

Harms to the enterprise: An enterprise is arguably a large and complex user, with broader concerns: data destruction, theft, espionage. As lower level protocols and basic defenses have improved, attacks have increased in complexity, often involving multiple steps, each of which circumvents some protection to gain another capability.

Harms to an undertaking: We distinguish an enterprise from an undertaking because for many attacks, the final harm may depend on actions far from the enterprise: to exploit credit cards an attacker must not only steal them but sell them, which may involve overseas transactions with other untrustworthy actors. To assess appropriate investment in security, it is necessary to analyze possible harms, and where those harms can best be mitigated.[4]

Harms to the state: With respect to attacks on a state's military capacity or critical infrastructure, the CIA triad (confidentiality, integrity and availability) may not enable useful assessment of the severity of the threat. Economic harms can include attempts to destabilize or degrade an economy, or espionage that hurts the competitive advantage of the private sector relative to other states. Since different societies have different values, the definition of harm is not universal: one nation may value protection of free speech, another protection of its citizens from exposure to unacceptable content. Similar conflicts can occur even within a society: methods used to pursue national security (surveillance) may conflict with the use of encryption as means to enhance CIA.

Harms to trustworthiness of the Internet: Making a system *trustworthy* is a broader goal: the system will not behave in ways that are adverse to the interests of its users, whether or not the behavior is illegal. Making a system trustworthy requires a focus on harms, not on correct (or legal) operation of a component. Lower layers alone cannot make applications trustworthy; applications and the context in which they are embedded must be designed for trustworthiness. As we have described, the application layer is more complex, and offers more opportunities for malice, misunderstanding and mis-alignment of interests. Many popular applications that invoke controversy about their trustworthiness today are created by commercial actors whose primary motivation is profitability (Facebook, Twitter), which might (for example) involve selling demographic information about users. Balancing this tension between the needs of users and the need to

[4]Current practice with credit card fraud is to punish the victim–the merchant who failed to exercise "sufficient" care.

generate revenue for investors are subjects of law and regulation, as well as a changing landscape of norms and expectations. Another example is spam, which some view as an exercise of free speech, and thus attempts to block it as censorship. This contention reminds us that part of what defines the experience of living in cyberspace is trying to create a trustworthy experience in a context where we must tolerate actors who do not have aligned interests. In this respect, cyberspace and the real world are similar.

The *correct-operation* focus and the *harm-based* focus are not inherently aligned. Attackers are nimble at moving among layers to achieve their goals, so one cannot assign responsibility for good security to a single layer–there is no *security* layer in the Internet. However, understanding security tradeoffs among the layers, some fundamental, some a matter of cost, is essential to understanding how to achieve improved security. This fact captures a key issue with respect to public and private sector investment in cybersecurity R&D. Different layers are controlled by different institutions and actors. Standardization, operation and governance organizations tend to align with layers, so resolution of security problems requires cross-institution conversation, not just cross-layer technical design.

- **Recommendation: The security community should be encouraged to explore the inherently multi-disciplinary harms-based conceptions of cybersecurity, bringing in issues of economics, policy, and mis-aligned incentives.** For example, a project might develop technical mechanisms that limit behaviors highly correlated with a specific class of harm, such as exfiltration of large quantities of data. Another project might track current commercial and non-commercial software practices to mitigate specific harms, so that the research community understands the current landscape of cooperation, incentive, investment and technology. Funding agencies could encourage development of new applications motivated less by centralized economic motivations and thus potentially more trustworthy.

3 Case studies

3.1 Securing interdomain routing in the Internet

The Internet is made up of regions called Autonomous Systems, or ASes. Each AS announces to its neighboring ASes which IP addresses it provides service for, and (in some cases) which addresses it can reach via other ASes. The neighboring ASes in turn pass this information on to their neighbors, and so on. All participating ASes continually propagate these messages to establish and maintain reachability on the Internet. Each such message, as it flows across the global network, accumulates the list of ASes through which a packet can reach the original addresses. There may be many such paths, so a sender picks the path it prefers, or more precisely, each AS computes routes to a particular set of addresses, selects from among the options offered to it, and then offers that "best" option to its neighbors.

Despite over a decade of engineering and standardization effort, and many millions of dollars of federal investment, there are no widespread technical security controls on this mechanism in place today. Any rogue AS can announce that it is a route to any other AS in the Internet.[5] If other ASes believe this announcement, traffic is deflected into that AS, where it can be dropped, examined, or modified. This type of event, called a *BGP hijack*, is not uncommon in the Internet today, and results in failures along all dimensions of CIA. Today, although ISPs attempt to monitor their systems for abnormal behavior, they cannot monitor reachability of their network from every other network on the Internet, so they must rely on end users reporting reachability problems (by phone calls, if their systems cannot influence routing); eventually, perhaps within a few hours, the offending AS is identified and isolated.

An obvious solution – cryptographic signatures on announcements so that they cannot be forged – has two formidable and non-technological barriers: incentive to migrate, and trust/key management. The migration problem is simple: there is no way that everyone will convert to the new scheme at once. Some actors will not invest in upgrading, and will continue to originate unsigned route assertions. Other actors can

[5]This vulnerability was known in 1982, but not then viewed as a security threat (RFC 827, p.32).

either disconnect them (unlikely), or accept the unsigned routes, in which case a malicious actor cannot be distinguished from a lazy actor, and we are essentially no better off. Until the last AS converts, we get little value from the scheme, unless we wrap it in complex high-level systems, such as globally distributed, trustworthy lists of ASes that have converted, so that a router knows which unsigned assertions to accept.

The issue of trust management is more complex. When an AS signs an assertion (for example, when MIT signs the assertion that it is AS 3, and that it hosts a particular set of addresses), it signs that assertion using an encryption key. The technical approach is to use a *public or asymmetric* key system, where MIT uses a private (secret) key to sign the assertion, and a public key it gives to everyone so they can decrypt the assertion and confirm that MIT signed it. But if MIT can issue itself a pair of keys and start signing assertions, a malicious actor could do the same, claiming to be MIT. The technical proposal (Resource Public Key Infrastructure or RPKI) was to create a trusted third party that could confirm, based on its own due diligence, which public key is associated with the real MIT. But why would anyone trust that third party? Such schemes require a *root of trust* – a single node we all trust to tell us which second-level parties to trust, and so on until we get to the trusted party that asserts who the real MIT is.

An engineer might consider this scheme simple and elegant, but it runs into insurmountable political obstacles in the real world. First, what single entity in the world would all the regions of the world agree to trust? When this scheme was proposed, several countries (including Russia) asserted that they would not assent to a common root of trust with the U.S. The agent who has the power to validate these assertions must, almost of necessity, have the power to revoke these assertions. Can we imagine a world in which some international organization, by some sort of vote, revokes its trust assertion about some nation and essentially ejects that region from the Internet? What about those second-level entities, that almost certainly are within some legal jurisdiction and thus presumably subject to the legal regime of that region? So does this scheme make the Internet more stable or less? Once people understood the social consequences of this scheme, there was substantial resistance to deployment. The problem with adding a "kill switch" to the Internet is to control who has access to it.

After ten years and many millions of dollars, we believe investors should consider hedging their bets against the fundamental political intractability of global Internet routing security schemes that assume a root of trust. We can imagine a different perspective – one closer to how we manage trust in the real world – that might serve as a more auspicious direction for federal R&D investment. A different design approach would allow actors (e.g., ASes) to make assertions about who they are, but validate these assertions in a way that makes them hard to revoke. That would solve the "jurisdiction" problem. But if a false assertion ever got started, how could it ever be revoked? Once we grasp the complexity of functioning among actors with quite different incentives and levels of trustworthiness, it is difficult to design a system that is robust at ejecting actors that are "bad" but also robust at not ejecting actors that are incorrectly judged "bad". Management of trust relationships, and the expression and manifestation of those relationships, becomes the defining feature of a successful scheme, not exactly how crypto technology is used. This framing recognizes that the landscape of security is a landscape of trust: regions of mutual trust will be more connected, more functional and more effective, and regions that don't trust each other may still communicate, but with more constraints, more limitations, and perhaps more failures, especially with respect to availability. This pattern will hold within any application that tries to tailor its behavior to the degree of trust among communicating parties, whether to exchange routing information or email.

Using this framing we can derive an alternative scheme to secure the AS routing system, a security architecture that explicitly recognizes its socio-technical embedding. The RPKI scheme described above, with a hierarchy of trusted certifiers and a single root of trust, is technically robust, in that it will always give the right answer *if the trust relations are valid and accepted by all the parties.* But this approach is not socially robust. Consider an alternative approach is less technically robust (one cannot prove that it will give the correct answer under certain assumptions) but more socially robust. Above, we rejected the idea that MIT make up a public-private key pair and start signing its assertion. What would happen if MIT did so? At first, various regions of the Internet might get conflicting assertions, if it happened that there was a malicious actor in the system at the time when the assertions started to be signed. That situation, while not desirable, is what we have today. But over time–days or weeks–it would become clear what key went

with the real MIT. Each AS in the network could learn this for itself, or groups of mutually trusting ASes could cooperate to learn it. Once the other ASes in the Internet have decided which key to trust, they have independent possession of that fact, and there is no authority that can compel a third party to invalidate it. Any AS can decide on its own to stop forwarding traffic to MIT, just as they can today.

What this scheme exploits is not a technical scheme for propagating trust, but a social protocol called "getting to know you", which humans have been running, probably for millions of years. We can be fooled, but in fact we are pretty good at it. And it is simple. It requires no trusted third parties, little administration (except that each AS should try very hard not to lose their own private key) and great adaptability to changes in the landscape of trust.

This case study illustrates that approaches to security that employ encryption must equally attend to larger issues of key management in the context of lack of mutual trust. Their design must be *socially robust*, not just technically robust.

3.2 Distributed Denial of Service attacks

Another problem, particularly for the enterprise, is an attack on availability by flooding a server (or the network immediately serving it) with so much traffic that valid traffic cannot get through. When such denial-of-service attacks use a large set of attacking machines to launch the attack, they are called Distributed Denial of Service (DDoS) attacks. From a technical perspective, they are challenging to mitigate, although an industry has sprung up to help provide this service. From a policy perspective, they are interesting due to the method the attackers use to obtain the attack machines: they insert malware into otherwise innocent end-nodes on the Internet, and then use them to carry out malicious behavior. Machines thus taken over are called *bots*, and the resulting collection a *botnet*. Why can this situation persist? Why are botnets so prevalent that an aspiring villain can rent time on them by the hour from their "owner"?

One answer is technical: across all layers of software on a modern end-node there will always be vulnerabilities that attackers can exploit. Another answer is the policy context: providers of software (e.g., end-node systems) disavow any liability for flaws that facilitate the creation of bots, shifting responsibility away from themselves. Attackers have developed sophisticated ways to exploit these vulnerabilities, including first penetrating poorly managed web sites, and then adding malicious content to web pages served by that web site. Then the so-called *botmaster* waits for end-nodes to make contact with these web sites, which will trigger attack by this content, which infects the end-node with malware. In economic terms, this problem is a *negative externality*. A web site manager who fails to keep his web server software up-to-date (with the latest security patches) may have his site infected with malware, which may cause him no harm unless he suffers loss of reputation or performance. Similarly, a home user may have little interest in becoming a security expert, and if his computer becomes infected with malware that turns it into a bot, it may not harm him if the bot is designed not to draw attention to itself by causing massive performance problems. Removing the malware would cost the user time and frustration, but offer no benefit (and could harm his machine if something goes wrong). So malware persists on the Internet, because the harm inflicted is not directed at the owner with the capability to remove it, but at the ultimate victim of the attack launched from it.

Policy could shift this landscape with more aggressive rules targeted at stamping out botnets. The government could require ISPs to disconnect machines infected with malware. But ISPs do not want responsibilities that impose excessive burdens on themselves, including bad relations with their customers and customer service support requirements that do not generate revenue. This space is a cascade of misaligned incentives and negative externalities.

This case study further illustrates our argument that attention to non-technical concerns must be central to real improvements in cybersecurity. While there may be technical innovations that help to mitigate DDoS attacks, to be effective they must be properly positioned in this larger context of economics, incentive and regulation and take into account the motivations of all the actors concerned in the solution.

4 Assessment of return on investment

We emphasize that any approach to improve cybersecurity should include some approach to measure return on investment–how can we argue that some proposed technology, policy, or combination, will improve, or has improved, the state of cybersecurity. Unfortunately, researchers attempting to understand vulnerabilities or measure the security state of the system face legal issues and threats of reprisal from actors uncomfortable with transparency on security issues. While recognizing the complexity of measurement and causality, federal investment should promote a high standard of practice with respect to these issues.

- **Recommendation: Policy-makers should strive to put in place data gathering protections that facilitate legitimate research in cybersecurity.** Funding for research and data gathering that undertakes to formally assess the state of the system should be a priority, as well as formal efforts to analyze what factors contribute to empirically observable improvements in security.

5 Summary of Recommendations

Below we repeat our list of recommendations covered in this comment.

1. **We believe the federal government can greatly increase the effectiveness of its investment in cybersecurity research and technology development by expanding its focus to include multi-disciplinary conceptions of cybersecurity, mapping and characterization of this larger context, recognizing the roles of mis-aligned economic incentives, externalities, and how informed and responsive policy can promote and measure progress.**
2. **The software engineering and application research community should be challenged to explore new concepts of application design that improve security without degrading other design dimensions.**
3. **Research should prioritize trust (and key) management frameworks and systems for use in different trust contexts**.
4. **Network architects and security experts should collaborate toward a "theory of availability" relevant to cybersecurity concerns, rather than just simple failures.**
5. **The security community should be encouraged to explore the inherently multi-disciplinary harms-based conceptions of cybersecurity, bringing in issues of economics, policy, and mis-aligned incentives.**
6. **Policy-makers should strive to put in place data gathering protections that facilitate legitimate research in cybersecurity.**