How Effective Are the Prevailing Attack-Defense Models for Cybersecurity Anyway?

Daojing He, *South China University of Technology and East China Normal University*

Sammy Chan, *City University of Hong Kong*

Yan Zhang, *Simula Research Laboratory*

Chunming Wu and Bing Wang, *Zhejiang University*

Attack-defense models play an important role in cybersecurity systems' design. After reviewing traditional and prevailing attack-defense models, the authors discuss recently proposed paradigm shifts and how to adopt new models.

The design guidance of cybersecurity attack-defense models has evolved and expanded slowly since the 1960s, and researchers have added more and more defense solutions into cyberspace to enforce those attack-defense models. Despite all these efforts, unfortunately, cyberspace is less secure than it was 50 years ago.[1] That isn't to say that no progress has been made in cybersecurity; on the contrary, its advancement has been tremendous.

In the security community, when we say that a system is secure, we usually mean that the system is secure as long as our adversary model and trust assumptions are satisfied. Thus, a careful balance must be kept when defining the adversary model and trust assumptions. For example, it's assumed in most existing handover authentication methods of mobile networks that the access points are trustworthy and would keep users' privacy-related information confidential. However, because such information is extremely sensitive and coveted by many companies to improve their business, in many cases such an assumption might not be valid. Recently, a novel handover authentication mechanism was proposed without such an assumption.[2]

In recent years, adversaries have found more opportunities for attacks due to the increasing scale of networked devices and new applications, including cloud computing and

1541-1672/14/$31.00 © 2014 IEEE
Published by the IEEE Computer Society

smart grids. Also, adversaries have become more sophisticated, and use more advanced techniques to exploit security vulnerabilities. Yet, new security systems are still designed based on some traditional attack-defense models. Therefore, the cybersecurity community needs to ask whether these models are still valid. With this in mind, here we review some prevailing attack-defense models. We show that each model just provides more opportunities for security failures and requires a paradigm shift to be useful for designing effective security strategies. Accordingly, we report on some recently suggested changes to these models. At the same time, we identify new challenges and suggest directions for future work on attack-defense models.

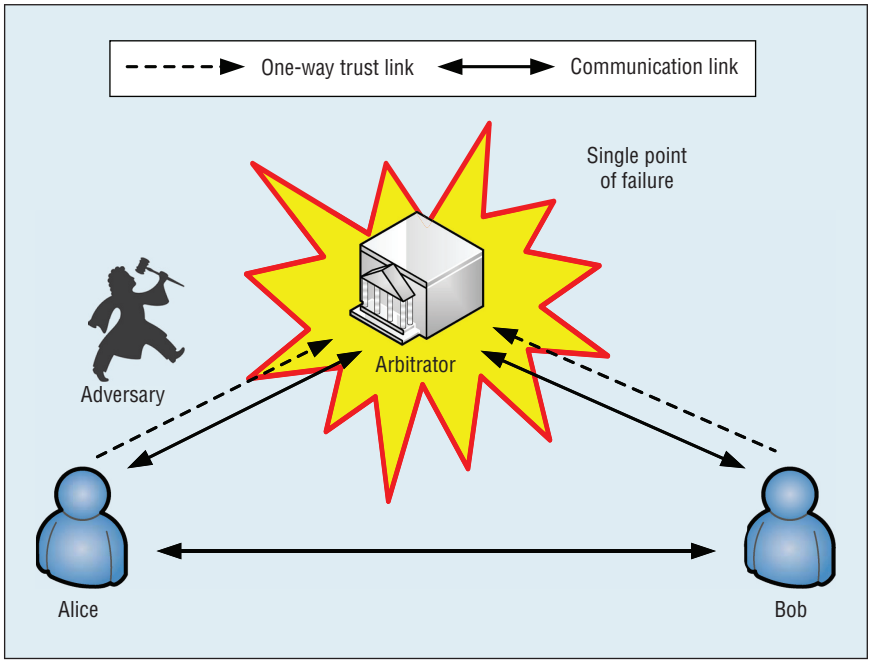

Figure 1. A schematic view of the arbitrated protocol. Because everyone in the network must trust the arbitrator, arbitrated protocols present a vulnerable point for anyone trying to subvert the network.

Single Point of Failure Model

Let's begin by looking at the single point of failure model. This problem occurs in a conventional setting when one operational mistake or a single vulnerability is sufficient to dismantle the entire system. This is the so-called *single point of failure* problem.

Single Point of Failure Problem from the Defender

In situations where the adversary predominates, the adversary can choose any weakness to attack while the defender needs to guard against all possibilities. Often there are many single points of failure in a system.

The centralized control architecture is commonly adopted by many cybersystems or cybersecurity systems due to its simplicity and effectiveness. However, single-point failure is a generic problem of the centralized control architecture. For example, a public key infrastructure suffers from the single point of failure due to the certification authority (CA). An adversary can affect the whole system by compromising the CA key. Another example is that, as Figure 1 shows, arbitrated protocols can present a vulnerable point for anyone trying to subvert the network, because everyone in the network must trust the arbitrator. Besides centralized architectures, distributed and peer-to-peer architectures might also suffer from the single point of failure problem. For example, in Napster (see www.rhapsody.com/napster), which is a peer-to-peer file-sharing system, the process of locating a file is centralized, and this makes it vulnerable to a single point of failure.

The common characteristic of single points of failure is that an attack only needs to exploit a single vulnerability to compromise a system. Thus, defenders have to find and eliminate every possible vulnerability to achieve perfect security. Accordingly, the prevailing assumption is that security must be integrated into every component, because components designed without security might become a point of failure. Consequently, security needs to pervade every aspect of system design. Software engineers have tried to achieve this for the last 50 years and yet the problem still persists, because system designers often fail in detecting all security vulnerabilities. That is, it's quite difficult to make sure security has pervaded every aspect of a specific system design. For example, many serious security vulnerabilities of existing systems, such as the Microsoft Windows operating system, are raised after they're installed. Moreover, with the addition of more and more security patches, the complex interactions between the new and existing codes just make the system more vulnerable to security threats. Also, typically each critical component is replicated to prevent their failures from affecting the entire system. However, replicating the critical component simply increases the number of attack points. Moreover, it also necessitates complex management tasks of keeping security information and strategies consistent. For instance, the keying materials

stored on the replicated components should be consistent.

Shifting Single Point of Failure Problem from Defenders and Adversaries

Such a single point of failure problem should be shifted from defenders to adversaries. That is, it's the adversary who should suffer from the single point of failure problem. In this case, the adversary must fight against any factor that prevents it from achieving its entire attack goal.

Recently, several researchers have begun to study how to shift the single-point failure.[3,4] For example, Sheng Xiao and his colleagues[3] suggest using the adversary's information loss to protect the system. In this approach, two wireless devices observe their regular link layer retransmissions and regularly update their shared secret key by hashing it with data packets that have been aired only once. This method relies on the fact that packet losses by an adversary are inevitable, and in the long run an adversary will miss a packet that has been received by the legitimate device; hence, the key will eventually be securely updated. In this case, the adversary must fight against any factor that might cause information loss. Associating system secrets with the communication process imposes a single point of failure problem to the adversary. In the error-prone wireless communications, the adversary must eavesdrop on all wireless transmissions without any error in order to trace the system secret updates. Xiao and his colleagues have built a prototype to implement their low-complexity algorithms over a wireless local area network (LAN) to show the efficiency in practice.[3] Another example is given in related work,[4] where Xiao and Weibo Gong suggest using the randomness of user mobility in user-cloud communication to create a new credential type as the proof of user identity.

Suggestions on Shifting the Single Point of Failure Problem

To explore the single point of failure of the adversary, the defender can analyze the adversary's possible attack strategy set. Unfortunately, the full range of potential attack scenarios is too rich to generate manually. Moreover, its complexity bars direct analysis and evaluation of the potential impact of alternative countermeasures. To address these challenges, many automated security protocol model checkers, such as On-the-Fly Model-Checker (OFMC), Failures-Divergences Refinement (FDR), Simple Promela Interpreter (SPIN), and Prism[5] have been developed. Based on the technique of model checking, these tools can effectively identify possible attacks on security protocols. For example, using a test-suite of 36 protocols, OFMC is able to identify almost all known attacks and discover a new one in less than one minute of CPU time.[6] Thus, a defender can first use such tools to automatically generate the possible attack strategies of adversaries, and then examine each attack strategy carefully to determine how to make it vulnerable to the single point of failure problem.

Moreover, all existing related works just focus on how to make use of communication randomness to shift the single-point failure to ensure the security of wireless communication systems. Hence, for future work, researchers are encouraged to study other ways to shift the single-point failure problem for different kinds of applications, and then based on these new models, to develop good security approaches for deployment in the real world. This is particularly important as we observe that, for efficiency, the new information and communication technologies tend to incur the single point of failure problem. For example, in the recently proposed software-defined networking (SDN) approach, network intelligence is centralized in the controller. This means that once adversaries can hack into to the control software, they can control the whole network.

For example, some aforementioned solutions[3,4] involve simple techniques to shift the single point of failure problem based on communication randomness. This randomness increases the work factor for an adversary. A weakness of such solutions is that the work factor of the defender is also proportionally increased. Here, we use a body sensor network (BSN) as an example to illustrate another way of shifting the single-point failure without increasing the defender's work factor. A BSN is a wireless network of biosensor nodes collecting personal health information (PHI) and delivering it to a handheld device, called the personal wireless hub (PWH). Secure transmission between biosensors and the PWH is important to ensure safe delivery and preserve privacy of PHI. However, the commonly adopted approach of using an encryption key for a long period of time suffers from the single point of failure problem. This is because once the encryption key is compromised, all previously transmitted PHI can be revealed. To address this issue, the single-point failure problem can be shifted to eavesdropping adversaries, because they face inevitable information loss for a couple of reasons. First, the transmission range of biosensors is small because the PWH is usually nearby but, at the same time, the adversary can't stay too close to the BSN in order not to be detected. Second, it's difficult for the adversary to follow a patient's random movement. A feasible way to shift the single point of failure problem is to constantly update an individual key using the collected PHI. This idea

is based on the observation that a patient's PHI pattern is quite unique and provides a sufficient degree of randomness. Because PHI are readily available, this approach doesn't increase the defender's work factor.

Interaction between Adversary and Defender: Passive Defensive or Active Defense?

Now, let's look at other attack-defense methods.

Passive and Static Defense

Many of the existing attack-defense methods tend to consider the adversary and defender separately. However, the decision to perform an attack is a tradeoff between the gain from a successful attack and the possible consequences of being caught. Therefore, cybersecurity isn't static, but a dynamic confronting process between attack and defense. Unfortunately, traditional defense technologies depend mostly on the knowledge of the adversary's intentions, and hence can't take into account the interactions between the adversary and defender's decisions. Especially, many current defense methods are passive in response to attacks and tend to be host-based. Technologies such as vulnerability scanners, certification and access control, traditional intrusion detection, firewalls, and antivirus software only passively record and prevent attacks by rules and configuration, but don't identify comprehensive intrusive planning and predict the next attack. For example, according to a set of rules, static packet filters on routers allow or deny specific traffic based on information found in packet headers. As a result, traditional defense measures are insufficient to secure a complex cybersystem, especially when current attacks are automated and intelligent.

Moving target defense has a better protection capability than static defense. It addresses the issue that a system operates with a relatively static layout, and the configuration gives adversaries a large window of opportunity to seek potential vulnerabilities by observing and probing the system. It changes the system's configuration dynamically while maintaining the functionality and availability to authorized users. However, existing moving target defense solutions are often designed to operate on a predefined mobility pattern or in a closed control loop. For example, the Intrusion Tolerance by Unpredictable Adaptation (ITUA) project makes use of a set of predefined control loops to respond to specific security events such as intrusion detection. Also, there are still many security weaknesses and efficiency problems in existing defense methods based on the concept of moving target. Here, we consider changing the Internet Protocol (IP) address as an example of moving target defense. To defeat distributed denial-of-service (DDoS) attacks, a simple approach is to invalidate the victim computer's IP address by changing it to a new one. This requires all Internet routers to be informed before the change of IP address is completed and edge routers can drop the attacking packets. Moreover, the adversary can render this solution a futile process through adding a domain name service tracing function to the DDoS attack tools.

Active and Adaptive Defense

In recent years, although there has been some work on active defense technologies, including active intrusion response systems, compared to passive defense, active defense receives considerably little attention owing to the inherent complexity in developing and deploying active defense solutions in an automated fashion. Different from passive defense systems, active systems aim to minimize the damage done by the adversary and/or attempt to locate or harm the adversary. Here, we consider an intrusion-response system as an example. The notification system mainly provides information about the intrusion response. A manual intrusion-response system provides a higher degree of automation than a notification-only system and allows the system administrator to execute a response from a predetermined set of defense actions based on the reported attack information. As opposed to manual and notification solutions, an automatic intrusion response system provides immediate response to the intrusion through an automated decision-making process. Although some current attack-defense systems (such as intrusion-detection systems) are greatly automated, automatic system support is still rather limited. On the other hand, some work has occurred on the development of adaptive defense solutions. During the attack time, these methods can dynamically adjust the defense selection to the changing environment in many ways, including adjustment of system resources devoted to defense, such as activation of additional defense, or consideration of success and failure of defense actions previously made by the system.

Suggestions on Active Defense

In many cases, current passive and active defense techniques act too late, after serious damage is already made. Attack prediction is critical for cybersecurity. It's better to make proactive, real-time defense decisions during an earlier stage of an attack. In such a way, damage can be well-controlled without consuming a lot for resources. Obviously, such prediction is generally difficult and often relies on probability measures and analysis of current users or system behaviors.

In the dynamic face-off of attacking and protecting, each side observes

the behaviors of the other side. To establish an effective security safeguard measure, defenders not only need to know vulnerabilities, but also to learn the intentions of adversaries from the observed behaviors. Only with complete security information and knowledge of intention can the security risk of overall cyberspace be roundly identified and evaluated, and effective defense strategy be made accordingly. For example, by constructing an attack scene and simulating the intruding process to estimate the adversary's actual motivation and predict possible continuing actions, defenders can master the current safety situation and then establish security policy. There has been some work on red teaming and cyberdefense competitions that attempt to achieve such goals. A security analyst's main goal is to look at the range of possible plans/actions an adversary might take. Unfortunately, often the collected intelligence is sparse and it's difficult for the analyst to initially find the adversary's specific intent. Thus, the analyst must look at the range of possible plans/actions an adversary may take. As we mentioned, to solve these challenges, many tools (including some standard automated model checkers such as SPIN and PRISM) have been developed. Although proactive defense is a desired feature, often it's difficult to guarantee 100 percent correctness of the triggered defense behavior.

In the attack and defense context, both the adversary and the defender make individual decisions for their strategies while seeking to maximize their conflicting objectives. This kind of decision is the optimal judgment for adversaries and defenders themselves in consideration of the opponents' strategy spaces. Based on game theory,[7] which is a method of studying strategic decision making, we can view the attack and defense behavior as a game process. Recently, theoretical gaming models have been proposed to evaluate cybersecurity and perform active defense. The interactions between an adversary and defender are modeled as a non-cooperative game. The adversary and the defender are the game players, in which each one tries to maximize its own payoff. The advantage of such a theoretical gaming approach over traditional ones is that it recognizes the strategic interaction between an adversary and a defender. Acquiring proper understanding of how to influence adversaries' behavior is therefore a necessary step toward better security practices.

Obviously, defenders can be active and dynamic to adjust their defensive policies so as to achieve the most effective defense, according to the adversary's different strategies. From game theory, we know that systems of multiple optimizing participants will converge, at best, to a state called Nash equilibrium, which denotes best strategies for both the adversary and the defender. Security decisions—which are arrived at by using such a theoretical gaming mechanism to help the defender allocate limited resources—balance perceived risks, and take the underlying incentive methods into account.

Future Work for Active Defense

You might wonder whether active defense is better than passive defense for all cases. The answer is no. One particular issue with active defense is that users need perfect knowledge of the adversary to make the defense system work better. For example, in most game theoretic approaches, it's assumed that both the adversary and defender know what the other can do. This requirement makes these methods inapplicable in some circumstances. On the other hand, the problem with a game theory point of view is that in some cases, the adversary can use deception such as feint and counter-feint, so that the defender could never be sure of the adversary's intention. In this case, similar to passive defense, current active defense technologies don't work well even if the defender has perfect knowledge of the adversary. Thus, future work on active defense should consider how to solve the aforementioned challenges.

Moreover, a good security protocol should follow multiple attack-defense models. This has the advantage of allowing the protocol to take the best features of different models while downplaying the weaknesses inherent in a given design. Therefore, another direction of future work is the marriage of active defense and shifting the single point of failure (from defenders to adversaries).

Is the Bastion Model or Defense-in-Depth Strategy Sufficient?

Next, let's consider the pros and cons of having a single, strong defensive line versus multiple defense lines.

Bastion Model

Currently, the most common model for cybersecurity is the Bastion model.[1] It uses a single monolithic solution to protect all critical assets. For example, deploying a single firewall between insiders and outsiders of an organization's computing resources is a Bastion design. Its supporters hope this firewall will be the ultimate security filter and prevent anything evil from ever getting to their critical systems. This model is attractive because only one perimeter needs to be the focus of all security efforts, and if the defense wall is strong enough, it can provide good security. An analogy is the crown jewels in the Tower of London—they've never been stolen, yet people know exactly where they are.

Unfortunately, industry experience has shown that in most cases the Bastion model suffers from the single point of failure problem. This is because even a strong defense strategy can only withstand targeted attacks up to a certain point—the point at which it becomes easier to achieve the attack goal by some other way. For example, no matter how good a data encryption system is, it won't prevent an adversary from going through someone's garbage to obtain the information. As stated by a recent report,[8] an analysis of 75 security incidents against control systems between 2002 and 2006 shows that more than half the external attacks come through secondary pathways such as dial-up connections, wireless systems, and mobile devices.

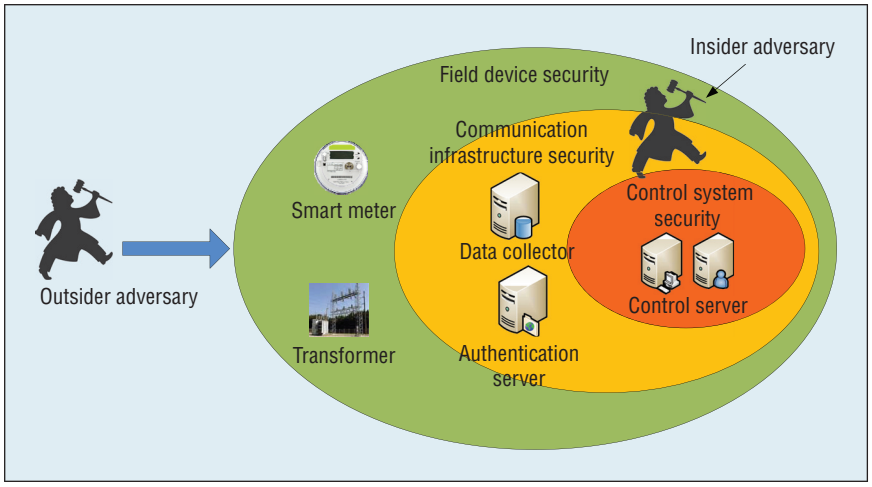

Figure 2. Insider threats to the defense-in-depth approach in a power grid system. A component at a lower level is more easily grasped by adversaries. Higher levels include a smaller number of components, but a fault at a high level results in greater damage than a fault at a lower level.

Defense-in-Depth Strategy

Instead of the Bastion model, some researchers have suggested that *defense-in-depth* is more appropriate. This refers to a defense strategy in which a unit forms multiple defense lines rather than defeating an adversary with a single, strong defensive line. From the point of view of "system reachability," there are various levels of depth in a cybersystem's architecture. Here, we consider a power grid system as an example (see Figure 2). Field devices such as meters and transformers are at the lowest level, communication infrastructures are at the intermediate level, and control systems are the highest level. A component at a lower level is more easily grasped by adversaries. Higher levels include a smaller number of components, but a fault at a high level results in greater damage than a fault at a lower level.

Compared to the Bastion model, a defense-in-depth approach has two advantages. On the one hand, as the adversary must spend many resources and a lot of time to occupy a territory, the attack will lose momentum during its launch. On the other hand, the method doesn't rely on one security mechanism, and thus there's no single point of failure. That is, multiple defense lines ensure that if one is broken, the others aren't, or at least not in the same way. Here, we consider a firewall as an example. The most common location for a firewall is the boundary separating a LAN from the Internet. Because studies show that a large portion of network intrusions are carried out by insiders, companies and organizations start to separate their LAN segments using internal firewalls in much the same way that they protect their LAN from the hostile Internet. The main disadvantage of this solution is that it requires maintaining multiple devices and rule sets. This can add to both complexity and cost. In general, the level of required security will determine the level of defense-in-depth.

Some Suggestions on Bastion Model or Defense-in-Depth Strategy

With the defense-in-depth strategy, a cybersystem will prevent the adversary who stays outside of the system from gaining access. However, this approach can't prevent an adversary residing inside the system from launching an attack. Referring to Figure 2, an adversary might not be an outsider. He or she could be the system administrator or a legitimate system user. Insider attacks are much more serious, especially in systems in which the different participants don't trust one another. For instance, although the defense-in-depth strategy is employed, system administrators still have unencumbered (often unaccountable) access to all functions that could shut down system operation or, worse yet, that could corrupt critical functions in an undetectable manner.

Many techniques such as firewalls aren't able to protect against internal users who might attack a cybersystem. More specifically, the most difficult attack to protect against is legitimate users using their access for illegitimate purposes. Firewalls are helpless against such an attack. An attack that can disguise itself to look like legitimate traffic will be able to bypass many firewalls. To defer and mitigate insider threats, the system should have a fine-grained and accountable access control framework to strengthen its workforce communications, workforce accountability, internal monitoring, privacy-preserving, and information management capabilities.

THE AUTHORS

Daojing He is currently a professor in the Software Engineering Institute, East China Normal University, China. He was an associate professor in the School of Computer Science and Engineering, South China University of Technology, from 2013 to 2014. His research interests include network and systems security, pervasive computing, smart grids, and wireless and sensor networks. He has a PhD in computer science from Zhejiang University. Contact him at hedaojinghit@gmail.com.

Sammy Chan is an associate professor in the Department of Electronic Engineering, City University of Hong Kong. His research interests include telecommunication networks, wireless networks, and performance evaluation. Chan has a PhD in communications engineering from the Royal Melbourne Institute of Technology. Contact him at eeschan@cityu.edu.hk.

Yan Zhang is working with Simula Research Laboratory, Norway. His research interests include resource, mobility, spectrum, energy, and data management in communication networks. Zhang has a PhD in electrical and electronics engineering from Nanyang Technological University, Singapore. Contact him at yanzhang@ieee.org.

Chunming Wu is a full professor in the College of Computer Science, Zhejiang University, China. His research interests include Internet QoS provisioning, reconfigurable networks, network virtualization, and Internet architectures. Wu has a PhD in computer applications from Zhejiang University. Contact him at wuchunming@zju.edu.cn.

Bin Wang is an associate professor in the College of Computer Science, Zhejiang University, China. His research interests include Internet network survivability, high-performance routing, network security, and Internet architectures. Wang has a PhD in communication and information system from the National Digital Switching System Engineering Technology Research Center of China. Contact him at binwang@zju.edu.cn.

Designing such a good access control framework is a challenging issue. Currently, only preliminary solutions are available.[9] A smart grid involves three parties—an electric utility company, consumers, and service providers. The creation of smart grids requires the household devices to be more intelligent and allows service providers to be involved in household power monitoring and management. These two factors bring new challenges in terms of security, privacy, and accountability. First, service providers must cope with free riders and malicious attacks. Second, it's critical to protect consumer privacy, especially in contexts such as smart meter privacy. Depending on the roles and available resources, adversaries can be outsiders (such as eavesdroppers or other consumers) or insiders (such as service providers and the electric utility company). Third, accountability is needed—that is, dishonest consumers and inside attackers should be pinpointed with the cooperation of the law authority. In previous work, we showed that the existing privacy-aware cryptographic primitives such as ring signature, blind signature, and group signature aren't applicable to achieve the aforementioned goal.[9] Further, to solve such a challenge, we proposed modifying the key generation and tracing phases of the existing construction of a group signature to achieve a good access control framework.

Further work should follow some basic rules of security, such as least privilege and separation of duties. For example, the principle of least privilege states that every program and every user of the system should only have access to the least amount of information and functions required to complete the job. Thus, we argue that a better approach might be a defense-in-depth approach against outsider attacks and an access control framework against insider attacks. It's difficult to maintain a system's security if most of the participants (especially system users and administrator) are involved in launching attacks, but with such an access control framework it's possible for legitimate participants to detect that an attack is going on, and which participant is responsible for a specific attack.

To a great extent, the development of security technologies is like an arms race. Because adversaries become more advanced, the attack-defense models used to protect cybersystems must advance at the same rate. This calls for a paradigm shift in the prevailing attack-defense models to make them more capable of repelling adversaries. Although no one can guarantee 100 percent security in cyberspace, we can work toward 100 percent risk acceptance. For example, frauds exist in current monetary systems: cash can be counterfeited and checks altered. Yet these systems are still successful because the benefits and conveniences outweigh the losses. Thus, a good attack-defense model must strike a balance between what's possible and what's acceptable. Work in this direction is still in its early stages, and we hope this article will stimulate further interest from researchers. ■

Acknowledgments

This research is partially supported by the European Commission's Seventh Framework Programme project called End-to-End Virtual Resource Management across Heterogeneous Networks and Services (EVANS; grant 2010-269323), a strategic research grant from the City University of Hong Kong (project 7004054), the National Grand Fundamental Research Program of China (grant 2012CB315903), the Pearl River Nova Program of Guangzhou (grant 2014J2200051), the Shanghai Knowledge Service Platform for Trustworthy Internet of Things (grant ZF1213), and the National Natural Science Foundation of China (grants 61103200, 51477056, and 61021004).

References

1. W. Wulf and A. Jones, "Reflections on Cybersecurity," *Science*, vol. 326, no. 5955, 2009, pp. 943–944.
2. D. He et al., "Secure and Efficient Handover Authentication Based on Bilinear Pairing Functions," *IEEE Trans. Wireless Comm.*, vol. 11, no. 1, 2012, pp. 48–53.
3. S. Xiao, W. Gong, and D. Towsley, "Secure Wireless Communication with Dynamic Secrets," *Proc. IEEE Infocom*, 2010, pp. 1–9.
4. S. Xiao and W. Gong, "Mobility Can Help: Protect User Identity with Dynamic Credential," *Proc. Mobile Data Management*, 2010, pp. 378–380.
5. M. Kwiatkowska, G. Norman, and D. Parker, "PRISM: Probabilistic Model Checking for Performance and Reliability Analysis," *ACM Sigmetrics Performance Evaluation Rev.*, vol. 36, no. 4, 2009, pp. 40–45.
6. D. Basin, S. Modersheim, and L. Vigan, "An On-the-Fly Model-Checker for Security Protocol Analysis," *Proc. Esorics*, 2003, pp. 253–270.
7. M. Wooldridge, "Does Game Theory Work?" *IEEE Intelligent Systems*, vol. 27, no. 6, 2012, pp. 76–80.
8. E. Byres, "Cyber Security and the Pipeline Control System," *Pipeline & Gas J.*, vol. 236, no. 2, 2009, pp. 58–59.
9. D. He et al., "Secure Service Provision in Smart Grid Communications," *IEEE Comm.*, vol. 50, no. 8, 2012, pp. 53–61.