DOI:10.1145/3277564

Steven M. Bellovin and Peter G. Neumann

Inside Risks
The Big Picture

A systems-oriented view of trustworthiness.

PREVIOUS *COMMUNICATIONS* INSIDE RISKS columns have discussed specific types of risks (to safety, security, reliability, and so on), and specific application areas (for example, critical national infrastructures, election systems, autonomous systems, the Internet of Things, artificial intelligence, machine learning, cybercurrencies and blockchains—all of which are riddled with security problems). We have also considered risks of deleterious misuses of social media, malware, malicious drones, risks to privacy, fake news, and the meaning of "truth." All of these and many more issues must be considered proactively as part of the development and operation of systems with requirements for trustworthiness.

We consider here certain overarching and underlying concepts that must be better understood and more systematically confronted, sooner rather than later. Some are more or less self-evident, some may be debatable, and others may be highly controversial.

- A preponderance of flawed hardware-software systems, which limits the development of trustworthy applications, which also impedes accountability and forensics-worthy rapid identification of culprits and causes failures.
- Lack of understanding of the properties of composed systems. Components that seem secure locally, when combined, may yield insecure systems.
- A lack of discipline and constructive uses of computer science, physical science, technology, and engineering, which hinders progress in trustworthiness, although new applications, widgets, and snake-oil-like hype continue apace without much concern for sound usability.
- A lack of appreciation for the wisdom that can be gained from science, engineering, and scientific methods, which impedes progress, especially where that wisdom is clearly relevant.
- A lack of understanding of the short-term and long-term risks by leaders in governments and business, which is becoming critical, as is their willingness to believe that today's sloppy systems are good enough for critical uses.
- A widespread failure to understand these risks is ominous, as history suggests they will pervasively continue to recur in the future.
- A general lack of awareness and education relating to all of these issues, requiring considerable rethinking of these issues.

Background

Progress toward trustworthy systems for critical security uses has been very spotty. For example, several National Academies of Science Computer Science and Technology Board studies have examined issues relating to computer and network security[4,6,11] and cryptography,[5] with extensive conclusions and recommendations that seem to have been widely ignored, or not farsighted

CREDIT TK

enough, or possibly both. Other studies have examined some of the implications of using cryptography,[1,2,7] where again related problems keep arising. Cryptography is an enormously useful concept for achieving trustworthy systems and networks; unfortunately, its effectiveness can be severely limited if it is not implemented in systems with sufficient trustworthiness. Thus, it is a trustworthiness enhancer, but cannot be relied on by itself to enable trustworthy systems and networks.

Total-System Trustworthiness

Trustworthiness is a total-system problem. That is, trustworthiness must consider not just attributes of individual elements, but also how they compose and interact. It is not uncommon for systems to fail even when every individual component is correct and seems locally secure. For example, the composition problem may be as simple as having different notions of the behavior of a particular interface—where each component might assume the other does input validation—or as complex as subtle, time- and input-dependent misbehavior under unusual circumstances. Dependencies on flawed hardware must also be considered, such as the recent speculative-execution and out-of-order execution attacks (for example, Spectre/Meltdown[14] and Foreshadow/Foreshadow-NG vulnerabilities.[15]

The so-called "Martin Luther King Day meltdown" of the AT&T long-distance network in 1990 is a classic example of the latter. There was a flaw in the recovery code when a phone switch rebooted and resumed normal operation. If a neighboring switch received two incoming calls within 1/100 of a second thereafter, it would crash. This, of course, triggered the same failures in its neighbors, iteratively throughout half a day.[8]

With so many known vulnerabilities, and new ones continually being discovered, it is obvious that defenses are often overwhelmed. For example, the Common Vulnerability Enumeration (mitre.cve) now includes more than 105,000 vulnerabilities—almost 11,000 since the beginning of 2018.

More recently, consider the Foreshadow/L1 Terminal attacks on SGX discussed at USENIX Security 2018, and subsequently discovered Foreshadow-NG vulnerabilities,[13,15] which broadly affect VMs, VMMs, operating systems, and SMM memory. The NG (next-generation) paper has attacks that "completely bypass the virtual memory abstraction by directly exposing cached physical memory contents to unprivileged applications and guest virtual machines." These attacks appear to be very serious.

Overall, there are no simple solutions. Precision in interface definition is one obvious approach, although obscure cases are difficult to specify—for example, call-arrival rate at a critical time.

Trustworthiness Also Must Respect Human Behavior

Achieving trustworthiness in complex systems also depends critically on the people involved throughout system development and use. Many systems have poorly defined functional and behavioral requirements—if any. System architectures seldom reflect critical requirements, and implementations seldom adhere to those requirements or design specifications. Formal methods have significant opportunities to improve trustworthiness, but are challenging to use coherently. In operation, user wisdom and sensible behavior are often assumed (instead of building people-tolerant systems), and the creativity and power of malicious misuse and malware are inadequately considered. Thus, trustworthiness must anticipate all sorts of human behavior, as well as environmental disruptions. In essence, achieving trustworthiness is very complex, and attempts to simplify it are generally fraught with vulnerabilities.

Future Directions for Systems Research and Development

A research program in systems poses many challenges. The most difficult is one of definition: What is systems research? What constitutes real innovation? Merely having multiple components is necessary, but not sufficient. Rather, what is needed is a demonstration that new techniques either contribute to the security of the full system or let us better evaluate security. Indeed, some early projects might simply be intended to

Achieving trustworthiness is very complex, and attempts to simplify it are generally fraught with vulnerabilities.

better define the problem and lay out a suitable research agenda.

One vital approach would be a unified theory of predictable subsystem composition that can be used to develop hardware-software systems for a wide range of applications out of demonstrably trustworthy components. Formal methods could be useful selectively. What is essential, though, is that the properties being composed are actually useful in real-world systems.

However, systems design is not a formal discipline today. Therefore, carefully documented open success stories that illustrate the power of an approach are also acceptable, especially if they enable constructive opportunities for the future.

On a smaller scale, developing mechanisms and tools that advance the goal of secure systems would also be useful. Thus, a scheme that provides strong protection for cryptographic keys while still leaving them useful for authorized uses is valuable.[3] This may be facilitated by specialized hardware—if that hardware is trustworthy (including available as needed). Thus, a variety of clean-slate hardware architecture specifications that can be implemented by multiple organizations and that can facilitate total systems that are much more trustworthy would also be useful. Again, formal methods could be useful selectively to prove critical properties of some of the specifications.

Conclusion

Research and its funding have often failed us. There is too much focus on narrow problems—point solutions to point problems—and too little effort devoted to systems aspects of solutions that include considerations of human behavior. Furthermore, many problems discussed long ago[8,9] still have not been adequately addressed today. In addition, underlying principles for trustworthy systems have been posited since the 1960s and recently revisited, but widely ignored in practice.[10] A recent book also has more relevant suggestions for the future.[12]

It is time to get serious about the dearth of trustworthy systems and the lack of deeper understanding of the risks that result from continuing on a business-as-usual course. C

References

1. Abelson, H. et al. The risks of key recovery, key escrow, and trusted third-party encryption. *World-Wide Web Journal 2*, 3 (Summer 1997), 241–257.
2. Abelson, H. et al. Keys under doormats: Mandating insecurity by requiring government access to all data and communications. *Journal of Cybersecurity 1*, 1 (Nov. 2015), Oxford University Press; http://www.cybersecurity.oxfordjournals.org/content/1/1/69
3. Bellovin, S.M. The key to the key. *IEEE Security and Privacy 13*, 6 (Nov.–Dec. 2015), 96–96.
4. Clark, D.D. et al. *Computers at Risk: Safe Computing in the Information Age.* National Research Council, National Academies Press, Washington, D.C., 1990.
5. Dam, K.W. and Lin, H.S., Eds. Cryptography's role in securing the information society. National Research Council, National Academies Press, Washington, D.C., 1996.
6. Goodman, S.E. and Lin, H.S., Eds. Toward a safer and more secure cyberspace. National Research Council, National Academies Press, Washington, D.C., 2007.
7. Landau, S. et al. *Codes, Keys, and Conflicts: Issues in U.S. Crypto Policy.* (ACM sponsored study), 1994.
8. Neumann P.G. *Computer-Related Risks.* Addison-Wesley and ACM Press, 1995.
9. Neumann, P.G. Principled assuredly trustworthy composable architectures, final report. SRI International, 2004; http://www.csl.sri.com/neumann/chats4.pdf
10. Neumann, P.G. Fundamental trustworthiness principles in CHERI. In *New Solutions for Cybersecurity*, MIT Press, Cambridge MA, 2018.
11. Schneider, F.B. and Blumenthal, M., Eds. *Trust in Cyberspace.* National Research Council, National Academies Press, 2101 Constitution Ave., Washington, D.C., 1998.
12. Shrobe, H. et al., Eds. *Solutions for Cybersecurity.* MIT Press, 2018.
13. Van Bulck et al. Foreshadow: Extracting the keys to the Intel SGX kingdom with transient out-of-order execution. *USENIX Security* (Aug. 14–17, 2018); http://foreshadowattack.eu/
14. Watson, R.N.M. et al. Capability hardware enhanced RISC instructions (CHERI): Notes on the Meltdown and Spectre attacks. University of Cambridge Technical Report 916, 2017; http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-916.pdf
15. Weisse, O. et al. Foreshadow-NG: Breaking the virtual memory abstraction with transient out-of-order execution (Aug. 14, 2018); http://foreshadowattack.eu/.

Steven M. Bellovin (smb@cs.columbia.edu) is a professor of Computer Science at Columbia University, and affiliate faculty at its law school.

Peter G. Neumann (neumann@csl.sri.com) is Chief Scientist of the SRI International Computer Science Lab, and moderator of the ACM Risks Forum. Both Peter and Steven have been co-authors of several of the cited NRC study reports, and co-authors of *Keys Under Doormats.*