Actual Causes of Security Violations (Extended Abstract)

Anupam Datta
Carnegie Mellon University
Email: danupam@cmu.edu

Dilsun Kaynar
Carnegie Mellon University
Email: dilsunk@cmu.edu

Divya Sharma
Carnegie Mellon University
Email: divyasharma@cmu.edu

Arunesh Sinha
Carnegie Mellon University
Email: aruneshs@cmu.edu

I. Extended Abstract

Accurate *blame assignment* for security violations is essential in a wide range of settings. For example, protocols for authentication and key exchange, electronic voting, auctions, and secure multiparty computation (in the semi-honest model) ensure desirable security properties if protocol parties follow their prescribed programs. However, if they deviate from their prescribed programs and a security property is violated, determining which agents should be blamed and appropriately punished is important to deter agents from committing future violations [1]. Our vision is that *actual causation* (i.e., identifying which agents' deviations caused a specific violation) is a useful building block for blame-assignment. The central contribution of this work is formalizing and reasoning about actual causation in decentralized multi-agent systems, in particular, to formalize programs (rather than events) as actual causes of security violations and to deal with non-deterministic systems.

Actual causation has been extensively studied in philosophy, law and artificial intelligence, where much of the definitional activity has centered around formalizing what it means for event c to be an actual cause of event e [2], [3]. However, prior work on blame-assignment in the cryptography and computer security literature has not formally studied this concept [4], [5], [6]. Many proposals regard deviation from protocol as a sufficient basis for blame-assignment (i.e., holding an agent accountable) while some others observe that deviation alone is not sufficient because some forms of deviance may be "irrelevant" to the violation.

Our thesis is that accountability is not a trace property since inferring whether an agent's deviation on a log is blameworthy requires analyzing alternative settings where the agent might have behaved differently. We can use the absence or presence of violations in those settings to make a judgment about blame. Counterfactuals defined in analytical philosophy provide a natural way to reason about alternative scenarios and these ideas have previously been applied for deterministic systems. However, executions in security settings involve interactions among concurrently running programs in the presence of adversaries, with non-deterministic scheduling of program reductions. This work proposes a formal definition of *programs as actual causes* considering such security settings.

Definition outline: Let l be a log of executed program expressions. All prescribed programs are supposed to collectively satisfy a security property. Let φ_V be the logical formulation of a violation of the security property. The basic idea is to partition the concurrently executing programs on l into two sets X and Y such that we can establish programs in X to be an actual cause of the violation φ_V, while implicitly establishing the irrelevance of programs in Y to the violation. Informally, the properties that need to be established to conclude that X is an actual cause of φ_V on log l, are:

1) *Occurrence*: The violation occurred on the log l.
2) a) *Necessity*: When prescribed programs X' and Y' are executed, instead of X and Y respectively, then no resulting trace satisfies φ_V.
   b) *Sufficiency*: If the programs in X and Y' are run concurrently, all of the resulting traces satisfy φ_V.
3) *Minimality*: No proper subset of X satisfies all three conditions above.

We are currently using our causal analysis techniques to analyze protocols for addressing weaknesses in the existing public key infrastructure [7].

References

[1] B. Lampson, "Computer security in the real world," *Computer*, vol. 37, no. 6, pp. 37 – 46, june 2004.

[2] J. Pearl, *Causality: models, reasoning, and inference*. Cambridge University Press, 2000.

[3] J. Y. Halpern and J. Pearl, "Causes and Explanations: A Structural-Model Approach. Part I: Causes," *British Journal for the Philosophy of Science*, vol. 56, no. 4, 2005.

[4] J. Feigenbaum, A. D. Jaggard, and R. N. Wright, "Towards a formal model of accountability," in *Proceedings of the workshop on New security paradigms*. ACM, 2011.

[5] R. Küsters, T. Truderung, and A. Vogt, "Accountability: Definition and Relationship to Verifiability," in *Proceedings of ACM Conference on Computer and Communications Security*, 2010.

[6] M. Backes, A. Datta, A. Derek, J. C. Mitchell, and M. Turuani, "Compositional analysis of contract-signing protocols," *Theor. Comput. Sci.*, vol. 367, no. 1-2, pp. 33–56, 2006.

[7] J. Clark and P. C. van Oorschot, "SoK: SSL and HTTPS: Revisiting Past Challenges and Evaluating Certificate Trust Model Enhancements," in *Proceedings of the IEEE Symposium on Security and Privacy*, 2013.