Cybersecurity as an Application Domain for Multiagent Systems

Munindar P. Singh
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206, USA
singh@ncsu.edu

ABSTRACT

The science of cybersecurity has recently been garnering much attention among researchers and practitioners dissatisfied with the ad hoc nature of much of the existing work on cybersecurity. Cybersecurity offers a great opportunity for multiagent systems research. We motivate cybersecurity as an application area for multiagent systems with an emphasis on normative multiagent systems. First, we describe ways in which multiagent systems could help advance our understanding of cybersecurity and provide a set of principles that could serve as a foundation for a new science of cybersecurity. Second, we argue how paying close attention to the challenges of cybersecurity could expose the limitations of current research in multiagent systems, especially with respect to dealing with considerations of autonomy and interdependence.

Categories and Subject Descriptors: I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence—[Multiagent systems] H.1.0 [Information Systems]: Models and Principles—[General] K.0 [Computing Milieux]: General

General Terms: Security.

Keywords: Cybersecurity, Norms, Trust, Organizations.

1. INTRODUCTION

Current cybersecurity practice conveys an ad hoc flavor—find a vulnerability; patch it; find the next vulnerability; and so on. The net result of such reactive practice is that even as our society comes to rely increasing upon computing, it suffers mounting losses from successful attacks as well as indirect losses in opportunity in dealing with potential attacks on an ad hoc, case-by-case basis. The unfortunate fact is that we lack in our scientific understanding of cybersecurity in order to tackle these challenges in a principled manner.

The last few years have seen a growing push within the research community to develop a *science of security*—to contrast with engineering of solutions to specific problems. Leading funding agencies, such as the US National Science Foundation and the US Department of Defense, have initiated research programs promoting the study of security as a science [36]. A new research symposium, HotSoS, has launched recently as well [23]. The motivation behind these efforts is to develop a systematic body of knowledge with strong theoretical and empirical underpinnings that would inform the engineering of secure information systems that can resist not only known but also unanticipated attacks.

Any science needs not only principles but also an approach to systematizing knowledge through empirical investigation. As a science of the artificial [33], cybersecurity needs principles that involve not only IT representations and architectures, but also the organizations and environments in which they are realized. What makes cybersecurity different from computing at large is, first, that security is inherently a human endeavor: not only does it concern humans, but humans are its active players. This recognition is leading to approaches that apply insights from psychology to understand user behavior, e.g., regarding privacy [12]. However, this body of work primarily seeks to map understanding of humans to user interface design [40] with so far only limited representation of their social relationships and interactions.

Second, cybersecurity fundamentally presupposes an open system. If a system could be perfectly circumscribed there would be no security challenges beyond ensuring its correctness or integrity: after all, every successful attack involves the violation of some assumption where the attacker does not play according to the rules. Of course, even a closed system can be so complicated that ensuring its correctness may not be feasible: its users may act in unexpected ways and take it outside its designed operating range: we consider these as violations of assumptions and thus a form of openness. The open nature of the system means that the participants and their actions are not known ahead of time. However, computing as a discipline carries a strong prejudice toward dealing with closed systems.

The foregoing presents an opportunity for the field of multiagent systems (MAS), especially the subfield dealing with norms and associated topics such as organizations. In particular, we propose that MAS seek to provide a foundation for the science of cybersecurity with a special emphasis on the subfield of normative MAS—or *NorMAS*, for short. In particular, despite progress in cybersecurity on the technical aspects, big gaps remain, especially at the social and human levels. MAS is well-placed to provide the requisite theories and methods and, arguably, more naturally than other computing disciplines. And, where MAS approaches need enhancement to support cybersecurity, it would benefit from stronger results obtained from the concomitant effort.

That is, the relationship between MAS and cybersecurity would be symbiotic and both fields would gain if the MAS field pursues cybersecurity as an application domain.

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May, 4–8, 2015, Istanbul, Turkey.*

2. UNDERSTANDING CYBERSECURITY

Cybersecurity is highly conducive to the multiagent systems way of thinking, arguably more so than any other branch of information technology. This is because not only does cybersecurity involve multiple autonomous participants, it arises as a challenge primarily because of the (potential) divergence of the interests of those participants. There is no viable reduction of cybersecurity problems to a single-agent problem even as a baseline, such as one might posit for cooperative settings.

Figure 1 illustrates the cybersecurity ecosystem in schematic terms. The three components in the middle form what we term the *system*. This figure highlights that there are two main kinds of autonomous entities that we deal with in security. On the left are stakeholders, whose perspective is reflected in the system. On the right are adversaries or attackers—the parties who interfere with the stakeholders. The stakeholders with respect to one system may be adversaries with respect to another: a company may be protecting its secrets even as it spies on others.

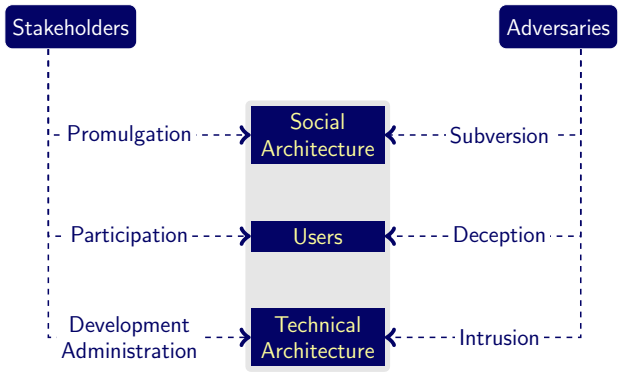

Figure 1: The cybersecurity ecosystem.

The *technical architecture* includes information technology resources as for computing, storage, and communications. The *users* provide the human element of the system: the people within the system who interact with one another as participants of the system and presumably provide value to one another. These include confused and uninformed individuals. The *social architecture* captures the rules of encounter that govern how the participants engage with one another within the system. It would describe the various roles, criteria for admittance, and what a participant may expect from another depending upon their respective roles. The schematic nature of this picture may hide the recursive or compositional nature of the problem. For example, it is possible for an adversary to gain entry into a software development organization via deception and then place Trojans in the technical architecture.

Stakeholders and adversaries interact with the system in three ways corresponding to the main components. Stakeholders develop and administer the technical architecture through activities such as programming, deploying, and configuring modules. Adversaries attack the technical architecture by intruding into it, e.g., by installing malicious software or eavesdropping on communications. Stakeholders participate as users whereas adversaries attack the system by deceptively participating as users, for example, through identity theft. Stakeholders promulgate the social architecture by establishing and revising the rules of encounter whereas adversaries attack the social architecture by subverting it in various ways. For example, a social architecture may involve voting on matters of governance, such as to prioritize resource allocations among users. A potential subversion would be when some users collude through a quid pro quo arrangement as a means to skew the election.

3. WHAT CAN MAS OFFER SECURITY?

Figure 1 highlights the essential idea that the notion of system must be far more expansive than how the term "system" is used in traditional computer science, where it refers usually to an artifact or machine (as in "the system is down"). The traditional conception falls within the technical architecture in Figure 1 and excludes the more subtle aspects, the users and social architectures. Traditional cybersecurity has concentrated on making the technical architecture secure but, though useful, those efforts are inadequate conception for the above reason. In the traditional conception, users reside outside the "system"—in contrast, here the users and their social relationships fall within the system.

The cybersecurity field has (arguably reluctantly) begun to consider challenges that lie outside of the technical architecture. For example, they now consider user training, e.g., to resist phishing attacks, as an important challenge [27, 41]. However, the most significant security challenges arise at the level of the social architecture. Although not articulated in the terms we propose here, the significance has been known from the early days of cybersecurity. For example, a US National Institute of Standards and Technology (NIST) document from 1994 states that "the primary threat to computer systems has traditionally been the insider attack" [5].

It is natural to define insider attacks as a failure to respect applicable norms. Although traditional formal techniques and improved programming discipline can reduce inadvertent vulnerabilities, even when a vulnerability is traced to the technical architecture, it is usually an insider "attack"—in the sense of failing to respect appropriate norms—that is the root cause. Similarly, vulnerabilities due to user behavior, e.g., sharing passwords or failing to patch a firewall, arise from failing to respect appropriate norms. A participant may violate norms because they are not stated, imperfectly stated, or with malice aforethought: a problem with norms that lies at the root of a cybersecurity vulnerability.

The above expansive view of a system makes clearer the avenues through which the AAMAS community can contribute to cybersecurity most naturally. We introduce some ideas next proposing how the normative concepts can provide the elements of a new potential foundations for security.

Governance via norms. We introduce some ideas about policy and governance in sociotechnical systems, approaching these topics from a normative standpoint. We describe how we can characterize a variety of security-relevant behaviors in normative terms touching upon the challenges of accountability [7, 20] and how accountability differs from, yet relates to, mechanisms for monitoring and sanctioning.

The foregoing leads us to advocate a normative view of systems as a basis for the science of security. Specifically, in this broad sense, users and malefactors alike are part of the system. A system thus corresponds to a society, whether the entire human society or, more often, a suitable microcosm. A security property is a norm in this system-as-a-society

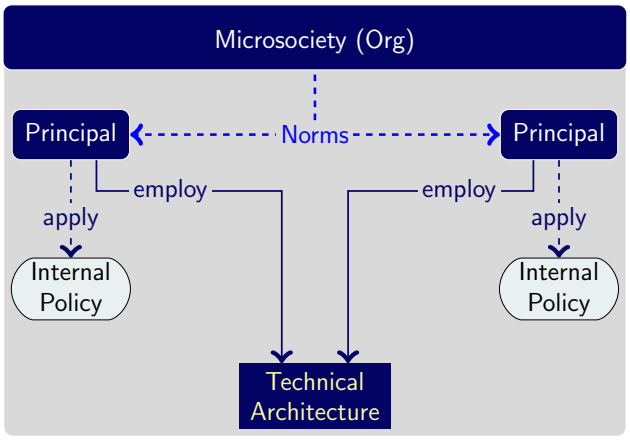

Figure 2: A system is a microsociety.

and a security violation is a violation of some norm. That is, a system's security and vulnerabilities arise not at its perimeter but in its very core.

Figure 2 models a system as a microsociety or an Org [35]. Singh [35] motivates the notion of a sociotechnical system as one involving principals. A principal is a social entity, i.e., an autonomous party. Each principal applies its private policy in determining how to interact. The principals are subject to the norms of their microsociety. They may violate these norms in light of their autonomy but would be subject to other norms, including sanctions, if they do so. The principals interact via and about the technical architecture of the system, which though shown conceptually as one box would in general be distributed.

Norms apply to potentially any security-relevant behavior and are an essential basis for accountability for (security-relevant and other) actions. Norms apply to all members of a microsociety, both stakeholders and adversaries, and can emerge from users sanctioning or imitating others. Although one function of norms and sanctioning is to deter adversaries, a more significant function is to help address the failures in human decision making and social interactions that are often the root causes of security threats. Specifically, the benefits of norms include supporting precise modeling of requirements for governance [15, 24], including specifying organizations with the correct authorizations and powers [35] and supporting accountability and trust [10]. A possible application of norms is in tackling even the minor—but nonetheless frequent and threat-prone—user behaviors, e.g., delaying patching computers or sharing accounts [17].

Governance arises from how the norms of a microsociety are created, manipulated, and exercised (through formal organizational processes or via emergence and diffusion), and considered by the principals as they interact. NorMAS can offer representations for norms and formal techniques for reasoning about norms, judging compliance of principals' actions, and verifying internal policies with respect to norms, e.g., [1, 3, 39]. Artikis et al. [4] specify institutions with associated normative propositions such as powers, permissions, prohibitions, and obligations, and support specifying enforcement policies for potential norm violations.

Trust. Cybersecurity relies upon an underpinning of trust [10]. However, trust is traditionally modeled via an oversimplified mechanism such as certificate chains [25], which merely presume trust and do not capture any of the subtle connotations of trust. NorMAS can offer a deeper understanding of trust incorporating sociotechnical aspects [11, 34] as well as a cognitive modeling of trust [8]. Other MAS research incorporates risk assessments in decision-making [9], discussing how to support policy (i.e., normative) violations when necessary as long as a responsible party would restore the norm after the fact.

Human models. Cybersecurity increasingly relies upon an understanding of user behavior. However, current work is limited to building psychological profiles, e.g., through surveys of users [40]. MAS can help map this idea to computational models of users that are richer in that they incorporate considerations of affect and personality, yet build on strong foundations of psychology research, e.g., [38]. Notable examples of MAS research include [13, 16, 22].

Mechanism design. In general terms, norms are a form of mechanism. However, mechanism design includes explicit economic models that seek to mold the behavior of autonomous principals to accomplish social goals. Some of same motivations apply to mechanism design as to norms. An example of an incentives-based approach is Feigenbaum et al.'s [19] approach based on deterrence for accountability.

4. HOW CAN SECURITY BENEFIT MAS?

In targeting cybersecurity as an application domain, MAS can gain from a motivation and clearer understanding of certain deep research challenges that lie at the heart of MAS.

Directed conception of norms. The study of norms in MAS builds on the informal understanding of norms as social conventions backed up by a social sanctioning process, e.g., [29, 32]. This conception leads to a view of norms as general conditions on system states [2]: liveness (something good happens) and safety (nothing bad happens).

However, it is important representationally and especially for security, that norms *not* be treated as general conditions [35]. Such traditional construals make sense when we are talking about a unitary system owned by one party and operated from the perspective of its owner. When we shift attention to open systems, general constraints make less sense: what is good or bad depends upon whom you ask. Moreover, to accommodate autonomy, we must ground the norms in a notion of accountability and traceability so when a norm is violated, we know who is to be held to account.

Correctness properties in open systems. MAS has largely adopted the formulation of correctness of behavior from conventional computer science for the purposes of verifying how agents deal with norms [26]. These approaches incorporate coalitions but conceptually are oriented toward closed systems where a machine is verified as generating only acceptable computations. In other words, the machine if correct cannot possibly violate the requirements.

However, in an open system, as needed for modeling security, we need to treat norms as requirements that an autonomous agent could violate. In particular, not every deviation from a norm would be undesirable. For example, an autonomous and cooperative physician may share a patient's private medical record with an unauthorized colleague in order to save the patient's life [28]. Not sharing the data would in fact be a security flaw—a form of misguided denial of service. Figure 3 illustrates this situation. The inner cone

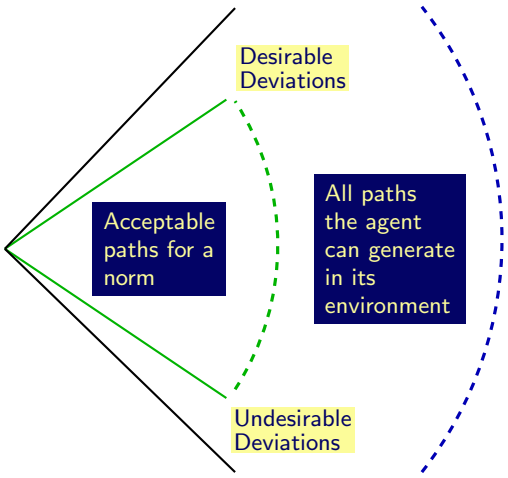

Figure 3: Correctness in the face of autonomy.

refers to the computations that comply with the norms; the outer cone refers to the computations that an autonomous agent can generate. The region between the cones is the set of potential deviations from the norms. Some of those deviations may be desirable, and blocking them a bad idea.

A related challenge is to quantify security in terms of the success and failure of norms, which could adapt notions of utility to produce metrics of value in cybersecurity.

Governance. The understanding of governance in MAS at present lacks an account of how norms can be created and adapted in a context-sensitive yet formal manner. For example, Figure 3 shows some desirable and some undesirable deviations from the norms. Observing such deviations may lead the stakeholders to promulgate new norms that incorporate what were previously desirable deviations and deter participants from making the undesirable deviations by altering the sanction structure.

In general terms, what we need are ways to characterize governance processes by which principals would evolve the relevant norms. However, recent work, e.g., [37], provides some bases that we could expand upon to address this challenge. Another relevant body of work is on formal argumentation, intensively studied in MAS, which has been applied to the configuration of technical policies [31]. Existing treatments of argumentation could be expanded to deal with norms as motivated here.

Trust and human decision making. Tackling cybersecurity from a MAS perspective would lead us to develop rich computational models of user affect [13, 14, 18] and decision making [21]. Recent works in MAS have formulated the bounded rationality of human decision making and how it is influenced by the computationally predictable affective states of users. In particular, a major challenge lies in, first, understanding how user traits (e.g., propensities in appraising situations and trusting others) and states (e.g., current affective state such as confusion or desire to make money) and decision context (e.g., make a purchase before getting on a plane). And, second, in producing agents that build and maintain real-time models of the users they assist and employ such models to guide user behavior. The net benefit would be to enhance cybersecurity.

5. DISCUSSION

In general, cybersecurity could involve virtually every part of MAS research. An incomplete list of relevant AAMAS topics includes agent-based simulation, game theory, social choice, constraints, multiagent learning, argumentation, human modeling, and multiagent system engineering. Our emphasis on norms is partly justified by their promoting autonomy and interdependence, marrying technical, social, and human concerns in cybersecurity, and enhancing requirements elicitation and explanation generation for stakeholders. And, of course, our focus on norms partly reflects our taste in research.

Some ideas relating to norms and institutions are recognized in cybersecurity but are usually not treated from a computational perspective. For example, Dong et al. [15] treat security itself (the concept of making a system secure) as either a good or a pool resource to be acquired through suitable investments. Viewing security as a club good [6] suggests that users, as members of a club, can impose requirements upon each other and on technology vendors to promote security upon pain of exclusion from the club. Viewing security as a common pool resource [30] suggests strategies for communities to enforce standards such as installing and sharing correct patches.

Outline of a plan. To apply MAS on cybersecurity will involve (1) developing mathematical (e.g., logical or decision-theoretic) modeling frameworks and formalizing properties (such as resilience) and associated decision procedures; (2) understanding real-life systems; (3) exercising those frameworks on those systems for validation and dissemination. Promulgation and subversion would involve addressing problems in, e.g., social choice and argumentation. Participation and deception would involve addressing problems in learning, emergence, and human decision making. We would deemphasize technical architecture as it is not MAS-specific.

Promoting closer ties between cybersecurity and MAS research will take a combination of demonstration (developing concrete MAS-centric problem formulations of security problems), dissemination (showing how the AAMAS and cybersecurity communities can help each other), and advocacy (encouraging participation by colleagues).

Essential tension. This paper's conceptual model brings out the essential interplay between the technical and social architectures. The main challenge is that (1) the social architecture relies upon the technical infrastructure not being corrupted, e.g., to reduce deception and promote traceability, and (2) the technical architecture relies upon the social architecture not being corrupted assure its integrity.

Conclusions. Both cybersecurity and MAS research and practice are at moments in history of their respective developments that they can each benefit from the confluence. MAS can fill the void of principles that cybersecurity needs filled to become a science; cybersecurity can bring a set of new challenges in research and practice to the fore that can help reinvigorate MAS research and help break down artificial disciplinary boundaries in MAS between formal methods, decision theory, norms, trust, and affective reasoning.

Acknowledgments

Thanks to Amit Chopra and the anonymous reviewers for comments. Thanks to the US Department of Defense for support under the Science of Security Lablet grant.

6. REFERENCES

[1] N. Alechina, M. Dastani, and B. Logan. Reasoning about normative update. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 20–26, Beijing, Aug. 2013.

[2] N. Alechina, M. Dastani, and B. Logan. Norm approximation for imperfect monitors. In *Proceedings of the 13th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 117–124. IFAAMAS, 2014.

[3] G. Andrighetto, G. Governatori, P. Noriega, and L. W. N. van der Torre, editors. *Normative Multi-Agent Systems*, volume 4 of *Dagstuhl Follow-Ups*. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013.

[4] A. Artikis, M. J. Sergot, and J. V. Pitt. Specifying norm-governed computational societies. *ACM Transactions on Computational Logic*, 10(1):1:1–1:42, Jan. 2009.

[5] L. E. Bassham and W. T. Polk. Threat assessment of malicious code and human threats. http://csrc.nist.gov/publications/nistir/threats/threats.html, Mar. 1994. Computer Security Division, National Institute of Standards and Technology.

[6] J. M. Buchanan. An economic theory of clubs. *Economica*, 32(125):1–14, Feb. 1965.

[7] B. Burgemeestre and J. Hulstijn. Designing for accountability and transparency: A value-based argumentation approach. In J. van den Hoven, P. E. Vermaas, and I. van de Poel, editors, *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, chapter 15, pages 1–28. Springer, Berlin, 2015.

[8] C. Castelfranchi and R. Falcone. *Trust Theory: A Socio-Cognitive and Computational Model*. Agent Technology. John Wiley & Sons, Chichester, United Kingdom, 2010.

[9] L. Chen, J. Crampton, M. Kollingbaum, and T. Norman. Obligations in risk-aware access control. In *Proceedings of the 10th Annual International Conference on Privacy, Security and Trust (PST)*, pages 145–152, Paris, July 2012. IEEE Computer Society.

[10] C. Cheshire. Online trust, trustworthiness, or assurance. *Dædalus, the Journal of the American Academy of Arts & Sciences*, 140(4):49–58, Fall 2011.

[11] A. K. Chopra, E. Paja, and P. Giorgini. Sociotechnical trust: An architectural approach. In *Proceedings of the 30th International Conference on Conceptual Modeling (ER)*, volume 6998 of *Lecture Notes in Computer Science*, pages 104–117, Brussels, 2011. Springer.

[12] L. F. Cranor and S. Garfinkel. *Security and Usability: Designing Secure Systems that People Can Use*. O'Reilly, Sebastopol, California, 2005.

[13] C. M. de Melo, P. Carnevale, S. Read, D. Antos, and J. Gratch. Bayesian model of the social effects of emotion in decision-making in multiagent systems. In *Proceedings of the 11th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 55–62, Valencia, Spain, June 2012.

[14] S. K. D'Mello, A. C. Strain, A. Olney, and A. C. Graesser. Affect, meta-affect, and affect regulation during complex learning. In R. Azevedo and V. Aleven, editors, *International Handbook of Metacognition and Learning Technologies*, volume 28 of *Springer International Handbooks of Education*, chapter 44, pages 669–681. Springer, Amsterdam, 2013.

[15] Z. Dong, V. Garg, L. J. Camp, and A. Kapadia. Pools, clubs and security: Designing for a party not a person. In *Proceedings of the New Security Paradigms Workshop (NSPW)*, pages 77–86, Bertinoro, Italy, Sept. 2012.

[16] H. Du and M. N. Huhns. Determining the effect of personality types on human-agent interactions. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT)*, pages 239–244, Atlanta, Nov. 2013. IEEE Computer Society.

[17] H. Du, B. Y. Narron, N. Ajmeri, E. Berglund, J. Doyle, and M. P. Singh. Understanding sanction under variable observability in a secure, collaborative environment. In *Proceedings of the International Symposium and Bootcamp on the Science of Security (HotSoS)*, Urbana, Illinois, Apr. 2015. ACM. To appear.

[18] J. R. Dunn and M. E. Schweitzer. Feeling and believing: The influence of emotion on trust. *Journal of Personality and Social Psychology*, 88(5):736–748, May 2005.

[19] J. Feigenbaum, J. Hendler, A. D. Jaggard, D. J. Weitzner, and R. N. Wright. Accountability and deterrence in online life (extended abstract). In *Proceedings of the 3rd International Web Science Conference*, pages 7:1–7:7, Koblenz, June 2011. ACM Press.

[20] J. Feigenbaum, A. D. Jaggard, and R. N. Wright. Open vs. closed systems for accountability. In *Proceedings of the Symposium and Bootcamp on the Science of Security (HotSoS)*, pages 4:1–4:11, Raleigh, North Carolina, Apr. 2014. ACM Press.

[21] Y. Gal, B. Grosz, S. Kraus, A. Pfeffer, and S. M. Shieber. Agent decision-making in open-mixed networks. *Artificial Intelligence*, 174(18):1460–1480, Dec. 2010.

[22] N. Guiraud, D. Longin, E. Lorini, S. Pesty, and J. Rivière. The face of emotions: A logical formalization of expressive speech acts. In *Proceedings of the 10th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 1031–1038, Taipei, May 2011.

[23] HotSoS. Symposium and Bootcamp on the Science of Security. http://cps-vo.org/group/hotsos2014, 2014.

[24] S. Ingolfo, I. Jureta, A. Siena, A. Perini, and A. Susi. Nòmos 3: Legal compliance of roles and requirements. In *Proceedings of the 33rd International Conference on Conceptual Modeling (ER)*, volume 8824 of *Lecture Notes in Computer Science*, pages 275–288, Atlanta, Oct. 2014. Springer.

[25] ITU. Recommendation X.509 – information technology - open systems interconnection - the directory: Public-key and attribute certificate frameworks. Article Number E 19454, International Telecommunication Union, Aug. 2001.

[26] P. Kazmierczak, T. Pedersen, and T. Ågotnes. NORMC: A norm compliance temporal logic model checker. In *Proceedings of the Sixth Starting AI Researchers' Symposium (STAIRS)*, pages 168–179, 2012.

[27] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. I. Hong. Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology*, 10(2):7:1–7:31, May 2010.

[28] S. Marinovic, N. Dulay, and M. Sloman. Rumpole: An introspective break-glass access control language. *ACM Transactions on Information and System Security (TISSEC)*, 17(1):2:1–2:31, Aug. 2014.

[29] J. Morales, M. López-Sánchez, J. A. Rodriguez-Aguilar, M. Wooldridge, and W. Vasconcelos. Automated synthesis of normative systems. In *Proceedings of the 12th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 483–489, St. Paul, Minnesota, May 2013. IFAAMAS.

[30] E. Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Number 4 in The Political Economy of Institutions and Decisions. Cambridge University Press, Cambridge, United Kingdom, 1990.

[31] J. Rowe, K. Levitt, S. D. Parsons, E. I. Sklar, A. Applebaum, and S. Jalal. Argumentation logic to assist in security administration. In *Proceedings of the New Security Paradigms Workshop (NSPW)*, pages 43–52, New York, 2012. ACM.

[32] B. T. R. Savarimuthu and S. Cranefield. Norm creation, spreading and emergence: A survey of simulation models of norms in multi-agent systems. *Multiagent and Grid Systems*, 7(1):21–54, 2011.

[33] H. Simon. *The Sciences of the Artificial*. MIT Press, Cambridge, Massachusetts, 3rd edition, 1996.

[34] M. P. Singh. Trust as dependence: A logical approach. In *Proceedings of the 10th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 863–870, Taipei, May 2011. IFAAMAS.

[35] M. P. Singh. Norms as a basis for governing sociotechnical systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):21:1–21:23, Dec. 2013.

[36] SoS. Science of Security: An online community to advanced cyber-security science. http://cps-vo.org/group/SoS, 2014.

[37] N. A. M. Tinnemeier, M. Dastani, and J.-J. C. Meyer. Programming norm change. In *Proceedings of the 9th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 957–964, Toronto, 2010. IFAAMAS.

[38] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, Jan. 1981.

[39] W. W. Vasconcelos, M. J. Kollingbaum, and T. J. Norman. Normative conflict resolution in multi-agent systems. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)*, 19(2):124–152, Oct. 2009.

[40] Y. Wang, P. G. Leon, A. Acquisti, L. F. Cranor, A. Forget, and N. M. Sadeh. A field trial of privacy nudges for Facebook. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, pages 2367–2376. ACM, 2014.

[41] L. Zhang-Kennedy, S. Chiasson, and R. Biddle. Stop clicking on "update later": Persuading users they need up-to-date antivirus protection. In A. Spagnolli, L. Chittaro, and L. Gamberini, editors, *Persuasive Technology*, volume 8462 of *Lecture Notes in Computer Science*, pages 302–322. Springer, 2014.