The Buck Stops Here: Trust Management in Multi-Agent Systems with Accountability

Bilal Khan* Dardo D. Kleiner† David Talmage*

Center for Computational Science
Naval Research Laboratory, Washington D.C.

Abstract

Much of security in multi-agent systems is based on models where each agent declares limits on what other agents are permitted to receive. Traditional systems are engineered to operate without violating their agents' cumulative declared constraints [14, 16].

In contrast, here we consider a trust model that is suited for use by ensembles of closely coupled agents operating in a system supporting agent accountability using audit trails for information flows. In such systems, an agent does not require enforcement of absolute limits on the what other agents receive, but instead seeks assurance that its personal liabilities will never exceed its declared risk tolerance. In short, each agent expects the system to behave in a manner which respects its declared accountability constraints—quantitative limits on what the agent agrees to be held accountable for sending.

This paper outlines a suite of protocols with which a multi-agent system can fulfill the cumulative accountability constraints of its constituent agents, and avoid subjecting any individual agent to greater liability than its declared risk tolerance. The protocols are shown to be efficient in a dynamic network setting, and are analyzed under a comprehensive set of failure models including link delay, link failure, and limited corruption in the control and data processing logic of agents.

1. Introduction

Traditionally, agent security has been approached at the microscopic scale, as a problem of reconciling *pairwise* inter-agent trust with the rendering of agent services [7, 13]. This work, in contrast, considers agent security at macroscopic scales, as a problem of dynamic data filtering in ensembles of cooperating agents [11] within a multi-agent system.

We consider a collection of agents V, cooperating by communicating over a dynamic network of logical connections $E \subset V \times V$. Within the agent network we assume that each datagram p is tagged with an immutable *sensitivity* classification, represented by an m-vector $\bar{\sigma}(p)$ of real values [12]. We shall define a partial order on sensitivity as follows: given two m-vectors $\bar{\sigma}, \bar{\sigma}'$ in $\mathbb{R}^m$, we write $\bar{\sigma} \leqslant \bar{\sigma}'$ if the corresponding ordering holds in *all* m coordinates of $\bar{\sigma}$ and $\bar{\sigma}'$. Larger sensitivity values implicitly mandate greater restrictions on the distribution of a packet.

In multi-agent systems which provide security contracts restricting the content of information flowing *to* individual agents, one interpretation of trust might be to associate with each agent w, the maximum sensitivity of information—say $\tau_1(w)$—that agent w is permitted to *receive* [17]. However, such a model implies that there is global consensus on τ_1—clearly an unreasonable assumption for a dynamic large-scale open multi-agent system. To support interpretation of trust that permits disagreement on trust levels, we could model trust using a pairwise function τ_2, where $\tau_2(v, w)$ specifies the maximum sensitivity of information that agent v wants agent w to *ever receive*. A priori, this model appears to be more flexible, but it is easy to see that a system that satisfies constraints of the second type of model is equivalent to a system based on the first type of model—where for each agent w, $\tau_1(w)$ has been taken to be $\min\{\tau_2(v, w) \mid v \in V\}$. In the dynamic setting then, the difference between the two models amounts to nothing more than maintaining a distributed consensus on the values of τ_1 in terms of the values τ_2. In short, the expressive power of the two trust models is equivalent—no additional flexibility is gained by adopting a pairwise trust function τ_2.

In this paper, we consider a trust model that is suited to open multi-agent systems supporting auditing and accounting of information flows [1, 5, 15].

* Advanced Engineering & Sciences, ITT Industries
† Computer Integration & Programming Solutions, Corp.

is a procedure to for agents to adjust their contextual-accountabilities, even in the presence of dynamic local-accountabilities and a dynamic agent network topology. Finally, we augmented the protocols to withstand fail-stop crashes of agents and total failures of links. We gave a synopsis of possible phenomena that can occur in situations of agent corruption.

Our present software development efforts [8, 9, 10] extend earlier initiatives seeking to harness mobile agents for network management [3, 4, 6] and general information management [2]. In our mobile agent framework, we incrementally construct networks of agents by dragging them from a palette onto a canvas and connecting agents together into a directed graph. We want the agents to decide the sensitivity level of information that flows from producers to consumers and we want them to adjust their contextual-accountability functions so that the system as a whole satisfies accountability-based trust constraints. The protocols described here have been successfully used in the context of our software.

References

[1] K. Aberer and Z. Despotovic. Managing trust in a peer-2-peer information system. In CIKM, pages 310–317, 2001.

[2] Y. Arens, C. Y. Chee, C.-N. Hsu, and C. A. Knoblock. Retrieving and integrating data from multiple information sources. International Journal of Cooperative Information Systems, 2(2):127–158, 1993.

[3] A. Bieszczad, S. K. Raza, B. Pagurek, and T. White. Agent-based schemes for plug-and-play network components. In Proceedings of the 3rd International Workshop on Agents in Telecommunications Applications IATA'98, AgentWorld'98, Paris, France, 1998.

[4] M. M. Cheikhrouhou, P. Conti, and J. Labetoulle. Intelligent agents in network management, a state-of-the-art. Networking and Information Systems, 1(1):9–38, 1998.

[5] Z. Despotovic, K. Aberer, and M. Hauswirth. Trust-aware cooperation.

[6] C. Frei and B. Faltings. A dynamic hierarchy of intelligent agents for network management. In Workshop on Artifi cial Intelligence in Distributed Information Networks (IJCAI'97), 1997.

[7] R. Grimm and B. N. Bershad. Providing policy neutral and transparent access control in extensible systems. Lecture Notes in Computer Science, 1603:311–338, 1999.

[8] B. Khan, D. D. Kleiner, and D. Talmage. CHIME: The Cellular Hierarchy Information Modeling Environment. In Proceedings of International Conference on Parallel and Distributed Computing and Systems 2000, Las Vegas, Nevada, 2000.

[9] B. Khan, D. D. Kleiner, and D. Talmage. Optiprism: A distributed hierarchical network management system for all-optical networks. In Proceedings of IEEE GLOBECOMM, San Antonio, Texas, 2001.

[10] B. Khan, D. D. Kleiner, and D. Talmage. The mother of all databases: A case study in universal situational awareness. In Submitted to IEEE MILCOM, Monterey, California, 2004.

[11] V. R. Lesser. Cooperative multiagent systems: A personal view of the state of the art. Knowledge and Data Engineering, 11(1):133–142, 1999.

[12] G. Pernul, A. M. Tjoa, and W. Winiwarter. Modelling data secrecy and integrity. Data Knowledge Engineering, 26(3):291–308, 1998.

[13] V. Roth. Mutual protection of cooperating agents. Lecture Notes in Computer Science, 1603:275–288, 1999.

[14] V. Swarup and J. T. Fbrega. Trust: Benefi ts, models, and mechanisms. Lecture Notes in Computer Science, 1603:3–18, 1999.

[15] G. Vigna. Protecting mobile agents through tracing. In Third Workshop on Mobile Object Systems, 1997.

[16] J. Vitek and C. J. (Eds.). Secure internet programming - security issues for distributed and mobile objects. Lecture Notes in Computer Science, 1603, 1999.

[17] U. G. Wilhelm, L. Buttyán, and S. Staamann. On the problem of trust in mobile agent systems. In Symposium on Network and Distributed System Security. Internet Society, 1998.