PerAdaMagazine
towards pervasive adaptation
10.2417/2200809.1303

Information sharing is enhanced using trust models

Stephen Marsh

Understanding trust, forgiveness, and regret may enable adaptive reasoning for agents that have to account for the risks of sharing information with others.

In the pervasive environment, information is shared to facilitate adaptation, to ascertain the best solutions to problems, and to better serve human needs. A major consideration in such an environment is privacy—not sharing what shouldn't be shared—as well as use and availability: we must ensure that information is used properly, and not stored for any longer than needed. In an ideal world, it may be possible to build closed systems that can enforce privacy and security 100% of the time. Yet in the real world it is naïve to assume lack of malicious behaviour and adherence to rules. In order to handle these risky situations, humans evolved a sense of trust.

For pervasive environments to be acceptable to people, they must be trusted in a very deep way. But we must know how trust works in order to enable systems to behave properly, share information correctly, and explain decisions. At its root, our model has a formalisation of the phenomena of trust, distrust, regret and forgiveness, allowing an artificial agent to reason with and about the concepts. The agent can make and justify decisions about its situation while accounting for the trust it has in others around it, and vice versa.

A word about risk. A trust-reasoning system is not a trust*worthy* one, in the trustworthy/trusted computing sphere. Trust is, in fact, a deep acknowledgement of the risk inherent in a situation where the truster is putting themselves in the hands of another. In this instance, the risk is that incorrect, potentially harmful information may be shared with others whether we like it or not. This kind of trust is application-agnostic, it applies in many areas, including security, mobile networks, and reputation systems. Since it incorporates an acceptance of risk, trust holds within it the strength to accept and adapt to mistakes, either perceived or actual, and handle malevolence. Remember that even security-based trust*worthy* system, effectively given *blind* trust, can still fail, and can do so in a most dramatic way. Little wonder trustworthiness is seen as one of the grand challenges in information security.[1] Additionally, even if such a truly trustworthy system were possible, it cannot enforce the behaviour of the *people* who work with it, hence the expanded need for trust-*reasoning*.

Over the years, we have developed and refined a generic model[2,3] that focuses on two broad questions: *how much do I trust you?* and *how much do I need to trust you?* Both are applied in a specific context. The questions, and their answers, are largely automatically calculated based on past experience (including other's experience via reputation), current context and user preferences. The model uses many inputs, including classic considerations of utility and risk. These considerations result in values along a continuum. They are also threshold based: if I don't have enough confidence in you, I won't cooperate. Additionally, the model notices emotional states.[4] Finally, trust is inherently adaptive, and can grow with the good results and shrink with the bad.

In the context of information sharing in pervasive technology, the two questions are the same, but the result is the exchange of different quantities and qualities of data. Information is labeled according to importance and privacy, and some can be obfuscated (for example, location could be specific points or general areas). The more I trust you, the more, and better, I'm willing to share.

The beauty of using trust for such technologies is that decisions can be justified to people in an intuitive way. No advanced knowledge, beyond being a social person, is required. Thus, 'I told Alice your address because I trust her *this* much, and you said your address was *this* private', is a valid and easily understood explanation. It's also just as easily corrected (for instance, by making sure your address is more private, or helping your trust agent revise its view of Alice).

In pervasive environments knowledge of both good and bad performers can be shared and acted upon, allowing better trust decision information in the future. Bad behaviour also results in regret,[3] which can ultimately be managed to assure the behaviour doesn't happen again[5]. Finally, reparations, time, and regret can lead to forgiveness[3,6,7] and the restoration of trust. It is this restoration that can allow the society to continue to function in the presence of noise and misunderstanding.

Continued on next page

Trust is a basic human judgment, without which society breaks down. A digital society is no different. In artificial societies, knowledge and application of trust, suspicion, regret and forgiveness can help adaptive technologies better interact with humans and each other. In addition, modelling these concepts allows adaptation and decision making based on ideas people intuitively understand, which should enhance their comfort with artificial societies. They will also allow better information sharing decisions and privacy protection. Our ongoing work is investigating the efficacy of integrating our system into information flow management, amalgamation and assurance in a dynamically networked environment.

Author Information

Stephen Marsh
Information Security Group
National Research Council Canada
Institute for Information Technology
Ottawa, Ontario, Canada
http://www.iit.nrc.gc.ca

References

1. S. Smith and E. Spafford, *Grand challenges in information security: process and output*, **IEEE Security and Privacy 2** (1), pp. 69–71, January–February 2004. doi:10.1109/MSECP.2004.1264859
2. S. Marsh, **Formalising trust as a computational concept**, Department of Computing Science, University of Stirling, 1994. doi:10.1.1.102.8227
3. S. Marsh and P. Briggs, *Examining trust, forgiveness and regret as computational concepts*, in J. Golbeck ed., **Computing with social trust and reputation (to appear)**, Springer Verlag, 2009. ISBN 978-1-84800-355-2
4. S. Marsh, *Optimism and pessimism in trust*, October 1994.
5. S. Etalle, J. den Hartog, and S. Marsh, *Trust and punishment*, **Autonomics '07: Proc. First Int'l Conf. on Autonomic Computing and Communication Systems, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering)**, pp. 1–6, 2007.
6. A. Vasalou and J. Pitt, *Reinventing forgiveness: a formal investigation of moral facilitation*, in P. Herrmann, V. Issarny, and S. Shiu (eds.), **Trust Management: Third Internation Conference, iTrust 2005, Proc., Lecture Notes in Computer Sci. 3477**, Springer Verlag, Berlin Heidelberg, 2005.
7. A. Vasalou, A. Hopfensitzb, and J. Pitt, *In praise of forgiveness: ways for repairing trust breakdowns in one-off online interactions*, **Int'l J. Human-Computer Studies 66**, pp. 466–480, 2008. doi:10.1016/j.ijhcs.2008.02.001