Privacy-intimacy tradeoff in self-disclosure

(Extended Abstract)

J. M. Such, A. Espinosa,
A. Garcia-Fornes
Dep. de Sistemes Informàtics i Computació
Universitat Politècnica de València
Camí de Vera s/n, València, Spain
{jsuch,aespinos,agarcia}@dsic.upv.es

C. Sierra
Institut d'Investigació en Intel·ligència Artificial, IIIA
Spanish Scientific Research Council, CSIC
08193 Bellaterra, Catalonia, Spain
sierra@iiia.csic.es

ABSTRACT

In this paper, we introduce a self-disclosure decision-making mechanism based on information-theoretic measures. This decision-making mechanism uses an intimacy measure between agents and the privacy loss that a particular disclosure may cause.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms

Theory

Keywords

Privacy, Intimacy, Information Theory

1. INTRODUCTION

Westin [5] defined privacy as a "personal adjustment process" in which individuals balance "the desire for privacy with the desire for disclosure and communication". Westin proposed his definition for privacy long before the explosive growth of the Internet. As far as we are concerned, it also applies to autonomous agents that engage in online interactions that require the disclosure of their principals' personal data attributes (PDAs). Agents, then, need to incorporate self-disclosure decision-making mechanisms allowing them to autonomously decide whether disclosing PDAs to other agents is acceptable or not.

Current self-disclosure decision-making mechanisms are usually based on a privacy-utility tradeoff ([2]). This tradeoff considers the direct benefit of disclosing a PDA and the privacy loss it may cause. There are many cases where the direct benefit of disclosing PDAs is not known in advance. This is the case in human relationships, where the disclosure of PDAs in fact plays a crucial role in the building of these relationships [1]. In such environments, the privacy-utility tradeoff is not appropriate and other more *social* approaches are needed. We present a self-disclosure decision-making mechanism based on what we call the privacy-intimacy tradeoff. This tradeoff considers the increase in intimacy to another agent rather than considering a direct benefit when disclosing a PDA.

Cite as: Privacy-intimacy tradeoff in self-disclosure (Extended Abstract), J. M. Such, A. Espinosa, A. Garcia-Fornes and C. Sierra, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1189-1190.

2. UNCERTAIN AGENT IDENTITIES

We assume a Multiagent System composed of a set of intelligent autonomous agents $Ag = \{\alpha_1, \ldots, \alpha_M\}$ that interact with one another through message exchanges. Agents in Ag are described using the same finite set of PDAs, $A = \{a_1, \ldots, a_N\}$. Each PDA $a \in A$ has a finite domain of possible values $V_a = \{v_1, \ldots, v_{K_a}\}$.

Definition 1. Given a set of PDAs $A = \{a_1, \ldots, a_N\}$, each one with domain $V_a = \{v_1, \ldots, v_{K_a}\}$, an uncertain agent identity (UAI), $I = \{P_1, \ldots, P_N\}$ is a set of discrete probability distributions P_i over the values V_{a_i} of each PDA a_i.

We thus denote P_a as the probability distribution of a over V_a and $p_a(\cdot)$ as its probability mass function, so that $p_a(v)$ is the probability for the value of a being equal to $v \in V_a$.

An agent $\alpha \in Ag$ manages its own UAI and two UAIs associated to each agent $\beta \in Ag \setminus \{\alpha\}$. We will refer to the UAI of an agent α as I_α. We denote $I_{\alpha,\beta}$ as the UAI that α believes that β has, i.e., what α knows (or thinks it knows) about I_β. Finally, we denote $I_{\alpha,\beta,\alpha}$ as the UAI that α believes that β believes that α has. This UAI is crucial for an agent α to model what agent β may know about its own UAI I_α for measuring privacy loss.

2.1 Uncertainty Measures

An agent needs to measure how much uncertainty there is in the probability distribution of a PDA. Taking into account this uncertainty, the agent may decide, for instance, whether to take specific actions to reduce this uncertainty under a desired threshold or not. A well-known measure of the uncertainty in a probability distribution is Shannon entropy:

$$H(P_a) \quad = \quad -\sum_{v \in V_a} p_a(v) \log_2 p_a(v) \qquad (1)$$

A method for aggregating the uncertainties of all of the probability distributions in an UAI is needed. In this paper, we use a simple computational method that is the mean of the uncertainties in each of the probability distributions in an UAI:

$$H(I) = \frac{1}{|A|}\sum_{a\in A} H(P_a) \quad (2)$$

With this measure an agent is able to know how certain it is about an UAI. We assume that at initialization time the entropy of an UAI I is the highest possible, i.e., the uncertainty in I will decrease as the agent obtains more information related to the PDAs being modeled.

2.2 Updating UAIs

UAIs are supposed to be dynamic, i.e., they may change as time goes by. These changes will potentially reduce the uncertainty in an UAI. An agent α may update the UAIs that it manages as it gets more information about the probability distributions for the PDAs in these UAIs. PDA values are private to each agent. We assume that α *discloses* its PDA values for a to β by sending a message$\mu = \langle\alpha, \beta, \langle\alpha, a, P_a\rangle\rangle$, where α represents the sender, β represents the receiver, and $\langle\alpha, a, P_a\rangle$ represents the claim "the probability distribution for the PDA a of α is P_a".

UAIs are updated with the disclosures that agents carry out. The update process of an UAI has two steps: (i) updating the probability distribution of the PDA being disclosed; and (ii) inferring updates of probability distributions of other PDAs based on the PDA being disclosed and other information already known. We denote that an UAI I is updated with a message μ as I^{μ}. Moreover, we denote that an UAI I is updated sequentially and in order considering a tuple of messages $M = (\mu_1, \ldots, \mu_P)$ as I^M.

Details about the updating process are obviated due to space restrictions.

3. INTIMACY

According to [3], intimate human partners have extensive personal information about each other. They usually share information about their PDAs, including preferences, feelings, and desires that they do not reveal to most of the other people they know. Indeed, self-disclosure and partner disclosure of PDAs play an important role in the development of intimacy[1].

Definition 2. Given an UAI I and a message μ, the information gain of message μ is:

$$\mathcal{I}(I, \mu) = H(I) - H(I^{\mu}) \quad (3)$$

Definition 3. Given an UAI I and a tuple of messages M, the information gain of M is:

$$\mathcal{I}(I, M) = H(I) - H(I^{M}) \quad (4)$$

Sierra and Debenham [4] defined the intimacy between α and β considering the amount of information that α knows about β and vice versa. We adapt this definition for the case of UAIs. Thus, we define intimacy as follows.

Definition 4. Given the UAIs $I_{\alpha,\beta}$ and $I_{\alpha,\beta,\alpha}$, a tuple of messages M from β to α and a tuple of messages M' from α to β, the intimacy between α and β is:

$$\mathcal{Y}_{\alpha,\beta} = \mathcal{I}(I_{\alpha,\beta}, M) \oplus \mathcal{I}(I_{\alpha,\beta,\alpha}, M') \quad (5)$$

Where $\oplus$ is an appropriate aggregation function.

4. PRIVACY LOSS

Disclosing PDAs always comes at a loss of privacy because personal information is made known. Therefore, it is crucial for agents to estimate the privacy loss that a disclosure may imply before deciding whether they actually carry it out.

Agent α may estimate (from its point of view) the extent to which β knows I_α by measuring the distance between I_α and $I_{\alpha,\beta,\alpha}$. Agent α can calculate this distance by measuring the Kullback-Leibler divergence between each probability distribution for each PDA in these UAIs.

Definition 5. Given two agents α and β, the message μ, and considering $Q_a \in I_{\alpha,\beta,\alpha}$, $Q_a^{\mu} \in I^{\mu}_{\alpha,\beta,\alpha}$ and $P_a \in I_\alpha$, the privacy loss for agent α if it sends μ to agent β is:

$$\mathcal{L}(I_{\alpha,\beta,\alpha}, \mu) = \sum_{a\in A} w_\alpha(a) \cdot (\mathrm{KL}(Q_a \parallel P_a) - \mathrm{KL}(Q_a^{\mu} \parallel P_a)) \quad (6)$$

$\mathrm{KL}(\cdot)$ is the Kullback-Leibler divergence. $w_\alpha(\cdot)$ is the sensitivity function for agent α that is defined as $w_\alpha : A \rightarrow [0, 1]$, such that $w_\alpha(a)$ is the *subjective* valuation that α attaches to the sensitivity for disclosing a.

5. DECISION MAKING

We consider the estimation of intimacy gain between two agents and the privacy loss. To estimate the increase in intimacy that the sending of a message μ may cause between α and β, we consider the information gain of μ, i.e. $\mathcal{I}(I_{\alpha,\beta,\alpha}, \mu)$. We consider that $\mathcal{I}(I_{\alpha,\beta,\alpha}, \mu)$ also acts as an estimation for $\mathcal{I}(I_{\alpha,\beta}, \nu)$, considering ν as a future message received by α from β as the reciprocation to μ. Then, α estimates that after sending μ to β and receiving ν from β, $\mathcal{Y}_{\alpha,\beta} \approx \mathcal{I}(I_{\alpha,\beta,\alpha}, \mu) \oplus \mathcal{I}(I_{\alpha,\beta,\alpha}, \mu)$. This assumption is grounded on the *disclosure reciprocity* phenomenon [1].

Disclosing PDAs always comes at a privacy loss. Then, α may choose to disclose a PDA that maximizes the estimation of the increase in intimacy while at the same time minimizing the privacy loss.

6. ACKNOWLEDGMENTS

This work has been partially supported by CONSOLIDER-INGENIO 2010 under grant CSD2007-00022, and projects TIN2008-04446, TIN2009-13839-C03-01, TIN2010-16306 and AGAUR (Generalitat de Catalunya) 2009-SGR-1434.

7. REFERENCES

[1] K. Green, V. J. Derlega, and A. Mathews. *The Cambridge Handbook of Personal Relationships*, chapter Self-Disclosure in Personal Relationships, pages 409–427. Cambridge University Press, 2006.

[2] A. Krause and E. Horvitz. A utility-theoretic approach to privacy and personalization. In *AAAI'08: Proceedings of the 23rd national conference on Artificial intelligence*, pages 1181–1188, 2008.

[3] R. Miller, D. Perlman, and S. Brehm. *Intimate relationships*. McGraw-Hill Higher Education, 2007.

[4] C. Sierra and J. Debenham. The LOGIC negotiation model. In *AAMAS '07: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–8, 2007.

[5] A. Westin. *Privacy and Freedom*. New York Atheneum, 1967.